Lab Module Questions

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
o   Open up Cytoscape.
o   Go to Apps >Reactome FI>Reactome Pathways.
o   Unfurl the "Signal Transduction" events, by clicking the triangle to the left of the event name, in the "Reactome" tab on the left.
o   Click on "Signaling by EGFR" or your favourite pathway.
o   Right-click on highlighted pathway name to display drop-down menu, select "Show Diagram" to display Signaling by EGFR pathway.
o   Right-click on highlighted pathway name to display drop-down menu, select "Analyze Pathway Enrichment"
o   Upload/Browse "GBM_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?
        o   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
o   Open up Cytoscape.
o   Go to Apps>Reactome FI and Select "Gene Set/Mutational Analysis".
o   Choose "2016 (Latest)" Version.
o   Upload/Browse "GBM_genesample.txt" file.
o   Select "Gene/sample number pair" and Choose sample cutoff value of 4.
o   Select "Fetch FI annotations".
o   Click OK.

   1.   Describe the size and composition of the GBM sub-network?
   2.   What are the driver mutations?
   3.   Describe the TP53-PEG3 interaction, and the source information to support this interaction?
   4.   Describe the data sources for the TAF1-TAF7L FI?
   5.   After clustering, how many modules are there?
   6.   How many pathway gene sets are there in Module 3 when the FDR Filter is set to 0.0001 and Module Size Filter to 10?
        o   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, 3?
        o   Hint: You don't need to list them all!

Example 3: Network-based analysis of OvCa somatic mutation
        o   Open up Cytoscape.

- Go to Apps>Reactome FI and Select "Gene Set/Mutational Analysis".
- Choose "2016 (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 network?
2. What are the most frequently mutated genes?
3. After clustering, how many modules are there?
4. How many pathway gene sets are there in Module 1 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, 2, 3 and 4?
6. Do the GO Biological Process annotations correlate with the significant pathway annotations for Module 2?
   - Hint: Analyze Module Functions>GO Biological Process. Select appropriate filters at each step.
7. What are the most significant GO Cell Component gene sets in Module 1 when the FDR Filter is set to 0.005 and Module Size Filter to 10? [Optional]
   - Hint: Analyze Module Functions>GO Cell Component. Select appropriate filters at each step.
8. 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>…….
9. What are the targets of Imatinib Mesylate?
   - Hint: Overlay Cancer Drugs>Fetch Cancer Drugs. Maybe apply filters?
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 (green line), and samples having no genes mutated in the module (red line).
12. Taking into what you have learned about module 4, what is your hypothesis?