

Package ‘clusterProfiler’

April 5, 2014

Type Package

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

Version 1.10.1

Author Guangchuang Yu, Li-Gen Wang

Maintainer Guangchuang Yu <guangchuangyu@gmail.com>

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

Depends R (>= 2.10), ggplot2

Imports

methods, stats4, DBI, plyr, AnnotationDbi, GO.db, KEGG.db, org.Hs.eg.db, DOSE, GOSemSim

Suggests ReactomePA, pathview, knitr

VignetteBuilder knitr

biocViews Clustering, GO, Pathways, Visualization, MultipleComparisons, GeneSetEnrichment

Collate 'AllGenerics.R' 'clusterProfiler-package.R' 'compareCluster.R' 'enrichGO.R' 'enrichKEGG.R' 'GFFparser.R' 'groupGO.R' 'utilities.R' 'zzz.R'

License Artistic-2.0

URL <http://online.liebertpub.com/doi/abs/10.1089/omi.2011.0118>

R topics documented:

clusterProfiler-package	2
buildGOMap	3
compareCluster	3
compareClusterResult-class	4
DataSet	5
enrichGO	5
enrichKEGG	6

getGOLevel	7
Gff2GeneTable	7
groupGO	8
groupGOResult-class	9
plot	9
plotting.clusterProfile	10
show	10
summary	11
viewKEGG	11

Index 13

clusterProfiler-package

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

Description

This package is designed to compare gene clusters functional profiles.

Details

Package: clusterProfiler
 Type: Package
 Version: 1.9.
 Date: 06-13-2013
 biocViews: GO, Clustering, Visulization
 Depends: AnnotationDbi, GO.db, org.Hs.eg.db, ggplot2, plyr, methods
 Suggests: GOSemSim
 License: Artistic-2.0

Author(s)

Guangchuang Yu <guangchuangyu@gmail.com>

Maintainer: Guangchuang Yu <guangchuangyu@gmail.com>

See Also

[compareClusterResult](#), [groupGOResult](#) [enrichResult](#)

buildGOMap	<i>buildGOMap</i>
------------	-------------------

Description

building GO mapping files

Usage

```
buildGOMap(gomap, compress = TRUE)
```

Arguments

gomap	data.frame with two columns names "entrezgene", and "go_accession"
compress	logical, indicate file save in compress or not.

Details

provided by a data.frame of gene and GO directly annotation file this function will building gene to GO and GO to gene mapping, with directly and undirectly annotation.

Value

files save in the the working directory

Author(s)

Yu Guangchuang

compareCluster	<i>Compare gene clusters functional profile Given a list of gene set, this function will compute profiles of each gene cluster.</i>
----------------	---

Description

Compare gene clusters functional profile Given a list of gene set, this function will compute profiles of each gene cluster.

Usage

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

Arguments

geneClusters a list of entrez gene id.
fun One of "groupGO", "enrichGO", "enrichKEGG", "enrichDO" or "enrichPathway" .
... Other arguments.

Value

A clusterProfResult instance.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[compareClusterResult-class](#), [groupGO](#) [enrichGO](#)

Examples

```
data(gcSample)
xx <- compareCluster(gcSample, fun="enrichKEGG", organism="human", pvalueCutoff=0.05)
#summary(xx)
#plot(xx, type="dot", caption="KEGG Enrichment Comparison")
```

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.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[groupGOResult](#) [enrichResult](#) [compareCluster](#)

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

Description

Datasets gcSample contains a sample of gene clusters.

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

Description

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

Usage

```
enrichGO(gene, organism = "human", ont = "MF",
         pvalueCutoff = 0.05, pAdjustMethod = "BH", universe,
         qvalueCutoff = 0.2, minGSSize = 5, readable = FALSE)
```

Arguments

gene	a vector of entrez gene id.
organism	Currently, only "human", "mouse" and "yeast" supported.
ont	One of "MF", "BP", and "CC" subontologies.
pvalueCutoff	Cutoff value of pvalue.
pAdjustMethod	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe	background genes
qvalueCutoff	qvalue cutoff
minGSSize	minimal size of genes annotated by Ontology term for testing.
readable	whether mapping gene ID to gene Name

Value

A enrichResult instance.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[enrichResult-class](#), [compareCluster](#)

Examples

```
#data(gcSample)
#yy <- enrichGO(gcSample[[1]], organism="human", ont="BP", pvalueCutoff=0.01)
#head(summary(yy))
#plot(yy)
```

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 = "human", pvalueCutoff = 0.05,
  pAdjustMethod = "BH", universe, minGSSize = 5,
  qvalueCutoff = 0.2, readable = FALSE)
```

Arguments

gene	a vector of entrez gene id.
organism	Currently, only "human" and "mouse" supported.
pvalueCutoff	Cutoff value of pvalue.
pAdjustMethod	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe	background genes
qvalueCutoff	qvalue cutoff
minGSSize	minimal size of genes annotated by Ontology term for testing.
readable	whether mapping gene ID to gene Name

Value

A `enrichResult` instance.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[enrichResult-class](#), [compareCluster](#)

Examples

```
data(gcSample)
yy = enrichKEGG(gcSample[[5]], pvalueCutoff=0.01)
head(summary(yy))
#plot(yy)
```

getGOLevel	<i>get GOIDs at a specific level</i>
------------	--------------------------------------

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://ygc.name>

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 the GFF file, this function will extract information and save it in working directory

Value

file save.

Author(s)

Yu Guangchuang

groupGO	<i>Functional Profile of a gene set at specific GO level. Given a vector of genes, this function will return the GO profile at specific level.</i>
---------	--

Description

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

Usage

```
groupGO(gene, organism = "human", ont = "CC", level = 2,
        readable = FALSE)
```

Arguments

gene a vector of entrez gene id.
 organism Currently, only "human" and "mouse" supported.
 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]], organism="human", 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.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

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

plot *plot method*

Description

plot method generics

Arguments

... Additional argument list

Value

plot

Author(s)

Guangchuang Yu <http://ygc.name>

plotting.clusterProfile

plotting-clusterProfile

Description

Internal plot function for plotting compareClusterResult

Usage

```
plotting.clusterProfile(clProf.reshape.df, type = "dot",  
  by = "percentage", title = "", font.size = 12)
```

Arguments

clProf.reshape.df	data frame of compareCluster result
type	one of dot and bar
by	one of percentage and count
title	graph title
font.size	graph font size

Value

ggplot object

Author(s)

Guangchuang Yu <http://ygc.name>

show

show method

Description

show method for compareClusterResult instance

show method for groupGOResult instance

Arguments

object	A compareClusterResult instance.
--------	----------------------------------

object	A groupGOResult instance
--------	--------------------------

Value

message
message

Author(s)

Guangchuang Yu <http://ygc.name>
Guangchuang Yu <http://ygc.name>

summary	<i>summary method</i>
---------	-----------------------

Description

summary method for compareClusterResult instance

Arguments

object A compareClusterResult instance.

Value

A data frame

Author(s)

Guangchuang Yu <http://ygc.name>

viewKEGG	<i>viewKEGG function is for visualize KEGG pathways works with enrichResult object to visualize enriched KEGG pathway</i>
----------	---

Description

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

Usage

```
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

Luo et al. (2013) Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* (Oxford, England), 29:14 1830–1831, 2013. ISSN 1367-4803 <http://bioinformatics.oxfordjournals.org/content/abstract/29/14/1830.abstract> PMID: 23740750

Index

- *Topic **classes**
 - compareClusterResult-class, 4
 - groupGOResult-class, 9
- *Topic **datasets**
 - DataSet, 5
- *Topic **manip**
 - compareCluster, 3
 - enrichGO, 5
 - enrichKEGG, 6
 - groupGO, 8
- *Topic **package**
 - clusterProfiler-package, 2

buildGOMap, 3

clusterProfiler

- (clusterProfiler-package), 2

clusterProfiler-package, 2

compareCluster, 3, 4, 6, 7, 9

compareClusterResult, 2, 9

compareClusterResult-class, 4

DataSet, 5

enrichGO, 4, 5

enrichKEGG, 6

enrichResult, 2, 4

gcSample (DataSet), 5

getGOLevel, 7

Gff2GeneTable, 7

groupGO, 4, 8, 9

groupGOResult, 2, 4

groupGOResult-class, 9

plot, 9

plot, compareClusterResult, ANY-method

- (plot), 9

plot, compareClusterResult-method

- (compareClusterResult-class), 4

plotting.clusterProfile, 10

show, 10

show, compareClusterResult-method

- (compareClusterResult-class), 4

show, groupGOResult-method

- (groupGOResult-class), 9

summary, 11

summary, compareClusterResult-method

- (compareClusterResult-class), 4

viewKEGG, 11