

Package ‘DOSE’

April 23, 2016

Type Package

Title Disease Ontology Semantic and Enrichment analysis

Version 2.8.3

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

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

VignetteBuilder knitr

License Artistic-2.0

URL <https://github.com/GuangchuangYu/DOSE>

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

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

NeedsCompilation no

R topics documented:

DOSE-package	3
ALLEXTID	3
barplot.enrichResult	4

clusterSim	4
cnetplot	5
cnetplot_internal	6
computeIC	6
DataSet	7
doSim	7
dotplot	8
enrich.internal	9
enrichDO	10
enrichMap	11
enrichNCG	11
enrichResult-class	12
EXTID2NAME	13
EXTID2TERMID	13
fortify.enrichResult	14
fortify.gseaResult	14
gene2DO	15
geneSim	15
getALLEG	16
getGeneSet	17
gsea	17
gseaAnalyzer	18
gseaplot	19
gseaResult-class	19
list2graph	20
mclusterSim	20
netplot	21
plot	22
rebuildAnnoData	23
scaleNodeColor	23
setReadable	24
setting.graph.attributes	24
show	25
simplot	26
summary	27
TERM2NAME	27
TERMID2EXTID	28
theme_dose	28
Index	29

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](#)

ALLEXTID

ALLEXTID

Description

Get all background External ID.

Usage

ALLEXTID(organism, ...)

Arguments

organism	organism
...	additional parameter

barplot.enrichResult *barplot*

Description

barplot

Usage

```
## S3 method for class 'enrichResult'
barplot(height, font.size = 12, title = "", ...)
```

Arguments

height	enrichResult object
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 method

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://ygc.name>

cnetplot_internal	<i>cnetplot_internal</i>
-------------------	--------------------------

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	<i>compute information content</i>
-----------	------------------------------------

Description

compute information content

Usage

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

Arguments

ont "DO"
organism "human"

Author(s)

Guangchuang Yu <http://ygc.name>

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 method

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

enrich.internal	<i>enrich.internal</i>
-----------------	------------------------

Description

internal method for enrichment analysis

Usage

```
enrich.internal(gene, organism, pvalueCutoff, pAdjustMethod = "BH", ont,  
               universe, minGSSize = 5, qvalueCutoff = 0.2, readable = FALSE, ...)
```

Arguments

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

Details

using the hypergeometric model

Value

A enrichResult instance.

Author(s)

Guangchuang Yu <http://ygc.name>

`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 = 5, qvalueCutoff = 0.2, readable = FALSE)
```

Arguments

<code>gene</code>	a vector of entrez gene id.
<code>ont</code>	one of DO or DOLite.
<code>pvalueCutoff</code>	Cutoff value of pvalue.
<code>pAdjustMethod</code>	one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
<code>universe</code>	background genes
<code>minGSSize</code>	minimal size of genes annotated by Ontology term for testing.
<code>qvalueCutoff</code>	qvalue Cutoff
<code>readable</code>	whether mapping gene ID to gene Name

Value

A `enrichResult` instance.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[enrichResult-class](#)

Examples

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

`enrichMap`*enrichMap*

Description

enrichment map

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

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

Details

enrichment map

Value

figure

Author(s)

G Yu

`enrichNCG`*enrichNCG*

DescriptionEnrichment analysis based on the Network of Cancer Genes database (<http://ncg.kcl.ac.uk/>)**Usage**`enrichNCG(gene, pvalueCutoff = 0.05, pAdjustMethod = "BH", universe,
minGSSize = 5, 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
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	<i>Class "enrichResult" This class represents the result of enrichment analysis.</i>
--------------------	--

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
 universe background gene
 geneInCategory gene and category association
 geneSets gene sets
 readable logical flag of gene ID in symbol or not.

Author(s)

Guangchuang Yu <http://ygc.name>

See Also

[enrichD0](#)

EXTID2NAME	<i>EXTID2NAME</i>
------------	-------------------

Description

mapping gene ID to gene Symbol

Usage

EXTID2NAME(geneID, organism)

Arguments

geneID	entrez gene ID
organism	one of "human", "mouse" and "yeast"

Value

gene symbol

Author(s)

Guangchuang Yu <http://ygc.name>

EXTID2TERMID	<i>EXTID2TERMID</i>
--------------	---------------------

Description

Mapping External ID to Ontology Term ID

Usage

EXTID2TERMID(gene, organism, ...)

Arguments

gene	gene ID vector
organism	organism
...	additional parameter

fortify.enrichResult *fortify*

Description

fortify

Usage

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

Arguments

model	enrichResult object
data	not use here
showCategory	Category numbers to show
order	logical
drop	logical
...	additional parameter

fortify.gseaResult *fortify.gseaResult*

Description

fortify.gseaResult

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	<i>convert Gene ID to DO Terms</i>
---------	------------------------------------

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://ygc.name>

geneSim	<i>geneSim</i>
---------	----------------

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", "rmax", "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>

getALLEG

getALLEG

Description

get all entrezgene ID of a specific organism

Usage

```
getALLEG(organism)
```

Arguments

organism species

Value

entrez gene ID vector

Author(s)

Yu Guangchuang

getGeneSet	<i>getGeneSet</i>
------------	-------------------

Description

preparing geneSets for gene set enrichment analysis

Usage

```
getGeneSet(setType, organism, ...)
```

Arguments

setType	type of gene sets
organism	organism
...	additional parameter

<i>gsea</i>	<i>gsea</i>
-------------	-------------

Description

generic function for gene set enrichment analysis

Usage

```
gsea(geneList, geneSets, setType, organism, exponent, nPerm, minGSSize,
     pvalueCutoff, pAdjustMethod, verbose, seed = FALSE, ...)
```

Arguments

geneList	order ranked geneList
geneSets	gene sets
setType	Type of geneSet
organism	organism
exponent	weight of each step
nPerm	permutation numbers
minGSSize	minimal 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.
...	additional parameter

Value

gseaResult object

Author(s)

Yu Guangchuang

gseAnalyzer

Gene Set Enrichment Analysis

Description

perform gsea analysis

Usage

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

Arguments

geneList	order ranked geneList
setType	Type of geneSet
organism	organism
exponent	weight of each step
nPerm	permutation numbers
minGSSize	minimal size of each geneSet for analyzing
pvalueCutoff	pvalue Cutoff
pAdjustMethod	p value adjustment method
verbose	print message or not
...	additional parameters

Value

gseaResult object

Author(s)

Yu Guangchuang

gseaplot	<i>visualize analyzing result of GSEA</i>
----------	---

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	<i>Class "gseaResult" This class represents the result of GSEA analysis</i>
------------------	---

Description

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

Slots

result	GSEA analysis
setType	setType
geneSets	geneSets
geneList	order rank geneList
permScores	permutation scores
params	parameters

Author(s)

Guangchuang Yu <http://ygc.name>

See Also[gseaplot](#)

list2graph	<i>convert gene IDs to igraph object</i>
------------	--

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>

mclusterSim	<i>mclusterSim</i>
-------------	--------------------

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", "rmax", "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

<code>g</code>	igraph object
<code>vertex.label.font</code>	font size
<code>vertex.label.color</code>	font text color
<code>vertex.label.cex</code>	cex of vertex label
<code>layout</code>	layout
<code>foldChange</code>	fold change
<code>fixed</code>	logical
<code>col.bin</code>	number of legend color bin
<code>legend.x</code>	x-axis position of legend
<code>legend.y</code>	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 <code>enrichResult</code> instance
type	one of <code>bar</code> , <code>cnet</code> or <code>enrichMap</code>
...	Additional argument list

Value

plot

plot

Author(s)

Guangchuang Yu <http://ygc.name>

Yu Guangchuang

rebuildAnnoData	<i>rebuilding annotation data</i>
-----------------	-----------------------------------

Description

rebuilding entrez and DO mapping datasets

Usage

```
rebuildAnnoData(file)
```

Arguments

file	do_rif.human.txt
------	------------------

Author(s)

Guangchuang Yu <http://ygc.name>

scaleNodeColor	<i>scaleNodeColor</i>
----------------	-----------------------

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)
```

Arguments

x enrichResult Object

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 edge

Value

igraph object

Author(s)

Yu Guangchuang

show	<i>show method</i>
------	--------------------

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://ygc.name>

`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, readable = FALSE)
```

Arguments

<code>sim</code>	similarity matrix
<code>xlab</code>	xlab
<code>ylab</code>	ylab
<code>color.low</code>	color of low value
<code>color.high</code>	color of high value
<code>labs</code>	logical, add text label or not
<code>digits</code>	round digit numbers
<code>labs.size</code>	lable size
<code>font.size</code>	font size
<code>readable</code>	TRUE or FALSE

Value

ggplot object

Author(s)

Yu Guangchuang

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

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

TERM2NAME	<i>TERM2NAME</i>
-----------	------------------

Description

Mapping Ontology Term ID to Name Symbol or Description

Usage

```
TERM2NAME(term, organism, ...)
```

Arguments

term	term ID vector
organism	organism
...	additional parameter

TERMID2EXTID	<i>TERMID2EXTID</i>
--------------	---------------------

Description

Mapping Ontology Term ID to External ID

Usage

```
TERMID2EXTID(term, organism, ...)
```

Arguments

term	term ID vector
organism	organism
...	additional parameter

theme_dose	<i>theme_dose</i>
------------	-------------------

Description

ggplot theme of DOSE

Usage

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

Arguments

font.size	font size
-----------	-----------

Index

- *Topic **classes**
 - enrichResult-class, [12](#)
 - gseaResult-class, [19](#)
- *Topic **datasets**
 - DataSet, [7](#)
- *Topic **manip**
 - enrich.internal, [9](#)
 - enrichDO, [10](#)
 - gseAnalyzer, [18](#)
- *Topic **package**
 - DOSE-package, [3](#)
- ALLEXID, [3](#)
- barplot.enrichResult, [4](#)
- clusterSim, [4](#)
- cnetplot, [5](#)
- cnetplot, enrichResult-method (cnetplot), [5](#)
- cnetplot_internal, [6](#)
- computeIC, [6](#)
- DataSet, [7](#)
- DO2ALLEG (DataSet), [7](#)
- DO2EG (DataSet), [7](#)
- DOLite2EG (DataSet), [7](#)
- DOLiteTerm (DataSet), [7](#)
- DOSE (DOSE-package), [3](#)
- DOSE-package, [3](#)
- DOSEEnv (DataSet), [7](#)
- doSim, [7](#)
- dotplot, [8](#)
- dotplot, enrichResult, ANY-method (dotplot), [8](#)
- dotplot, enrichResult-method (dotplot), [8](#)
- EG2ALLDO (DataSet), [7](#)
- EG2DO (DataSet), [7](#)
- EG2DOLite (DataSet), [7](#)
- enrich.internal, [9](#)
- enrichDO, [10](#), [13](#)
- enrichMap, [11](#)
- enrichNCG, [11](#)
- enrichResult, [3](#)
- enrichResult-class, [12](#)
- EXTID2NAME, [13](#)
- EXTID2TERMINID, [13](#)
- fortify.enrichResult, [14](#)
- fortify.gseaResult, [14](#)
- gene2DO, [15](#)
- geneList (DataSet), [7](#)
- geneSim, [15](#)
- getALLEG, [16](#)
- getGeneSet, [17](#)
- gsea, [17](#)
- gseahResult-class (gseaResult-class), [19](#)
- gseAnalyzer, [18](#)
- gseaplot, [19](#), [20](#)
- gseaResult-class, [19](#)
- IC (DataSet), [7](#)
- Info_Contents_human_DO (DataSet), [7](#)
- list2graph, [20](#)
- mclusterSim, [20](#)
- NCG_DOSE_Env (DataSet), [7](#)
- netplot, [21](#)
- plot, [22](#)
- plot, enrichResult, ANY-method (plot), [22](#)
- plot, enrichResult-method (enrichResult-class), [12](#)
- plot, gseaResult, ANY-method (plot), [22](#)
- plot, gseaResult-method (gseaResult-class), [19](#)
- rebuildAnnoData, [23](#)

scaleNodeColor, [23](#)
setReadable, [24](#)
setting.graph.attributes, [24](#)
show, [25](#)
show, enrichResult-method
 (enrichResult-class), [12](#)
show, gseaResult-method
 (gseaResult-class), [19](#)
simplot, [26](#)
summary, [27](#)
summary, enrichResult-method
 (enrichResult-class), [12](#)
summary, gseaResult-method
 (gseaResult-class), [19](#)

TERM2NAME, [27](#)
TERMID2EXTID, [28](#)
theme_dose, [28](#)