Appendix A5: Cytospace Methods

Networks can be comprised of proteins, genes, and other molecular entities, and the relationships (edges) might represent interactions, co-expression, or any other kind of association. Cytoscape is an open-source software platform for visualising complex networks and integrating them with any attribute data. While it was initially designed for biological research to visualise molecular interaction networks and biological pathways, it has since evolved to serve many other disciplines. In biology, researchers often use Cytoscape to understand the relationships between proteins in a signalling pathway, to visualise changes in gene expression, or to map out complex interactions in cellular processes.

Key Features of Cytoscape:

1. Visualization: Users can customise network views with styles and useful graphics, like gradient, node pie charts, and graphics mapping.

2. Data Integration: Cytoscape supports importing network/data files in various formats. Also, it has tools to map annotations, ontology, or other metadata on the network.

3. Analysis: Cytoscape core distribution provides a basic set of features for data integration, analysis, and visualisation. Additional features are available as plugins.

4. Plugins: Extensibility is one of Cytoscape’s core features. There are many available plugins which can be installed from within Cytoscape.

How Cytoscape is used:

1. Importing Data: One starts by importing network data, which can be edge lists, adjacency matrices, and other formats. Cytoscape also supports the importing of node/edge attributes.

2. Network Visualization: Once the data is imported, one can visualise the data as a network where nodes represent entities (like proteins or genes) and edges represent relationships or interactions.

3. Customisation: Customize the appearance of nodes and edges based on attributes. For instance, one could colour nodes based on certain attributes or adjust node size according to some quantitative data.

4. Analysis: Built-in tools can compute various network metrics, like degree distribution, shortest path, or clustering coefficient. One can also find network motifs or patterns, perform network enrichment analysis, etc.

5. Plugins: Cytoscape's capabilities can be extended by installing plugins. For example, there are plugins for network clustering, advanced visualisation, and even integrating with external databases.

6. Export: Once the network is visualised and analysed, the network data can be exported in various formats, both as an image and as data.

7. Online Platforms: While Cytoscape is primarily a desktop application, there is an associated online platform called CyNetShare, which allows users to share their Cytoscape networks online. Additionally, the Cytoscape.js library allows for network visualisation in web applications.

Cytoscape Add ons- EnrichmentMap, Autoannotate

For the purposes of analysis of gene expression data in this thesis, the Apps ‘EnrichmentMap’ and ‘Autoannotate’ were used.

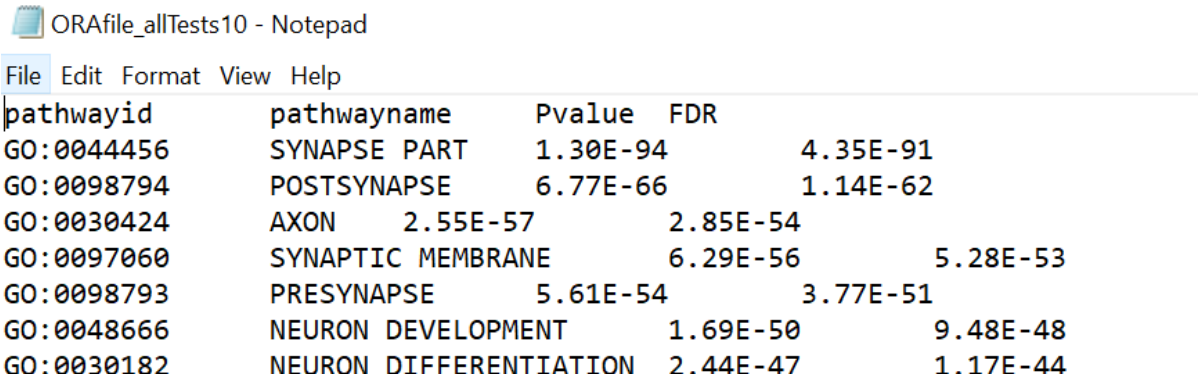
Input file requirement

Input files required: allGOnKEGG.gmt, geneid_inputfile.csv

'geneid_inputfile.csv' is the input file with the header 'GeneID'.

STEP 1

The pathway enrichment file was format as below into a text file (.txt) and imported

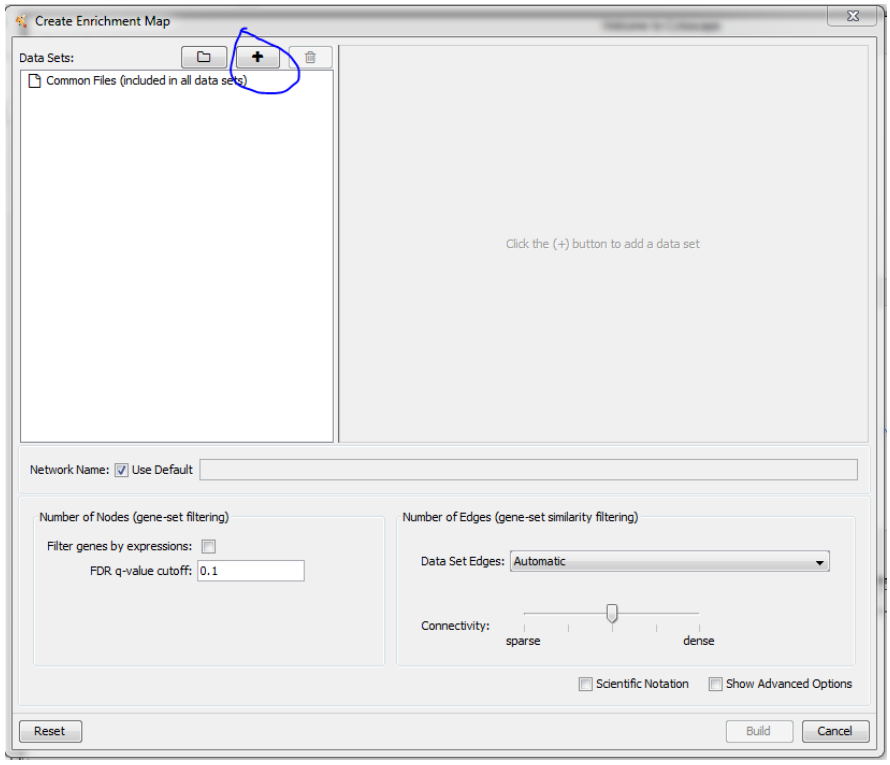


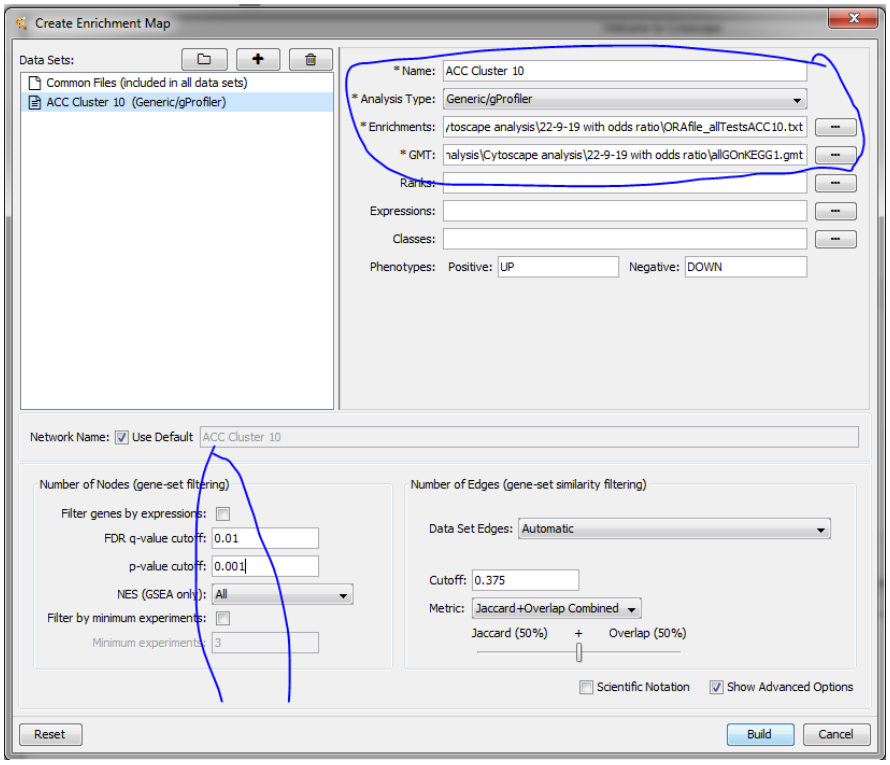
STEP 2

Configure a suitable *GMT file

STEP 3

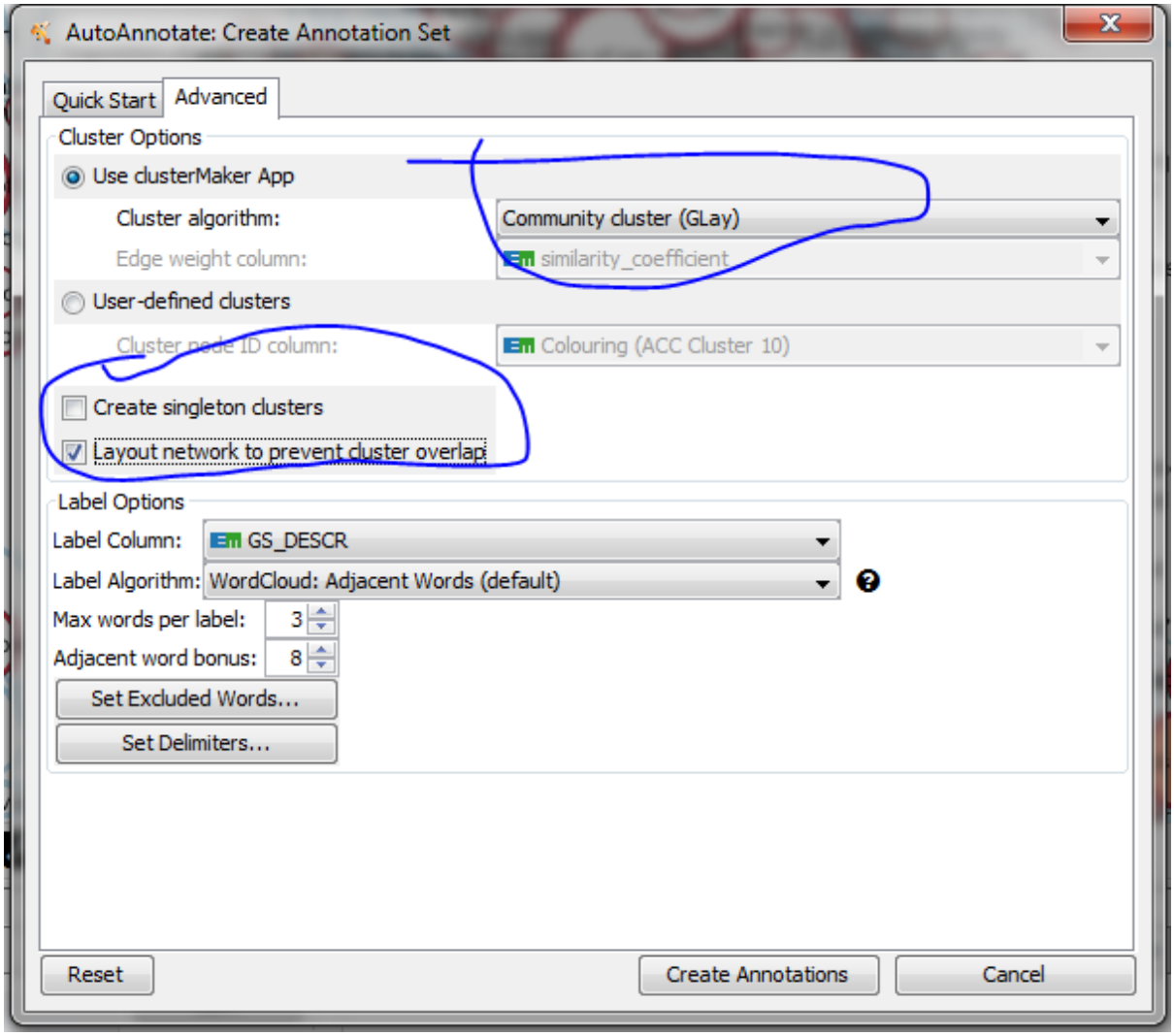
The App ‘EnrichmentMAP’ was used to import the relevant data, FDR set to 0.1 and p-value cut-off set to 0.001 for the GMT file. The number of edges was set to a cutoff of 0.375, and the ‘Build’ button was activated. Settings are shown in the diagram below.



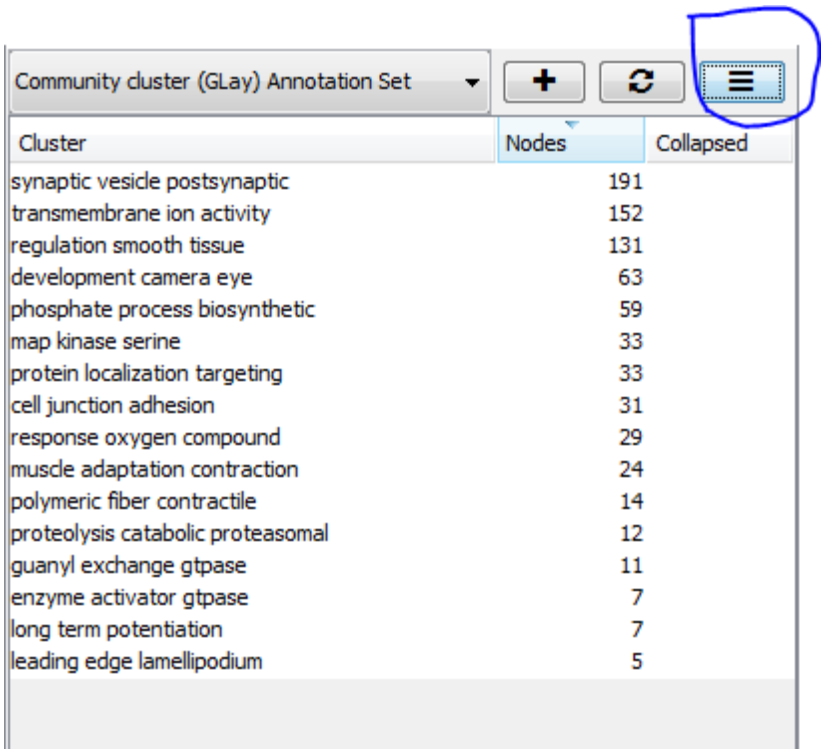


STEP 4: ANNOTATION

The Apps ‘AutoAnnotate’ was used to create annotation sets. Settings are shown in the diagrams below.



To Create Annotationa activate the button shown



STEP 5: Generate Layout clusters using the Grid or Cose layout (multiple times), thereby achieving the desired layout.

STEP 6: Rename cluster names after reviewing the pathways in each node

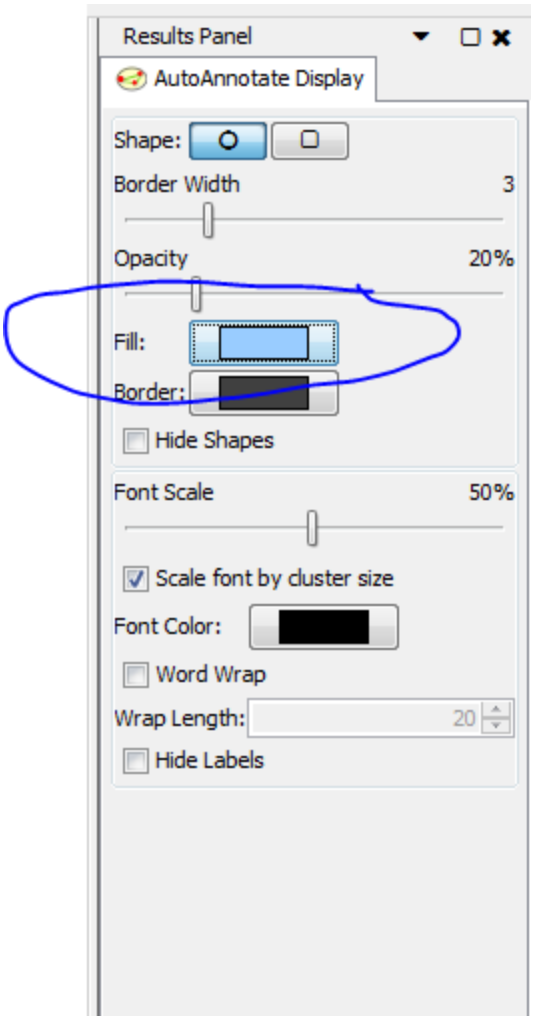
STEP 7: Manual arrangement of layout

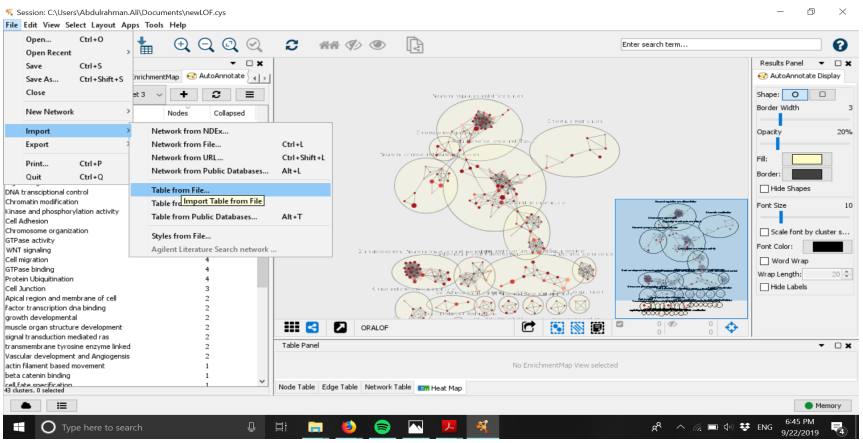
STEP 8: Style setting is Minimal, and other settings are shown below

Edge-stroke colour- gene overlap blue (sat 30)

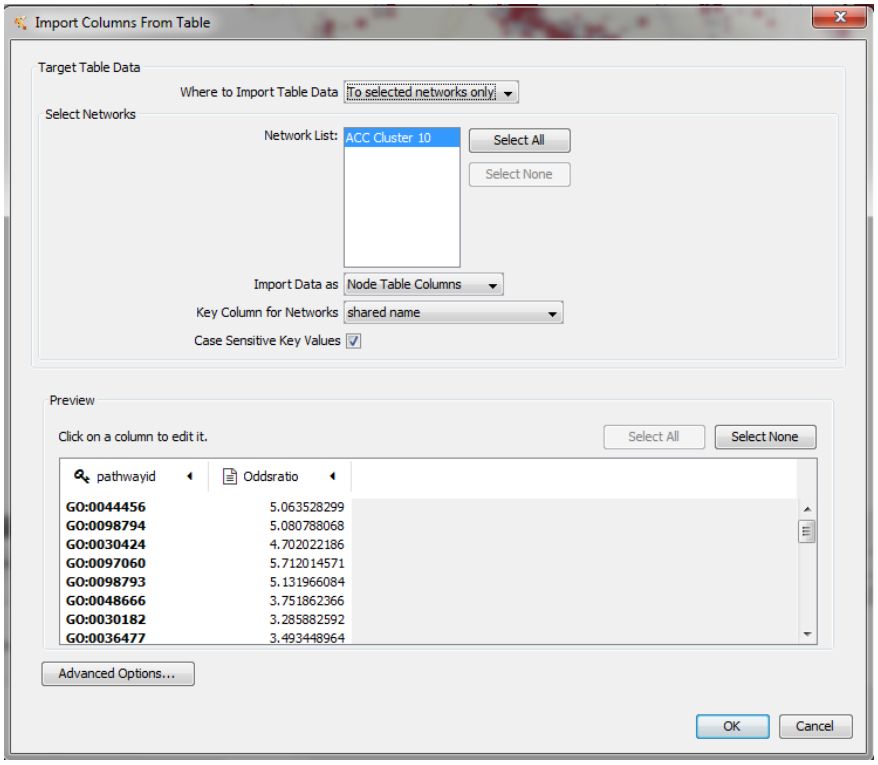
Change cluster background colour to blue (saturation 20)

Remove node names---Label-ES

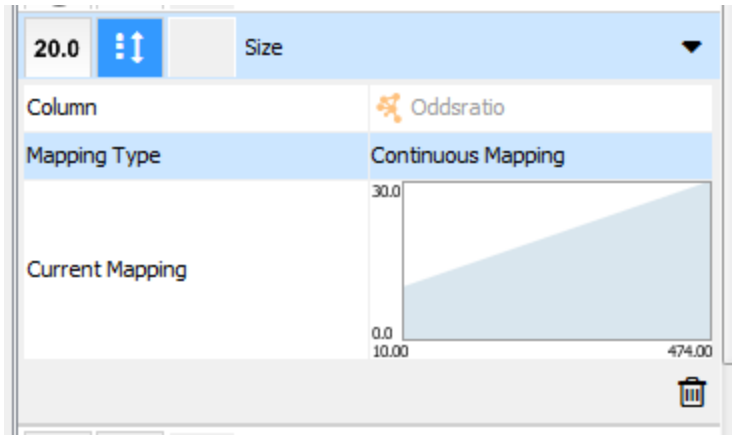




Import Table with pathway id and odds ratio as column

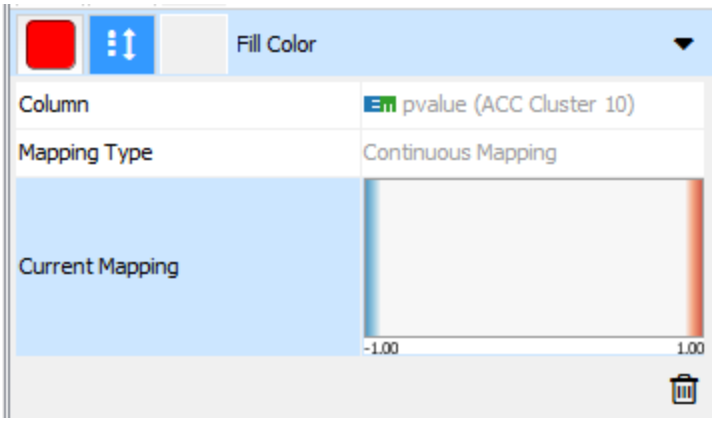


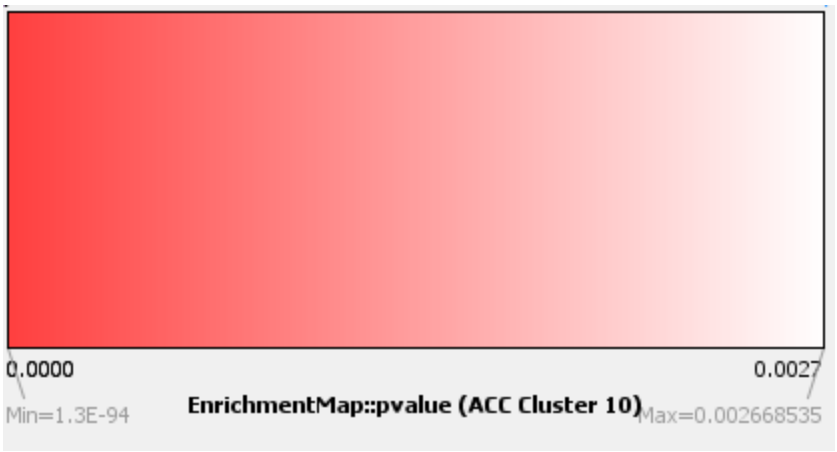
Size:



Manually check value

Set Fill color→





Check values manually

To set the background, select Style→ Network— and stipulate Background white with Node shape round

STEP 9: Save as SVG file

Export to TIFF, JPG using Inkscape or Adobe Illustrator

Save table data