CBW Module 4 Lab Exercise

Aim: This exercise will provide you with an opportunity to perform pathway and network analysis using the Reactome Functional Interaction (FI) and the ReactomeFIViz app.

Goal: Analyze gene lists and somatic mutation data to identify biology that contributes to GBM and ovarian cancer.

Example 1: Pathway-based analysis of GBM genelist

- Open up Cytoscape.
- Go to Apps >Reactome FI>Reactome Pathways.
- Unfurl the "Signal Transduction" events, by clicking the triangle to the left of the event name, in the "Reactome" tab on the left.
- Click on "RAF/MAP kinase cascade" or your favourite pathway.
- Right-click on highlighted pathway name to display drop-down menu, select "Show Diagram" to display RAF/MAP kinase cascade pathway.
- Right-click on highlighted pathway name to display drop-down menu, select "Analyze Pathway Enrichment"
- Upload/Browse "GBM_TCGA_GeneList.txt" into Reactome Pathway Enrichment Analysis, and click "OK".

1. *What are the most significant biological pathways when the FDR Filter is set to 0.05?*
   - Hint: Right-click on selected pathway in Table Panel, and click "View in Diagram". Purple-coloured nodes reflect hits in the dataset. Right-click on highlighted nodes to invoke additional features.

Example 2: Network-based analysis of GBM gene-sample data

- Open up Cytoscape.
- Go to Apps>Reactome FI and Select "Gene Set/Mutational Analysis".
- Choose "2014 (Latest)" Version.
- Upload/Browse "GBM_genesample.txt" file.
- Select "Gene/sample number pair" and Choose sample cutoff value of 4.
- Select "Fetch FI annotations".
- Click OK.

1. *Describe the size and composition of the GBM sub-network?*
2. *What are the driver genes?*
3. *Describe the TP53-PEG3 interaction, and the source information to support this interaction?*
4. *Describe the data sources for the RB1-PIK3R1 FI?*
5. *After clustering, how many modules are there?*
6. *How many pathway gene sets are there in Module 2 when the FDR Filter is set to 0.005 and Module Size Filter to 10?*
   - Hint: Analyze Module Functions>Pathway Enrichment. Select appropriate filters at each step.
7. *What are the most significant pathway gene sets in Module 0, 1, 2?*
   - *Hint: You don't need to list them all!*
8. *Is there any evidence to support NF1 expression in breast cancer cells?*
   - Hint: Select Node > Fetch Cancer Gene Index > Select Appropriate Filter and Keyword>etc.

Example 3: Network-based analysis of OvCa somatic mutation

- Open up Cytoscape.
- Go to Apps>Reactome FI and Select "Gene Set/Mutational Analysis".
- Choose "2014 (Latest)" Version.
- Upload/Browse "OVCA_TCGA_MAF.txt" file.
- Select "NCI MAF" (Mutation Annotation File) and Choose sample cutoff value of 4.
- Do not select "Fetch FI annotations".
- Click OK.

1. *Describe the size and composition of the OvCa subnetwork?*
2. *What is the driver gene?*
3. *After clustering, how many modules are there?*
4. *How many pathway gene sets are there in Module 0 when the FDR Filter is set to 0.005 and Module Size Filter to 10?*
   - Hint: Analyze Module Functions>Pathway Enrichment. Select appropriate filters at each step.
5. *What are the most significant pathway gene sets in Module 0, 1, 2, 3?*
6. *Do the GO Biological Process annotations correlate with the significant pathway annotations for Module 0?*
   - Hint: Analyze Module Functions>GO Biological Process. Select appropriate filters at each step.
7. *What are the most significant GO Cellular Component gene sets in Module 1, and do they correlate with the pathway annotations?*
   - Hint: Analyze Module Functions>GO Cellular Component. Select appropriate filters at each step.
8. *What are the most significant GO Molecular Function gene sets in Module 2? [Optional]*
   - Hint: Analyze Module Functions>GO Molecular Function. Select appropriate filters at each step.
9. *Are any of the modules annotated with the NCI Disease term: "Stage_IV_Breast_Cancer" [malignant cancer]?*
   - Hint: Load Cancer Gene Index>Neoplasm>Neoplasm_by_Site>Breast Neoplasm.
10. *How many modules are statistically significant in the CoxPH analysis?*
    - Hint: Analyze Module Functions>Survival Analysis>Upload/Browse "OVCA_TCGA_Clinical.txt". Click OK.
11. *What does the Kaplan-Meyer plot show for the most clinically significant modules?*
    - Hint: Click the most statistically significant module link [blue line] from the CoxPH results panel. Click OK. Click #_plot.pdf to display Kaplan-Meyer plot. Repeat this for the other significant module links. KM plot: samples having genes mutated in a module (red line), and samples having no genes mutated in the module (green line).
12. *Taking into what you have learned about module 2 (ie. pathway and GO annotations), what is your hypothesis?*