Package 'clusterProfiler'

April 9, 2015

Type Package

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

Version 2.0.1

Author Guangchuang Yu, Li-Gen Wang

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 (>= 3.0.0)

Imports methods, stats4, plyr, ggplot2, AnnotationDbi, GO.db, KEGG.db,
DOSE, GOSemSim

Suggests org.Hs.eg.db, ReactomePA, pathview, knitr

VignetteBuilder knitr

License Artistic-2.0

URL https://github.com/GuangchuangYu/clusterProfiler

biocViews Clustering, GO, Pathways, Visualization, MultipleComparison,
GeneSetEnrichment

R topics documented:

clusterProfiler-package

> *statistical analysis and visulization of functional profiles for genes and
> gene clusters The package implements methods to analyze and visual-
> ize 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	*buildGOmap*

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

---

cnetplot	*cnetplot*

Description

category-gene-net plot

Usage

```
cnetplot(x, ...)
```

Arguments

x	enrichResult object
...	additional parameter

Details

category gene association

Value

figure

Author(s)

ygc

---

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

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.

Slots

compareClusterResult cluster comparing result

geneClusters a list of genes

fun one of groupGO, enrichGO and enrichKEGG

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

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

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

Description

Datasets gcSample contains a sample of gene clusters.

---

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

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	One of "anopheles", "arabidopsis", "bovine", "canine", "chicken", "chimp", "coeli-color", "ecolik12","ecsakai", "fly", "gondii","human", "malaria", "mouse", "pig", "rat","rhesus", "worm", "xenopus", "yeast" and "zebrafish".
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 | *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.

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   One of "anopheles", "arabidopsis", "bovine", "canine", "chicken", "chimp", "eco-lik12","ecsakai", "fly", "human", "malaria", "mouse", "pig", "rat","rhesus", "worm", "xenopus", "yeast" and "zebrafish".

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)
```

---

enrichMap   *enrichMap*

Description

enrichment map

Usage

```
enrichMap(x, ...)
```

Arguments

x	enrichResult or gseaResult
...	additional parameter

Details

enrichMap

Value

figure

Author(s)

ygc

---

getGeneSet.BP   *getGeneSet.BP*

Description

getGeneSet.BP

Usage

```
## S3 method for class BP
getGeneSet(setType = "BP", organism)
```

Arguments

setType	gene set type
organism	organism

---

getGeneSet.CC *getGeneSet.CC*

Description

getGeneSet.CC

Usage

```
## S3 method for class CC
getGeneSet(setType = "CC", organism)
```

Arguments

setType   gene set type

organism   organism

---

getGeneSet.GO *getGeneSet.GO*

Description

getGeneSet.GO

Usage

```
## S3 method for class GO
getGeneSet(setType = "GO", organism)
```

Arguments

setType   gene set type

organism   organism

---

getGeneSet.KEGG　　　　　*getGeneSet.KEGG*

Description

getGeneSet.KEGG

Usage

```
## S3 method for class KEGG
getGeneSet(setType = "KEGG", organism)
```

Arguments

setType	gene set type
organism	organism

---

getGeneSet.MF　　　　　*getGeneSet.MF*

Description

getGeneSet.MF

Usage

```
## S3 method for class MF
getGeneSet(setType = "MF", organism)
```

Arguments

setType	gene set type
organism	organism

---

getGOLevel   *get GOIDs at a specific level*

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

Value

file save.

Author(s)

Yu Guangchuang

---

groupGO	*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.*

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	One of "anopheles", "arabidopsis", "bovine", "canine", "chicken", "chimp", "coeli-color", "ecolik12","ecsakai", "fly", "gondii","human", "malaria", "mouse", "pig", "rat","rhesus", "worm", "xenopus", "yeast" and "zebrafish".
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.

Slots

result GO classification result

ontology Ontology

level GO level

organism one of "human", "mouse" and "yeast"

gene Gene IDs

geneInCategory gene and category association

readable logical flag of gene ID in symbol or not.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

compareClusterResult compareCluster groupGO

---

gseaplot   *gseaplot*

Description

visualize analyzing result of GSEA

Usage

```
gseaplot(x, ...)
```

Arguments

x	gseaResult object
...	additional parameters

Details

plotting function for gseaResult

Value

figure

Author(s)

ygc

---

gseGO   *gseGO*

Description

Gene Set Enrichment Analysis of Gene Ontology

Usage

```
gseGO(geneList, ont = "BP", organism = "human", exponent = 1,
  nPerm = 1000, minGSSize = 10, pvalueCutoff = 0.05,
  pAdjustMethod = "BH", verbose = TRUE)
```

Arguments

geneList	order ranked geneList
ont	one of "BP", "MF", "CC" or "GO"
organism	One of "anopheles", "arabidopsis", "bovine", "canine", "chicken", "chimp", "coeli-color", "ecolik12","ecsakai", "fly", "gondii","human", "malaria", "mouse", "pig", "rat","rhesus", "worm", "xenopus", "yeast" and "zebrafish".
exponent	weight of each step
nPerm	permutation numbers
minGSSize	minimal size of each geneSet for analyzing
pvalueCutoff	pvalue Cutoff
pAdjustMethod	pvalue adjustment method
verbose	print message or not

Value

gseaResult object

Author(s)

Yu Guangchuang

gseKEGG	*gseKEGG*

Description

Gene Set Enrichment Analysis of KEGG

Usage

```
gseKEGG(geneList, organism = "human", exponent = 1, nPerm = 1000,
  minGSSize = 10, pvalueCutoff = 0.05, pAdjustMethod = "BH",
  verbose = TRUE)
```

Arguments

geneList	order ranked geneList
organism	One of "anopheles", "arabidopsis", "bovine", "canine", "chicken", "chimp", "ecolik12","ecsakai", "fly", "human", "malaria", "mouse", "pig", "rat","rhesus", "worm", "xenopus", "yeast" and "zebrafish".
exponent	weight of each step
nPerm	permutation numbers
minGSSize	minimal size of each geneSet for analyzing
pvalueCutoff	pvalue Cutoff
pAdjustMethod	pvalue adjustment method
verbose	print message or not

Value

gseaResult object

Author(s)

Yu Guangchuang

---

plot   *plot method*

Description

plot method generics

Usage

```
## S4 method for signature compareClusterResult,ANY
plot(x, type = "dot", title = "",
  font.size = 12, showCategory = 5, by = "geneRatio",
  colorBy = "p.adjust")
```

Arguments

...	Additional argument list
x	compareClusterResult object
type	one of bar or dot
title	figure title
font.size	font size
showCategory	category numbers
by	one of geneRatio, Percentage or count
colorBy	one of pvalue or p.adjust

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 = "geneRatio",
    colorBy = "p.adjust", 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
colorBy	one of pvalue or p.adjust

Value

ggplot object

Author(s)

Guangchuang Yu <http://ygc.name>

---

viewKEGG	*viewKEGG function is for visualize KEGG pathways works with en-richResult object to visualize enriched KEGG pathway*

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