

# Package ‘clusterProfiler’

December 5, 2021

**Type** Package

**Title** A universal enrichment tool for interpreting omics data

**Version** 4.2.0

**Maintainer** Guangchuang Yu <guangchuangyu@gmail.com>

**Description** This package supports functional characteristics of both coding and non-coding genomics data for thousands of species with up-to-date gene annotation. It provides a universal interface for gene functional annotation from a variety of sources and thus can be applied in diverse scenarios. It provides a tidy interface to access, manipulate, and visualize enrichment results to help users achieve efficient data interpretation. Datasets obtained from multiple treatments and time points can be analyzed and compared in a single run, easily revealing functional consensus and differences among distinct conditions.

**Depends** R (>= 3.5.0)

**Imports** AnnotationDbi, downloader, DOSE (>= 3.13.1), dplyr, enrichplot (>= 1.9.3), GO.db, GOSemSim, magrittr, methods, plyr, qvalue, rlang, stats, tidy, utils, yulab.utils

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

**VignetteBuilder** knitr

**ByteCompile** true

**License** Artistic-2.0

**URL** <https://yulab-smu.top/biomedical-knowledge-mining-book/> (docs),  
<https://doi.org/10.1016/j.xinn.2021.100141> (paper)

**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\_14

**git\_last\_commit** aae0e16

**git\_last\_commit\_date** 2021-10-26

**Date/Publication** 2021-12-05

**Author** Guangchuang Yu [aut, cre, cph]  
 (<<https://orcid.org/0000-0002-6485-8781>>),  
 Li-Gen Wang [ctb],  
 Erqiang Hu [ctb],  
 Meijun Chen [ctb],  
 Giovanni Dall'Olio [ctb]

## R topics documented:

|                                   |    |
|-----------------------------------|----|
| clusterProfiler-package . . . . . | 3  |
| bitr . . . . .                    | 3  |
| bitr_kegg . . . . .               | 4  |
| browseKEGG . . . . .              | 4  |
| buildGOMap . . . . .              | 5  |
| compareCluster . . . . .          | 6  |
| DataSet . . . . .                 | 7  |
| download_KEGG . . . . .           | 7  |
| dropGO . . . . .                  | 8  |
| enrichDAVID . . . . .             | 8  |
| enricher . . . . .                | 9  |
| enrichGO . . . . .                | 10 |
| enrichKEGG . . . . .              | 12 |
| enrichMKEGG . . . . .             | 13 |
| enrichWP . . . . .                | 14 |
| get_wp_organisms . . . . .        | 15 |
| Gff2GeneTable . . . . .           | 15 |
| go2ont . . . . .                  | 16 |
| go2term . . . . .                 | 16 |
| gofilter . . . . .                | 17 |
| groupGO . . . . .                 | 17 |
| groupGOResult-class . . . . .     | 18 |
| GSEA . . . . .                    | 19 |
| gseGO . . . . .                   | 20 |
| gseKEGG . . . . .                 | 21 |
| gseMKEGG . . . . .                | 22 |
| gseWP . . . . .                   | 23 |
| idType . . . . .                  | 24 |
| ko2name . . . . .                 | 25 |
| merge_result . . . . .            | 25 |
| plotGOgraph . . . . .             | 26 |
| read.gmt . . . . .                | 27 |
| search_kegg_organism . . . . .    | 27 |
| simplify . . . . .                | 28 |
| uniprot_get . . . . .             | 29 |

**Index**

**30**

---

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.

---

bitr

*bitr*

---

### 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

---

|           |                  |
|-----------|------------------|
| bitr_kegg | <i>bitr_kegg</i> |
|-----------|------------------|

---

**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

**Author(s)**

Guangchuang Yu

---

|            |                   |
|------------|-------------------|
| browseKEGG | <i>browseKEGG</i> |
|------------|-------------------|

---

**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 | <i>buildGOMap</i> |
|------------|-------------------|

---

**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)**

Yu Guangchuang

---

|                |   |
|----------------|---|
| compareCluster | <i>Compare gene clusters functional profile</i> |
|----------------|---|

---

### 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 <code>Entrez~group</code> |
| 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='groupG0', OrgDb='org.Hs.eg.db')
as.data.frame(xx.formula.twogroups)

## End(Not run)
```

---

|         |  |
|---------|--|
| DataSet | <i>Datasets gcSample contains a sample of gene clusters.</i> |
|---------|--|

---

### Description

Datasets gcSample contains a sample of gene clusters.

Datasets kegg\_species contains kegg species information

Datasets DE\_GSE8057 contains differential expressed genes obtained from GSE8057 dataset

---

|               |                      |
|---------------|----------------------|
| download_KEGG | <i>download_KEGG</i> |
|---------------|----------------------|

---

### Description

download the latest version of KEGG pathway/module

### Usage

```
download_KEGG(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 | <i>dropGO</i> |
|--------|---------------|

---

**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 | <i>enrichDAVID</i> |
|-------------|--------------------|

---

**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  | adjusted 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

---

|          |                 |
|----------|-----------------|
| enricher | <i>enricher</i> |
|----------|-----------------|

---

**Description**

A universal enrichment analyzer

**Usage**

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

**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

---

|          |   |
|----------|---|
| enrichGO | <i>GO Enrichment Analysis of a gene set. Given a vector of genes, this function will return the enrichment GO categories after FDR control.</i> |
|----------|---|

---

**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(
  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  | 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.  |
| 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 | <i>KEGG Enrichment Analysis of a gene set. Given a vector of genes, this function will return the enrichment KEGG categories with FDR control.</i> |
|------------|--|

---

### Description

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 ' <a href="http://www.genome.jp/kegg/catalog/org_list.html">http://www.genome.jp/kegg/catalog/org_list.html</a> '   |
| 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**

```
## Not run:
data(geneList, package='DOSE')
de <- names(geneList)[1:100]
yy <- enrichKEGG(de, pvalueCutoff=0.01)
head(yy)

## End(Not run)
```

---

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**

```
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 | <i>enrichWP</i> |
|----------|-----------------|

---

**Description**

ORA analysis for WikiPathways

**Usage**

```
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 | <i>get_wp_organism</i> |
|------------------|------------------------|

---

**Description**

list supported organism of WikiPathways

**Usage**

```
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 | <i>Gff2GeneTable</i> |
|---------------|----------------------|

---

**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`*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>

**See Also**

[compareClusterResult](#) [compareCluster](#) [groupGO](#)

---

GSEA

*GSEA*

---

**Description**

a universal gene set enrichment analysis tools

**Usage**

```
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

---

gseGO

*gseGO*

---

**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

*gseKEGG*

---

**Description**

Gene Set 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

---

gseMKEGG

*gseMKEGG*

---

**Description**

Gene Set Enrichment Analysis of KEGG Module

**Usage**

```
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",
    ...
)

```

### 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

---

gseWP

*gseWP*

---

### Description

GSEA analysis for WikiPathways

### Usage

```
gseWP(geneList, organism, ...)
```

**Arguments**

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

**Details**

This function performs GSEA using WikiPathways

**Value**

A `gseaResult` instance

**Author(s)**

Guangchuang Yu

---

|                     |               |
|---------------------|---------------|
| <code>idType</code> | <i>idType</i> |
|---------------------|---------------|

---

**Description**

list ID types supported by `annoDb`

**Usage**

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

**Arguments**

|       |               |
|-------|---------------|
| OrgDb | annotation db |
|-------|---------------|

**Value**

character vector

**Author(s)**

Guangchuang Yu



---

|         |                |
|---------|----------------|
| ko2name | <i>ko2name</i> |
|---------|----------------|

---

**Description**

convert ko ID to descriptive name

**Usage**

```
ko2name(ko)
```

**Arguments**

ko ko ID

**Value**

data.frame

**Author(s)**

guangchuang yu

---

|              |                     |
|--------------|---------------------|
| merge_result | <i>merge_result</i> |
|--------------|---------------------|

---

**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

*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                                  |

**Value**

GO DAG graph

**Author(s)**

Guangchuang Yu

---

|          |                 |
|----------|-----------------|
| read.gmt | <i>read.gmt</i> |
|----------|-----------------|

---

**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 | <i>search_kegg_organism</i> |
|----------------------|-----------------------------|

---

**Description**

search kegg organism, listed in [http://www.genome.jp/kegg/catalog/org\\_list.html](http://www.genome.jp/kegg/catalog/org_list.html)

**Usage**

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

**Arguments**

|                   |  |
|-------------------|--|
| str               | string   |
| by                | one of 'kegg.code', 'scientific_name' and 'common_name'  |
| ignore.case       | TRUE or FALSE  |
| use_internal_data | logical, use kegg_species.rda or latest online KEGG data |

**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  
)  
  
## S4 method for signature 'compareClusterResult'  
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>  
 issue #162 <https://github.com/GuangchuangYu/clusterProfiler/issues/162>

---

|             |                    |
|-------------|--------------------|
| uniprot_get | <i>uniprot_get</i> |
|-------------|--------------------|

---

**Description**

retrieves annotation data from uniprot

**Usage**

```
uniprot_get(taxID)
```

**Arguments**

|       |             |
|-------|-------------|
| taxID | taxonomy ID |
|-------|-------------|

**Value**

gene table data frame

**Author(s)**

guangchuang yu

# Index

- \* **classes**
  - groupGOResult-class, 18
- \* **datasets**
  - DataSet, 7
- \* **manip**
  - compareCluster, 6
  - enrichGO, 10
  - enrichKEGG, 12
  - groupGO, 17
  
- bitr, 3
- bitr\_kegg, 4
- browseKEGG, 4
- buildGOMap, 5
  
- clusterProfiler
  - (clusterProfiler-package), 3
- clusterProfiler-package, 3
- compareCluster, 6, 11, 13, 18, 19
- compareClusterResult, 19
  
- DataSet, 7
- DE\_GSE8057 (DataSet), 7
- download\_KEGG, 7
- dropGO, 8
  
- enrichDAVID, 8
- enricher, 9
- enrichGO, 6, 10
- enrichKEGG, 12
- enrichMKEGG, 13
- enrichWP, 14
  
- gcSample (DataSet), 7
- get\_wp\_organisms, 15
- Gff2GeneTable, 15
- go2ont, 16
- go2term, 16
- gofilter, 17
- groupGO, 6, 17, 19
- groupGOResult-class, 18
  
- GSEA, 19
- gseGO, 20
- gseKEGG, 21
- gseMKEGG, 22
- gseWP, 23
  
- idType, 24
  
- kegg\_species (DataSet), 7
- ko2name, 25
  
- merge\_result, 25
  
- plotGOgraph, 26
  
- read.gmt, 27
  
- search\_kegg\_organism, 27
- show,groupGOResult-method
  - (groupGOResult-class), 18
- simplify, 28
- simplify,compareClusterResult-method
  - (simplify), 28
- simplify,enrichResult-method
  - (simplify), 28
- simplify,gseaResult-method (simplify), 28
  
- uniprot\_get, 29