Package ‘clusterProfiler’

April 15, 2020

Type Package

Title statistical analysis and visualization of functional profiles for genes and gene clusters

Version 3.14.3

Maintainer Guangchuang Yu <guangchuangyu@gmail.com>

Description This package implements methods to analyze and visualize functional profiles (GO and KEGG) of gene and gene clusters.

Depends R (>= 3.4.0)

Imports AnnotationDbi, DOSE (>= 3.5.1), enrichplot (>= 0.99.7), ggplot2, GO.db, GOSemSim, magrittr, methods, plyr, qvalue, rvcheck, stats, tidyr, utils

Suggests AnnotationHub, dplyr, KEGG.db, knitr, org.Hs.eg.db, prettydoc, ReactomePA, testthat

VignetteBuilder knitr

ByteCompile true

License Artistic-2.0

URL https://guangchuangyu.github.io/software/clusterProfiler

BugReports https://github.com/GuangchuangYu/clusterProfiler/issues

biocViews Annotation, Clustering, GeneSetEnrichment, GO, KEGG, MultipleComparison, Pathways, Reactome, Visualization

RoxygenNote 7.0.2

git_url https://git.bioconductor.org/packages/clusterProfiler

git_branch RELEASE_3_10

git_last_commit d9752bc

git_last_commit_date 2020-01-08

Date/Publication 2020-04-14

Author Guangchuang Yu [aut, cre, cph]
(https://orcid.org/0000-0002-6485-8781),
Li-Gen Wang [ctb],
Giovanni Dall’Olio [ctb] (formula interface of compareCluster)
clusterProfiler-package

R topics documented:

clusterProfiler-package .................................................. 2
bit ..................................................... 3
bit_kegg .................................................. 3
browseKEGG ............................................. 4
buildGOmap .................................................. 4
compareCluster ............................................. 5
compareClusterResult-class ........................................ 6
DataSet ................................................... 6
dotplot.compareClusterResult-method ................................ 7
download_KEGG ............................................ 7
dropGO .................................................... 8
enrichDAVID ............................................. 9
enricher ................................................... 10
enrichGO .................................................. 11
enrichKEGG .............................................. 12
enrichMKEGG ............................................ 13
fortify.compareClusterResult ...................................... 14
getGOLevel ................................................ 15
Gff2GeneTable .............................................. 15
go2ont ..................................................... 16
go2term .................................................... 16
gofilter ..................................................... 17
groupGO ................................................... 17
groupGOResult-class ........................................ 18
GSEA ..................................................... 19
gseGO ..................................................... 20
gseKEGG .................................................. 21
gseMKEGG ................................................ 22
idType ..................................................... 23
KEGG_path2extid .......................................... 23
ko2name ................................................... 24
merge_result .............................................. 24
plotGOgraph ............................................... 25
plotting.clusterProfile ...................................... 25
read.gmt .................................................. 26
search_kegg_organism ......................................... 27
simplify ................................................... 27
uniprot_get ............................................... 29
viewKEGG ................................................ 29

Index 31
Description

statistical analysis and visualization of functional profiles for genes and gene clusters. The package implements methods to analyze and visualize functional profiles of gene and gene clusters.

bitr

Biological Id TRanslator

Usage

bitr(geneID, fromType, toType, OrgDb, drop = TRUE)

Arguments

geneID input gene id
fromType input id type
toType output id type
OrgDb annotation db
drop drop NA or not

Value
data.frame

Author(s)

Guangchuang Yu

bitr_kegg

convert biological ID using KEGG API

Usage

bitr_kegg(geneID, fromType, toType, organism, drop = TRUE)

Arguments

geneID input gene id
fromType input id type
toType output id type
organism supported organism, can be search using search_kegg_organism function
drop drop NA or not
browseKEGG

**Description**
open KEGG pathway with web browser

**Usage**
```r
browseKEGG(x, pathID)
```

**Arguments**
- `x`: an instance of enrichResult or gseaResult
- `pathID`: pathway ID

**Value**
`url`

**Author(s)**
Guangchuang Yu

buildGOmap

**Description**
building GO mapping files

**Usage**
```r
buildGOmap(gomap)
```

**Arguments**
- `gomap`: data.frame with two columns of GO and gene ID

**Details**
provided by a data.frame of GO (column 1) and gene (column 2) direct annotation this function will building gene to GO and GO to gene mapping, with directly and indirectly (ancestor GO term) annotation.
Value

data.frame, GO annotation with indirect annotation

Author(s)

Yu Guangchuang

compareCluster  Compare gene clusters functional profile

Description

Given a list of gene set, this function will compute profiles of each gene cluster.

Usage

compareCluster(geneClusters, fun = "enrichGO", data = "", ...)

Arguments

geneClusters  a list of entrez gene id. Alternatively, a formula of type Entrez~group
fun           One of "groupGO", "enrichGO", "enrichKEGG", "enrichDO" or "enrichPathway".
data          if geneClusters is a formula, the data from which the clusters must be extracted.
...            Other arguments.

Value

A clusterProfResult instance.

Author(s)

Guangchuang Yu https://guangchuangyu.github.io

See Also

compareClusterResult-class, groupGO enrichGO

Examples

## Not run:
data(gcSample)
xx <- compareCluster(gcSample, fun="enrichKEGG",
                     organism="hsa", pvalueCutoff=0.05)
as.data.frame(xx)
# plot(xx, type="dot", caption="KEGG Enrichment Comparison")

## formula interface
mydf <- data.frame(Entrez=c("1", "100", "1000", "100101467",
                           "100127206", "100128071"),
                    group = c("A", "A", "A", "B", "B", "B"),
                    othergroup = c("good", "good", "bad", "bad", "good", "bad"))
xx.formula <- compareCluster(Entrez~group, data=mydf, 
  fun='groupGO', OrgDb='org.Hs.eg.db')
as.data.frame(xx.formula)

## formula interface with more than one grouping variable
xx.formula.twogroups <- compareCluster(Entrez~group+othergroup, data=mydf, 
  fun='groupGO', OrgDb='org.Hs.eg.db')
as.data.frame(xx.formula.twogroups)

## End(Not run)

---

compareClusterResult-class

Class "compareClusterResult" This class represents the comparison result of gene clusters by GO categories at specific level or GO enrichment analysis.

Description

Class "compareClusterResult" This class represents the comparison result of gene clusters by GO categories at specific level or GO enrichment analysis.

Slots

compareClusterResult cluster comparing result
geneClusters a list of genes
fun one of groupGO, enrichGO and enrichKEGG
.call function call

Author(s)

Guangchuang Yu https://guangchuangyu.github.io

See Also

groupGOResult enrichResult compareCluster

---

DataSet

Datasets gcSample contains a sample of gene clusters.

Description

Datasets gcSample contains a sample of gene clusters.

Datasets kegg_species contains kegg species information
Description

dot plot method

Usage

```r
## S4 method for signature 'compareClusterResult'
dotplot(  
  object,  
  x = ~Cluster,  
  color = "p.adjust",  
  showCategory = 5,  
  split = NULL,  
  font.size = 12,  
  title = "",  
  by = "geneRatio",  
  includeAll = TRUE  
)
```

Arguments

- `object`: compareClusterResult object
- `x`: x variable
- `color`: one of pvalue or p.adjust
- `showCategory`: category numbers
- `split`: ONTOLOGY or NULL
- `font.size`: font size
- `title`: figure title
- `by`: one of geneRatio, Percentage or count
- `includeAll`: logical

Description

download the latest version of KEGG pathway/module

Usage

```r
download_KEGG(species, keggType = "KEGG", keyType = "kegg")
```
dropGO

Arguments

- **species**: species
- **keggType**: one of 'KEGG' or 'MKEGG'
- **keyType**: supported keyType, see bitr_kegg

Value

- list

Author(s)

Guangchuang Yu

---

dropGO  
dropGO

Description

drop GO term of specific level or specific terms (mostly too general).

Usage

```r
dropGO(x, level = NULL, term = NULL)
```

Arguments

- **x**: an instance of 'enrichResult' or 'compareClusterResult'
- **level**: GO level
- **term**: GO term

Value

- modified version of x

Author(s)

Guangchuang Yu
enrichDAVID

Description
enrichment analysis by DAVID

Usage
enrichDAVID(
  gene, idType = "ENTREZ_GENE_ID",
  universe, minGSSize = 10,
  maxGSSize = 500,
  annotation = "GOTERM_BP_FAT",
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  qvalueCutoff = 0.2,
  species = NA,
  david.user
)

Arguments
gene input gene
idType id type
universe background genes. If missing, the all genes listed in the database (eg TERM2GENE table) will be used as background.
minGSSize minimal size of genes annotated for testing
maxGSSize maximal size of genes annotated for testing
annotation david annotation
pvalueCutoff pvalue cutoff on enrichment tests to report
pAdjustMethod one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
qvalueCutoff qvalue cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted pvalues, ii) pvalueCutoff on adjusted pvalues and iii) qvalueCutoff on qvalues to be reported.
species species
david.user david user

Value
A enrichResult instance

Author(s)
Guangchuang Yu
Description

A universal enrichment analyzer

Usage

enricher(
  gene,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  universe,
  minGSize = 10,
  maxGSize = 500,
  qvalueCutoff = 0.2,
  TERM2GENE,
  TERM2NAME = NA
)

Arguments

gene a vector of gene id
pvalueCutoff pvalue cutoff on enrichment tests to report
pAdjustMethod one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe background genes. If missing, the all genes listed in the database (eg TERM2GENE table) will be used as background.
minGSize minimal size of genes annotated for testing
maxGSize maximal size of genes annotated for testing
qvalueCutoff qvalue cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted pvalues, ii) pvalueCutoff on adjusted pvalues and iii) qvalueCutoff on qvalues to be reported.
TERM2GENE user input annotation of TERM TO GENE mapping, a data.frame of 2 column with term and gene
TERM2NAME user input of TERM TO NAME mapping, a data.frame of 2 column with term and name

Value

A enrichResult instance

Author(s)

Guangchuang Yu
enrichGO

**GO Enrichment Analysis of a gene set.** Given a vector of genes, this function will return the enrichment GO categories after FDR control.

**Description**

GO Enrichment Analysis of a gene set. Given a vector of genes, this function will return the enrichment GO categories after FDR control.

**Usage**

```r
enrichGO(
  gene,
  OrgDb,
  keyType = "ENTREZID",
  ont = "MF",
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  universe,
  qvalueCutoff = 0.2,
  minGSSize = 10,
  maxGSSize = 500,
  readable = FALSE,
  pool = FALSE
)
```

**Arguments**

- **gene**: a vector of entrez gene id.
- **OrgDb**: OrgDb
- **keyType**: keytype of input gene
- **ont**: One of "BP", "MF", and "CC" subontologies, or "ALL" for all three.
- **pvalueCutoff**: pvalue cutoff on enrichment tests to report
- **pAdjustMethod**: one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
- **universe**: background genes. If missing, the all genes listed in the database (eg TERM2GENE table) will be used as background.
- **qvalueCutoff**: qvalue cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted pvalues, ii) pvalueCutoff on adjusted pvalues and iii) qvalueCutoff on qvalues to be reported.
- **minGSSize**: minimal size of genes annotated by Ontology term for testing.
- **maxGSSize**: maximal size of genes annotated for testing
- **readable**: whether mapping gene ID to gene Name
- **pool**: If ont='ALL', whether pool 3 GO sub-ontologies

**Value**

An enrichResult instance.
enrichKEGG

KEGG Enrichment Analysis of a gene set. Given a vector of genes, this function will return the enrichment KEGG categories with FDR control.

Usage

enrichKEGG(
  gene,
  organism = "hsa",
  keyType = "kegg",
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  universe,
  minGSSize = 10,
  maxGSSize = 500,
  qvalueCutoff = 0.2,
  use_internal_data = FALSE,
)

Arguments

gene a vector of entrez gene id.
organism supported organism listed in 'http://www.genome.jp/kegg/catalog/org_list.html'
keyType one of "kegg", 'ncbi-geneid', 'ncib-proteinid' and 'uniprot'
pvalueCutoff pvalue cutoff on enrichment tests to report
pAdjustMethod one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe background genes. If missing, the all genes listed in the database (eg TERM2GENE table) will be used as background.
**enrichMKEGG**

MinGSSize  minimal size of genes annotated by Ontology term for testing.
MaxGSSize  maximal size of genes annotated for testing
QvalueCutoff  qvalue cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted pvalues, ii) pvalueCutoff on adjusted pvalues and iii) qvalueCutoff on qvalues to be reported.
Use_internal_data  logical, use KEGG.db or latest online KEGG data

**Value**

A enrichResult instance.

**Author(s)**

Guangchuang Yu [http://ygc.name](http://ygc.name)

**See Also**

enrichResult-class, compareCluster

**Examples**

```r
data(geneList, package='DOSE')
de <- names(geneList)[1:100]
yy <- enrichMKEGG(de, pvalueCutoff=0.01)
head(yy)
```

**Description**

KEGG Module Enrichment Analysis of a gene set. Given a vector of genes, this function will return the enrichment KEGG Module categories with FDR control.

**Usage**

```r
enrichMKEGG(
gene,
organism = "hsa",
keyType = "kegg",
pvalueCutoff = 0.05,
pAdjustMethod = "BH",
universe,
minGSSize = 10,
maxGSSize = 500,
qvalueCutoff = 0.2
)
```
Arguments

gene  a vector of entrez gene id.
organism supported organism listed in ‘http://www.genome.jp/kegg/catalog/org_list.html’
keyType one of "kegg", 'ncbi-geneid', 'ncib-proteinid' and 'uniprot'
pvalueCutoff pvalue cutoff on enrichment tests to report
pAdjustMethod one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe background genes. If missing, the all genes listed in the database (eg TERM2GENE table) will be used as background.
minGSSize minimal size of genes annotated by Ontology term for testing.
maxGSSize maximal size of genes annotated for testing
qvalueCutoff qvalue cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted pvalues, ii) pvalueCutoff on adjusted pvalues and iii) qvalueCutoff on qvalues to be reported.

Value

A enrichResult instance.

Description

convert compareClusterResult to a data.frame that ready for plot

Usage

## S3 method for class 'compareClusterResult'
fortify(
  model,
  data,
  showCategory = 5,
  by = "geneRatio",
  split = NULL,
  includeAll = TRUE
)

Arguments

model compareClusterResult object
data not use here
showCategory category numbers
by one of geneRatio, Percentage or count
split ONTOLOGY or NULL
includeAll logical
getGOLevel

Value
data.frame

Author(s)
Guangchuang Yu

Description
query GOIDs at a specific level.

Usage
getGOLevel(ont, level)

Arguments
ont Ontology
level GO level

Value
a vector of GOIDs

Author(s)
Guangchuang Yu http://guangchuangyu.github.io

Gff2GeneTable

Description
read GFF file and build gene information table

Usage
Gff2GeneTable(gffFile, compress = TRUE)

Arguments
gffFile GFF file
compress compress file or not

Details
given a GFF file, this function extracts information from it and save it in working directory
Value

file save.

Author(s)

Yu Guangchuang

---

go2ont
go2ont

Description

convert goid to ontology (BP, CC, MF)

Usage

go2ont(goid)

Arguments

goid

a vector of GO IDs

Value
data.frame

Author(s)

Guangchuang Yu

---

go2term
go2term

Description

convert goid to descriptive term

Usage

go2term(goid)

Arguments

goid

a vector of GO IDs

Value
data.frame

Author(s)

Guangchuang Yu
**gofilter**

**Description**
filter GO enriched result at specific level

**Usage**
gofilter(x, level = 4)

**Arguments**
x: output from enrichGO or compareCluster
level: GO level

**Value**
updated object

**Author(s)**
Guangchuang Yu

---

**groupGO**

*Functional Profile of a gene set at specific GO level. Given a vector of genes, this function will return the GO profile at a specific level.*

**Description**
Functional Profile of a gene set at specific GO level. Given a vector of genes, this function will return the GO profile at a specific level.

**Usage**
groupGO(
gene,
OrgDb,
keyType = "ENTREZID",
ont = "CC",
level = 2,
readable = FALSE
)
Arguments

gene a vector of entrez gene id.
OrgDb OrgDb
keyType key type of input gene
ont One of "MF", "BP", and "CC" subontologies.
level Specific GO Level.
readable if readable is TRUE, the gene IDs will mapping to gene symbols.

Value
A groupGOResult instance.

Author(s)
Guangchuang Yu http://ygc.name

See Also
groupGOResult-class, compareCluster

Examples

data(gcSample)
yy <- groupGO(gcSample[[1]], 'org.Hs.eg.db', ont="BP", level=2)
head(summary(yy))
#plot(yy)

---

Class "groupGOResult" This class represents the result of functional Profiles of a set of gene at specific GO level.

Description
Class "groupGOResult" This class represents the result of functional Profiles of a set of gene at specific GO level.

Slots
result GO classification result
ontology Ontology
level GO level
organism one of "human", "mouse" and "yeast"
gene Gene IDs
readable logical flag of gene ID in symbol or not.

Author(s)
Guangchuang Yu https://guangchuangyu.github.io
See Also

compareClusterResult compareCluster groupGO

---

**Description**

A universal gene set enrichment analysis tool.

**Usage**

```r
GSEA(geneList, exponent = 1, nPerm = 1000, minGSSize = 10, maxGSSize = 500, pvalueCutoff = 0.05, pAdjustMethod = "BH", TERM2GENE, TERM2NAME = NA, verbose = TRUE, seed = FALSE, by = "fgsea")
```

**Arguments**

- `geneList`: order ranked geneList
- `exponent`: weight of each step
- `nPerm`: number of permutations
- `minGSSize`: minimal size of each geneSet for analyzing
- `maxGSSize`: maximal size of genes annotated for testing
- `pvalueCutoff`: pvalue cutoff
- `pAdjustMethod`: p value adjustment method
- `TERM2GENE`: user input annotation of TERM TO GENE mapping, a data.frame of 2 column with term and gene
- `TERM2NAME`: user input of TERM TO NAME mapping, a data.frame of 2 column with term and name
- `verbose`: logical
- `seed`: logical
- `by`: one of 'fgsea' or 'DOSE'

**Value**

gseaResult object
Description

Gene Set Enrichment Analysis of Gene Ontology

Usage

gseGO(
geneList,
ont = "BP",
OrgDb,
keyType = "ENTREZID",
exponent = 1,
nPerm = 1000,
minGSSize = 10,
maxGSSize = 500,
pvalueCutoff = 0.05,
pAdjustMethod = "BH",
verbose = TRUE,
seed = FALSE,
by = "fgsea"
)

Arguments

geneList order ranked geneList
ont one of "BP", "MF", and "CC" subontologies, or "ALL" for all three.
OrgDb OrgDb
keyType keytype of gene
exponent weight of each step
nPerm permutation numbers
minGSSize minimal size of each geneSet for analyzing
maxGSSize maximal size of genes annotated for testing
pvalueCutoff pvalue Cutoff
pAdjustMethod pvalue adjustment method
verbose print message or not
seed logical
by one of ‘fgsea’ or ‘DOSE’

Value
gseaResult object
Author(s)
Yu Guangchuang

gseKEGG

description
Gene Set Enrichment Analysis of KEGG

Usage

gseKEGG(
geneList,
organism = "hsa",
keyType = "kegg",
exponent = 1,
nPerm = 1000,
minGSSize = 10,
maxGSSize = 500,
pvalueCutoff = 0.05,
pAdjustMethod = "BH",
verbose = TRUE,
use_internal_data = FALSE,
seed = FALSE,
by = "fgsea"
)

Arguments

geneList order ranked geneList
organism supported organism listed in 'http://www.genome.jp/kegg/catalog/org_list.html'
keyType one of "kegg", 'ncbi-geneid', 'ncib-proteinid' and 'uniprot'
exponent weight of each step
nPerm permutation numbers
minGSSize minimal size of each geneSet for analyzing
maxGSSize maximal size of genes annotated for testing
pvalueCutoff pvalue Cutoff
pAdjustMethod pvalue adjustment method
verbose print message or not
use_internal_data logical, use KEGG.db or latest online KEGG data
seed logical
by one of 'fgsea' or 'DOSE'

Value
gseaResult object
Author(s)

Yu Guangchuang

Description

Gene Set Enrichment Analysis of KEGG Module

Usage

gseMKEGG(
geneList,
organism = "hsa",
keyType = "kegg",
exponent = 1,
nPerm = 1000,
minGSSize = 10,
maxGSSize = 500,
pvalueCutoff = 0.05,
pAdjustMethod = "BH",
verbose = TRUE,
seed = FALSE,
by = "fgsea"
)

Arguments

geneList order ranked geneList
organism supported organism listed in 'http://www.genome.jp/kegg/catalog/org_list.html'
keyType one of "kegg", 'ncbi-geneid', 'ncib-proteinid' and 'uniprot'
exponent weight of each step
nPerm permutation numbers
minGSSize minimal size of each geneSet for analyzing
maxGSSize maximal size of genes annotated for testing
pvalueCutoff pvalue Cutoff
pAdjustMethod pvalue adjustment method
verbose print message or not
seed logical
by one of 'fgsea' or 'DOSE'

Value

gseaResult object

Author(s)

Yu Guangchuang
**idType**

**Description**
list ID types supported by annoDb

**Usage**

```r
idType(OrgDb = "org.Hs.eg.db")
```

**Arguments**

- **OrgDb**: annotation db

**Value**
character vector

**Author(s)**
Guangchuang Yu

---

**KEGG_path2extid**

**Description**
query all genes in a KEGG pathway or module

**Usage**

```r
KEGG_path2extid(
    keggID,
    species = sub("\d+$", "", keggID),
    keggType = "Path",
    keyType = "kegg"
)
```

**Arguments**

- **keggID**: KEGG ID, path or module ID
- **species**: species
- **keggType**: one of 'Path' or 'Module'
- **keyType**: KEGG gene type, one of "ncbi-proteinid", "ncbi-geneid", "uniprot", or "kegg"

**Value**
extid vector
**ko2name**

**Description**
convert ko ID to descriptive name

**Usage**
ko2name(ko)

**Arguments**
ko ko ID

**Value**
data.frame

**Author(s)**
guangchuang yu

---

**merge_result**

**Description**
merge a list of enrichResult objects to compareClusterResult

**Usage**
merge_result(enrichResultList)

**Arguments**
enrichResultList a list of enrichResult objects

**Value**
a compareClusterResult instance

**Author(s)**
Guangchuang Yu
plotGOgraph

**Description**

plot GO graph

**Usage**

```r
plotGOgraph(
  x,
  firstSigNodes = 10,
  useInfo = "all",
  sigForAll = TRUE,
  useFullNames = TRUE,
  ...
)
```

**Arguments**

- `x`: output of enrichGO or gseGO
- `firstSigNodes`: number of significant nodes (rectangle nodes in the graph)
- `useInfo`: additional info
- `sigForAll`: if TRUE the score/p-value of all nodes in the DAG is shown, otherwise only score will be shown
- `useFullNames`: logical
- `...`: additional parameter of showSigOfNodes, please refer to topGO

**Value**

GO DAG graph

**Author(s)**

Guangchuang Yu

plotting.clusterProfile

**Description**

Internal plot function for plotting compareClusterResult
Usage

```r
plotting.clusterProfile(
    clProf.reshape.df,
    x = "Cluster",
    type = "dot",
    colorBy = "p.adjust",
    by = "geneRatio",
    title = "",
    font.size = 12
)
```

Arguments

- `clProf.reshape.df`: data frame of compareCluster result
- `x`: x variable
- `type`: one of dot and bar
- `colorBy`: one of pvalue or p.adjust
- `by`: one of percentage and count
- `title`: graph title
- `font.size`: graph font size

Value

ggplot object

Author(s)

Guangchuang Yu [https://guangchuangyu.github.io](https://guangchuangyu.github.io)

---

Description

parse gmt file to a data.frame

Usage

```r
read.gmt(gmtfile)
```

Arguments

- `gmtfile`: gmt file

Value

data.frame

Author(s)

Guangchuang Yu
search_kegg_organism

Description

search kegg organism, listed in http://www.genome.jp/kegg/catalog/org_list.html

Usage

search_kegg_organism(str, by = "scientific_name", ignore.case = FALSE)

Arguments

str string
by one of 'kegg.code', 'scientific_name' and 'common_name'
ignore.case TRUE or FALSE

Value
data.frame

Author(s)
Guangchuang Yu

simplify

simplify method

Description

simplify output from enrichGO and gseGO by removing redundancy of enriched GO terms
simplify output from compareCluster by removing redundancy of enriched GO terms

Usage

## S4 method for signature 'enrichResult'
simplify(
x,  
cutoff = 0.7,  
by = "p.adjust",  
select_fun = min,  
measure = "Wang",  
semData = NULL  
)

## S4 method for signature 'gseaResult'
simplify(
x,  
cutoff = 0.7,  
by = "p.adjust",  
select_fun = min,  
measure = "Wang",  
semData = NULL  
)
Arguments

- `x`: output of enrichGO
- `cutoff`: similarity cutoff
- `by`: feature to select representative term, selected by `select_fun` function
- `select_fun`: function to select feature passed by `by` parameter
- `measure`: method to measure similarity
- `semData`: GOSemSimDATA object

Value

- updated enrichResult object
- updated compareClusterResult object

Author(s)

Guangchuang Yu

Gwang-Jin Kim and Guangchuang Yu

Guangchuang Yu

References

issue #28 [https://github.com/GuangchuangYu/clusterProfiler/issues/28](https://github.com/GuangchuangYu/clusterProfiler/issues/28)

issue #162 [https://github.com/GuangchuangYu/clusterProfiler/issues/162](https://github.com/GuangchuangYu/clusterProfiler/issues/162)
**uniprot_get**

---

**Description**

Retrieve annotation data from uniprot

**Usage**

`uniprot_get(taxID)`

**Arguments**

- `taxID`: taxonomy ID

**Value**

gene table data frame

**Author(s)**

guangchuang yu

---

**viewKEGG**

`viewKEGG` function is for visualize KEGG pathways works with enrichResult object to visualize enriched KEGG pathway

---

**Description**

`viewKEGG` function is for visualize KEGG pathways works with enrichResult object to visualize enriched KEGG pathway

**Usage**

```r
viewKEGG(
    obj,
    pathwayID,
    foldChange,
    color.low = "green",
    color.high = "red",
    kegg.native = TRUE,
    out.suffix = "clusterProfiler"
)
```
Arguments

- obj: enrichResult object
- pathwayID: pathway ID or index
- foldChange: fold change values
- color.low: color of low foldChange genes
- color.high: color of high foldChange genes
- kegg.native: logical
- out.suffix: suffix of output file

References

Index

*Topic **classes**
  - compareClusterResult-class, 6
  - groupGOResult-class, 18

*Topic **datasets**
  - DataSet, 6

*Topic **manip**
  - compareCluster, 5
  - enrichGO, 11
  - enrichKEGG, 12
  - groupGO, 17
  - bitr, 3
  - bitr_kegg, 3
  - browseKEGG, 4
  - buildGOnmap, 4
  - clusterProfiler
    - (clusterProfiler-package), 2
    - clusterProfiler-package, 2
    - compareCluster, 5, 6, 12, 13, 18, 19
    - compareClusterResult, 19
    - compareClusterResult-class, 6
  - DataSet, 6
  - dotplot,compareClusterResult,ANY-method
    - (dotplot,compareClusterResult-method), 7
  - dotplot,compareClusterResult-method, 7
  - download_KEGG, 7
  - dropGO, 8
  - enrichDAVID, 9
  - enricher, 10
  - enrichGO, 5, 11
  - enrichKEGG, 12
  - enrichMKEGG, 13
  - enrichResult, 6
  - fortify.compareClusterResult, 14
  - gcSample (DataSet), 6
  - getGOlevel, 15
  - Gff2GeneTable, 15
  - go2ont, 16
  - go2term, 16
  - gofilter, 17
  - groupGO, 5, 17, 19
  - groupGOREsult, 6
  - groupGOREsult-class, 18
  - GSEA, 19
  - gseGO, 20
  - gseKEGG, 21
  - gseMKEGG, 22
  - idType, 23
  - KEGG_path2extid, 23
  - kegg_species (DataSet), 6
  - ko2name, 24
  - merge_result, 24
  - plot,compareClusterResult-method
    - (compareClusterResult-class), 6
  - plotGOgraph, 25
  - plotting.clusterProfile, 25
  - read.gmt, 26
  - search_kegg_organism, 27
  - show,compareClusterResult-method
    - (compareClusterResult-class), 6
  - show,groupGOREsult-method
    - (groupGOREsult-class), 18
  - simplify, 27
  - simplify,compareClusterResult-method
    - (simplify), 27
  - simplify,enrichResult-method
    - (simplify), 27
  - simplify,gseaResult-method(simplify), 27
  - summary,compareClusterResult-method
    - (compareClusterResult-class), 6
  - uniprot_get, 29
  - viewKEGG, 29