R documentation

of 'man/analyze_mycobacterium_data.Rd' etc.

March 6, 2018

---

analyze_mycobacterium_data

Analyzes the data(CSV file with protien and gene information) from a user specified file and carries out GO and KEGG enrichment analysis and finally outputs the results as tables at the user-specified location. This function is essentially a wrapper around the rest of the functions provided, so that standard analysis can be done easily.

Description

Analyzes the data(CSV file with protien and gene information) from a user specified file and carries out GO and KEGG enrichment analysis and finally outputs the results as tables at the user-specified location. This function is essentially a wrapper around the rest of the functions provided, so that standard analysis can be done easily.

Usage

```
analyze_mycobacterium_data(input_filename, output_prefix, species,
  do_KEGG = TRUE, do_GO = FALSE)
```

Arguments

input_filename	The location of file you want to analyze. It should be in CSV format with a protein column and gene column
output_prefix	The output prefix you want added to all output files.
species	The species you are analyzing (either 'avium' or 'abscessum')
do_KEGG	Whether you want KEGG enrichment analysis (optional; default TRUE)
do_GO	Whether you want to do GO enrichment analysis (optional; default FALSE)

get_data	*import clusterProfiler import UniProt.ws Reads in data from the disk. This data should be in CSV format with the first column as the protein names and the second column as the gene names.*

Description

import clusterProfiler import UniProt.ws Reads in data from the disk. This data should be in CSV format with the first column as the protein names and the second column as the gene names.

Usage

```
get_data(filename)
```

Arguments

filename   The filename of the CSV file you want to read in.

Value

A dataframe with two columns. The first labeled "protein" the second "gene"

get_entrez_IDs	*Retrieves entrez IDs for a given species that are compatible for GO enrichment analysis.*

Description

Retrieves entrez IDs for a given species that are compatible for GO enrichment analysis.

Usage

```
get_entrez_IDs(data, species)
```

Arguments

data   The orignal protein data composed of a 'gene' column and a 'protein' column

species   The species you are analyzing (either 'avium' or 'abscessus')

Value

A vector of entrez IDs

get_KEGG_enrichment	*Calculates and returns results for KEGG enrichment given a species and corresponding UniProt IDs.*

Description

Calculates and returns results for KEGG enrichment given a species and corresponding UniProt IDs.

Usage

```
get_KEGG_enrichment(uniprot_IDs, species, pvalueCutoff = 0.05)
```

Arguments

uniprot_IDs	The corresponding UniProt IDs you want to test
species	the species you are analyzing (either 'avium' or 'abscessus')
pvalueCutoff	The p-value threshold that is considered significant (optional, default=0.05)

Value

A results object from clusterProfiler

get_uniprot_IDs	*Retrieves the UniProt IDs for a given species that are compatible for KEGG enrichment.*

Description

Retrieves the UniProt IDs for a given species that are compatible for KEGG enrichment.

Usage

```
get_uniprot_IDs(data, species)
```

Arguments

data	The orignal protein data composed of a 'gene' column and a 'protein' column
species	The species you are analyzing (either 'avium' or 'abscessus')

write_KEGG_enrichment	*Writes the results from the KEGG enrichment object to disk as a comma separated value format file.*

Description

Writes the results from the KEGG enrichment object to disk as a comma separated value format file.

Usage

```
write_KEGG_enrichment(KEGG_enrichment_object, filename)
```

Arguments

KEGG_enrichment_object

The clusterProfiler KEGG enrichment results

filename   The location where the file should be written

Value

No return value

Index