Package 'DRviaSPCN'

January 13, 2024

Type Package

Title Drug Repurposing in Cancer via a Subpathway Crosstalk Network

Version 0.1.4

Maintainer Junwei Han <hanjunwei1981@163.com>

Description A systematic biology tool was developed to repurpose drugs via a subpathway crosstalk network. The operation modes include 1) calculating centrality scores of SPs in the context of gene expression data to reflect the influence of SP crosstalk, 2) evaluating drug-disease reverse association based on disease- and drug-induced SPs weighted by the SP crosstalk, 3) identifying cancer candidate drugs through perturbation analysis. There are also several functions used to visualize the results.

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Imports GSVA, clusterProfiler, graphics, igraph, pheatmap, stats, sp, ChemmineR

Depends R (>= 3.6)

Suggests rmarkdown, knitr, DT

VignetteBuilder knitr

NeedsCompilation no

Author Junwei Han [aut, cre, cph], Jiashuo Wu [aut]

Repository CRAN

Date/Publication 2024-01-13 05:00:02 UTC

R topics documented:

`DRviaSPCN-package` *The DRviaSPCN package*

Description

DRviaSPCN is a R package for repurposing drugs via subpathway crosstalk network.

Details

DRviaSPCN

The main goals of the DRviaSPCN package is to provide main functions and some data.

The main functions includes: 1) CalCentralityScore 2) Optimaldrugs and four visualization functions: 1) plotSPW 2) getMolecularFM 3) Disease2SPheatmap 4) Drug2SPheatmap.

For more details, please see browseVignettes("DRviaSPCN")

`CalCentralityScore` *Calculating eigenvector centrality of subpathways*

Description

The function "CalCentralityScore" is used to calculate the eigenvector centrality of subpathways.

Usage

```
CalCentralityScore(ExpData,Label,nperm=1000)
```

Arguments

`ExpData`	A gene expression profile of interest (rows are genes, columns are samples).
`Label`	A character vector consist of "0" and "1" which represent sample class in gene expression profile. "0" means normal sample and "1" means disease sample.
`nperm`	Number of random permutations (default: 1000).

Value

A dataframe with seven columns those are subpathway ID, subpathway name, subpathway size, genes in subpathway, centralscore (eigenvector centrality), Pvalue and FDR.

Examples

```
library(igraph)
#Obtain input data
GEP<-GetExample('GEP')
Slabel<-GetExample('Slabel')
#Run the function
#CentralityScoreResult<-CalCentralityScore(ExpData=GEP,Label=Slabel,nperm=1000)
```

Disease2SPheatmap *Plot a heat map of the subpathways activity regulated by disease*

Description

The "Disease2SPheatmap" function plots a heat map of the subpathways that are regulated by disease. We map subpathways to the disease gene expression through ssgsea to get a subpathway abundance matrix. Then we visualize the matrix by heatmap.

Usage

```
Disease2SPheatmap(CentralityScore,ExpData,Label,pcut=0.05,bk=c(-2,2),
                  cluster.rows=FALSE,cluster.cols=FALSE,show.rownames=TRUE,
                  show.colnames=FALSE,col=c("navy","firebrick3"),
                  cell.width=NA,cell.height=NA,scale="row",fontsize=7,
                  fontsize.row=9,fontsize.col=10)
```

Arguments

CentralityScore	A dataframe with seven columns those are subpath ID, subpathway name, subpathway size, genes in subpathway, centralscore (eigenvector centrality), Pvalue and FDR.
ExpData	A gene expression profile of interest.
Label	A character vector consist of "0" and "1" which represent sample class in gene expression profile. "0" means normal sample and "1" means disease sample.
pcut	A numeric value which represent threshold. Subpathways with p-value less than this threshold will be screened out and visualized.
bk	A numeric vector that covers the range of values. Users could adjust color depth through this parameter.
cluster.rows	Boolean values determining if rows should be clustered or hclust object.
cluster.cols	Boolean values determining if columns should be clustered or hclust object.

show.rownames	Boolean specifying if row names are be shown.
show.colnames	Boolean specifying if column names are be shown.
col	Vector of colors used in heatmap.
cell.width	Individual cell width in points. If left as NA, then the values depend on the size of plotting window.
cell.height	Individual cell height in points. If left as NA, then the values depend on the size of plotting window.
scale	Character indicating if the values should be centered and scaled in either the row direction or the column direction, or none. Corresponding values are "row", "column" and "none".
fontsize	Base fontsize for the plot (default: 10).
fontsize.row	Fontsize for rownames (default: 10).
fontsize.col	Fontsize for colnames (default: 10).

Value

A heat map

Examples

```
#Load depend package
library(GSVA)
library(pheatmap)
#Obtain input data (The "CentralityScoreResult" is the result of function "CalCentralityScore")
GEP<-GetExample('GEP')
Slabel<-GetExample('Slabel')
CentralityScoreResult<-GetExample('CentralityScoreResult')
#Run the function
#Disease2SPheatmap(CentralityScore=CentralityScoreResult,ExpData=GEP,Label=Slabel,
#                  pcut=0.05,bk=c(-2,2),cluster.rows=FALSE,cluster.cols=FALSE,
#                  show.rownames=TRUE,show.colnames=FALSE,col=c("navy","firebrick3"),
#                  cell.width=NA,cell.height=NA,scale="row",fontsize=7,
#                  fontsize.row=9,fontsize.col=10)
```

Drug2SPheatmap *Plot a heat map of the subpathways activity regulated by drugs*

Description

The function "Drug2SPheatmap" plots heatmaps of the subpathways that are regulated by drugs. We map subpathways to the disease gene expression through ssgsea to get a subpathway abundance matrix. Then we visualize the matrix by heatmap.

Usage

```
Drug2SPheatmap(drugname="",DrugSPPvalue,ExpData,Label,pcut=0.05,
         bk=c(-2,2),cluster.rows=FALSE,cluster.cols=FALSE,show.rownames=TRUE,
         show.colnames=FALSE,col=c("navy","firebrick3"),
         cell.width=NA,cell.height=NA,scale="row",fontsize=7,
         fontsize.row=9,fontsize.col=10)
```

Arguments

drugname	A character which represent interest drug name with specific concentration, cell line and duration.
DrugSPPvalue	A matrix which colunms represent drugs and rows respresent subpathways. Values in this matrix is the pvalue of subpathways centrality score regulated by drugs.
ExpData	A gene expression profile of interest.
Label	A character vector consist of "0" and "1" which represent sample class in gene expression profile. "0" means normal sample and "1" means disease sample.
pcut	A numeric value which represent threshold. Subpathways with p-value less than this threshold will be screened out and visualized.
bk	A numeric vector that covers the range of values. Users could adjust color depth through this parameter.
cluster.rows	Boolean values determining if rows should be clustered or hclust object.
cluster.cols	Boolean values determining if columns should be clustered or hclust object.
show.rownames	Boolean specifying if row names are be shown.
show.colnames	Boolean specifying if column names are be shown.
col	Vector of colors used in heatmap.
cell.width	Individual cell width in points. If left as NA, then the values depend on the size of plotting window.
cell.height	Individual cell height in points. If left as NA, then the values depend on the size of plotting window.
scale	Character indicating if the values should be centered and scaled in either the row direction or the column direction, or none. Corresponding values are "row", "column" and "none".
fontsize	Base fontsize for the plot (default: 10).
fontsize.row	Fontsize for rownames (default: 10).
fontsize.col	Fontsize for colnames (default: 10).

Value

A heat map

Examples

```
#Load depend package
library(GSVA)
library(pheatmap)
##Obtain input data
#Statistic significance of centrality score of subpathways
#induced by each drug were stored in package "DRviaSPCNData".
#"DRviaSPCNData" has been uploaded to the github repository.
#Users can download and install through "install_github" function and
#set parameter url="hanjunwei-lab/DRviaSPCNData".
#After installing and loading package "DRviaSPCNData",
#users can use the following command to get the data:
#DrugSPPvalueMatrix<-GetData('DrugSPPvalueMatrix')
GEP<-GetExample('GEP')
Slabel<-GetExample('Slabel')

#Run the function
#Drug2SPheatmap(drugname = "methotrexate_HL60_6_8.8e-06",
#                 DrugSPPvalue=DrugSPPvalueMatrix,
#                 ExpData=GEP,Label=Slabel,pcut=0.05,bk=c(-2,2),
#                 cluster.rows=FALSE,cluster.cols=FALSE,show.rownames=TRUE,
#                 show.colnames=FALSE,col=c("navy","firebrick3"),
#                 cell.width=NA,cell.height=NA,scale="row",fontsize=7,
#                 fontsize.row=9,fontsize.col=10)
```

`envData` *An environment variable which includes some example data*

Description

An environment variable which includes some example data. CentralityScoreResult: The result of function "CalCentralityScore". Drugs_CID: PubCham Database ID of drugs. GEP: An example gene expression profile. GOInfo: Biological Process data from Gene Ontology. GoSubPconGene: Genes symble shared by a pair of subpathways with Biological Process. Jaccardscore: Jaccard score shared by a pair of subpathways with Biological Process. Slabel: A character vector consist of "0" and "1" which represent sample class in gene expression profile. Opdrugresult: The result of function "Optimaldrugs". SubPathwayInfo: Subpathway information from SubpathwayMiner. SubPathwaymapdata: Subpathway structure data.

Usage

```
envData
```

Format

An environment variable

getDEscore *Calculating Log2 Fold Change of genes*

Description

Function "getDEscore" uses gene expression profile to calculate Log2 Fold Change of genes.

Usage

```
getDEscore(inexpData, Label)
```

Arguments

inexpData	A gene expression profile of interest (rows are genes, columns are samples).The data in the expression profile is best not be log2 converted.
Label	A character vector consist of "0" and "1" which represent sample class in gene expression profile. "0" means normal sample and "1" means disease sample.

Value

A one-column matrix of Log2 Fold Change which rownames is gene.

GetExample *Get example data*

Description

This function is used to achieve exxample data.

Usage

```
GetExample(exampleData)
```

Arguments

exampleData	A character, should be one of"GEP","Slabel", "CentralityScoreResult" and "Opdrugresult".

Value

example data

`getMolecularFm` *Plot chemical molecular formula of drugs*

Description

The function "getMolecularFm" outputs the chemical molecular formula of a drug or compound . The results can be visualized by the "plot" function.

Usage

```
getMolecularFm(drugid = NULL, drugname = NULL,sdfSET,main = "", sub = "")
```

Arguments

`drugid`	A character string of DrugBank ID.
`drugname`	A character string of drug name.
`sdfSET`	Sdf data of drug structure.
`main`	An overall title for the chemical structure graph.
`sub`	A sub title for the chemical structure graph.

Value

Chemical molecular formula of the drug or compound.

Examples

```
#"sdfSET" has been uploaded to the
#github repository.Users can download and install through "install_github"
#function and set parameter url="hanjunwei-lab/DRviaSPCNData".
#After installing and loading package "DRviaSPCNData",
#users can use the following command to get the data.
# Obtain molecular formula and visualize it.

#Get the sdf data of drug structure from DRviaSPCNData package
#library("Chemminer")
#sdf<-GetData('sdfSET')
#Run the function
#Mole_formula<-getMolecularFm(drugname ="methotrexate",sdfSET=sdf)
#plot(Mole_formula)
```

Optimaldrugs *Identifying the optimal drugs*

Description

Function "Optimaldrugs" used to identify the optimal drugs for specific disease.

Usage

```
Optimaldrugs(ExpData,Label,DrugSPESC,CentralityScore,nperm=1000,
                    topcut=10,pcut=0.01,weight=FALSE)
```

Arguments

ExpData	A gene expression profile of interest (rows are genes, columns are samples).
Label	A character vector consist of "0" and "1" which represent sample class in gene expression profile. "0" means normal sample and "1" means disease sample.
DrugSPESC	A matrix with n rows and m columns. n is the number of subpathways and m is the number of all drugs. The values in this matrix is weighted enrichmentscore of subpathways induced by each drug. The users could obtain this matrix from our example data.
CentralityScore	The result of function "CalCentralityScore".
nperm	Number of random permutations (default: 1000).
topcut	The parameter "topcut" represents the number of selected SPs from the top or bottom of the ranked SP list. The topcut defaults to 10.
pcut	The parameter "pcut" represents the threshold of statistical significance for screen SPs. The pcut defaults to 0.01.
weight	A boolean value determines the method for calculating the drug-disease association score of the drug. "weight=FALSE"(default): Similar to "CMap" (Lamb et al., 2006), no weight is needed. "weight=TRUE": KS random walk statistic with individualized subpathway activity score as weight was used to calculate the drug-disease reverse association score.

Value

A dataframe with four columns which are "Drug"(drug names),"DES"(drug enrichment score),"p-value"(statistical significance),"FDR"(adjusted statistical significance).

Examples

```
##Obtain input data
#Weighted enrichmentscore of subpathways induced by each drug were stored
#in package "DRviaSPCNData". "DRviaSPCNData" has been uploaded to the
#github repository.Users can download and install through "install_github"
#function and set parameter url="hanjunwei-lab/DRviaSPCNData".
```

```
#After installing and loading package "DRviaSPCNData",
#users can use the following command to get the data.
#DrugSPESCMatrix<-GetData('DrugSPESCMatrix')
CentralityScoreResult<-GetExample("CentralityScoreResult")
GEP<-GetExample("GEP")
Slabel<-GetExample("Slabel")
#Run the function
#Opdrugresult<-Optimaldrugs(ExpData=GEP,Label=Slabel,DrugSPESC=DrugSPESCMatrix,
#CentralityScore=CentralityScoreResult,nperm=1000,topcut=10,pcut=0.01,weight=FALSE)
```

PackageLoaded *PackageLoaded*

Description

Determine if the package is loaded, if no package is loaded.

Usage

```
PackageLoaded(name)
```

Arguments

name	A character which is the name of package.

Value

A boolean value.

---

plotSPW *Plot subpathway network graph*

Description

The function plotSPW can visualize subpathway network graph.

Usage

```
plotSPW(
  subpathwayID,
  layout = NULL,
  margin = 0,
  vertex.label.cex = 0.8,
  vertex.label.font = 0.8,
  vertex.label.dist = 1,
  vertex.size = 12,
```

```
  edge.arrow.width = 3,
  edge.label.cex = 0.6,
  vertex.label.color = "black",
  vertex.color = "#F08080",
  vertex.frame.color = "dimgray",
  edge.color = "grey70",
  edge.label.color = "dimgray",
  sub = NULL,
  main = NULL
)
```

Arguments

subpathwayID	Subpathway id .A character vector.
layout	A matrix of x-y coordinates with two dims. Determine the placement of the nodes for drawing a graph.
margin	A numeric. The value is usually between -0.5 and 0.5, which is able to zoom in or out a subpathway graph. The default is 0.
vertex.label.cex	A numeric vector of node label size.
vertex.label.font	A numeric vector of label font.
vertex.label.dist	A numeric vector of label dist.
vertex.size	A numeric vector of Node size. See plot.igraph.
edge.arrow.width	Edge arrow width. The default is 3.
edge.label.cex	Edge label size.
vertex.label.color	A vector of node label colors. The default is black.
vertex.color	A vector of node colors. The default is the KEGG node color.
vertex.frame.color	A vector of node frame color. The default is dimgray.
edge.color	A vector of edge color. The default is dimgray
edge.label.color	A vector of edge label color. The default is dimgray.
sub	A character string of subtitle.
main	A character string of main title.

Value

a subpathway map

Examples

```
# load depend package
library(igraph)
# plot network graph of the subpathway "00020_4"
plotSPW("00020_4")
```

Index