Package ‘DOSE’

March 18, 2024

Type Package
Title Disease Ontology Semantic and Enrichment analysis
Version 3.28.2
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 (>= 3.5.0)
Imports AnnotationDbi, HDO.db, BiocParallel, fgsea, ggplot2, GOSemSim (>= 2.27.1), methods, qvalue, reshape2, stats, utils, yulab.utils
Suggests prettydoc, clusterProfiler, gson (>= 0.0.5), knitr, rmarkdown, org.Hs.eg.db, MPO.db, HPO.db, testthat
VignetteBuilder knitr
ByteCompile true
License Artistic-2.0
Encoding UTF-8
BugReports https://github.com/GuangchuangYu/DOSE/issues
biocViews Annotation, Visualization, MultipleComparison, GeneSetEnrichment, Pathways, Software
RoxygenNote 7.2.3
git_url https://git.bioconductor.org/packages/DOSE
git_branch RELEASE_3_18
git_last_commit 9e236d9
git_last_commit_date 2023-12-10
Repository: Bioconductor 3.18
Date/Publication: 2024-03-18
Author: Guangchuang Yu [aut, cre],
        Li-Gen Wang [ctb],
        Vladislav Petyuk [ctb],
        Giovanni Dall'Olio [ctb],
        Erqiang Hu [ctb]

R topics documented:

- DOSE-package
- clusterSim
- compareClusterResult-class
- computeIC
- DataSet
- doseSim
- doSim
- enrichDGN
- enrichDO
- enricher_internal
- enrichHPO
- enrichMPO
- enrichNCG
- enrichResult-class
- EXTID2NAME
- gene2DO
- geneID
- geneInCategory
- geneSim
- gseaResult-class
- GSEA_internal
- gseDGN
- gseDO
- gseHPO
- gseMPO
- gseNCG
- gsfilter
- hpoSim
- mclusterSim
- mpoSim
- parse_ratio
- rebuildAnnoData
- reexports
- setReadable
- show
- simplot
**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, HDO.dbb
- **Suggests:** clusterProfiler, GOSemSim
- **License:** Artistic-2.0

**Author(s)**

Guangchuang Yu, Li-Gen Wang

Maintainer: Guangchuang Yu <guangchuangyu@gmail.com>

**See Also**

enrichResult
Description

semantic similarity between two gene clusters

Usage

clusterSim(
  cluster1,
  cluster2,
  ont = "DO",
  organism = "hsa",
  measure = "Wang",
  combine = "BMA"
)

Arguments

cluster1 a vector of gene IDs
cluster2 another vector of gene IDs
ont one of "DO" and "MPO"
organism organism
measure One of "Resnik", "Lin", "Rel", "Jiang" and "Wang" methods.
combine One of "max", "avg", "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")
clusterSim(cluster1, cluster2)
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
- `gene2Symbol` gene ID to Symbol
- `keytype` Gene ID type
- `readable` logical flag of gene ID in symbol or not.
- `.call` function call
- `termsim` Similarity between term
- `method` method of calculating the similarity between nodes
- `dr` dimension reduction result

Author(s)

Guangchuang Yu [https://yulab-smu.top](https://yulab-smu.top)

See Also

- `enrichResult`

---

**computeIC**

*compute information content*

Description

compute information content

Usage

`computeIC(ont = "DO")`
Arguments

ont one of "DO" and "MPO"

Author(s)

Guangchuang Yu [https://yulab-smu.top](https://yulab-smu.top)

---

**Data Set**

**Datasets**

**Description**

Information content and DO term to entrez gene IDs mapping

---

**doseSim**

**doSim**

**Description**

measuring similarities between two DO term vectors.

**Usage**

doseSim(DOID1, DOID2, measure = "Wang", ont = "DO")

**Arguments**

- **DOID1**
  - DO term, MPO term or HPO term vector

- **DOID2**
  - DO term, MPO term or HPO term vector

- **measure**
  - one of "Wang", "Resnik", "Rel", "Jiang", "Lin", and "TCSS".

- **ont**
  - one of "DO" and "MPO"

**Details**

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

**Value**

score matrix
doSim

Description
measuring similarities between two MPO term vectors.

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

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOID1</td>
<td>DO term vector</td>
</tr>
<tr>
<td>DOID2</td>
<td>DO term vector</td>
</tr>
<tr>
<td>measure</td>
<td>one of &quot;Wang&quot;, &quot;Resnik&quot;, &quot;Rel&quot;, &quot;Jiang&quot;, &quot;Lin&quot;, and &quot;TCSS&quot;.</td>
</tr>
</tbody>
</table>

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

Value
score matrix

Author(s)
Guangchuang Yu https://guangchuangyu.github.io

enrichDGN
Enrichment analysis based on the DisGeNET (http://www.disgenet.org/)

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

Usage

enrichDGN(
  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

Value

A enrichResult instance

Author(s)

Guangchuang Yu

References


enrichDGNv enrichDGN

Description

Enrichment analysis based on the DisGeNET (http://www.disgenet.org/)
enrichDGNv

Usage

enrichDGNv(
    snp,
    pvalueCutoff = 0.05,
    pAdjustMethod = "BH",
    universe,
    minGSSize = 10,
    maxGSSize = 500,
    qvalueCutoff = 0.2,
    readable = FALSE
)

Arguments

snp a vector of SNP
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

References

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", 
  organism = "hsa", 
  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 and DOLite.
organism one of "hsa" and "mmu"
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

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

Description

`enricher_internal` method for enrichment analysis

Usage

```r
enricher_internal(
gene,
pvalueCutoff,
pAdjustMethod = "BH",
universe = NULL,
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.
Enrichment analysis based on the DisGeNET (http://www.disgenet.org/)

Description

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

Usage

enrichHPO(
  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

Value

A enrichResult instance

Author(s)

Erqiang Hu
References


enrichMPO  
Enrichment analysis based on the DisGeNET (http://www.disgenet.org/)

Description

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

Usage

enrichMPO(
  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

Value

A enrichResult instance

Author(s)

Erqiang Hu
References
Janet et al. (2015) DisGeNET: a discovery platform for the dynamical exploration of human
diseases and their genes. Database bav028 http://database.oxfordjournals.org/content/
2015/bav028.long

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
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**: p-value cutoff
- **pAdjustMethod**: p-value adjust method
- **qvalueCutoff**: q-value cutoff
- **organism**: only "human" supported
- **ontology**: biological ontology
- **gene**: Gene IDs
- **keytype**: Gene ID type
- **universe**: background gene
- **gene2Symbol**: mapping gene to Symbol
- **geneSets**: gene sets
- **readable**: logical flag of gene ID in symbol or not.
- **termsim**: Similarity between term
- **method**: method of calculating the similarity between nodes
- **dr**: dimension reduction result

Author(s)

Guangchuang Yu [https://yulab-smu.top](https://yulab-smu.top)

See Also

- enrichDO
**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 [https://yulab-smu.top](https://yulab-smu.top)

---

**gene2DO**

*convert Gene ID to DO Terms*

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

**Usage**

```
gene2DO(gene, organism = "hsa", ont = "DO")
```

**Arguments**

- **gene**: entrez gene ID
- **organism**: organism
- **ont**: ont

**Value**

DO Terms
geneID

Author(s)

Guangchuang Yu https://yulab-smu.top

geneID
geneID generic

description
geneID generic

Usage
geneID(x)

Arguments
x enrichResult object

Value
'geneID' return the 'geneID' column of the enriched result which can be converted to data.frame via 'as.data.frame'

Examples
data(geneList, package="DOSE")
de <- names(geneList)[1:100]
x <- enrichDO(de)
geneID(x)

geneInCategory
geneInCategory generic

description
geneInCategory generic

Usage
geneInCategory(x)

Arguments
x enrichResult

geneSim

Value

'geneInCategory' return a list of genes, by splitting the input gene vector to enriched functional categories.

Examples

data(geneList, package="DOSE")
de <- names(geneList)[1:100]
x <- enrichDO(de)
geneInCategory(x)

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc

desc
gseaResult-class

Description

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

Slots

- result: GSEA analysis
- organism: organism
- setType: setType
- geneSets: geneSets
- geneList: order rank geneList
- keytype: ID type of gene
- permScores: permutation scores
- params: parameters
- gene2Symbol: gene ID to Symbol
- readable: whether convert gene ID to symbol
- dr: dimension reduction result

Author(s)

Guangchuang Yu http://ygc.name

GSEA_internal

Description

generic function for gene set enrichment analysis
Usage

GSEA_internal(
    geneList,
    exponent,
    minGSSize,
    maxGSSize,
    eps,
    pvalueCutoff,
    pAdjustMethod,
    verbose,
    seed = FALSE,
    USER_DATA,
    by = "fgsea",
    ...
)

Arguments

geneList       order ranked geneList
exponent       weight of each step
minGSSize      minimal size of each geneSet for analyzing
maxGSSize      maximal size of each geneSet for analyzing
eps            This parameter sets the boundary for calculating the p value.
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
by             one of 'fgsea' or 'DOSE'
...            other parameter

Value

gseaResult object

Author(s)

Yu Guangchuang
DisGeNET Gene Set Enrichment Analysis

Description

perform gsea analysis

Usage

gseDGN(
  geneList,
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  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 each geneSet for analyzing
pvalueCutoff pvalue Cutoff
pAdjustMethod p value adjustment method
verbose     print message or not
seed        logical
by          one of 'fgsea' or 'DOSE'
...          other parameter

Value

gseaResult object

Author(s)

Yu Guangchuang
**gseDO**

**DO Gene Set Enrichment Analysis**

**Description**

perform gsea analysis

**Usage**

```r
gseDO(
  geneList,
  organism = "hsa",
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  verbose = TRUE,
  seed = FALSE,
  by = "fgsea",
  ...
)
```

**Arguments**

- `geneList` order ranked geneList
- `organism` one of "hsa" and "mmu"
- `exponent` weight of each step
- `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
- `by` one of 'fgsea' or 'DOSE'
- `...` other parameter

**Value**

gseaResult object

**Author(s)**

Yu Guangchuang
**gseHPO**

**MPO Gene Set Enrichment Analysis**

**Description**
perform gsea analysis

**Usage**

```r
gseHPO(
  geneList,
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  verbose = TRUE,
  seed = FALSE,
  by = "fgsea",
  ...
)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>geneList</td>
<td>order ranked geneList</td>
</tr>
<tr>
<td>exponent</td>
<td>weight of each step</td>
</tr>
<tr>
<td>minGSSize</td>
<td>minimal size of each geneSet for analyzing</td>
</tr>
<tr>
<td>maxGSSize</td>
<td>maximal size of each geneSet for analyzing</td>
</tr>
<tr>
<td>pvalueCutoff</td>
<td>pvalue Cutoff</td>
</tr>
<tr>
<td>pAdjustMethod</td>
<td>p value adjustment method</td>
</tr>
<tr>
<td>verbose</td>
<td>print message or not</td>
</tr>
<tr>
<td>seed</td>
<td>logical</td>
</tr>
<tr>
<td>by</td>
<td>one of 'fgsea' or 'DOSE'</td>
</tr>
<tr>
<td>...</td>
<td>other parameter</td>
</tr>
</tbody>
</table>

**Value**
gseaResult object

**Author(s)**
Erqiang Hu
Description

perform gsea analysis

Usage

gseMPO(
  geneList,
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  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 each geneSet for analyzing
pvalueCutoff pvalue Cutoff
pAdjustMethod p value adjustment method
verbose print message or not
seed logical
by one of 'fgsea' or 'DOSE'
... other parameter

Value

gseaResult object

Author(s)

Erqiang Hu
gseNCG

**Description**

perform gsea analysis

**Usage**

```r
gseNCG(
  geneList,
  exponent = 1,
  minGSSize = 10,
  maxGSSize = 500,
  pvalueCutoff = 0.05,
  pAdjustMethod = "BH",
  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 each geneSet for analyzing
- `pvalueCutoff`: pvalue Cutoff
- `pAdjustMethod`: p value adjustment method
- `verbose`: print message or not
- `seed`: logical
- `by`: one of 'fgsea' or 'DOSE'
- `...`: other parameter

**Value**

`gseaResult` object

**Author(s)**

Yu Guangchuang
gsfilter

description
filter enriched result by gene set size or gene count

Usage

```r
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

hpoSim

dosim

description
measuring similarities between two MPO term vectors.

Usage

```r
hpoSim(DOID1, DOID2, measure = "Wang")
```

Arguments

- `DOID1`: HPO term vector
- `DOID2`: HPO term vector
- `measure`: one of "Wang", "Resnik", "Rel", "Jiang", "Lin", and "TCSS". 
**Details**

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

**Value**

cscore matrix

---

**Description**

Pairwise semantic similarity for a list of gene clusters

**Usage**

```r
mclusterSim(
  clusters,
  ont = "DO",
  organism = "hsa",
  measure = "Wang",
  combine = "BMA"
)
```

**Arguments**

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

**Value**

similarity matrix

**Author(s)**

Yu Guangchuang

**Examples**

```r
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")
```
**mpoSim**  
*doSim*

**Description**
measuring similarities between two MPO term vectors.

**Usage**
```r
mpoSim(DOID1, DOID2, measure = "Wang")
```

**Arguments**
- **DOID1**: MPO term vector
- **DOID2**: MPO term vector
- **measure**: one of "Wang", "Resnik", "Rel", "Jiang", "Lin", and "TCSS".

**Details**
provide two MPO term vectors, this function will calculate their similarities.

**Value**
score matrix

---

**parse_ratio**

**Description**
parse character ratio to double value, such as 1/5 to 0.2

**Usage**
```r
parse_ratio(ratio)
```

**Arguments**
- **ratio**: character vector of ratio to parse

**Value**
A numeric vector (double) of parsed ratio

**Author(s)**
Guangchuang Yu
rebuildAnnoData

Description
rebuilding entrez and DO mapping datasets

Usage
rebuildAnnoData(file)

Arguments
file
do_rif.human.txt

Author(s)
Guangchuang Yu https://yulab-smu.top

---

setReadable

Description
mapping geneID to gene Symbol

Usage
setReadable(x, OrgDb, keyType = "auto")

Arguments
x
enrichResult Object
OrgDb
OrgDb
keyType
keyType of gene
show

Value

enrichResult Object

Author(s)

Yu Guangchuang

show

show method

Description

show method for gseaResult instance
show method for enrichResult instance

Usage

show(object)
show(object)

Arguments

object A enrichResult instance.

Value

message
message

Author(s)

Guangchuang Yu https://yulab-smu.top
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  label size
font.size  font size

Value
ggplot object

Author(s)
Yu Guangchuang
**Description**

summary method for gseaResult instance
summary method for enrichResult instance

**Usage**

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

**Arguments**

- `object` A `enrichResult` instance.
- `...` additional parameter

**Value**

A data frame
A data frame

**Author(s)**

Guangchuang Yu [https://guangchuangyu.github.io](https://guangchuangyu.github.io)
Guangchuang Yu [http://guangchuangyu.github.io](http://guangchuangyu.github.io)

---

**theme_dose**

**Description**

ggplot theme of DOSE

**Usage**

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

**Arguments**

- `font.size` font size
theme_dose

Value

ggplot theme

Examples

library(ggplot2)
qplot(1:10) + theme_dose()
Index

* classes
  compareClusterResult-class, 5
  enrichResult-class, 15
  gseaResult-class, 19

* datasets
  DataSet, 6

* internal
  reexports, 29

* manip
  enrichDO, 10
  enricher_internal, 11
  gseDGN, 21
  gseDO, 22
  gseHPO, 23
  gseMPO, 24
  gseNCG, 25

* package
  DOSE-package, 3

clusterSim, 4
compareClusterResult-class, 5
computeIC, 5

DataSet, 6
DGN_EXTID2PATHID (DataSet), 6
DGN_PATHID2EXTID (DataSet), 6
DGN_PATHID2NAME (DataSet), 6
DO2ALLEG (DataSet), 6
DO2EG (DataSet), 6
DOIC (DataSet), 6
DOLite2EG (DataSet), 6
DOLiteTerm (DataSet), 6
DOSE / DOSE-package, 3
DOSE-package, 3
doseSim, 6
doSim, 7
dotbl (DataSet), 6

EG2ALLDO (DataSet), 6
EG2DO (DataSet), 6

EG2DOLite (DataSet), 6
enrichDGN, 7
enrichDGNv, 8
enrichDO, 10, 15
enricher_internal, 11
enrichHPO, 12
enrichMPO, 13
enrichNCG, 14
enrichResult, 3, 5
enrichResult-class, 15
EXTID2NAME, 16

facet_grid, 29
facet_grid (reexports), 29
gene2DO, 16
geneID, 17
geneInCategory, 17
geneList (DataSet), 6
geneSim, 18
GSEA_internal, 19
gseahResult-class (gseaResult-class), 19
gseaResult-class, 19
gseDGN, 21
gseDO, 22
gseHPO, 23
gseMPO, 24
gseNCG, 25
gsfilter, 26

hpoSim, 26
hpotbl (DataSet), 6

mclusterSim, 27
mpoSim, 28
mpotbl (DataSet), 6

NCG_EXTID2PATHID (DataSet), 6
NCG_PATHID2EXTID (DataSet), 6
NCG_PATHID2NAME (DataSet), 6
INDEX

parse_ratio, 28
plot,compareClusterResult-method
  (compareClusterResult-class), 5
rebuildAnnoData, 29
reexports, 29
setReadable, 29
show, 30
show,compareClusterResult-method
  (compareClusterResult-class), 5
show,enrichResult-method
  (enrichResult-class), 15
show,gseaResult-method
  (gseaResult-class), 19
simplot, 31
summary, 32
summary,compareClusterResult-method
  (compareClusterResult-class), 5
summary,enrichResult-method
  (enrichResult-class), 15
summary,gseaResult-method
  (gseaResult-class), 19
theme_dose, 32

VDGN_EXTID2PATHID (DataSet), 6
VDGN_PATHID2EXTID (DataSet), 6
VDGN_PATHID2NAME (DataSet), 6