

## Lab Module Answers

### Example 1

1. SCF-KIT, EGFR, Downstream Signal Transduction, PDGF, PI3K/AKT signaling, INSR, NGF, FGFR etc;

### Example 2

1. The overall sub-network consists of 276 nodes and 704 edges. The largest component of the subnetwork consists of 240 nodes and 618 edges, with the remainder of nodes and edges distributed amongst 7 other small subnetworks and interactions.
2. Couple of ways to answer this. The driver mutations are probably the frequently mutated gene in the samples. The node size is proportional to the number of samples where the gene is mutated. Method 1- Look for the largest nodes in the diagram. Method 2 – Click Node Table and sort by “sampleNumber”. The largest node is TP53, ie. mutations in the TP53 gene are highly prevalent, occurring in 100 samples. Other driver mutations include EGFR (95) and PTEN (93). Additional mutations of interest include NF1, PIK3R1, PIK3CA, PIK3R1, RYR2, RB1.
3. Search for “TP53 PEG3” in search bar in top right of Cytoscape tool. Annotated Functional Interaction based upon data from the TRED database. This targeted interaction describes an interaction between TP53 (regulator) and PEG3 (target). An immunoprecipitation experiment demonstrates the interaction, and the supporting evidence has been published in the paper with a PubMed ID: 11679586.
4. Search for “TAF1 TAF7L” in search bar in top right of Cytoscape tool. Predicted Functional Interaction based upon data (2/9 sources are true) from a mouse interaction database and GO (GO BP sharing). FI score: 0.61
5. 19 modules, with 11 modules of  $10 \geq$  genes.
6. 14 modules, depending on the results of the enrichment analysis. Some pathways gene sets at the cutoff threshold may come or go but those highly significant gene sets are always there.
7. 0: RTK signalling, 1: ECM and Integrin signalling, and 3: Cell Cycle.

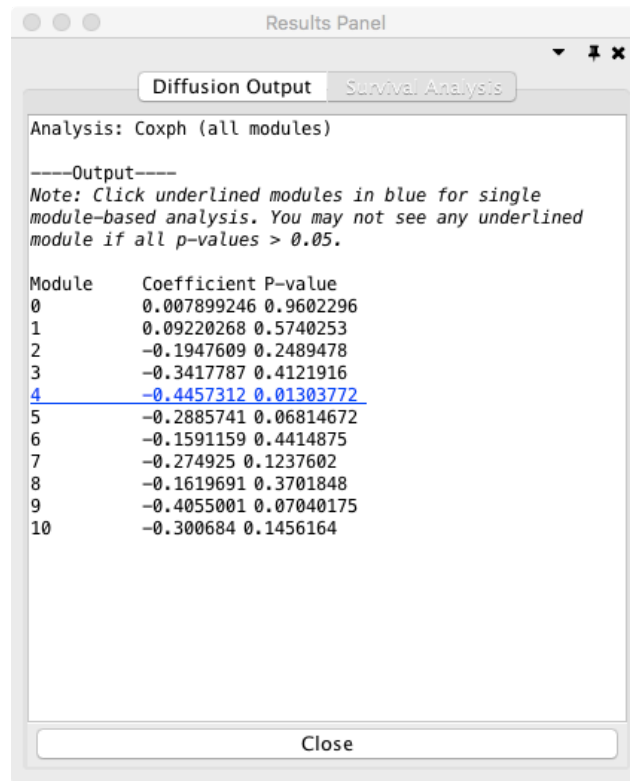
### Example 3

1. The overall sub-network consists of 245 nodes and 489 edges. The largest component of the subnetwork consists of 201 nodes and 418 edges, with the remainder of nodes and edges distributed amongst 14 other small subnetworks and interactions.
2. The largest node is TP53, ie. mutations in the TP53 gene are highly prevalent, occurring in at least 96% of HGS-OvCa samples.
3. After clustering, there are 27 modules with 11 modules of  $10 \geq$  genes.
4. 7 modules, depending on the results of the enrichment analysis. Some pathways gene sets at the cutoff threshold may come or go but those highly significant gene sets are always there.
5. 0: RTK Signaling, 2: ECM and Integrin signalling, 3: Cell Cycle – Meiotic and Mitotic, 4: Calcium signalling-Adrenergic Signaling-Cardiac Muscle Contraction,
6. Yes, ECM organization and Cell adhesion.
7. Nuclear components - Nucleoplasm, nuclear membrane, nuclear pore, chromatin, etc.
8. Modules 0 and 3 will be highlighted. Navigate through hierarchy. Neoplasm > Neoplasm\_by\_Site > Breast Neoplasm > Malignant\_Breast\_Neoplasm > Breast Carcinoma > Stage\_IV\_Breast\_Cancer. Go

back to the Network Module Browser. Genes in the modules that have 'Stage IV Breast Cancer' annotations will be yellow-highlighted: BRCA1, NRG1, TP53, INSR, EGFR.

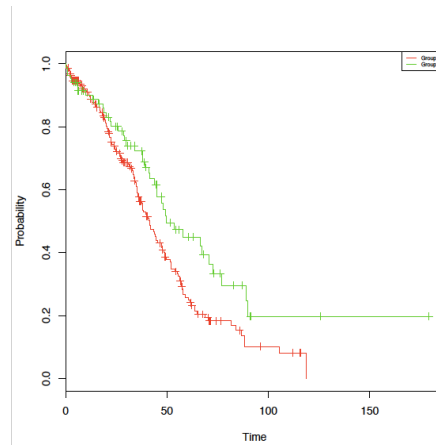
9. EGFR and KIT

10. 1 module: 4



11. The ReactomeFIViz app splits samples into two groups: samples having genes mutated in a module (green line), and samples having no genes mutated in the module (red line). The plugin uses the log-rank test to compare the two survival curves, and estimates p-values. In Modules X (KM: p= 0.0124), patient with genes mutated (green line) have a better prognosis than patients with no

gene mutations (red line). Module 4 is most statistically significant modules from the CoxPH and KM analysis.



- 12.** In Module 4, the Calcium signaling, Chemical Synapse/Neurotransmission and Muscle Contraction annotations reflect a shared set of genes. These genes represent voltage-gated ion channels, which are a group of transmembrane ion channels that activated by changes in electrical potential difference. Even though ion channels are especially critical in neurons and muscle tissue, they are common in many types of cells, controlling the influx and outflux of ions. There are a number of genetic disorders, which disrupt normal functioning of ion channels. Calcium homeostasis is essential for cell migration, and tumor metastasis in particular. It may be that mutations in Module 4 genes disrupt calcium homeostasis, thereby impairing the tumour's ability to metastasize, and extending patient's overall survival.