Package 'DOSE'

October 12, 2016

Type Package

Title Disease Ontology Semantic and Enrichment analysis

Version 2.10.7

Author Guangchuang Yu <guangchuangyu@gmail.com> with contributions from Li-Gen Wang, Vladislav Petyuk and Giovanni Dall'Olio.

Maintainer Guangchuang Yu <guangchuangyu@gmail.com>

Description This package implements five methods proposed by Resnik, Schlicker, Jiang, Lin and Wang respectively for measuring semantic similarities among DO terms and gene products. Enrichment analyses including hypergeometric model and gene set enrichment analysis are also implemented for discovering disease associations of high-throughput biological data.

Depends R (>= 3.1.0)

Imports AnnotationDbi, DO.db, ggplot2, GOSemSim, graphics, grDevices, grid, igraph, methods, plyr, qvalue, reshape2, scales, stats4, utils

Suggests org.Hs.eg.db, clusterProfiler, knitr, BiocStyle

VignetteBuilder knitr

License Artistic-2.0

URL http://guangchuangyu.github.io/DOSE

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

biocViews Annotation, Visualization, MultipleComparison, GeneSetEnrichment, Pathways, Software

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

DOSE-package *Disease Ontology Semantic and Enrichment analysis Implemented five methods proposed by Resnik, Schlicker, Jiang, Lin and Wang respectively for measuring DO semantic similarities, and hypergeometric test for enrichment analysis.*

Description

This package is designed to estimate DO-based semantic similarity measurement and enrichment analysis.

Details

Package:	DOSE
Type:	Package
Version:	2.3.5
Date:	2-27-2012
biocViews:	Bioinformatics, Annotation
Depends:	
Imports:	methods, AnnotationDbi, DO.db
Suggests:	clusterProfiler, GOSemSim
License:	Artistic-2.0

Author(s)

Guangchuang Yu, Li-Gen Wang

Maintainer: Guangchuang Yu <guangchuangyu@gmail.com>

See Also

enrichResult

barplot.enrichResult *barplot*

Description

barplot of enrichResult

Usage

```
## S3 method for class 'enrichResult'
barplot(height, x = "Count", colorBy = "pvalue",
  showCategory = 5, font.size = 12, title = "", ...)
```

Arguments

height	enrichResult object
x	one of 'Count' and 'GeneRatio'
colorBy	one of 'pvalue', 'p.adjust', 'qvalue'
showCategory	number of categories to show
font.size	font size
title	plot title
...	other parameter, ignored

---

clusterSim *clusterSim*

Description

semantic similarity between two gene clusters

Usage

```
clusterSim(cluster1, cluster2, measure = "Wang", combine = "BMA")
```

Arguments

cluster1	a vector of gene IDs
cluster2	another vector of gene IDs
measure	One of "Resnik", "Lin", "Rel", "Jiang" and "Wang" methods.
combine	One of "max", "average", "rcmax", "BMA" methods, for combining

Details

given two gene clusters, this function calculates semantic similarity between them.

Value

similarity

Author(s)

Yu Guangchuang

Examples

```
## cluster1 <- c("835", "5261","241", "994")
## cluster2 <- c("307", "308", "317", "321", "506", "540", "378", "388", "396")
## clusterSim(cluster1, cluster2, ont="MF", organism="human", measure="Wang")
```

---

cnetplot *cnetplot method*

Description

cnetplot

Usage

```
cnetplot(x, showCategory = 5, categorySize = "geneNum", foldChange = NULL,
  fixed = TRUE, ...)

## S4 method for signature 'enrichResult'
cnetplot(x, showCategory = 5,
  categorySize = "geneNum", foldChange = NULL, fixed = TRUE, ...)
```

Arguments

x	enrichResult object
showCategory	number of category plotted
categorySize	one of geneNum or pvalue
foldChange	fold change of expression value
fixed	logical
...	additional parameters

Value

plot

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

cnetplot_internal *cnetplot_internal*

Description

plot function of gene Concept Net.

Usage

```
cnetplot_internal(inputList, categorySize = "geneNum", showCategory = 5,
  pvalue = NULL, foldChange = NULL, fixed = TRUE, DE.foldChange = NULL,
  ...)
```

Arguments

inputList	a list of gene IDs
categorySize	setting category size
showCategory	number of categories to plot
pvalue	pvalue
foldChange	fold Change
fixed	logical
DE.foldChange	logical
...	additional parameters

Value

plotted igraph object.

Author(s)

Guangchuang Yu http://ygc.name

computeIC *compute information content*

Description

compute information content

Usage

```
computeIC(ont = "DO", organism = "human")
```

Arguments

ont	"DO"
organism	"human"

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

DataSet *Datasets*

Description

Information content and DO term to entrez gene IDs mapping

---

doSim *doSim*

Description

measuring similarities between two DO term vectors.

Usage

```
doSim(DOID1, DOID2, measure = "Wang")
```

Arguments

DOID1	DO term vector
DOID2	DO term vector
measure	one of "Wang", "Resnik", "Rel", "Jiang", and "Lin".

Details

provide two DO term vectors, this function will calculate their similarities.

Value

score matrix

Author(s)

Guangchuang Yu http://ygc.name

dotplot *dotplot method*

Description

dotplot

dotplot for enrichResult

Usage

```
dotplot(object, ...)

## S4 method for signature 'enrichResult'
dotplot(object, x = "geneRatio",
  colorBy = "p.adjust", showCategory = 10, font.size = 12, title = "")
```

Arguments

object	an instance of enrichResult
...	additional parameter
x	variable for x axis
colorBy	one of 'pvalue', 'p.adjust' and 'qvalue'
showCategory	number of category
font.size	font size
title	plot title

Value

plot

Author(s)

Guangchuang Yu

Guangchuang Yu

enrichDO *DO Enrichment Analysis*

Description

Given a vector of genes, this function will return the enrichment DO categories with FDR control.

Usage

```
enrichDO(gene, ont = "DO", pvalueCutoff = 0.05, pAdjustMethod = "BH",
  universe, minGSSize = 10, maxGSSize = 500, qvalueCutoff = 0.2,
  readable = FALSE)
```

Arguments

gene	a vector of entrez gene id
ont	one of DO or DOLite.
pvalueCutoff	pvalue cutoff
pAdjustMethod	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe	background genes
minGSSize	minimal size of genes annotated by NCG category for testing
maxGSSize	maximal size of each geneSet for analyzing
qvalueCutoff	qvalue cutoff
readable	whether mapping gene ID to gene Name

Value

A enrichResult instance.

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

See Also

enrichResult-class

Examples

```
data(geneList)
gene = names(geneList)[geneList > 1]
yy = enrichDO(gene, pvalueCutoff=0.05)
summary(yy)
```

enricher_internal *enrich.internal*

Description

interal method for enrichment analysis

Usage

```
enricher_internal(gene, pvalueCutoff, pAdjustMethod = "BH", universe,
  minGSSize = 10, maxGSSize = 500, qvalueCutoff = 0.2, USER_DATA)
```

Arguments

gene	a vector of entrez gene id.
pvalueCutoff	Cutoff value of pvalue.
pAdjustMethod	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe	background genes
minGSSize	minimal size of genes annotated by Ontology term for testing.
maxGSSize	maximal size of each geneSet for analyzing
qvalueCutoff	cutoff of qvalue
USER_DATA	ontology information

Details

using the hypergeometric model

Value

A enrichResult instance.

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

enrichMap *enrichMap*

Description

enrichment map

Usage

```
enrichMap(x, n = 50, fixed = TRUE, vertex.label.font = 1, ...)
```

Arguments

x	gseaResult or enrichResult object
n	maximum number of category to shown
fixed	if set to FALSE, will invoke tkplot
vertex.label.font	font size of vertex label
...	additional parameter

Details

enrichment map

Value

figure

Author(s)

G Yu

enrichNCG *enrichNCG*

Description

Enrichment analysis based on the Network of Cancer Genes database (http://ncg.kcl.ac.uk/)

Usage

```
enrichNCG(gene, pvalueCutoff = 0.05, pAdjustMethod = "BH", universe,
  minGSSize = 10, maxGSSize = 500, qvalueCutoff = 0.2, readable = FALSE)
```

Arguments

gene	a vector of entrez gene id
pvalueCutoff	pvalue cutoff
pAdjustMethod	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe	background genes
minGSSize	minimal size of genes annotated by NCG category for testing
maxGSSize	maximal size of each geneSet for analyzing
qvalueCutoff	qvalue cutoff
readable	whether mapping gene ID to gene Name

Details

given a vector of genes, this function will return the enrichment NCG categories with FDR control

Value

A enrichResult instance

Author(s)

Guangchuang Yu

---

enrichResult-class *Class "enrichResult" This class represents the result of enrichment analysis.*

Description

Class "enrichResult" This class represents the result of enrichment analysis.

Slots

result enrichment analysis
pvalueCutoff pvalueCutoff
pAdjustMethod pvalue adjust method
qvalueCutoff qvalueCutoff
organism only "human" supported
ontology biological ontology
gene Gene IDs
keytype Gene ID type
universe background gene
geneInCategory gene and category association
gene2Symbol mapping gene to Symbol
geneSets gene sets
readable logical flag of gene ID in symbol or not.

Author(s)

Guangchuang Yu http://ygc.name

See Also

enrichDO

EXTID2NAME *EXTID2NAME*

Description

mapping gene ID to gene Symbol

Usage

```
EXTID2NAME(OrgDb, geneID, keytype)
```

Arguments

OrgDb	OrgDb
geneID	entrez gene ID
keytype	keytype

Value

gene symbol

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

---

fortify.enrichResult *fortify*

Description

convert enrichResult object for ggplot2

Usage

```
## S3 method for class 'enrichResult'
fortify(model, data, showCategory = 5, by = "Count",
  order = FALSE, drop = FALSE, ...)
```

Arguments

model	enrichResult object
data	not use here
showCategory	Category numbers to show
by	one of Count and GeneRatio
order	logical
drop	logical
...	additional parameter

---

fortify.gseaResult *fortify.gseaResult*

Description

convert gsea result for ggplot2

Usage

```
## S3 method for class 'gseaResult'
fortify(model, data, geneSetID, ...)
```

Arguments

model	gseaResult object
data	not used.
geneSetID	gene set ID
...	additional parameter

Value

figure

Author(s)

G Yu

gene2DO *convert Gene ID to DO Terms*

Description

provide gene ID, this function will convert to the corresponding DO Terms

Usage

```
gene2DO(gene)
```

Arguments

gene	entrez gene ID

Value

DO Terms

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

geneSim *geneSim*

Description

measuring similarities bewteen two gene vectors.

Usage

```
geneSim(geneID1, geneID2 = NULL, measure = "Wang", combine = "BMA")
```

Arguments

geneID1	entrez gene vector
geneID2	entrez gene vector
measure	one of "Wang", "Resnik", "Rel", "Jiang", and "Lin".
combine	One of "max", "average", "rcmax", "BMA" methods, for combining semantic similarity scores of multiple DO terms associated with gene/protein.

Details

provide two entrez gene vectors, this function will calculate their similarity.

Value

score matrix

Author(s)

Guangchuang Yu http://ygc.name

`gseaplot` *visualize analyzing result of GSEA*

Description

plotting function for gseaResult

Usage

```
gseaplot(gseaResult, geneSetID, by = "all")
```

Arguments

`gseaResult`	gseaResult object
`geneSetID`	geneSet ID
`by`	one of "runningScore" or "position"

Value

ggplot2 object

Author(s)

Yu Guangchuang

`gseaResult-class` *Class "gseaResult" This class represents the result of GSEA analysis*

Description

Class "gseaResult" This class represents the result of GSEA analysis

Slots

`result` GSEA anaysis

`organism` organism

`setType` setType

`geneSets` geneSets

`geneList` order rank geneList

`keytype` ID type of gene

`permScores` permutation scores

`params` parameters

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

See Also

gseaplot

`GSEA_internal` *GSEA_internal*

Description

generic function for gene set enrichment analysis

Usage

```
GSEA_internal(geneList, geneSets, exponent, nPerm, minGSSize, maxGSSize,
  pvalueCutoff, pAdjustMethod, verbose, seed = FALSE, USER_DATA)
```

Arguments

`geneList`	order ranked geneList
`geneSets`	gene sets
`exponent`	weight of each step
`nPerm`	permutation numbers
`minGSSize`	minimal size of each geneSet for analyzing
`maxGSSize`	maximal size of each geneSet for analyzing
`pvalueCutoff`	p value Cutoff
`pAdjustMethod`	p value adjustment method
`verbose`	print message or not
`seed`	set seed inside the function to make result reproducible. FALSE by default.
`USER_DATA`	annotation data

Value

gseaResult object

Author(s)

Yu Guangchuang

`gseDO` *DO Gene Set Enrichment Analysis*

Description

perform gsea analysis

Usage

```
gseDO(geneList, exponent = 1, nPerm = 1000, minGSSize = 10,
  maxGSSize = 500, pvalueCutoff = 0.05, pAdjustMethod = "BH",
  verbose = TRUE, seed = FALSE)
```

Arguments

`geneList`	order ranked geneList
`exponent`	weight of each step
`nPerm`	permutation numbers
`minGSSize`	minimal size of each geneSet for analyzing
`maxGSSize`	maximal size of each geneSet for analyzing
`pvalueCutoff`	pvalue Cutoff
`pAdjustMethod`	p value adjustment method
`verbose`	print message or not
`seed`	logical

Value

gseaResult object

Author(s)

Yu Guangchuang

gseNCG *NCG Gene Set Enrichment Analysis*

Description

perform gsea analysis

Usage

```
gseNCG(geneList, exponent = 1, nPerm = 1000, minGSSize = 10,
  maxGSSize = 500, pvalueCutoff = 0.05, pAdjustMethod = "BH",
  verbose = TRUE, seed = FALSE)
```

Arguments

geneList	order ranked geneList
exponent	weight of each step
nPerm	permutation numbers
minGSSize	minimal size of each geneSet for analyzing
maxGSSize	maximal size of each geneSet for analyzing
pvalueCutoff	pvalue Cutoff
pAdjustMethod	p value adjustment method
verbose	print message or not
seed	logical

Value

gseaResult object

Author(s)

Yu Guangchuang

gsfilter *gsfilter*

Description

filter enriched result by gene set size or gene count

Usage

```
gsfilter(x, by = "GSSize", min = NA, max = NA)
```

Arguments

x	instance of enrichResult or compareClusterResult
by	one of 'GSSize' or 'Count'
min	minimal size
max	maximal size

Value

update object

Author(s)

Guangchuang Yu

---

list2graph *convert gene IDs to igraph object*

Description

convert a list of gene IDs to igraph object.

Usage

```
list2graph(inputList)
```

Arguments

inputList	a list of gene IDs

Value

a igraph object.

Author(s)

Guangchuang Yu http://ygc.name

load_OrgDb *load_OrgDb*

Description

load OrgDb

Usage

```
load_OrgDb(OrgDb)
```

Arguments

OrgDb	OrgDb object or OrgDb name

Value

OrgDb object

Author(s)

Guangchuang Yu

mclusterSim *mclusterSim*

Description

Pairwise semantic similarity for a list of gene clusters

Usage

```
mclusterSim(clusters, measure = "Wang", combine = "BMA")
```

Arguments

clusters	A list of gene clusters
measure	one of "Wang", "Resnik", "Rel", "Jiang", and "Lin".
combine	One of "max", "average", "rcmax", "BMA" methods, for combining semantic similarity scores of multiple DO terms associated with gene/protein.

Value

similarity matrix

Author(s)

Yu Guangchuang

Examples

```
## cluster1 <- c("835", "5261","241")
## cluster2 <- c("578","582")
## cluster3 <- c("307", "308", "317")
## clusters <- list(a=cluster1, b=cluster2, c=cluster3)
## mclusterSim(clusters, measure="Wang")
```

`netplot` *netplot*

Description

plot network

Usage

```
netplot(g, vertex.label.font = 2, vertex.label.color = "#666666",
  vertex.label.cex = 1.5, layout = layout.fruchterman.reingold,
  foldChange = NULL, fixed = TRUE, col.bin = 10, legend.x = 1,
  legend.y = 1)
```

Arguments

`g`	igraph object
`vertex.label.font`	font size
`vertex.label.color`	font text color
`vertex.label.cex`	cex of vertex label
`layout`	layout
`foldChange`	fold change
`fixed`	logical
`col.bin`	number of legend color bin
`legend.x`	x-axis position of legend
`legend.y`	y-axis position of legend

Details

plot network of igraph object

Value

plot

Author(s)

Yu Guangchuang

`plot` *plot method*

Description

plot method generics

plot method for gseaResult

Usage

```
## S4 method for signature 'enrichResult,ANY'
plot(x, type = "bar", ...)

## S4 method for signature 'gseaResult,ANY'
plot(x, type = "gseaplot", ...)
```

Arguments

`x`	A `enrichResult` instance
`type`	one of bar, cnet or enrichMap
`...`	Additional argument list

Value

plot

plot

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

Yu Guangchuang

rebuildAnnoData *rebuiding annotation data*

Description

rebuilding entrez and DO mapping datasets

Usage

```
rebuildAnnoData(file)
```

Arguments

file	do_rif.human.txt

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

scaleNodeColor *scaleNodeColor*

Description

scale color nodes

Usage

```
scaleNodeColor(g, foldChange, node.idx = NULL, DE.foldChange)
```

Arguments

g	igraph object
foldChange	fold Change
node.idx	index of node to color
DE.foldChange	logical

Details

color nodes based on fold change of expression

Value

igraph object

Author(s)

Yu Guangchuang

setReadable *setReadable*

Description

mapping geneID to gene Symbol

Usage

```
setReadable(x, OrgDb, keytype = "auto")
```

Arguments

x	enrichResult Object
OrgDb	OrgDb
keytype	keytype of gene

Value

enrichResult Object

Author(s)

Yu Guangchuang

---

setting.graph.attributes *setting.graph.attributes*

Description

setting basic attributes of a graph

Usage

```
setting.graph.attributes(g, node.size = 8, node.color = "#B3B3B3",
  edege.width = 2, edege.color = "#8DA0CB")
```

Arguments

g	igraph object
node.size	size of node
node.color	color of node
edege.width	edege width
edege.color	color of edege

Details

setting size and color of node and edege

Value

igraph object

Author(s)

Yu Guangchuang

---

show *show method*

Description

show method for `enrichResult` instance

show method for `gseaResult` instance

Usage

```
show(object)
```

```
show(object)
```

Arguments

`object`	A `enrichResult` instance.

Value

message

message

Author(s)

Guangchuang Yu http://ygc.name

Guangchuang Yu http://guangchuangyu.github.io

simplot *simplot*

Description

plotting similarity matrix

Usage

```
simplot(sim, xlab = "", ylab = "", color.low = "white",
  color.high = "red", labs = TRUE, digits = 2, labs.size = 3,
  font.size = 14)
```

Arguments

sim	similarity matrix
xlab	xlab
ylab	ylab
color.low	color of low value
color.high	color of high value
labs	logical, add text label or not
digits	round digit numbers
labs.size	lable size
font.size	font size

Value

ggplot object

Author(s)

Yu Guangchuang

summary *summary method*

Description

summary method for enrichResult instance

summary method for gseaResult instance

Usage

```
summary(object, ...)
```

```
summary(object, ...)
```

Arguments

object	A enrichResult instance.
...	additional parameter

Value

A data frame

A data frame

Author(s)

Guangchuang Yu http://guangchuangyu.github.io

Guangchuang Yu http://guangchuangyu.github.io

---

theme_dose *theme_dose*

Description

ggplot theme of DOSE

Usage

```
theme_dose(font.size = 14)
```

Arguments

font.size	font size

upsetplot *upsetplot method*

Description

upsetplot method generics

upsetplot

Usage

```
upsetplot(x, ...)

## S4 method for signature 'enrichResult'
upsetplot(x, n = 10, ...)
```

Arguments

x	object
...	additional parameter
n	number of categories to be plotted

Value

plot

Author(s)

Guangchuang Yu

Examples

```
## Not run:
require(DOSE)
data(geneList)
de=names(geneList)[1:100]
x <- enrichDO(de)
upsetplot(x, 8)

## End(Not run)
```

Index