Functional Analysis

Gain greater biological insight into the differential expression results.
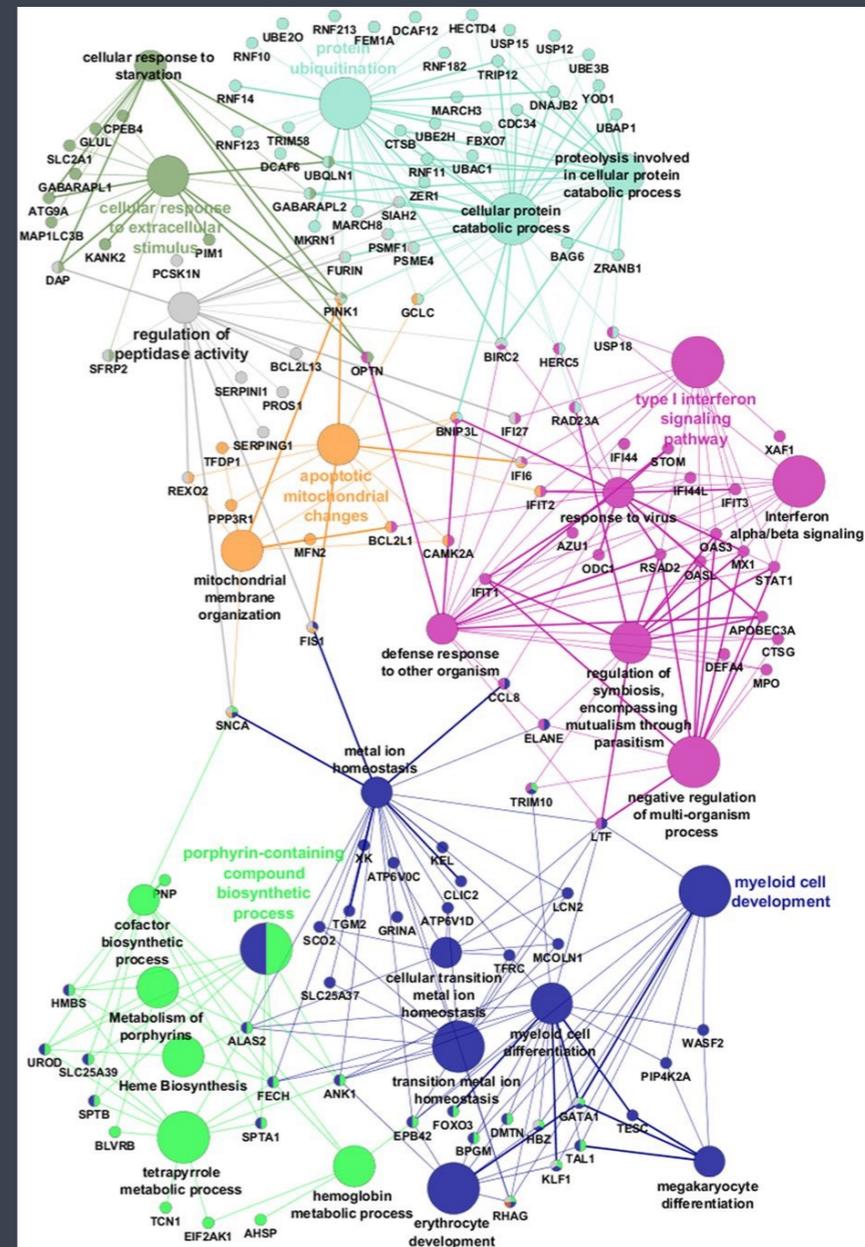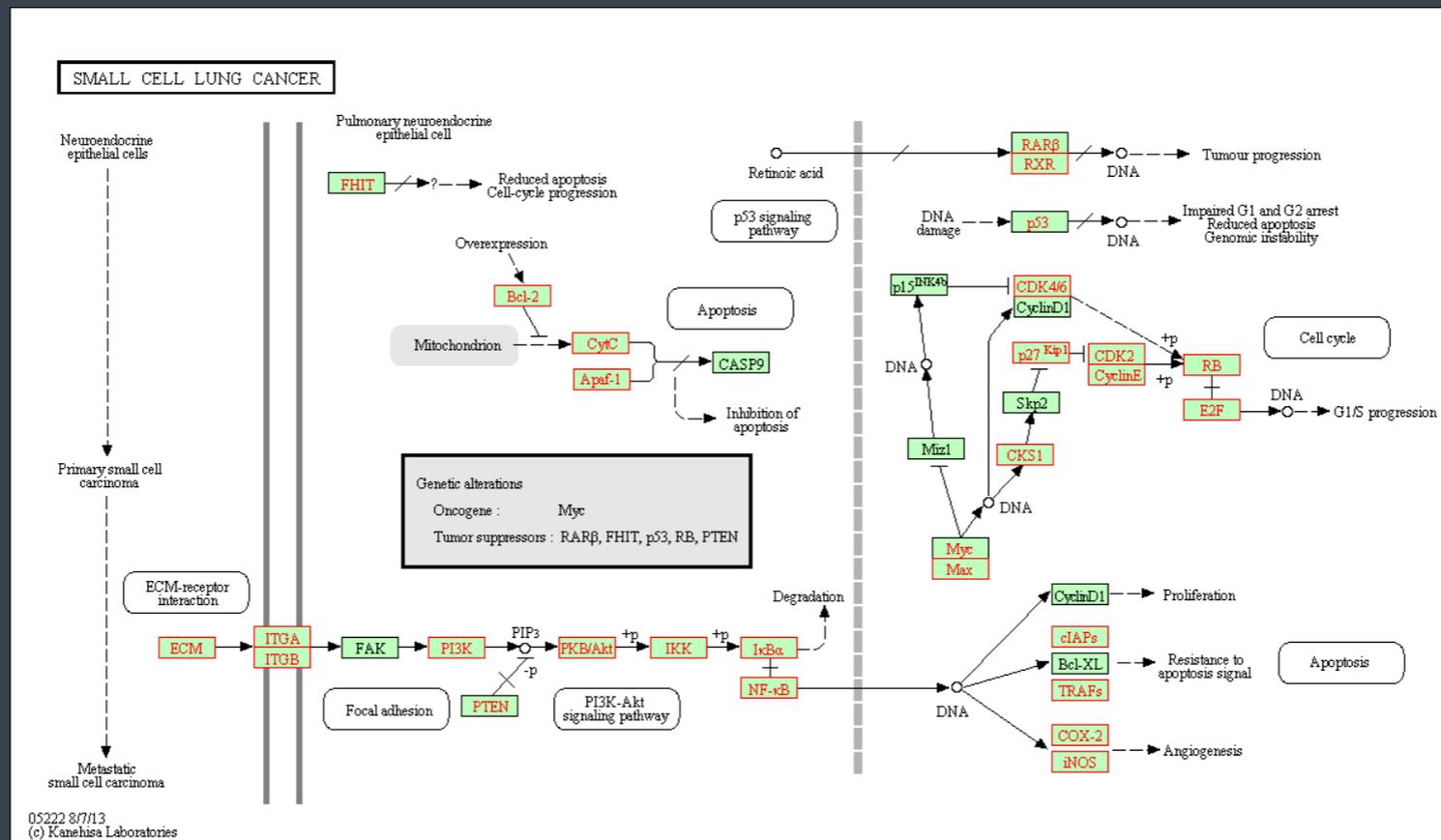
Databases

To explore which biological pathways are enriched, need to know which genes are associated with each of the pathways.

Databases

Online databases annotate, store, and share experimentally- and electronically-inferred information about which genes are associated with particular processes and/or pathways.
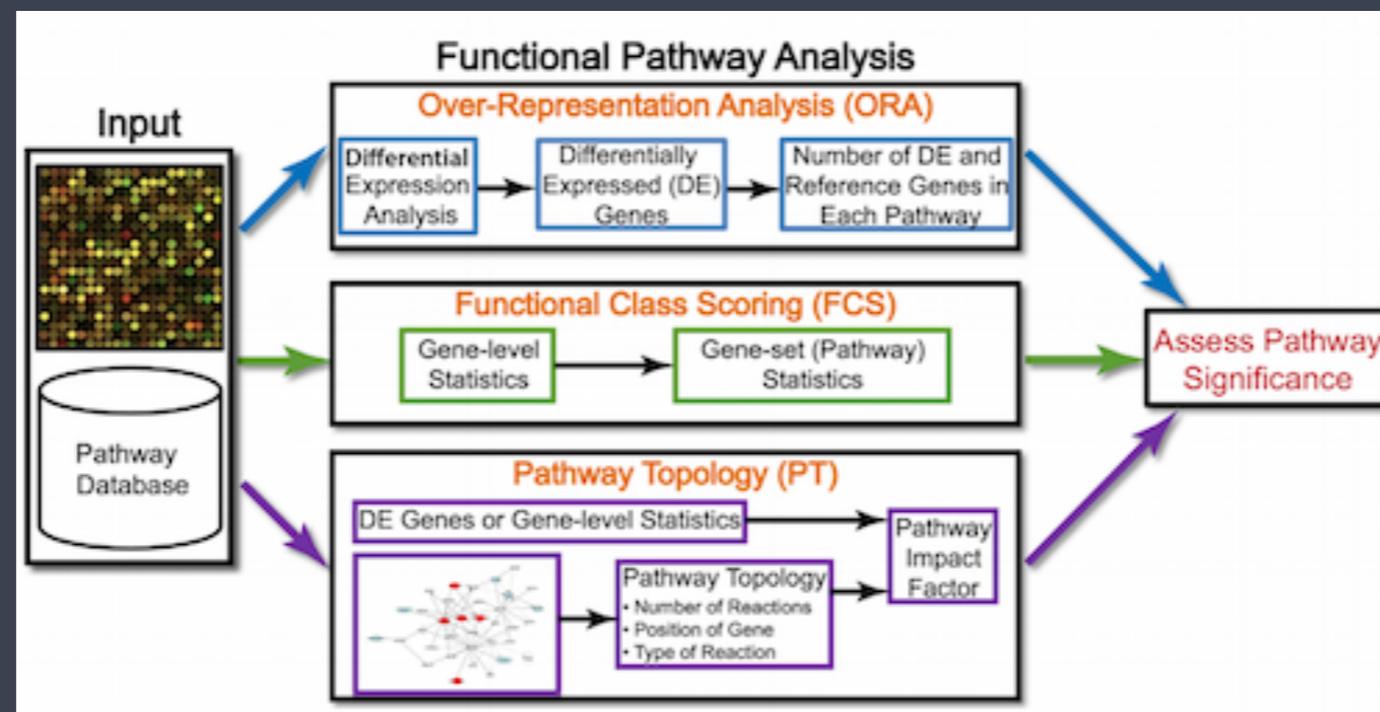
Databases

Popular databases for functional analysis include:

- Gene Ontology: genes associated with particular biological processes, cellular components, and molecular functions

- KEGG: genes associated with particular biological pathways

- Reactome: genes associated with particular biological pathways

- MSigDB: genes associated with particular biological states/processes, motifs, perturbations, or pathways (including KEGG and Reactome pathways)

- Human Phenotype Ontology: genes associated with phenotypic abnormalities
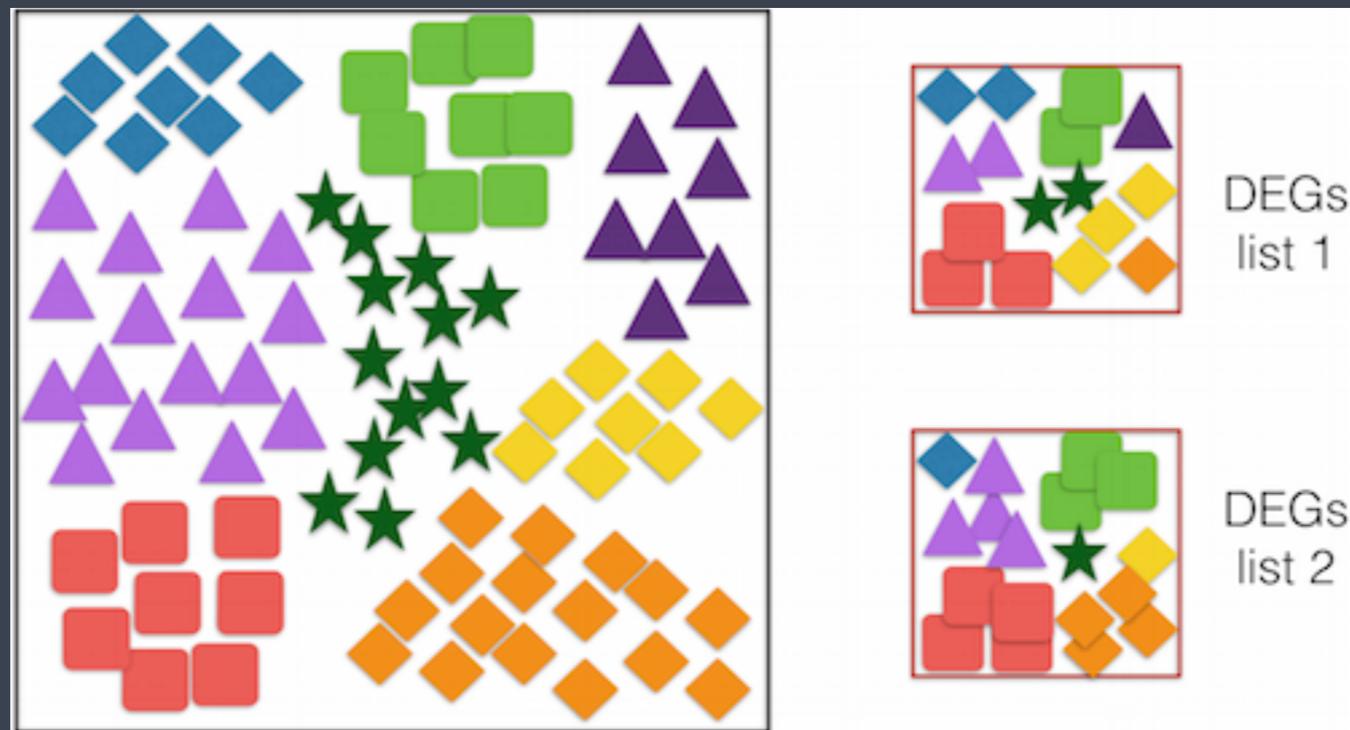
Functional Analysis Methods

To gain greater biological insight into the differential expression results, functional analyses cover a range of techniques.

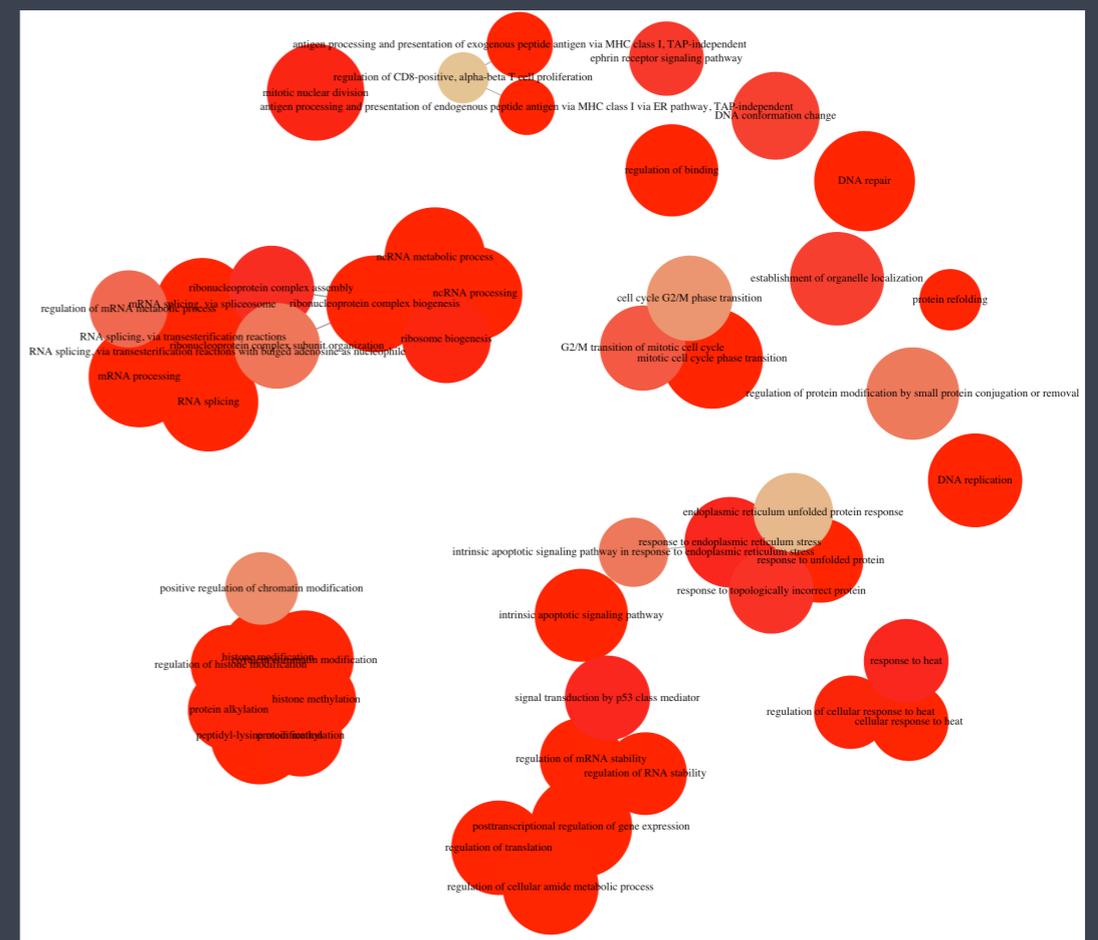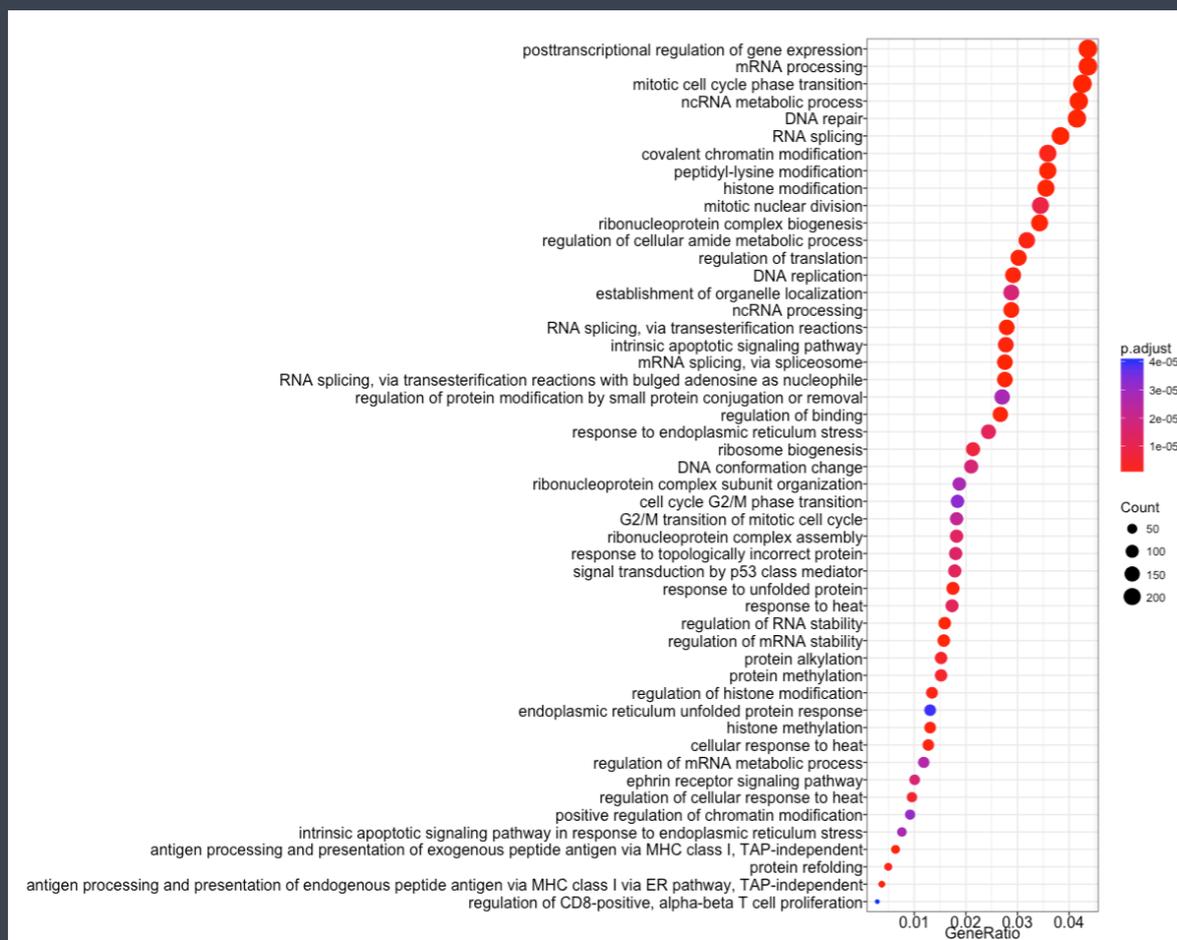However, they can loosely be categorized into three main types:

Over-representation Analysis

Compares the proportion of genes associated with a particular process/pathway in the list of differentially expressed genes to the proportion of genes associated with that pathway in a background list (genes tested for DE).
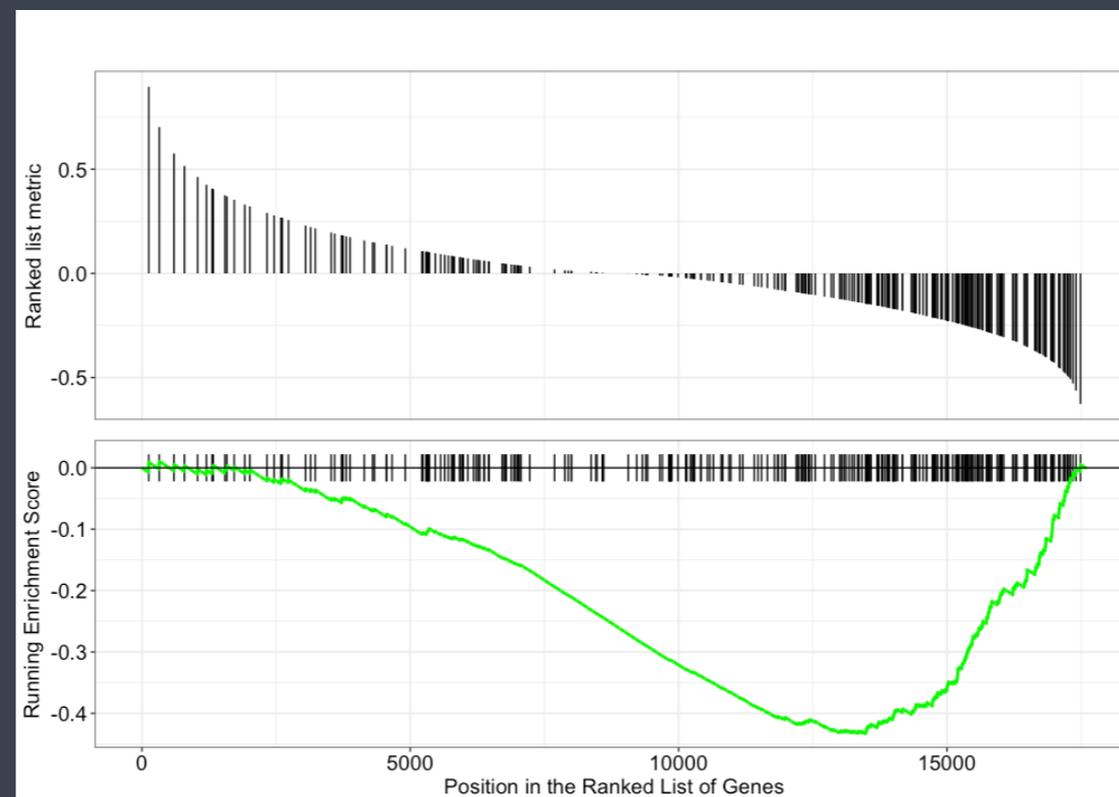
Over-representation Analysis

Over-representation analyses identify processes/pathways related to genes exhibiting larger changes in expression between the conditions.
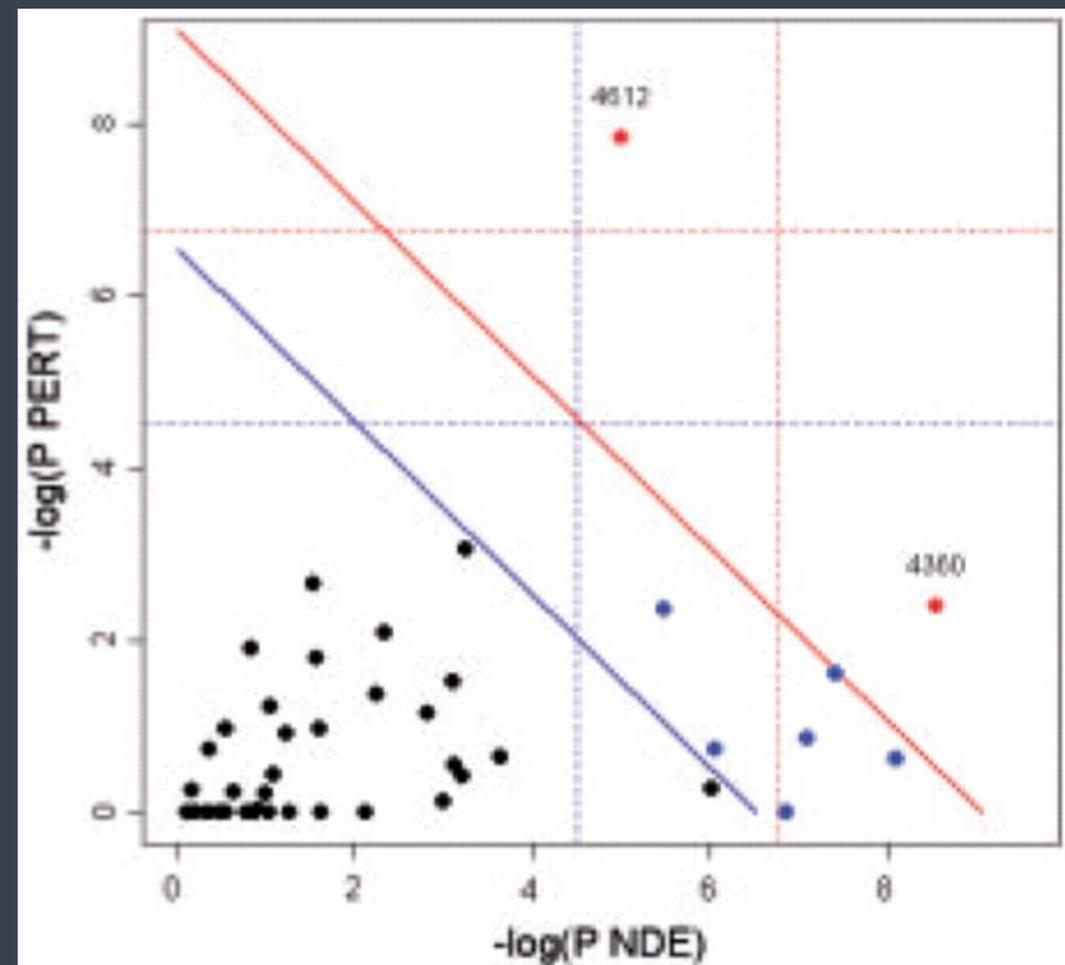
Functional Class Scoring

Functional class scoring tools, such as GSEA, are particularly helpful when there are few DE genes or when processes have genes exhibiting weaker but coordinated expression changes.

Functional Class Scoring

The GSEA method uses the log2 fold changes for ALL genes from the differential expression results to determine whether any biological pathways are enriched among the genes with positive or negative fold changes.
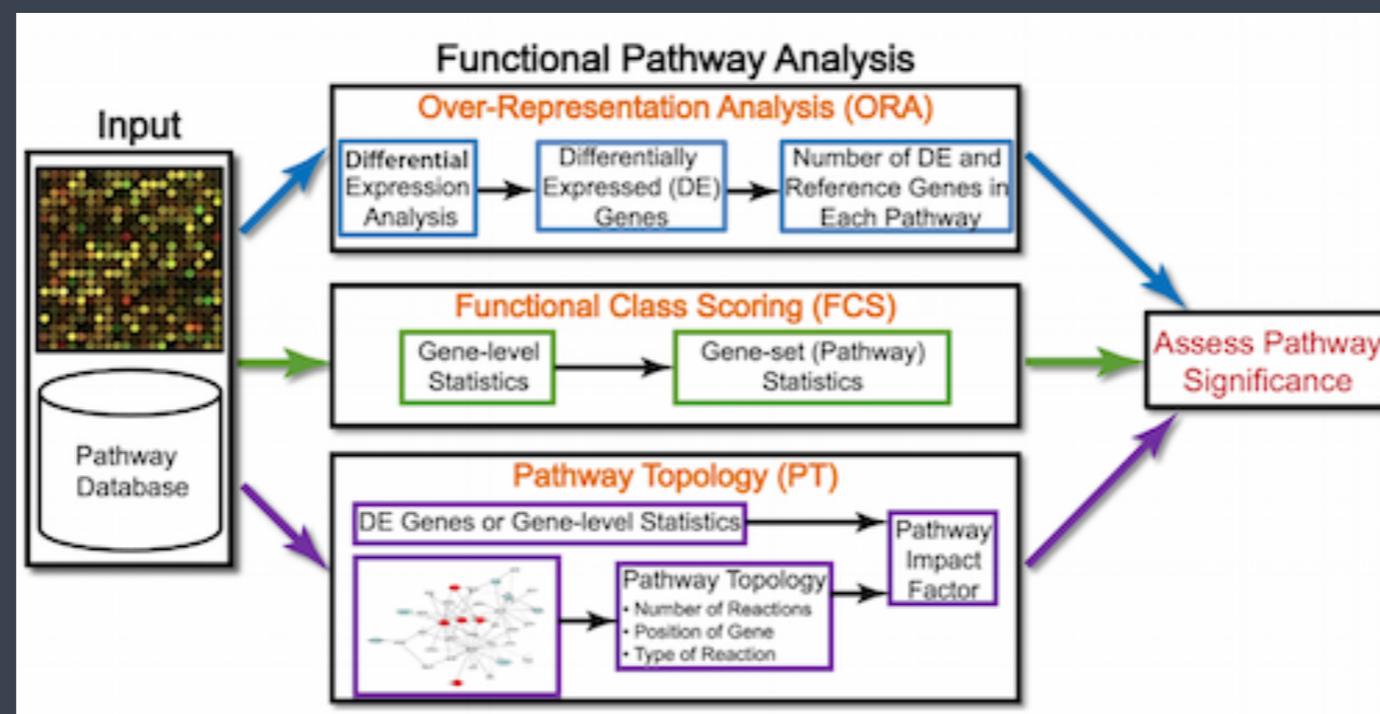
Pathway Topology Analysis

Pathway topology analysis tools often use gene interaction information along with the fold changes and adjusted p-values from differential expression analysis to identify significantly dysregulated pathways.

Tarca, AL, et. al. *Bioinformatics*, **25**(1):75–82

Functional Analysis Methods

There are many other methods for functional analysis, including co-expression clustering (WGCNA) and network-based analyses. Your desired output will help with the choice of a method.

Functional Pathway Analysis