

Functional enrichment analysis plots

Mikhail Dozmorov

2018-12-16

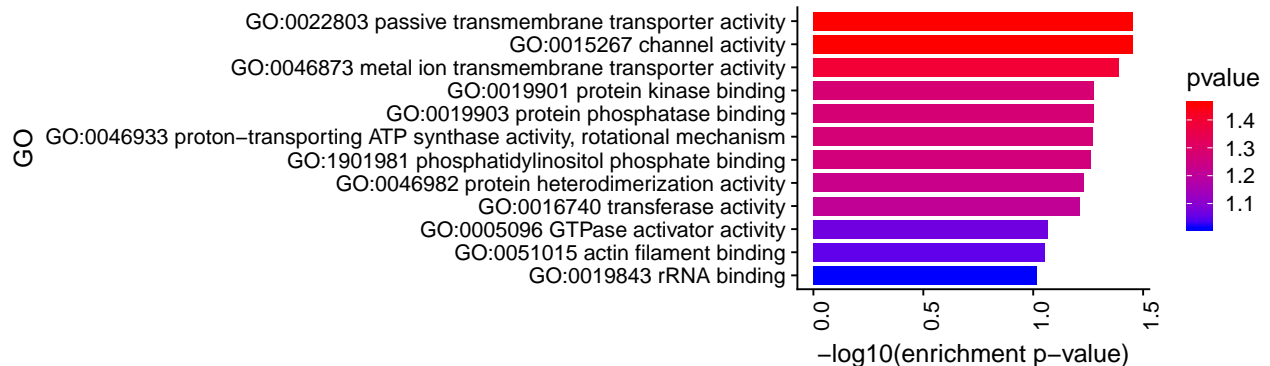
Settings

```
# File name with the results of edgeR analysis
fileName <- "results/DEGs_C231_CRISPR1_vs_231_noCRISPR_hg38.xlsx"
logFC <- TRUE # Are we using column 3: 'logFC' or 6: 'PValue'?
# Analysis settings
max_kegg_genes = 2000 # Maximum number of genes to run enrichment analysis on
max_GO_to_plot = 12 # Maximum number of GOs to plot
nperm = 10000 # Number of permutations
```

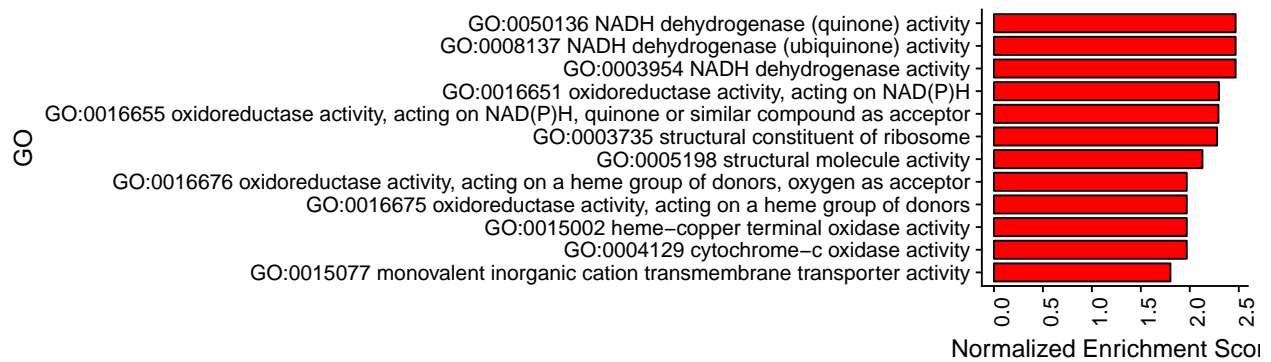
Brief Methods: Gene Set Enrichment Analysis (GSEA [PMID: 16199517]) was performed on differentially expressed genes ranked by log-Fold Change using `clusterProfiler` 3.10.0. 10^4 permutations were performed to estimate permutation-based enrichment p-value.

Gene Ontology has three domains - “Molecular Function”, “Biological Process”, and “Cellular Component”. Enrichment analyses is done separately for each.

Gene Ontology: Molecular Function, all genes

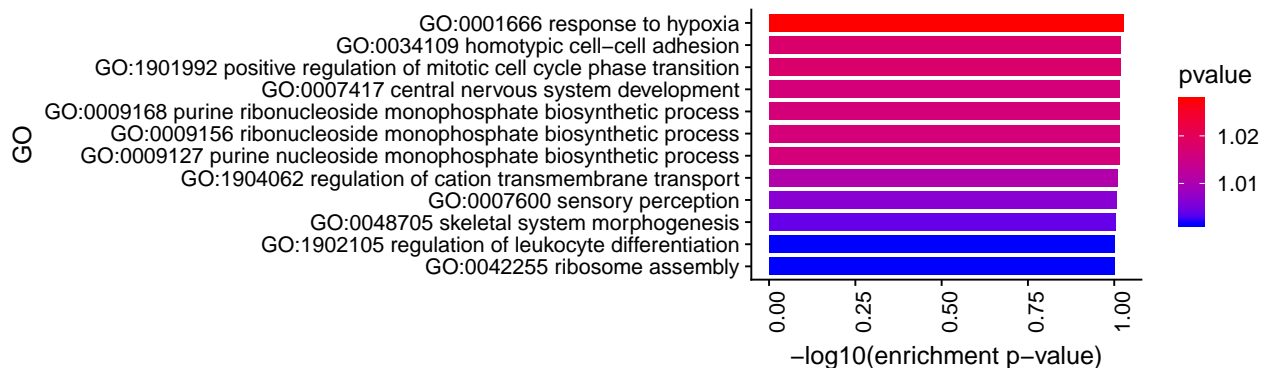


Gene Ontology: Molecular Function, upregulated genes

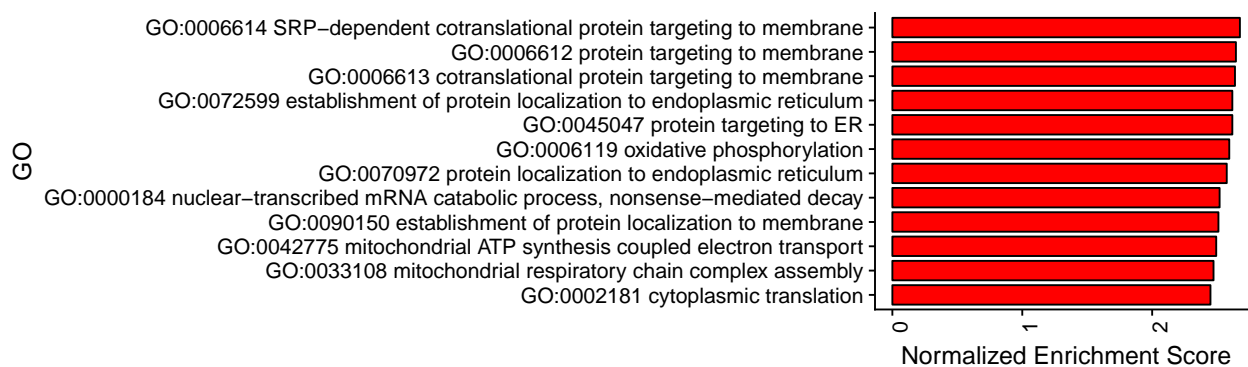


Gene Ontology: Molecular Function, downregulated genes

Gene Ontology: Biological processes, all genes

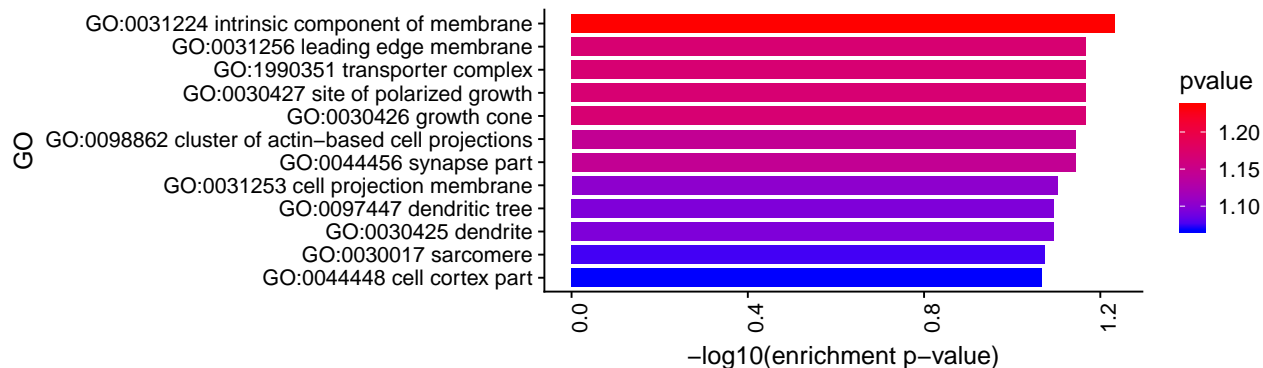


Gene Ontology: Biological processes, upregulated genes

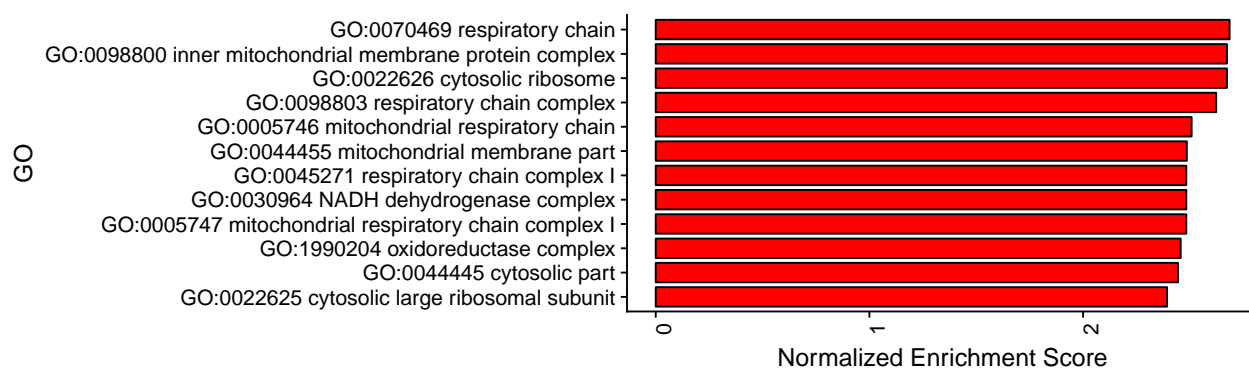


Gene Ontology: Biological processes, downregulated genes

Gene Ontology: Cellular Component, all genes

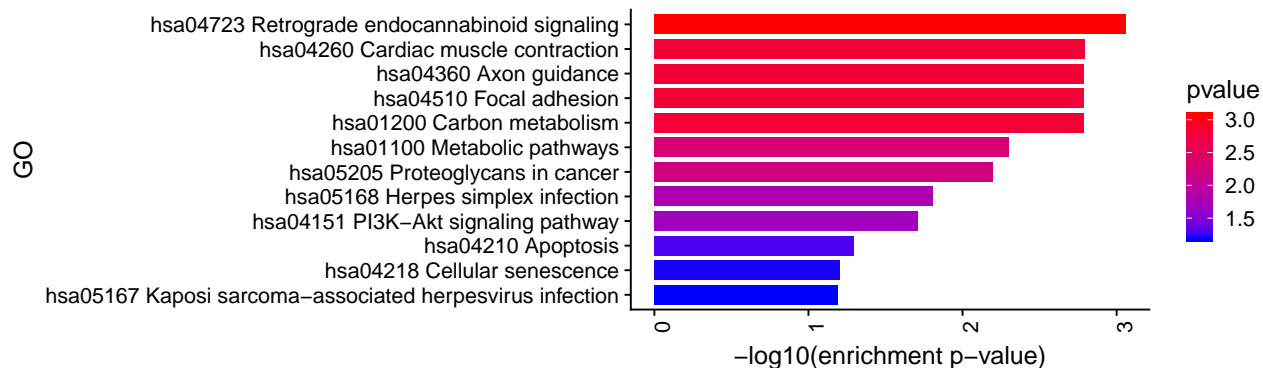


Gene Ontology: Cellular Component, upregulated genes

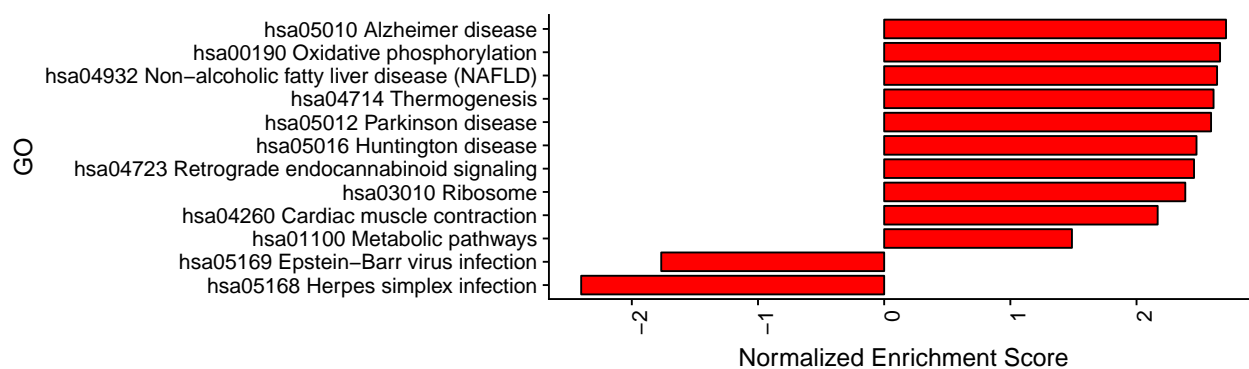


Gene Ontology: Cellular Component, downregulated genes

KEGG canonical pathways, all genes



KEGG canonical pathways, upregulated genes



KEGG canonical pathways, downregulated genes