

Toy dataset and R code for: Gene behaviors-based network enrichment analysis.

Title of paper: Gene behaviors-based network enrichment analysis and its application to reveal immune disease pathways enriched with COVID-19 severity specific gene networks.

■ Input datasets

1. NW_C.csv and NW_N.csv: Gene networks of phenotypes C and N.

The 1st ~ 4th columns are regulator gene, target genes, effect of regulator on target genes (i.e., $\hat{\beta}$ in (1)) and regulatory effect of the regulator genes (i.e., r_j^C and r_j^N in (3)), respectively.

2. PW_genes.csv: Genes involved in a specific pathway

■ Implementation of R-code

1. Compute Regulatory effect

Compute the difference of regulatory effect (i.e., $d_j^{(1)}(C, N) = r_j^C - r_j^N$ in (6)).

2. Compute Enrichment score based on regulatory effect

Compute maximum deviation from the zero of $Hit(V^T, i) - Miss(V^T, i)$ based on $d_j^{(1)}(C, N)$.

3. Compute Jaccard Distance

Compute Jaccard Distance (i.e., $Jl_j = \frac{|N_j^C \cap N_j^N|}{|N_j^C \cup N_j^N|}$ in (4))

4. Compute Enrichment Score based on Regulatory effect and Jaccard distance

Compute maximum deviation from the zero of $Hit(V^T, i) - Miss(V^T, i)$ based on $Wd_j^{(1)}(C, N)$.

■ Output datasets

Output 1: Enrichment score based on regulatory effect $d_j^{(1)}(C, N)$.

: ES_RG

Output 2: Enrichment score based on regulatory effect and Jaccard distance $Wd_j^{(1)}(C, N)$.

: ES_RGJD