

Introduction to Pathway Analysis and Visualization

Additional exercise instructions


In this exercise, we will use the results from ORA analysis to select a GO term, and create a network from that GO term in Cytoscape using the [String app](#), from the STRING database of protein-protein interactions. Finally, we will visualize our expression data on the network.

Data visualization in Cytoscape

Getting the list of genes for a GO term

1. Open the data table results from the ORA analysis, corresponding to GO terms. Find the top GO term.
2. In the **Genes** column and copy the contents for the top term.

Cytoscape

1. Launch Cytoscape. Using the App Manager, install the STRING app.
2. In the Network panel, select **STRING protein query** in the drop-down and paste the list of genes. Click the **Options** icon  and set the confidence score to 0.40 and the maximum number of interactors to 0. Click the search icon to search.
3. You may see a **Resolve Ambiguous Terms** interface, which identifies symbols in your input who have more than one potential match in the STRING database. You can resolve the issues by choosing which to import one or all of the matches.
4. When the network loads, it will have the STRING style applied. To remove this and allow for better data visualization, go to the **STRING Results Panel** to the right of the network. Deselect the following options:
 - a. Glass ball effect
 - b. String colors
5. Before we can visualize the data on the network, the network has to be annotated with Ensembl IDs (which is what the data is annotated with). In the **Node Table** (below the network), right-click on the **stringdb – canonical name** column header, and select **Map column...**
6. In the mapping interface, select **Uniprot** as the **Map from** column and **Ensembl** as the **To** column. Click **OK** to proceed. You should get a popup informing you of how many of the identifiers were mapped.
7. Now we can import the lung cancer dataset like in the previous tutorial and color the nodes by Log2FC. Go to **File > Import table from File** and select the “lung.expr.csv” data file.

8. Under **Where to Import Table Data**, select **To a Network Collection** and select the relevant network collection in the second drop-down.
9. In the **Key Column for Network** drop-down, select the new **Ensembl** column.
10. In the **Preview** table, note that the **Gene ID** column is already selected as the key.
11. Click **OK** to import.
12. We can now setup a style using the imported data. In the **Style** tab of the **Control Panel**, click the **Map.** box for **Fill Color**. In the Column drop-down, select "**log2FC**".
13. In the **Mapping Type** drop-down, select **Continuous Mapping**. This will create a default color gradient from blue (negative fold change) to red (positive fold change).

Additional down-stream analysis

[Extend the STRING network with pathway associations](#)