LinkedOmics & WikiPathways Demo

Eric Jaehnig

10/08/2018

I. LinkFinder to identify associations between TP53 mutations and changes in gene expression (RNA-seq)

1. go to http://linkedomics.com

2. Click enter as guest (or register and sign in)

3. STEP-1: Select cancer cohort >> Breast Invasive Carcinoma (TCGA_BRCA)

4. STEP-2: Select query dataset >> Show all, select Data type: Mutation – Analysis Level: Gene; this is the dataset containing the query data

5. skip STEP-2b (if you wanted to select a subset of patients to focus on, this is where you would do that)

6. STEP-3: Select query attribute; type TP53 in the box; should also be in pulldown menu

7. STEP-4: Select target dataset >> Show all, select Data type: RNAseq – Platform: HiSeq RNA; this is the gene expression from which we will identify genes that are associated with p53 mutations

8. STEP-5: Select statistical method >> T-test (this test allows us to identify genes with different levels of expression in p53 mutant cancer patients relative to non-mutant patients

9. Submit Query

10. Click on View button >> LinkFinder results will be shown (results may vary depending on analysis):

A. Upper left: Downloadable table of genes from the target dataset and change in expression of each in TP53 mutant compared to wild-type (Log2F(old)C(hange)) along with T test results indicating significance of difference between mutant and wild-type (P-value and FDR)

B. Upper right: Volcano plot of -LogPvalue vs. LogFC

C. Lower left: Heat map of expression of the top 50 genes upregulated in mutant (black bar at top) vs. wild-type (grey bar at top) for each patient in cohort

D. Lower right: Heat map of expression of the top 50 genes downregulated in mutant (black bar at top) vs. wild-type (grey bar at top) for each patient in cohort

II. LinkInterpreter to perform WikiPathway enrichment analysis on genes that are both positively (upregulated in mutant) and negatively (downregulated in mutant) associated with TP53 mutation

1. Click on LinkInterpreter tab

2. Select Tool >> Gene Set Enrichment Analysis (GSEA)

    GSEA considers expression values of all of the genes in data; ORA just considers overlap of genes that meet specified significance threshold with genes in pathways

3. Rank Criteria >> P-Value

4. Mininum Number of Genes >> 3

5. Simulations >> 500 (1000 is better but takes longer)

5. Submit Query >> wait…

6. Results from WebGestalt are displayed

7. Scroll down to Enrichment Results; top 2 WikiPathways are related: Retinoblastoma in Cancer and Cell Cycle

8. Scroll down to Select an enriched gene set… >> select WP2446: Retinoblastoma (RB) in Cancer

9. GSEA results for WP2446 are displayed; table at the bottom shows “leading edge” genes in the pathway

10. Click on the WP2446 link next to box with arrow (teal color) to view the pathway; leading edge genes are highlighted in red