Functional Enrichment Analysis

Gene set enrichment analysis.

Enrichment analysis is a computational method used to determine whether a particular set of genes, proteins, or other biological entities is enriched for specific functional annotations, such as gene ontology terms, biological pathways, or disease associations. It helps researchers gain insights into the functional characteristics and underlying biological processes associated with a given set of entities.

Enrichment analysis typically involves the following steps:

Input data: The analysis begins with a list of genes, proteins, or other entities of interest. This list is usually derived from experimental data, such as differentially expressed genes, a set of genes associated with a specific phenotype, or a group of proteins identified through a high-throughput experiment.

Background set: A background set is defined as a reference population against which the enrichment is assessed. It typically represents the entire genome, proteome, or a specific subset of entities with similar characteristics to the input set. The background set provides a context for determining whether the observed enrichment is statistically significant.

Functional annotation databases: Enrichment analysis relies on curated databases or resources that contain functional annotations associated with genes, proteins, or other biological entities. Examples of such databases include Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and others. These databases link biological entities to various functional categories, such as biological processes, molecular functions, cellular components, or pathways.

Statistical analysis: Various statistical methods are used to assess the enrichment of functional annotations within the input set compared to the background set. The choice of statistical test depends on the nature of the data and the specific enrichment analysis algorithm employed. Commonly used statistical tests include hypergeometric test, Fisher's exact test, chi-square test, and binomial test. These tests calculate a p-value or another statistical metric to determine the significance of the observed enrichment.

Multiple testing correction: Since enrichment analysis typically involves testing multiple functional categories simultaneously, multiple testing correction methods are applied to account for the possibility of false positives. The most commonly used correction method is the Benjamini-Hochberg procedure, which controls the false discovery rate (FDR).

Result interpretation: The output of enrichment analysis is a list of significantly enriched functional categories, pathways, or annotations associated with the input set. These results provide insights into the biological processes, molecular functions, or disease associations related to the analyzed entities. Researchers can explore and prioritize the enriched functional categories for further investigation or use them to generate hypotheses and guide experimental studies.

Enrichment analysis is widely used in genomics, transcriptomics, proteomics, and other 'omics' fields. It helps researchers uncover the functional implications of their data, identify underlying biological

processes, prioritize candidate genes or proteins, and gain a deeper understanding of the biological context of the studied entities.

Functional enrichment analysis is a computational method used to determine the biological functions, processes, or pathways that are overrepresented or significantly associated with a set of genes or proteins of interest. It helps researchers gain insights into the underlying biological meaning and potential roles of the genes or proteins under investigation.

Here's how functional enrichment analysis is typically performed:

Gene/protein set selection: The analysis begins with the selection of a set of genes or proteins for which functional enrichment is to be assessed. This set is typically obtained from experimental data, such as differentially expressed genes, a list of genes associated with a particular phenotype, or a group of proteins involved in a specific pathway.

Background selection: A background set is chosen as a reference to compare the gene/protein set of interest. The background set usually represents the entire genome or proteome, or a subset of genes/proteins with similar characteristics to the analyzed set. It provides a context for determining whether the observed enrichment is statistically significant.

Functional annotation databases: Functional enrichment analysis relies on curated databases that contain annotations linking genes/proteins to specific biological functions, processes, pathways, or molecular functions. Examples of such databases include Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and others.

Statistical analysis: Various statistical methods are employed to determine if any functional categories or terms are overrepresented or significantly enriched in the gene/protein set compared to the background. The choice of statistical test depends on the size of the gene/protein set, the nature of the data, and the specific enrichment analysis algorithm used. Commonly used statistical tests include hypergeometric test, Fisher's exact test, chi-square test, and binomial test.

Multiple testing correction: Since functional enrichment analysis typically involves testing multiple categories or terms simultaneously, multiple testing correction methods are applied to account for the possibility of false-positive results. The most commonly used correction method is the Benjamini-Hochberg procedure, which controls the false discovery rate (FDR).

Result interpretation: The output of functional enrichment analysis is a list of functional categories, terms, or pathways that are significantly enriched in the gene/protein set. These results provide insights into the biological processes or molecular functions associated with the analyzed set. Researchers can prioritize and explore the enriched terms to gain a better understanding of the underlying biological mechanisms or to generate hypotheses for further experimental validation.

Functional enrichment analysis is widely used in various fields of biological research, including genomics, transcriptomics, proteomics, and systems biology. It helps to uncover the functional significance of genes/proteins, unravel biological processes, identify potential disease-related pathways, and prioritize candidate genes/proteins for further investigation.

From:

<https://neurosurgerywiki.com/wiki/> - **Neurosurgery Wiki**

Permanent link:

https://neurosurgerywiki.com/wiki/doku.php?id=functional_enrichment_analysis

Last update: **2024/06/07 02:52**