reactome

EMBL-EBI NYU Langone MEDICAL CENTER OHSU OICR

CBW Lab Module 4 Answers

Example 1

1. Signaling by ERBB4, PI3K events in ERBB2 signaling, FLT3 signaling, GAB1 signalsome.

Example 2

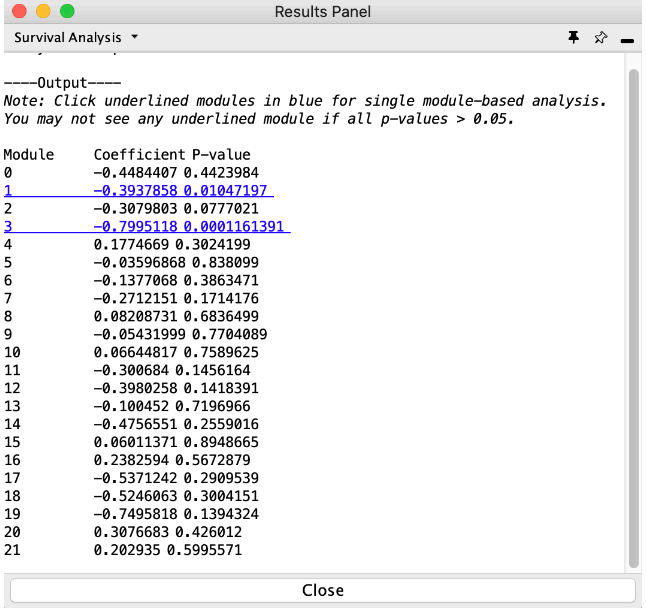
1. The overall sub-network consists of 292 nodes and 733 edges. The largest component of the subnetwork consists of 254 nodes and 633 edges, with the remainder of nodes and edges distributed amongst 10 other small subnetworks and interactions.
2. There are a couple of ways to answer this. The node size 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 gene nodes include EGFR (95) and PTEN (93). Additional nodes 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: 0.54
5. 21 modules, with 9 modules of 10 ≥ genes.
6. 17 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: TP53 signaling, 1:RTK Signaling, and 3: ECM and Integrin signalling.

Example 3

- The overall sub-network consists of 268 nodes and 576 edges. The largest component of the subnetwork consists of 241 nodes and 521 edges, with the remainder of nodes and edges distributed amongst 9 other small subnetworks and interactions.
- The largest node is TP53, ie. mutations in the TP53 gene are highly prevalent, occurring in at least 96% of HGS-OvCa samples.
- After clustering, there are 22 modules with 12 modules of 10 ≥ genes.
- 23 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.
- 0: Cell Cycle, 1: ECM and Integrin signalling, 2: RTK signalling, 3: Calcium signalling-Adreneric Signaling-Cardiac Muscle Contraction.
- Yes, ECM organization and Cell adhesion.
- Nuclear components - Nucleoplasm, nuclear membrane, nuclear pore, chromatin, etc.
- Modules 0 and 2 will be highlighted. Navigate through hierarchy. Neoplasm > Neoplasm_by_Site > Breast Neoplasm > Maligant_Breast_Neoplasm > Breast Carcinoma > Stage_IV_Breast_Cancer.

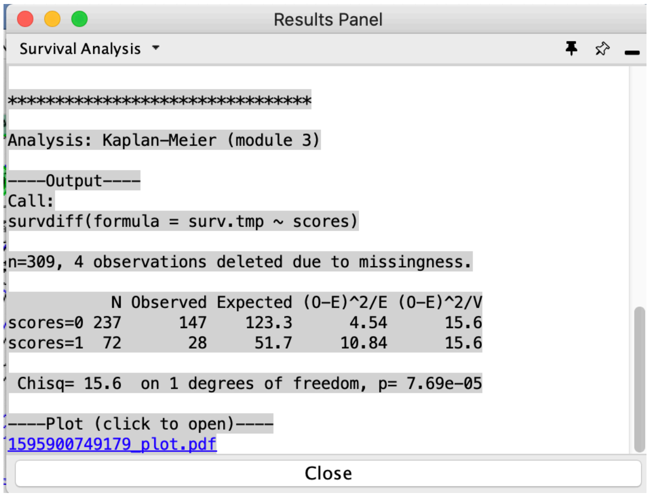
Go back to the Network Module Browser. Genes in the modules that have ‘Stage IV Breast Cancer’ annotations will be yellow-highlighted: BRCA1, NRG1, TP53, INSR, EGFR.

- EGFR and KIT
- 2 modules: 1 and 3

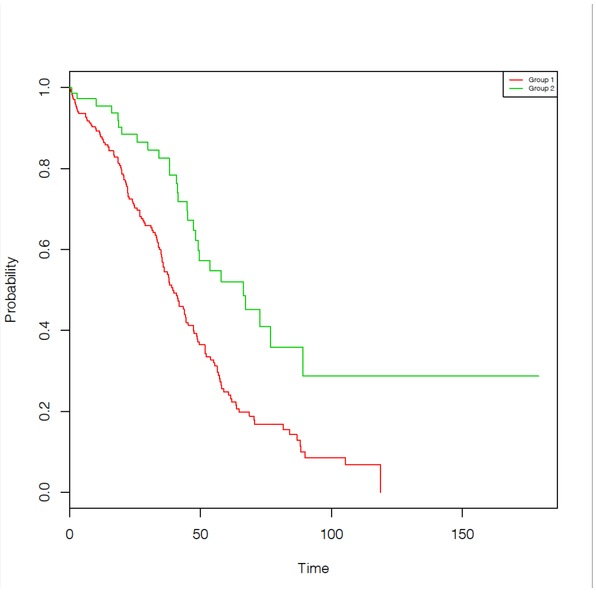


```
----Output----
Note: Click underlined modules in blue for single module-based analysis.
You may not see any underlined module if all p-values > 0.05.

Module     Coefficient P-value
0          -0.4484407 0.4423984
1          -0.3937858 0.01047197
2          -0.3079803 0.0777021
3          -0.7995118 0.0001161391
4          0.1774669 0.3024199
5          -0.03596868 0.838099
6          -0.1377068 0.3863471
7          -0.2712151 0.1714176
8          0.08208731 0.6836499
9          -0.05431999 0.7704089
10         0.06644817 0.7589625
11         -0.300684 0.1456164
12         -0.3980258 0.1418391
13         -0.100452 0.7196966
14         -0.4756551 0.2559016
15         0.06011371 0.8948665
16         0.2382594 0.5672879
17         -0.5371242 0.2909539
18         -0.5246063 0.3004151
19         -0.7495818 0.1394324
20         0.3076683 0.426012
21         0.202935 0.5995571
```

- 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 Module 3 (KM: p= 0.0000769), patient with genes mutated (green line) have a better prognosis than patients with no gene mutations (red line). Module 3 is most statistically significant modules from the CoxPH and KM analysis.



```
********************************

Analysis: Kaplan-Meier (module 3)

----Output----
Call:
survdiff(formula = surv.tmp ~ scores)

n=309, 4 observations deleted due to missingness.

           N Observed Expected (O-E)^2/E (O-E)^2/V
scores=0 237      147    123.3      4.54      15.6
scores=1  72       28     51.7     10.84      15.6

 Chisq= 15.6  on 1 degrees of freedom, p= 7.69e-05

----Plot (click to open)----
1595900749179_plot.pdf
```



- In Module 3, 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 3 genes disrupt calcium homeostasis, thereby impairing the tumour's ability to metastasize, and extending patient's overall survival.