

Gene Ontology and Pathway Enrichment Analysis Application

Introduction

This application, created using Shiny, performs Gene Ontology (GO) and KEGG Pathway enrichment analysis. GO and KEGG enrichment analyses help identify biological processes, molecular functions, and cellular components (GO) as well as relevant pathways (KEGG) that are statistically overrepresented in a set of genes. The application is designed to allow users to either analyze a predefined list of example genes or input their own gene list.

Functionality of the Application

1. Analysis: Allows users to set the analysis parameters. Users can select either the example gene list or input a custom list of genes. For custom genes, users specify the gene type (Gene Symbol or Ensembl ID). Once configured, clicking "Run Enrichment Analysis" initiates the analysis.

2. Results: Displays the results of the enrichment analysis across four sub-tabs:

- GO Plots: Visualizes the top enriched terms for Biological Process, Molecular Function, and Cellular Component ontologies as bar plots.

- Pathway Plots: Shows the results of KEGG pathway enrichment analysis as a bar plot.

- GO Tables: Provides detailed tables for enriched Biological Processes, Molecular Functions, and Cellular Components.

- oPathway Table: Shows the detailed table for enriched KEGG pathways.

Methods Used for Analysis

1. Gene List Conversion

The application accepts either Gene Symbols or Ensembl IDs. These IDs are converted to Entrez IDs, a standardized ID system used in many biological databases, via the `org.Hs.eg.db` package and the `mapIds` function from the `AnnotationDbi` package..

2. GO Enrichment Analysis

GO enrichment analysis is performed using the `enrichGO` function from the `clusterProfiler` package

- Biological Process (BP): Represents broad biological objectives, such as "cellular metabolic process" or "cell differentiation."

- Molecular Function (MF): Describes activities at the molecular level, like "DNA binding" or "enzyme regulator activity."

- Cellular Component (CC): Indicates locations, at the subcellular level, such as "nucleus" or "cytoplasm."

3. KEGG Pathway Enrichment Analysis

Using the `enrichKEGG` function from `clusterProfiler`, the analysis is set up for the organism 'hsa' (Homo sapiens).

Output Visualizations and Interactive Tables

The application provides results as both visual plots and tables:

- Plots: Displayed using barplot from enrichplot, these highlight the top 10 significant terms for each category.
- Tables: Rendered interactively using DT::datatable, the tables allow users to sort, filter, and page through the enriched terms.