# Package: GseaVis (via r-universe)

June 22, 2024

**Title** Implement for 'GSEA' Enrichment Visualization

**Version** 0.1.0

**Maintainer** Jun Zhang <`3219030654@stu.cpu.edu.cn`>

**Description** Mark your interesting genes on plot and support more parameters to handle your own gene set enrichment analysis plot.

**License** MIT + file LICENSE

**Encoding** UTF-8

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.2.3

**Imports** AnnotationDbi, aplot, circlize, clusterProfiler, DOSE, dplyr, ggplot2, ggpp, ggrepel, ggsci, GO.db, grDevices, gson, magrittr, methods, purrr, RColorBrewer, readr, reshape2, stringr, tibble, utils, yulab.utils

**URL** `https://github.com/junjunlab/GseaVis`

**BugReports** `https://github.com/junjunlab/GseaVis/issues`

**Depends** R (>= 2.10)

**LazyData** true

**Repository** https://junjunlab.r-universe.dev

**RemoteUrl** https://github.com/junjunlab/gseavis

**RemoteRef** HEAD

**RemoteSha** f20de046d93172f9f9e9f8e79eab9a3ab040714c

## Contents

---

## `dfGO2gseaResult` *Create gseaResult Object from Enrichment Analysis Output Dataframe*

---

### Description

This function creates a gseaResult object from the results of enrichment analysis using gene set enrichment analysis (GSEA) for a given gene list and Gene Ontology (GO) data.

### Usage

```
dfGO2gseaResult(
  enrich.df = NULL,
  geneList = NULL,
  OrgDb = NULL,
  keytype = "ENTREZID",
  setType = c("BP", "CC", "MF", "ALL"),
  pvalueCutoff = 0.05,
  eps = 1e-10,
  pAdjustMethod = "BH",
  exponent = 1,
  minGSSize = 120,
  maxGSSize = 500
)
```

### Arguments

| | |
|---|---|
| `enrich.df` | A data frame containing the results of enrichment analysis. |
| `geneList` | A decreasing sorted vector of gene list. |
| `OrgDb` | The organism-specific annotation database. |
| `keytype` | The type of gene identifier used in the analysis (default is "ENTREZID"). |
| `setType` | The type of GO ontology to use ("BP" for Biological Process, "CC" for Cellular Component, "MF" for Molecular Function, or "ALL" for all ontologies). |
| `pvalueCutoff` | The p-value cutoff for significance (default is 0.05). |
| `eps` | A small value to avoid division by zero (default is 1e-10). |
| `pAdjustMethod` | The p-value adjustment method (default is "BH" for Benjamini-Hochberg). |
| `exponent` | The exponent for weighting p-values (default is 1). |
| `minGSSize` | The minimum gene set size (default is 120). |
| `maxGSSize` | The maximum gene set size (default is 500). |

## Value

A gseaResult object containing the results of GSEA.

---

## `dfKEGG2gseaResult` *Create gseaResult Object from KEGG Enrichment Analysis Output Dataframe*

---

### Description

This function creates a gseaResult object from the results of KEGG enrichment analysis for a given gene list and organism.

### Usage

```
dfKEGG2gseaResult(
  enrich.df = NULL,
  geneList = NULL,
  organism = "hsa",
  keytype = "kegg",
  setType = "KEGG",
  use_internal_data = FALSE,
  pvalueCutoff = 0.05,
  eps = 1e-10,
  pAdjustMethod = "BH",
  exponent = 1,
  minGSSize = 120,
  maxGSSize = 500
)
```

### Arguments

| | |
|---|---|
| `enrich.df` | A data frame containing the results of KEGG enrichment analysis. |
| `geneList` | A decreasing sorted vector of gene identifiers. |
| `organism` | The organism for KEGG enrichment analysis (default is "hsa" for Homo sapiens). |
| `keytype` | The type of gene identifier used in the analysis (default is "kegg"). |
| `setType` | The type of enrichment analysis ("KEGG" for KEGG pathways). |
| `use_internal_data` | Logical value indicating whether to use internal KEGG data (default is FALSE). |
| `pvalueCutoff` | The p-value cutoff for significance (default is 0.05). |
| `eps` | A small value to avoid division by zero (default is 1e-10). |
| `pAdjustMethod` | The p-value adjustment method (default is "BH" for Benjamini-Hochberg). |
| `exponent` | The exponent for weighting p-values (default is 1). |
| `minGSSize` | The minimum gene set size (default is 120). |
| `maxGSSize` | The maximum gene set size (default is 500). |

### Value

A gseaResult object containing the results of KEGG enrichment analysis.

---

## `dotplotGsea` *dotplotGsea*

---

### Description

dotplotGsea

### Arguments

| | |
|---|---|
| `data` | GSEA enrich object from clusterProfiler, defalut is NULL. |
| `pval` | pvalue cutoff to select significant terms, defalut is NULL. |
| `pajust` | adjusted pvalue cutoff to select significant terms, defalut is 0.05. |
| `order.by` | the X axis, defalut is "GeneRatio". |
| `str.width` | the width of term name, defalut is 50. |
| `base_size` | theme base size, defalut is 12. |
| `topn` | show the top terms, defalut is NULL. |
| `scales` | facet scales, defalut is "free_x". |
| `add.seg` | whether add segment line to point, defalut is "FALSE". |
| `line.col` | segment line color, defalut is "grey80". |
| `line.size` | segment line size, defalut is 1.5. |
| `line.type` | segment line type, defalut is "solid". |

### Value

a ggplot object.

### Author(s)

Jun Zhang

---

## `GSEAmultiGP` *Perform GSEA Multi-group Plotting*

---

### Description

This function generates a multi-group plot for Gene Set Enrichment Analysis (GSEA) results.

### Usage

```
GSEAmultiGP(
  gsea_list = NULL,
  geneSetID = NULL,
  exp_name = NULL,
  addPval = FALSE,
  curve.col = NULL,
  curve.linewidth = 1,
  kegg = FALSE,
  lineSize = 1,
  base_size = 12,
  nesDigit = 2,
  pDigit = 2,
  pvalX = 0.9,
  pvalY = 0.9,
  rect.bm.col = c("#CC3333", "white", "#003366"),
  subplot.heights = c(0.4, 0.2, 0.08),
  legend.position = c(0.85, 0.85),
  rect.bm.label = c("Up regulated", "Down regulated"),
  breaks.n = 6
)
```

### Arguments

| | |
|---|---|
| `gsea_list` | A list of GSEA results for multiple experiments. |
| `geneSetID` | The ID of the gene set to be visualized. |
| `exp_name` | Names of the experiments corresponding to the GSEA results. |
| `addPval` | Logical, indicating whether to add NES (Normalized Enrichment Score) and p-value labels to the plot. Default is FALSE. |
| `curve.col` | A vector of colors for the curves representing different experiments. If NULL, random colors are assigned. |
| `curve.linewidth` | The curve linewidth. Default is 1. |
| `kegg` | Logical, indicating whether the gene set is a KEGG pathway. Default is FALSE. |
| `lineSize` | The size of the lines in the enrichment score curve plot. Default is 1. |
| `base_size` | The base font size for the plot. Default is 12. |

| | |
|---|---|
| `nesDigit` | The number of digits to round NES values to. Default is 2. |
| `pDigit` | The number of digits to round p-values to. Default is 2. |
| `pvalX` | The X-coordinate for placing p-value labels on the plot. Default is 0.9. |
| `pvalY` | The Y-coordinate for placing p-value labels on the plot. Default is 0.9. |
| `rect.bm.col` | A vector of colors for the bottom rectangle representing up-regulated and down-regulated genes. Default colors c("#CC3333","white","#003366") are provided. |
| `subplot.heights` | Heights of subplots in the multi-group plot. Default values c(0.4,0.2,0.08) are provided. |
| `legend.position` | The position of the legend in the plot. Default is (0.85, 0.85). |
| `rect.bm.label` | Labels for the bottom rectangle, specifying "Up regulated" and "Down regulated". Default labels are provided. |
| `breaks.n` | The number of X axis breaks. Default is 6. |

## Value

Returns a multi-panel plot for GSEA results.

## Author(s)

Jun Zhang

## See Also

`gsInfo`, `ggplot2`

## Examples

```
## Not run:
# Example Usage
result <- GSEAmultiGP(gsea_list = gsea_results,
                      geneSetID = "gene_set_1",
                      exp_name = c("Exp1", "Exp2"),
                      addPval = TRUE,
                      curve.col = c("red", "blue"),
                      pvalX = 0.9,
                      pvalY = 0.9,
                      rect.bm.col = c("#CC3333", "white", "#003366"),
                      subplot.heights = c(0.4, 0.2, 0.08),
                      legend.position = c(0.85, 0.85),
                      rect.bm.label = c("Up regulated", "Down regulated"))

## End(Not run)
```

---

`gseaNb` *gseaNb*

---

## Description

gseaNb

## Arguments

| | |
|---|---|
| `object` | GSEA enrich results. |
| `filePath` | filePath the path of the GSEA software enrichment outputs or "readGseaFile" object, defalut is NULL. |
| `subPlot` | which plot to show, 1/2/3, default is 3. |
| `lineSize` | curve line size. default is 0.8. |
| `geneSetID` | which pathway name to plot. |
| `rmSegment` | whether to remove segment on the curve plot, default is FALSE. |
| `termWidth` | the width or the term name, defalut is 40. |
| `segCol` | segment color on the curves, defalut is "red". |
| `addGene` | whether add gene name on the curve, defalut is FALSE. |
| `geneCol` | gene name label color, defalut is NULL. |
| `arrowAngle` | arrow angle, defalut is 20. |
| `arrowLength` | arrow line length, defalut is 0.2. |
| `arrowEnd` | arrow end, defalut is "last". |
| `arrowType` | arrow type, defalut is "closed". |
| `curveCol` | curve color, defalut is c("#76BA99", "#EB4747", "#996699"). |
| `htCol` | heatmap color, defalut is c("#08519C", "#A50F15"). |
| `rankCol` | gene rank fill color, defalut is c("#08519C", "white", "#A50F15"). |
| `rankSeq` | gene rank plot X axis breaks, defalt is 5000. |
| `htHeight` | the relative height when "subplot = 2" to the vertical line plot, defalut is 0.3. |
| `force` | the gene label force, refer to geom_text_repel function, defalut is 20. |
| `max.overlaps` | refer to geom_text_repel function, defalut is 50. |
| `geneSize` | gene label text size, defalut is 4. |
| `newGsea` | whether show new style of plot, defalut is FALSE. |
| `addPoint` | new style plot with point layer, defalut is TRUE. |
| `newCurveCol` | new style plot curve color, defalut is c("#336699", "white", "#993399"). |
| `newHtCol` | new style plot heatmap color, defalut is c("#336699", "white", "#993399"). |
| `rmHt` | whether remove new style plot heatmap, defalut is FALSE. |
| `addPval` | whether add pvalue and NES, defalut is FALSE. |

| | |
|---|---|
| `pvalX` | set pvalue label x position, defalut is 0.9. |
| `pvalY` | set pvalue label y position, defalut is 0.9. |
| `pvalSize` | set pvalue label text size, defalut is 4. |
| `pCol` | pvalue label color, defalut is "grey30". |
| `pHjust` | pvalue label hjust, defalut is 1. |
| `rmPrefix` | whether remove GO term prefix like "GOBP/KEGG/CC/MF_*", defalut is TRUE. |
| `nesDigit` | the NES score digits retained, defalut is 2. |
| `pDigit` | the pvalue and pajust value digits retained, defalut is 2. |
| `markTopgene` | whether add top n genes on plot, defalut is FALSE. |
| `topGeneN` | the number of genes to be marked on plot, defalut is 5. |
| `kegg` | whether input is gseKEGG object, defalut is FALSE. |
| `legend.position` | the legend position, defalut is "right". |
| `whether` | add target gene expression heatmap, defalut is FALSE. |
| `exp` | the expression matrix,tpm/fpkm/rpkm format, defalut is NULL. |
| `scale.exp` | whether scale the expression matrix, defalut is TRUE. |
| `sample.order` | the expression matrix sample orders, defalut is NULL. |
| `exp.col` | the expression colors, defalut is c('blue','white','red'). |
| `ht.legend` | whether show the heatmap legend, defalut is TRUE. |
| `ght.relHight` | the relative height to the main plot, defalut is 0.4. |
| `ght.geneText.size` | the gene lable text size, defalut is 6. |
| `ght.facet` | whether facet expression heatmap, defalut is FALSE. |
| `ght.facet.scale` | the facet plot scale argumrnt, defalut is "free". |
| `termID.order` | the facet term ID orders, defalut is NULL. |
| `rank.gene` | add your gene label on rank plot, defalut is NULL. |
| `rank.gene.nudgey` | the gene label nudge y on rank plot, defalut is 2. |
| `rm.newGsea.ticks` | whether remove right axis when you plot multiple terms with newGsea plot, defalut is TRUE. |
| `pFill` | the pvalue table fill color when you plot multiple terms with newGsea plot, defalut is transparent. |
| `base_size` | the plot theme font size, defalut is 12. |
| `ncol` | the columns for newGSEA plot with multiple terms, defalut is 1 |

## Value

ggplot2 object

## Author(s)

Jun Zhang

## Examples

```
# load data
test_data <- system.file("extdata", "gseaRes.RDS", package = "GseaVis")
gseaRes <- readRDS(test_data)

# all plot
gseaNb(object = gseaRes,
      geneSetID = 'GOBP_NUCLEOSIDE_DIPHOSPHATE_METABOLIC_PROCESS',
      subPlot = 2)
```

---

# `gsInfo` *gsInfo*

---

## Description

gsInfo

## Usage

```
gsInfo(object, geneSetID)
```

## Arguments

| | |
|---|---|
| `object` | gseaResult object |
| `geneSetID` | gene set ID |

## Value

data.frame

## Author(s)

Guangchuang Yu

---

## `gsInfoNew` *gsInfoNew*

### Description

gsInfoNew

### Usage

```
gsInfoNew(geneList = NULL, geneSetID = NULL, geneSet = NULL, exponent = 1)
```

### Arguments

| | |
|---|---|
| `geneList` | geneList for GSEA software outputs which is saved in "gseaOutputs/ranked_gene_list_treat_versus_contr |
| `geneSetID` | gene set ID. |
| `geneSet` | enrichment term sets for GSEA software outputs which is saved in "*/edb/gene_sets". |
| `exponent` | weight of each step, defalut is 1. |

### Value

a data.frame

### Author(s)

Guangchuang Yu, modified by JunZhang

---

## `intergrated` *This is a test data for this package test data describtion*

### Description

This is a test data for this package test data describtion

### Usage

```
intergrated
```

### Format

An object of class `list` of length 3.

### Author(s)

JunZhang

---

## `readGseaFile` *readGseaFile*

---

### Description

readGseaFile

### Arguments

| | |
|---|---|
| `filePath` | the path of the GSEA software enrichment outputs, defalut is NULL. |

### Value

a list contains meta(intergated enrichment results), glist(the ordered gene lists), gset(all background enrichment terms).

---

## `volcanoGsea` *volcanoGsea*

---

### Description

volcanoGsea

### Usage

```
volcanoGsea(
  data = NULL,
  NES.cutoff = 1,
  pvalue.cutoff = NULL,
  p.adjust.CUTOFF = 0.05,
  nudge.y = c(0, 0),
  topN = 5,
  point.size = 3,
  point.color = c("#CC3333", "#CCCCCC", "#0099CC"),
  ...
)
```

### Arguments

| | |
|---|---|
| `data` | GSEA enrich object from clusterProfiler, defalut is NULL. |
| `NES.cutoff` | NES cutoff to select significant terms, defalut is 1. |
| `pvalue.cutoff` | pvalue cutoff to select significant terms, defalut is NULL. |
| `p.adjust.CUTOFF` | adjusted pvalue cutoff to select significant terms, defalut is 0.05. |
| `nudge.y` | y shift to ajust label, defalut is c(0,0). |

| | |
|---|---|
| `topN` | top term to show, defalut is 5. |
| `point.size` | point size, defalut is 3. |
| `point.color` | point color, defalut is c('#CC3333','#CCCCCC','#0099CC'). |
| `...` | other arguments passed by geom_text_repel. |

## Value

a ggplot object.

## Author(s)

Jun Zhang

# Index