Package ‘clusterProfiler’

November 21, 2020

Type Package

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

Version 3.18.0

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, downloader, DOSE (>= 3.13.1), dplyr, enrichplot (>= 1.9.3), GO.db, GOSemSim, magrittr, methods, plyr, qvalue, rlang, rvcheck, stats, tidyr, utils

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

VignetteBuilder knitr

ByteCompile true

License Artistic-2.0


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

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

RoxygenNote 7.1.1

git_url https://git.bioconductor.org/packages/clusterProfiler
git_branch RELEASE_3_12
git_last_commit 064a6e6
git_last_commit_date 2020-10-27

Date/Publication 2020-11-20

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)
R topics documented:

clusterProfiler-package .................................................. 2
bitr ................................................................. 3
bitr_kegg ........................................................... 3
browseKEGG .......................................................... 4
buildGOmap ............................................................ 4
compareCluster .......................................................... 5
DataSet ................................................................. 6
download_KEGG ........................................................ 6
dropGO ................................................................. 7
enrichDAVID ........................................................... 7
enricher ................................................................. 8
enrichGO ............................................................... 9
enrichKEGG ............................................................ 10
enrichMKEGG ........................................................... 12
enrichWP ............................................................... 13
get_wp_organisms ....................................................... 13
Gff2GeneTable .......................................................... 14
go2ont ................................................................. 14
go2term ................................................................. 15
gofilter ................................................................. 15
groupGO ............................................................... 16
groupGOResult-class ................................................... 17
GSEA ................................................................. 17
gseGO ................................................................. 18
gseKEGG ............................................................... 19
gseMKEGG ............................................................. 20
gseWP ................................................................. 21
idType ................................................................. 22
KEGG_path2extid ....................................................... 22
ko2name ............................................................... 23
merge_result ........................................................... 24
plotGOgraph ............................................................ 24
read.gmt .............................................................. 25
search_kegg_organism .................................................. 25
simplify ............................................................... 26
uniprot_get ........................................................... 27
viewKEGG ............................................................. 28

Index 29

clusterProfiler-package

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

Description

The package implements methods to analyze and visualize functional profiles of gene and gene clusters.
Description

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

Description

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

Value

data.frame
browseKEGG

**Description**

open KEGG pathway with web browser

**Usage**

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**

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)**

Guangchuang Yu
compareCluster

Author(s)
Yu Guangchuang

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](https://guangchuangyu.github.io)

See Also
`compareClusterResult-class`, `groupGO`, `enrichGO`

Examples
```r
## 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)

### Dataset

**Datasets gcSample** contains a sample of gene clusters.

**Datasets kegg_species** contains kegg species information

### download_REEGG

**download_REEGG**

**Description**

download the latest version of KEGG pathway/module

**Usage**

download_REEGG(species, keggType = "KEGG", keyType = "kegg")

**Arguments**

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

**Value**

list

**Author(s)**

Guangchuang Yu
**dropGO**

**Description**

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

**Usage**

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 (e.g., 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 adjusted p-value cutoff on enrichment tests to report
pAdjustMethod one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
qvalueCutoff q-value cutoff on enrichment tests to report as significant. Tests must pass i) pvalueCutoff on unadjusted p-values, ii) pvalueCutoff on adjusted p-values and iii) qvalueCutoff on q-values to be reported.

Value

A enrichResult instance

Author(s)

Guangchuang Yu

Description

A universal enrichment analyzer

Usage

enricher(
  gene,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  universe,
  minGSSize = 10,
  maxGSSize = 500,
  qvalueCutoff = 0.2,
  TERM2GENE,
  TERM2NAME = NA
)
enrichGO

Arguments

gene a vector of gene id
pvalueCutoff adjusted 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 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.
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

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

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
enrichGO(********
enrichKEGG

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

Description

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

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 adjusted 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 (e.g. 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.

Author(s)

Guangchuang Yu https://guangchuangyu.github.io

See Also

enrichResult-class, compareCluster

Examples

## Not run:
data(geneList, package = "DOSE")
de <- names(geneList)[1:100]
yy <- enrichGO(de, ‘org.Hs.eg.db’, ont="BP", pvalueCutoff=0.01)
head(yy)

## End(Not run)
enrichKEGG

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 adjusted 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.
use_internal_data logical, use KEGG.db or latest online KEGG data

Value

A enrichResult instance.

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

See Also

enrichResult-class, compareCluster

Examples

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

**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.

### 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**: adjusted 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.
**enrichWP**

**Description**

ORA analysis for WikiPathways

**Usage**

```r
enrichWP(gene, organism, ...)
```

**Arguments**

- `gene`: a vector of entrez gene id
- `organism`: supported organisms, which can be accessed via the `get_wp_organisms()` function
- `...`: additional parameters, see also the parameters supported by the `enricher()` function

**Details**

This function performs over-representation analysis using WikiPathways

**Value**

A `enrichResult` instance

**Author(s)**

Guangchuang Yu

---

**get_wp_organisms**

**Description**

list supported organism of WikiPathways

**Usage**

```r
get_wp_organisms()
```

**Details**

This function extracts information from 'http://data.wikipathways.org/current/gmt/' and lists all supported organisms

**Value**

supported organism list
Author(s)
Guangchuang Yu

Gff2GeneTable

Description
read GFF file and build gene information table

Usage
Gff2GeneTable(gffFile, compress = TRUE)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gffFile</td>
<td>GFF file</td>
</tr>
<tr>
<td>compress</td>
<td>compress file or not</td>
</tr>
</tbody>
</table>

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

Value
file save.

Author(s)
Guangchuang Yu

go2ont

Description
convert goid to ontology (BP, CC, MF)

Usage
go2ont(goid)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>goid</td>
<td>a vector of GO IDs</td>
</tr>
</tbody>
</table>

Value
data.frame

Author(s)
Guangchuang Yu
**go2term**

**Description**
convert goid to descriptive term

**Usage**
g2o(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://guangchuangyu.github.io

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)
**groupGOResult-class**

**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](https://guangchuangyu.github.io)

**See Also**

`compareClusterResult`, `compareCluster`, `groupGO`

---

**GSEA**

**GSEA**

**Description**

a universal gene set enrichment analysis tools

**Usage**

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

Arguments

geneList: order ranked geneList
exponent: weight of each step
minGSSize: minimal size of each geneSet for analyzing
maxGSSize: maximal size of genes annotated for testing
eps: This parameter sets the boundary for calculating the p value.
pvalueCutoff: adjusted 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'
... other parameter

Value
gseaResult object

Author(s)

Guangchuang Yu

Description

Gene Set Enrichment Analysis of Gene Ontology

Usage

gseGO(
  geneList, 
  ont = "BP", 
  OrgDb, 
  keyType = "ENTREZID", 
  exponent = 1, 
  minGSSize = 10, 
  maxGSSize = 500, 
  eps = 1e-10, 
  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
- `minGSSize` minimal size of each geneSet for analyzing
- `maxGSSize` maximal size of genes annotated for testing
- `eps` This parameter sets the boundary for calculating the p value.
- `pvalueCutoff` pvalue Cutoff
- `pAdjustMethod` pvalue adjustment method
- `verbose` print message or not
- `seed` logical
- `by` one of 'fgsea' or 'DOSE'
- `...` other parameter

**Value**

gseaResult object

**Author(s)**

Yu Guangchuang

---

**gseKEGG**

geneSet Enrichment Analysis of KEGG

**Usage**

gseKEGG(
  geneList,
  organism = "hsa",
  keyType = "kegg",
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  eps = 1e-10,
  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
- minGSSize: minimal size of each geneSet for analyzing
- maxGSSize: maximal size of genes annotated for testing
- eps: This parameter sets the boundary for calculating the p value.
- 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'
- ... other parameter

Value

- gseaResult object

Author(s)

- Yu Guangchuang

Description

Gene Set Enrichment Analysis of KEGG Module

Usage

```r
gseMKEGG(
  geneList, 
  organism = "hsa", 
  keyType = "kegg", 
  exponent = 1, 
  minGSSize = 10, 
  maxGSSize = 500, 
  eps = 1e-10, 
  pvalueCutoff = 0.05, 
  pAdjustMethod = "BH", 
  verbose = TRUE, 
  seed = FALSE, 
  by = "fgsea", 
  ...
)
```
gseWP

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
minGSSize, minimal size of each geneSet for analyzing
maxGSSize, maximal size of genes annotated for testing
eps, This parameter sets the boundary for calculating the p value.
pvalueCutoff, pvalue Cutoff
pAdjustMethod, pvalue adjustment method
verbose, print message or not
seed, logical
by, one of `fgsea` or 'DOSE'
...
other parameter

Value

gseaResult object

Author(s)

Yu Guangchuang

Description

GSEA analysis for WikiPathways

Usage

gseWP(geneList, organism, ...)

Arguments

geneList, ranked gene list
organism, supported organisms, which can be accessed via the get_wp_organisms() function
...
additional parameters, see also the parameters supported by the enricher() function

Details

This function performs GSEA using WikiPathways
**Value**

A gseaResult instance

**Author(s)**

Guangchuang Yu

---

**idType**

**Value**

A gseaResult instance

**Author(s)**

Guangchuang Yu

---

**Description**

list ID types supported by annoDb

**Usage**

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

**Arguments**

0rgDb annotation db

**Value**

character vector

**Author(s)**

Guangchuang Yu

---

**KEGG_path2extid**

**Value**

query all genes in a KEGG pathway or module

**Usage**

KEGG_path2extid(
    keggID,
    species = sub("\d+$", ",", keggID),
    keggType = "Path",
    keyType = "kegg"
)
**ko2name**

**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

**Author(s)**

guangchuang yu

---

**ko2name**

**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
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
**read.gmt**

**Value**

GO DAG graph

**Author(s)**

Guangchuang Yu

---

**Description**

parse gmt file to a data.frame

**Usage**

read.gmt(gmtfile)

read.gmt.wp(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

---

**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**

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

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

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

### 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)

---

### Description
Retrieve annotation data from uniprot

### Usage
```r
uniprot_get(taxID)
```

### Arguments
- `taxID`: taxonomy ID

### Value
- gene table data frame

### Author(s)
- guangchuang yu
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

* classes
  groupGOResult-class, 17

* datasets
  DataSet, 6

* manip
  compareCluster, 5
  enrichGO, 9
  enrichKEGG, 10
  groupGO, 16
  bitr, 3
  bitr_kegg, 3
  browseKEGG, 4
  buildGOMap, 4
  gcSample (DataSet), 6
  gff2GeneTable, 14
  go2ont, 14
  go2term, 15
  gofilter, 15
  groupGO, 5, 16, 17
  groupGOResult-class, 17
  GSEA, 17
  gseGO, 18
  gseKEGG, 19
  gseMKEGG, 20
  gseWP, 21
  idType, 22
  KEGG_path2extid, 22
  kegg_species (DataSet), 6
  ko2name, 23
  merge_result, 24
  plotGOgraph, 24
  read.gmt, 25
  search_kegg_organism, 25
  show, groupGOResult-method (groupGOResult-class), 17
  simplify, 26
  simplify, compareClusterResult-method (simplify), 26
  simplify, enrichResult-method (simplify), 26
  simplify, gseaResult-method (simplify), 26
  uniprot_get, 27
  viewKEGG, 28