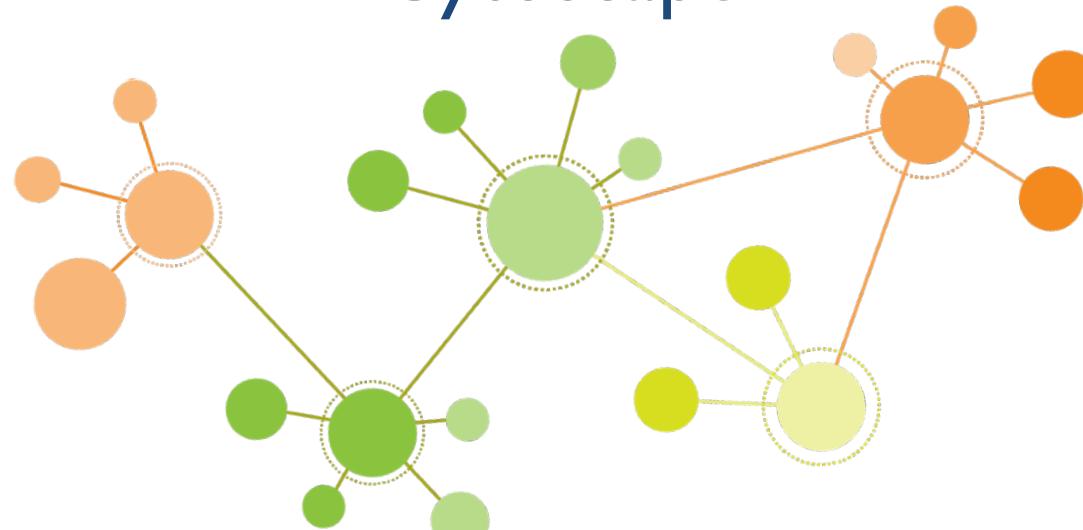


Introduction to Data Visualization

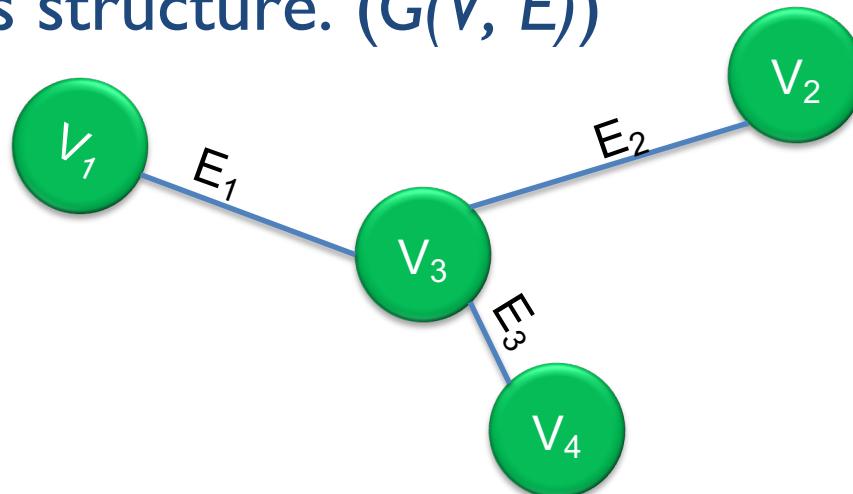
Network Visualization Cytoscape



Halil Bisgin, Ph.D.

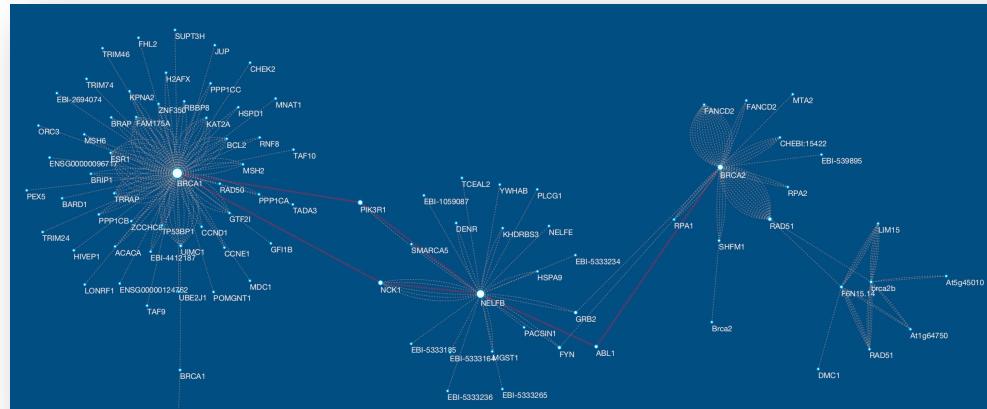
What is a network?

- A **network** consists of nodes (vertices) which are linked to each through well-defined relationships.
 - friend – friend (*social networks*)
 - gene – gene (*gene interaction network*)
 - protein – protein (*protein interaction network*)
- **Graph theory** provides a formal definition for networks structure. ($G(V, E)$)

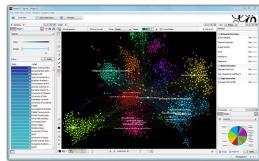


Network Visualization

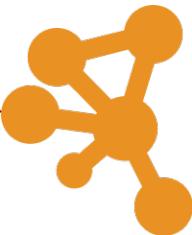
- Network analysis: we can perform a series of analyses on a network (graph) through several algorithms.
- However, visualizing a network offers more.
- We can use specialized tools, to visualize the relationships between nodes and add more features for a more informative image.



Network Visualization Softwares- I



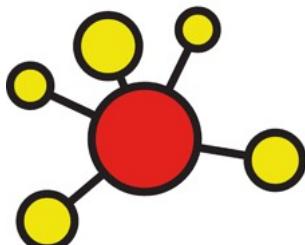
– **Gephi** is the leading visualization and exploration software for all kinds of graphs and networks.



– **Cytoscape** is an open source software platform for visualizing complex networks and integrating these with any type of attribute data.



– **Graphviz** is open source graph visualization software with web and interactive graphical interfaces, and auxiliary tools, libraries, and language bindings.

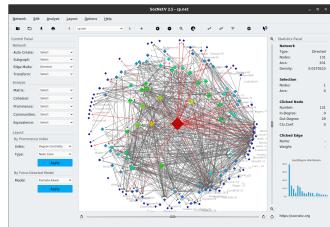


– **igraph**: a collection of network analysis tools with connectors in R, Python, Mathematica and C/C++.

Network Visualization Softwares-2



— *UCINET 6 for Windows* is a software package for the analysis of social network data. UCINET comes with the NetDraw network visualization tool.



— *The Social Network Analysis and Visualization Software (SocNetV)*.

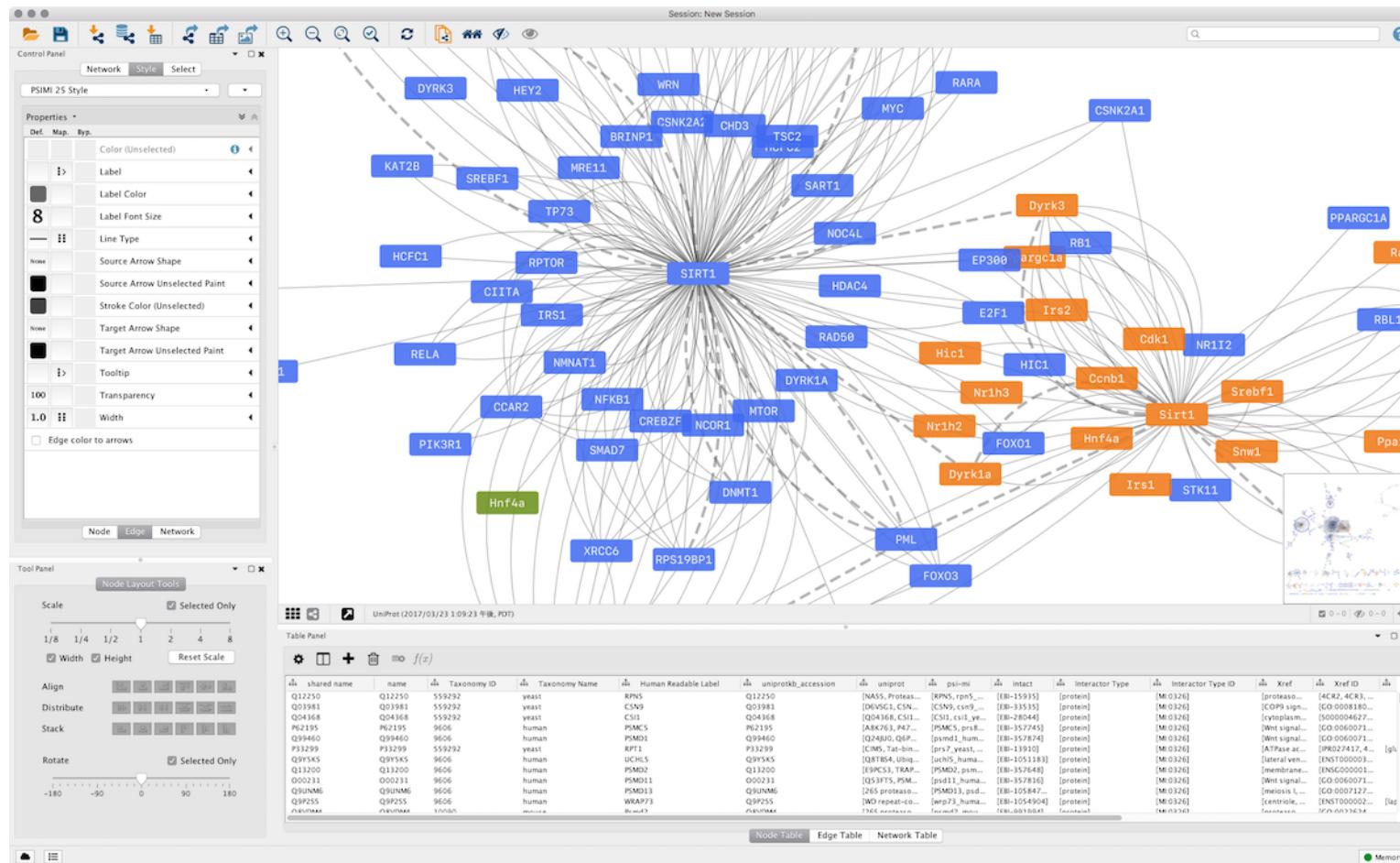


— *Pajek Analysis and visualization of very large networks on Windows*



- **Cytoscape** is an open source software platform for **visualizing** molecular interaction networks and biological pathways and **integrating** these networks with annotations, gene expression profiles and other state data.
- Although Cytoscape was originally designed for biological research, now it is a general platform for complex network analysis and visualization. .

Cytoscape Desktop



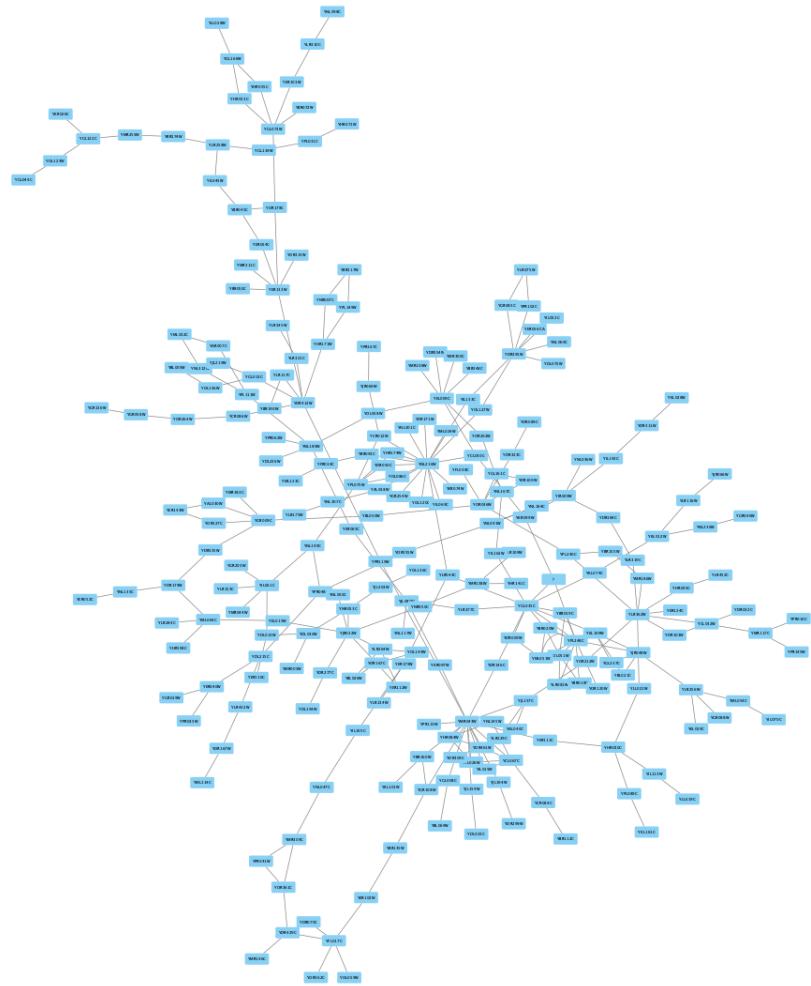
Cytoscape Tutorial

- The following content can also be found at
<https://github.com/cytoscape/cytoscape-tutorials/wiki>



Loading Network

- To get started, install and launch the latest version of **Cytoscape**.
 - We will use NDEx to find a relevant network. In the **Network Search** interface in the **Control Panel**, select NDEx from the drop-down, and type in "GAL1 GAL4 GAL80".
 - In the search results, find the **galFiltered network with data**. Click the **Import network to Cytoscape** (green arrow) button to the left of the network name.
 - The network will open with the default style, similar to the network on the right:



To learn more about **importing networks from local files**, see the [Loading Networks Tutorial](#).

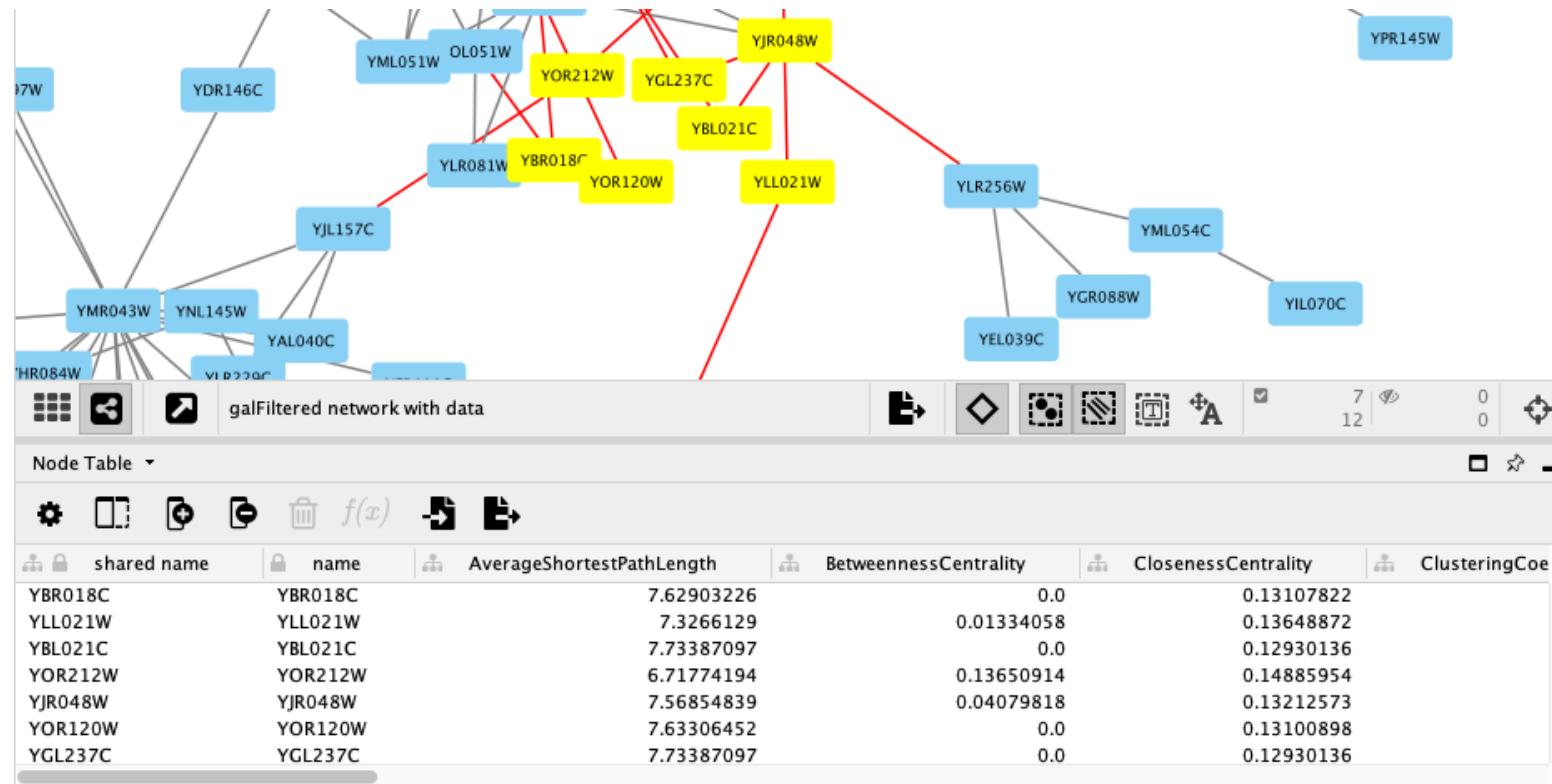
Visualizing Expression Data on Networks

Probably the most common use of expression data in Cytoscape is to set the **visual properties** of the nodes (color, shape, border) 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 show an example of doing this.

The data used in this example is from yeast, and represents an experiment of perturbations of the genes **Gal1**, **Gal4**, and **Gal80**, which are all yeast transcription factors.

Visualizing Expression Data on Networks

For this tutorial, the experimental data was part of the Cytoscape network file you loaded earlier, and is visible in the Node Table:



- Selecting nodes in the network (Shift + Click or Click-and-Drag) will update the **Node Table** to show only data for those nodes.
- Similarly, you can select one or more rows in the **Node Table**, right-click on the selected rows and click **Select nodes from selected rows** to highlight the corresponding nodes in the network.

To learn more about **importing data**, see the [Importing Data From Tables](#).

Visualizing Expression Data on Networks

We can now use the data to manipulate the visual properties of the network by mapping specific data columns to visual style properties:

- The **gal80Rexp** expression values will be mapped to **Node Fill Color**; nodes with low expression will be colored blue, nodes with high expression will be colored red.
- Significance for expression values will be mapped to **Node Border Width**, so nodes with significant changes will appear with a thicker border.

Set Node Fill Color

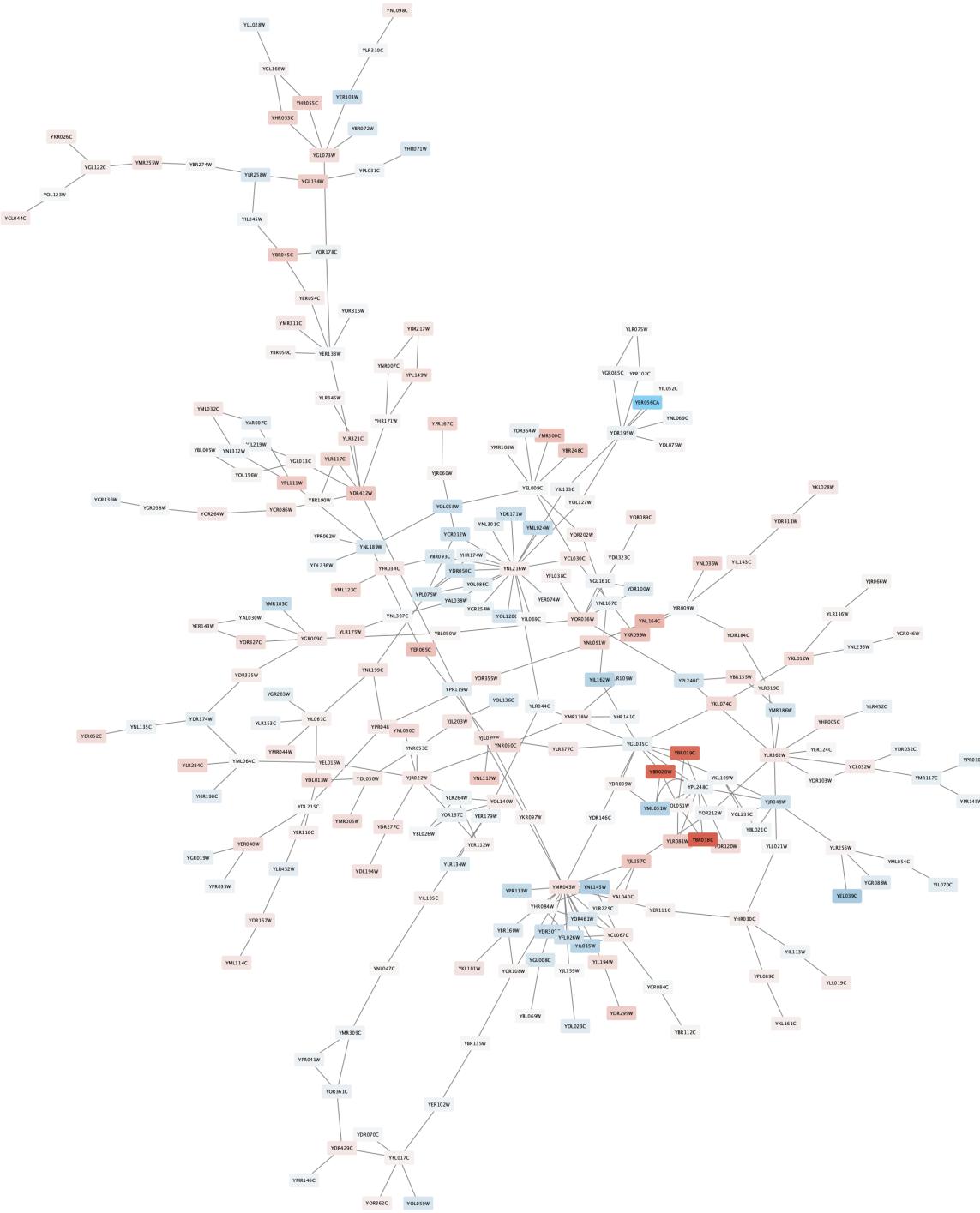
- Click on the **Style** tab in the **Control Panel**.
- Find **Fill Color** in the list of **Node** properties and expand it to view the mapping options.
- Click the **-- select value --** cell in the **Column** section, and select **gal80Rexp** in the drop-down.
- Click the **-- select value--** cell in the **Mapping Type** section, and select **Continuous Mapping** in the drop-down.
- This produces a default **ColorBrewer** gradient ranging from blue to red for expression values. For our purposes, this gradient works so we don't need to edit it.



To learn more about **changing node color**, see the [Visualizing Expression Data Tutorial](#).

Set Node Fill Color

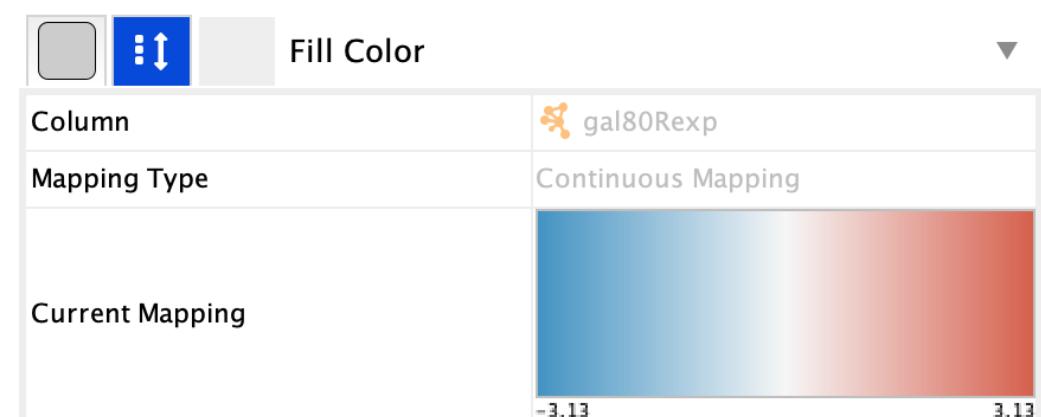
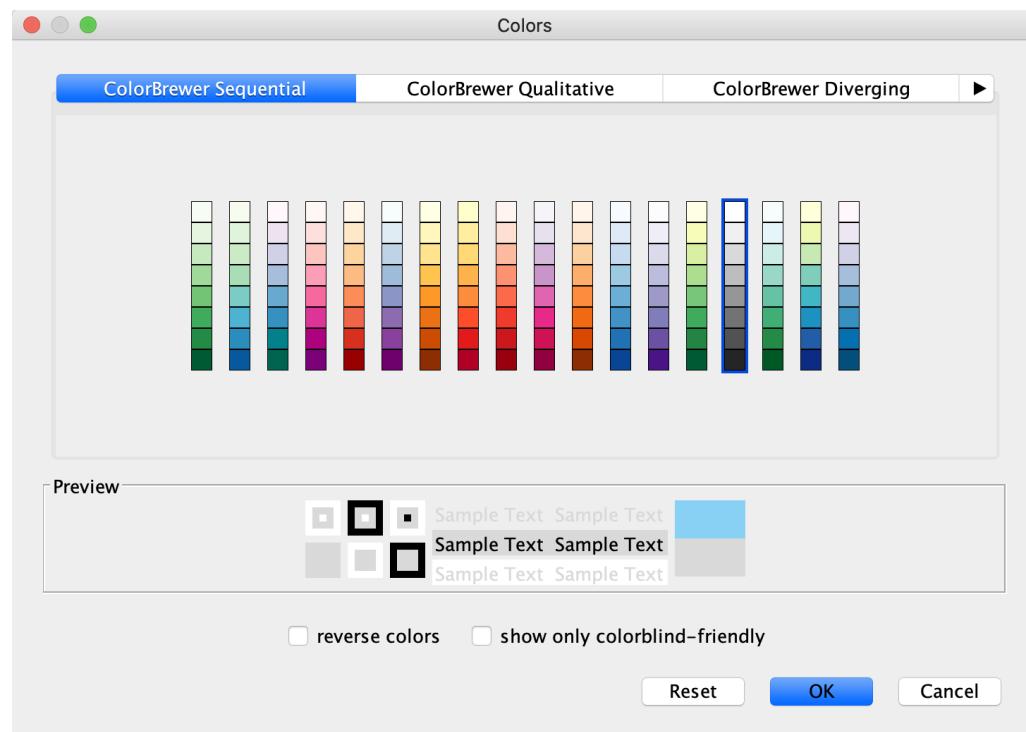
The nodes in the network are now colored based on the gal80RExp data column:



Set Default Node Color

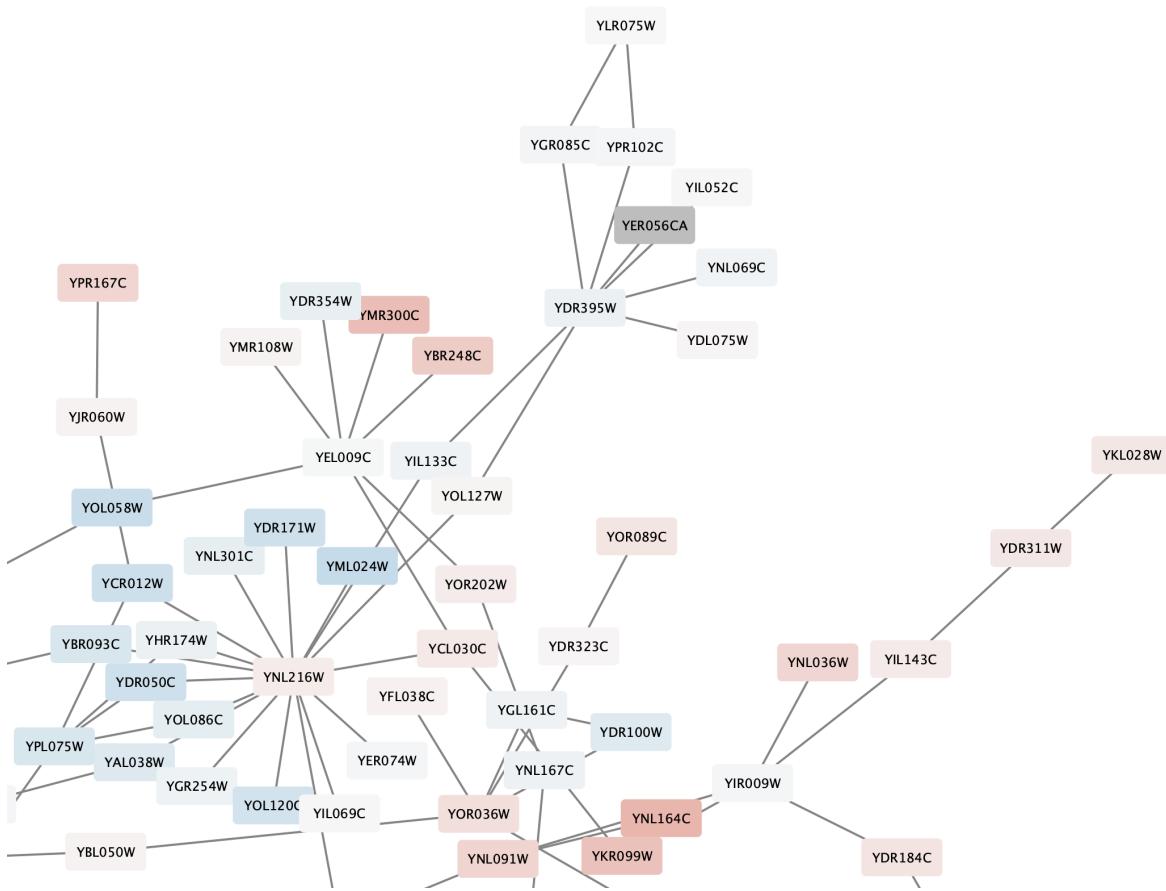
Some nodes in the network don't have any data, and for those nodes, the default color applies. In our case, the default color is blue, which falls within the spectrum of our blue-red gradient. This is not ideal for data visualization, so a useful trick is to choose a color outside the gradient spectrum to distinguish nodes with no defined expression value.

- Still in the **Style** tab, under **Node Fill Color**, click the **Def.** (leftmost) cell next to **Fill Color**.
- In the **Colors** interface, you can click on any single color tile in a color palette to choose it. In this case, choose a light gray color.



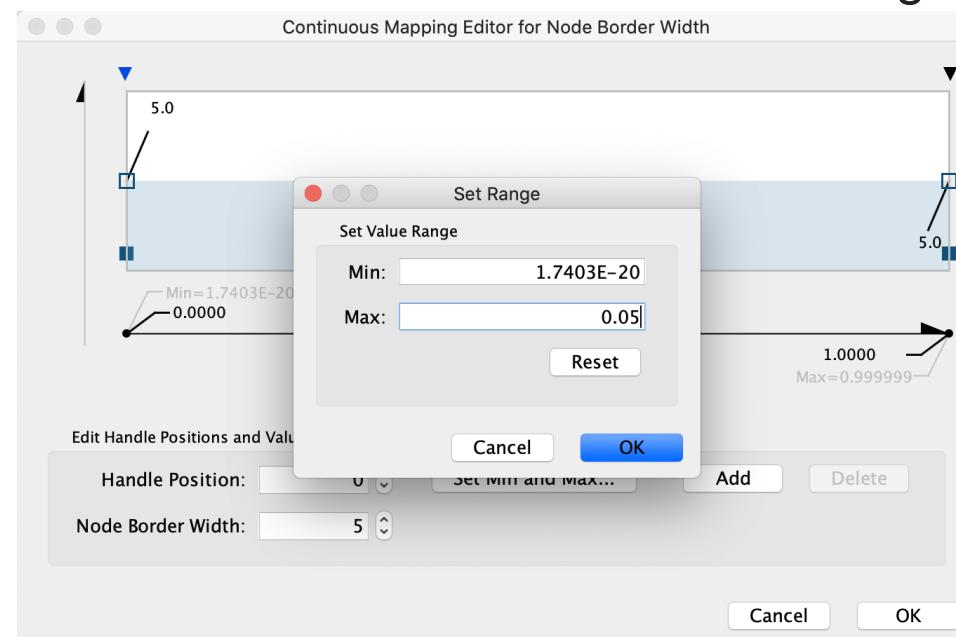
Set Default Node Color

Note that at least one node in the network is now colored grey:



Set Node Border Width

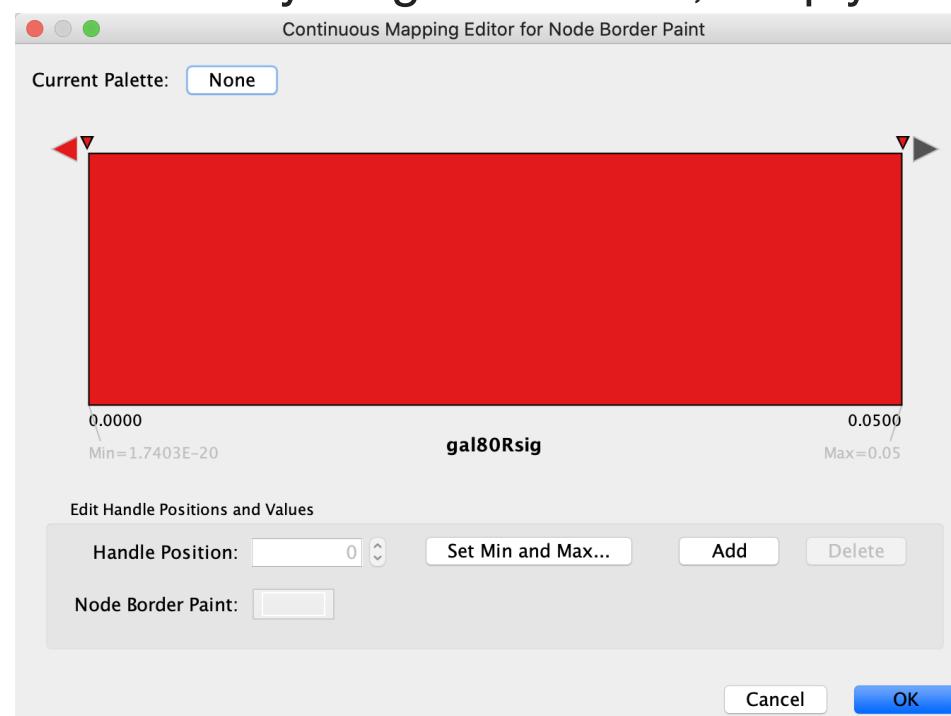
- Click on the **Map.** cell for the **Node Border Width** property in the **Style** panel.
- Click the **-- select value --** cell in the **Column** section, and select **gal80Rsig** in the drop-down.
- Click the **-- select value--** cell in the **Mapping Type** section, and select **Continuous Mapping** in the drop-down.
- Double-click on the gradient, which defines the node border width over the range of p values.
- First, we will define the min/max of the range of p values we are interested in, that is anything under 0.05. Click on **Set Min and Max...** and set the max to **0.05**. Click **OK** to save.
- Click on the handle for the max value (black triangle on the right-most end of the gradient) and set the value to **5** in the **Node Border Width** field. Click **OK** to update the mapping.
- Click the **Def.** value for **Node Border Paint** and select a dark grey color.



To learn more about **visualizing data**, see the [Visualizing Expression Data Tutorial](#).

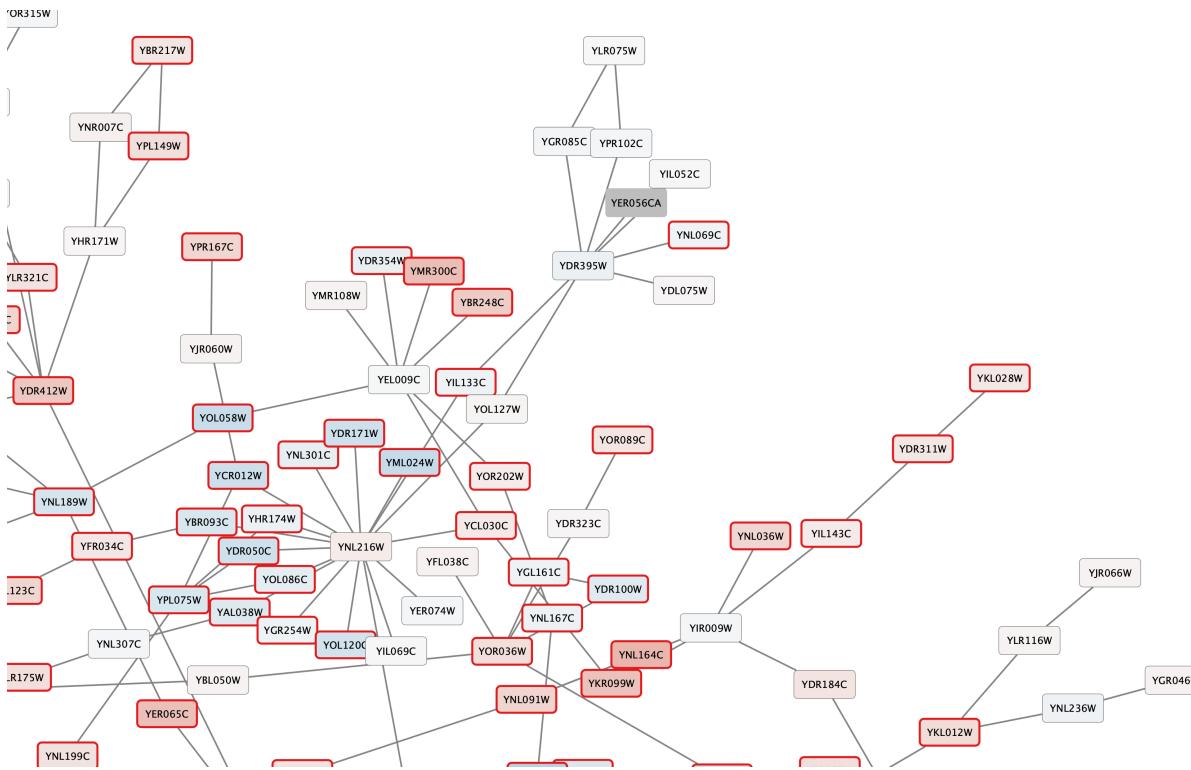
Set Node Border Color

- Click on the **Map.** cell for the **Node Border Paint** property in the **Style** panel.
- Click the **-- select value --** cell in the **Column** section, and select **gal80Rsig** in the drop-down.
- Click the **-- select value--** cell in the **Mapping Type** section, and select **Continuous Mapping** in the drop-down.
- Double-click on the gradient.
- Similar to the last step, we will define the min/max of the range of p values we are interested in, that is anything under 0.05. Click on **Set Min and Max...** and set the max to **0.05**. Click **OK** to save.
- Click on the large triangle handle at the right-most end of the gradient. This defines the border color for nodes with **gal80Rsig** > 0.05, so let's set it to dark grey.
- Next, set all the other handles to the same dark red color. We can delete the middle handle since we want the same color for anything under 0.05, simply select it and click **Delete**.



Set Node Border Width

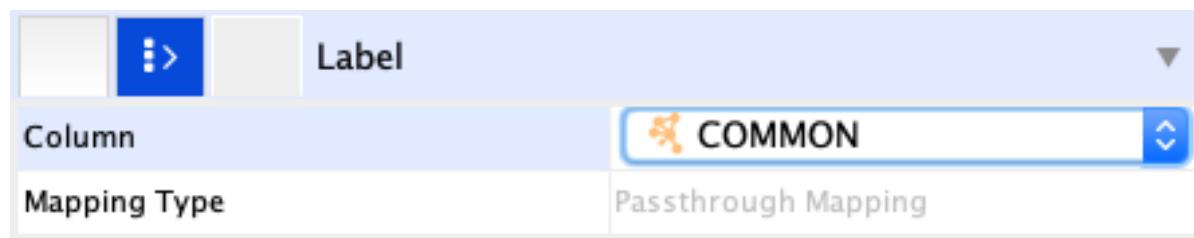
All nodes with a **gal80RSig** p-value of 0.05 or less will now have a thicker red border:



Set Node Label

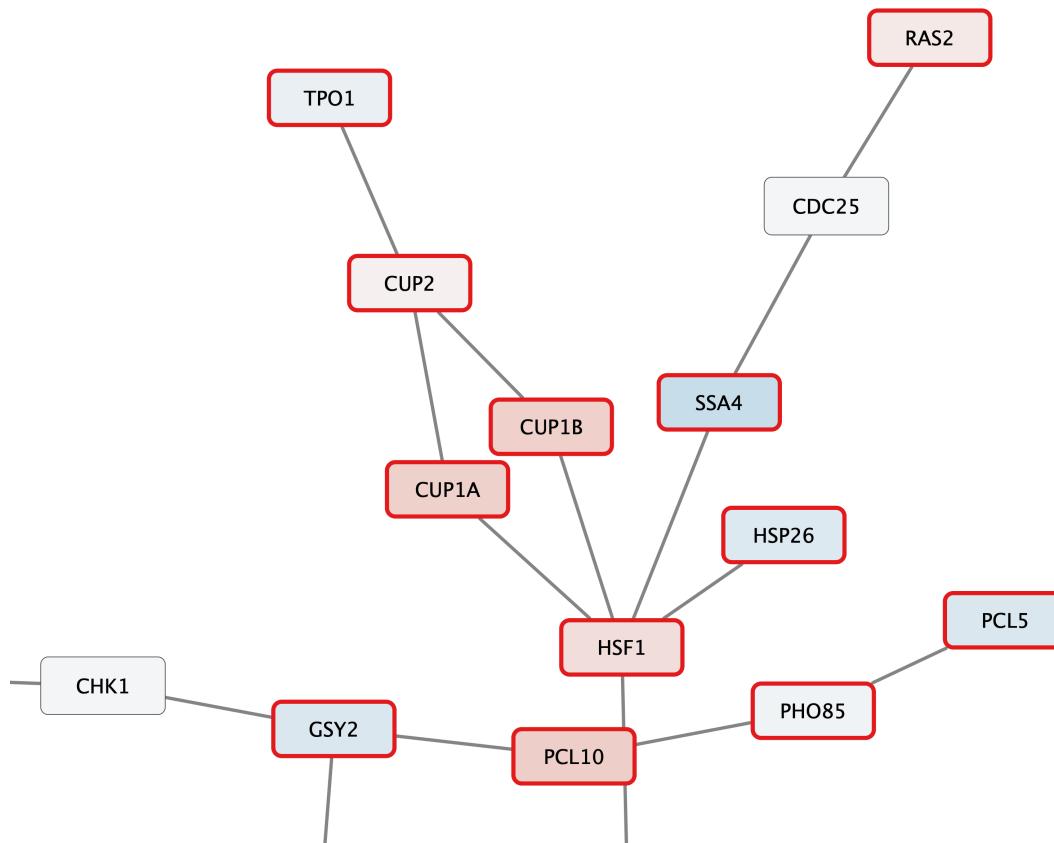
The network nodes are by default labeled with yeast ORF identifiers. We can change the label to something more readable, like the gene symbol.

- For the **Node Label**, change the **Column** selection for the mapping from **name** to **COMMON**.
- Click the **-- select value --** cell in the **Mapping Type** section, and select **Passthrough Mapping** in the drop-down.



Set Node Label

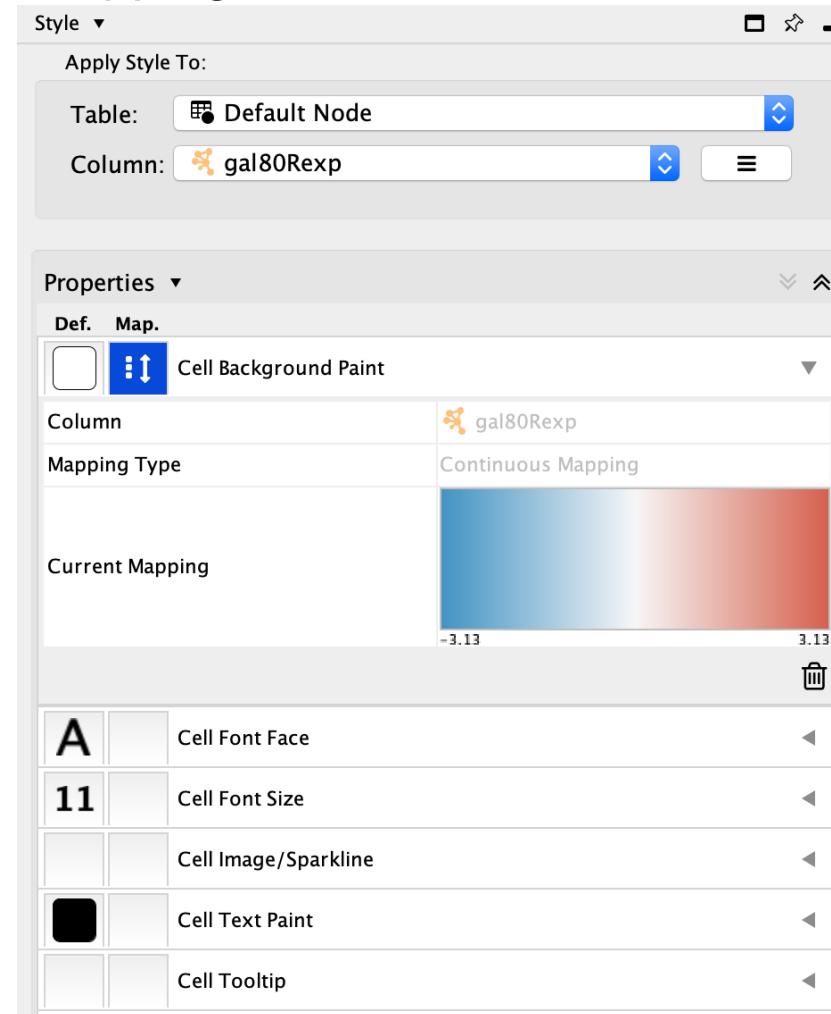
Zoom in to see the updated labels:



Set the Node Table Style

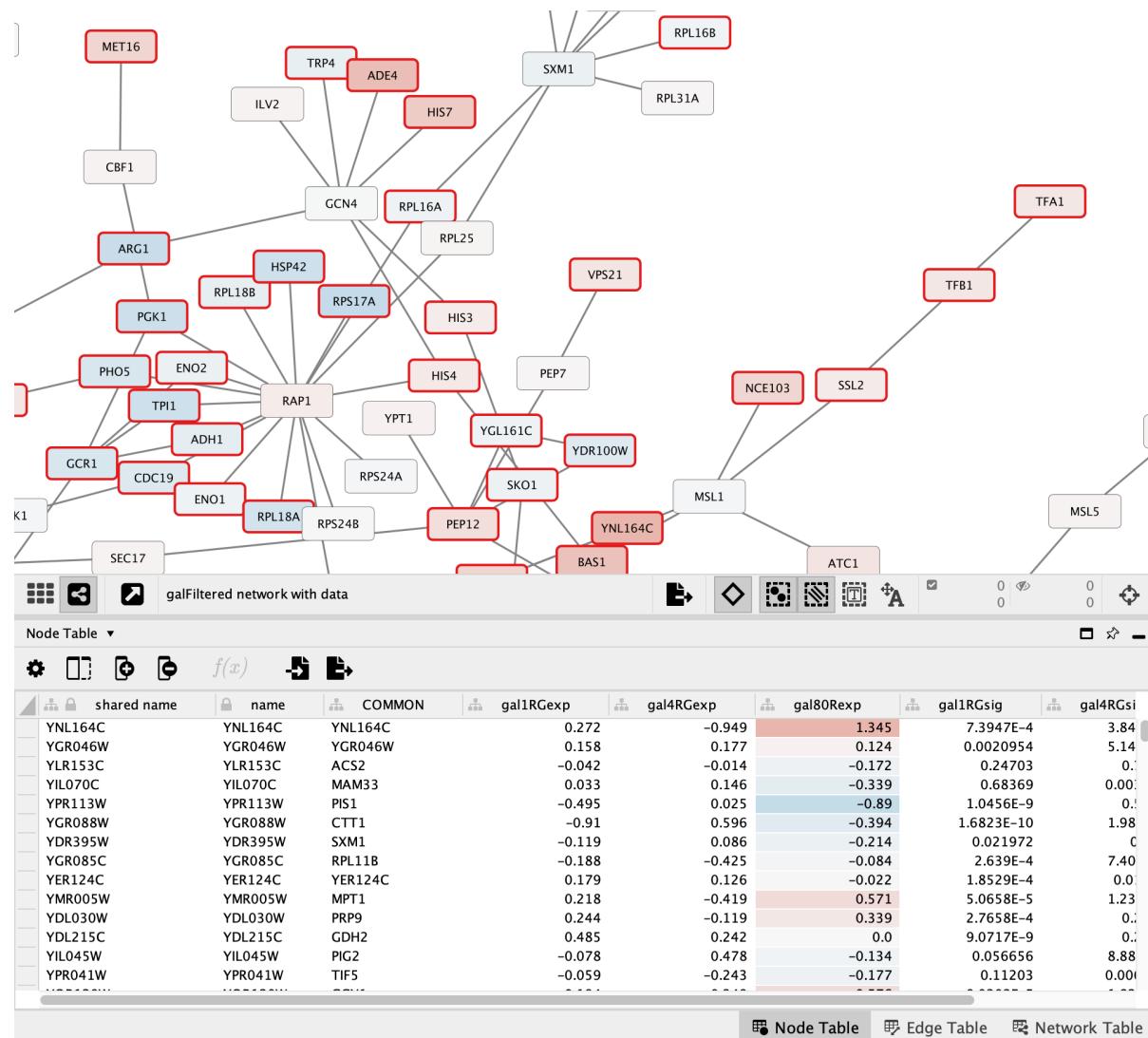
In addition to creating a visualization for the network, it is also possible to add a style to the **Node Table**. The **Table** tab of the **Style** interface offers default and mapping style options for a set of table cell attributes.

- At the top of the **Table** tab of the **Style** panel, select **Default Node** at the top.
- Next, select the data column you want to add the style to under the **Column** drop-down, in this case the **gal80Rexp** column.
- In the mapping column of **Cell Background Paint** select the **gal80Rexp** column and select a default red-blue continuous mapping.



Set the Node Table Style

The background of the cells in the **gal80RExp** column is now colored in the same way that the nodes in the network are colored:

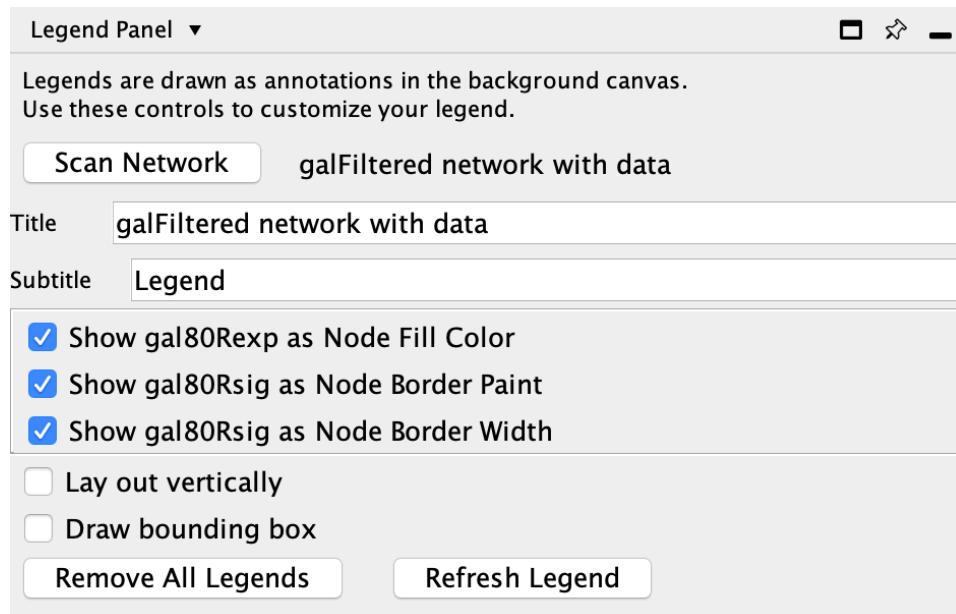


Creating the Legend

The Cytoscape app **Legend Creator** allows you to create a customized legend for your visualization.

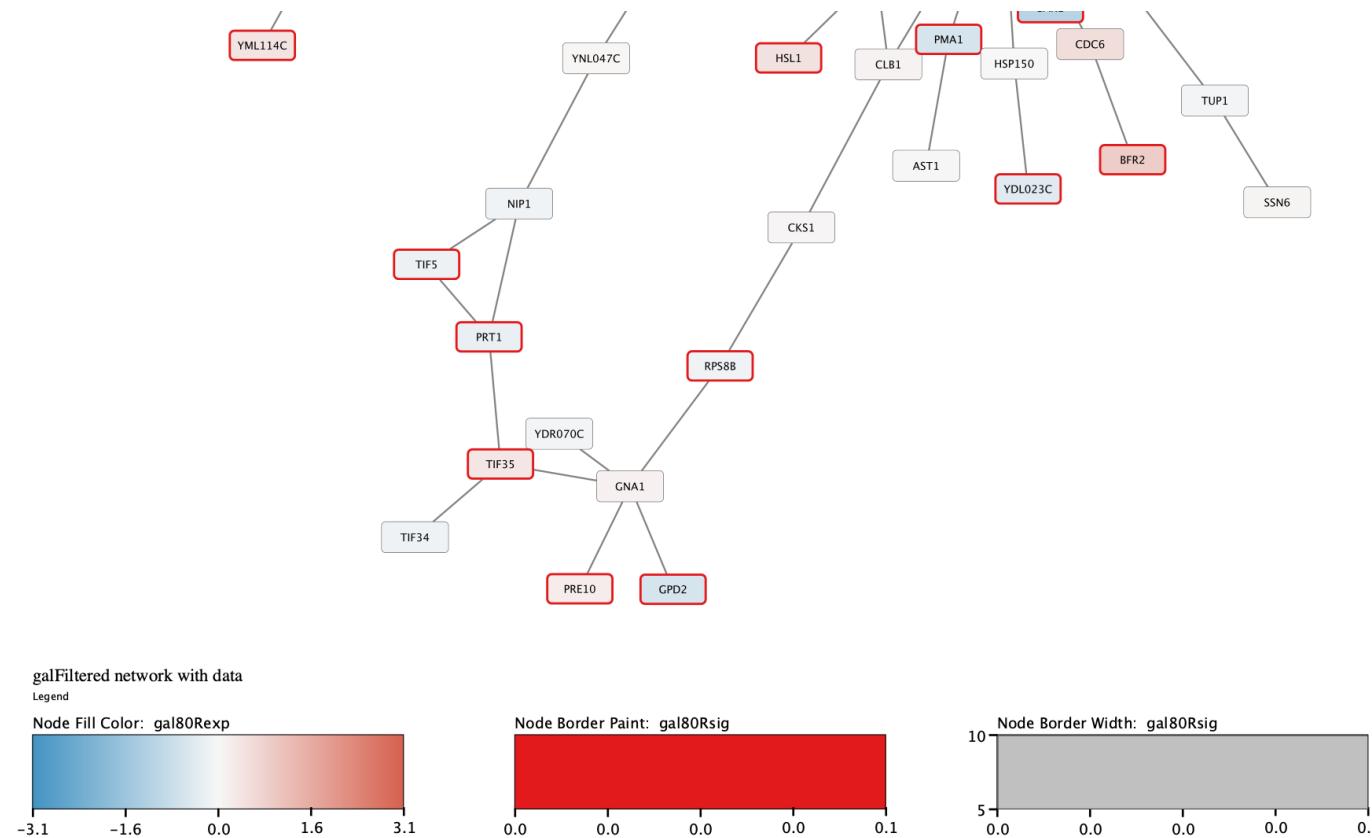
You can install Legend Creator from the [App Store](#) or via the [Cytoscape App Manager](#).

- In the **Control Panel**, go to the **Legend Panel**. This is the interface for the **Legend Creator** app, and will list multiple options for creating a legend.
- Note that the two attributes that we used for mapping, **gal80Rexp** and **gal80Rsig**, are selected by default. For our purposes this is actually what we want. Click **Refresh Legend** to continue. The legend will appear in the lower left of the network view.



Moving a Legend

- To move the legend, first click the **Toggle Annotation Selection** icon  at the bottom of the **Network View Window**. You can now move the legend like you would any object, for example by click and drag.



To learn more about **creating legends**, see the [Legend Creator](#).

Layouts

An important aspect of network visualization is the layout, meaning the positioning of nodes and edges. Our network had a preset layout in the original file you imported, but this can be changed.

- Let's change the layout to **Degree Sorted Circle Layout** by selecting **Layout → Degree Sorted Circle Layout**. In this layout, nodes are sorted by degree (connectedness), with the highest degree node at the 6 o'clock position, and remaining nodes are sorted counter clockwise based on decreasing degree.

For this network, a degree-sorted circle layout may not be the most effective. Instead, let's try a force-directed layout instead.

- By default, **Prefuse Force-directed Layout** is set as the **Preferred Layout** in Cytoscape, so an easy way to apply the **Prefuse Force-directed Layout** is by clicking the **Preferred Layout**  in the toolbar. The preferred layout algorithm can be changed under **Layout → Settings...**

Cytoscape supports many different layout algorithms, described in detail in the [Cytoscape manual](#).

Filtering Nodes

Cytoscape allows you to easily filter and select nodes and edges based on data attributes. Next, we will select a subset of nodes with high expression in the gal80 knockout:

- Go to the **Filter** tab in the **Control Panel**.
- Click on the **+** button and select **Column Filter**.
- Under **Choose column...**, select **Node: gal80Rexp**. Notice how all the nodes in the network are selected at first.
- Using the slider or the input fields, specify values of **2** or higher by changing the minimum value to 2.

You should now see only a few nodes in the network selected, highlighted yellow.

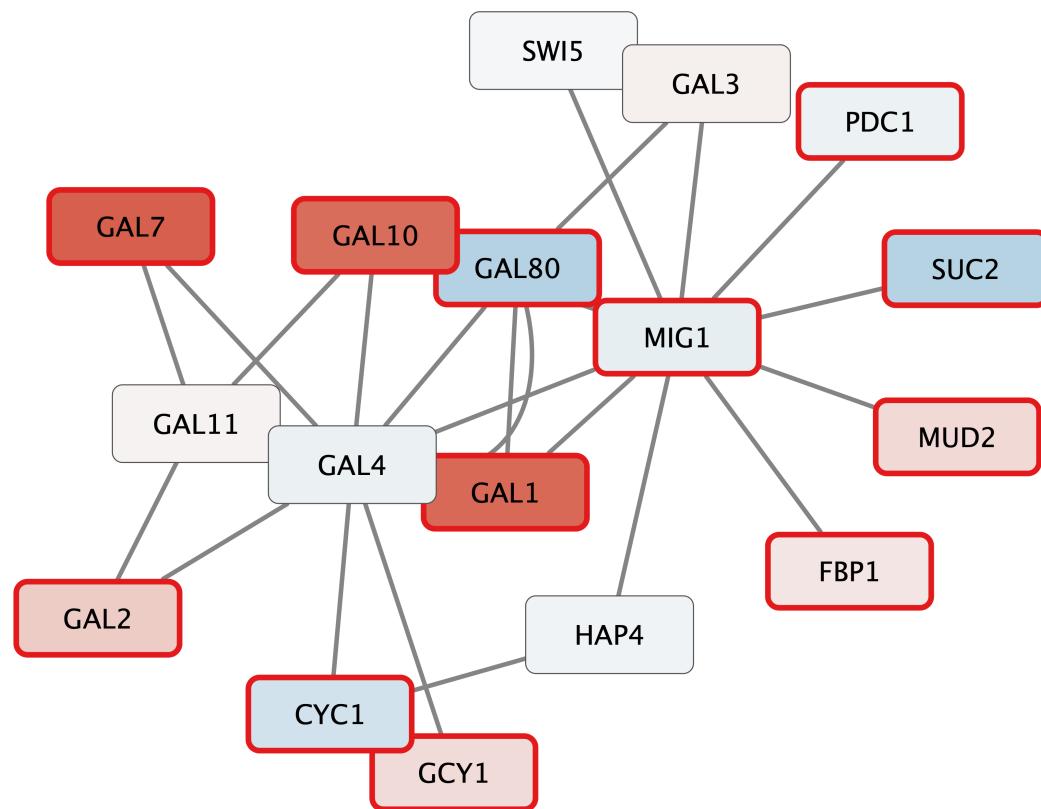
To learn more about **filtering and selection**, see the [Filtering by Selection Tutorial](#).

Expand Selection and Create New Network

We have now selected only the few top expressing nodes. To see the context of these nodes in the larger network, we can expand the selection of nodes to include the nodes connecting to the selected nodes, i.e. the first neighbors. Once we have that larger selection, we can create a new network.

- Select the first neighbors of selected nodes by clicking the **First Neighbors of Selected Nodes → Undirected** button  in the toolbar.
- To expand the selection to second-degree nodes, repeat the last step of clicking the **First Neighbors of Selected Nodes → Undirected** button  again.
- Create a new network from the current selection by clicking the **New Network From Selection, All edges** button  in the toolbar.
- Click on the **Preferred Layout** button  in the toolbar to apply a force-directed layout.

Your new network will look similar to this:



Saving Results

Cytoscape provides a number of ways to save results and visualizations:

- As a session: **File → Save, File → Save As...**
- As an image: **File → Export → Network to Image...**
- To the web: **File → Export → Network to Web Page...** ([Example](#))
- To a public repository: **File → Export → Network to NDEx**, or **File → Export → Collection to NDEx**
- As a graph format file: **File → Export → Network to File.**

Formats:

- CX JSON
- Cytoscape.js JSON
- GraphML
- PSI-MI
- XGMML
- SIF