Package 'DrInsight'

June 3, 2020

Title Drug Repurposing: Integration and Systematic Investigation of
Genomic High Throughput Data

Version 0.1.2

Author Jinyan Chan [aut, cre], Jinghua Gu [aut]

Maintainer Jinyan Chan <Jinyan_Chan@baylor.edu>

Description
This is a connectivity mapping-based drug repurposing tool that identifies drugs that can potentially reverse query disease phenotype or have similar functions with query drugs.

Depends R (>= 3.4.0),igraph (>= 1.1.2), qusage (>= 2.12.0)

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Suggests knitr (<= 1.18),rmarkdown (<= 1.8)

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-03 04:10:03 UTC

R topics documented:

---

brca.tcga	*A sample breast cancer gene differential expression analysis t statistic score data*

Description

This data set gives the TCGA breast cancer tumor versus normal t-test scores of each gene.

Usage

```
brca.tcga
```

Format

A matrix containing the t statistics of 17960 genes

Source

TCGA

CEG.network.graph	*Plot The Pathway-CEG Network for A Drug*

Description

This function allows user to plot the pathway-CEG network for a specified pathway in a specified drug.

Usage

```
CEG.network.graph(
  path.analysis.res = path.analysis.res,
  pathway.FDR.cutoff = 0.1,
  drug.name = NULL,
  pathway.name = NULL,
  show.plot = TRUE
)
```

Arguments

`path.analysis.res`
> The pathway analysis results. Output of function pathway.analysis().

`pathway.FDR.cutoff`
> The FDR threshold to select significant drug specific pathways and the default is 0.1.

`drug.name`
> The name of the drug user would like to analyze.The specified drug name should among those that are listed in the output table of function network.graph().

`pathway.name`
> The name of the pathway user would like to analyze. The specified pathway name should among those that are listed in the output table of function network.graph().

`show.plot`
> True or False, specifying if you want to show the pathway-CEG plot for a specific drug and pathway.

Examples

```
## get the Dr. Insight drug identification results
drug.ident.res = drug.ident(query.data = example.disease, cmap.ref.profiles = example.drug.profiles,
                repurposing.unit = "treatment", connectivity = "negative")

## load in example pathway data
data("example.pathway")

## Performe pathway analysis (for the drugs that are identified by ident.drug())
path.analysis.res = pathway.analysis(drug.ident.res = drug.ident.res,
                    pathway.list = example.pathway,drug.FDR.cutoff = 0.5)

path.CEG.network = CEG.network.graph(path.analysis.res = path.analysis.res,
                pathway.FDR.cutoff = 0.5,drug.name = "drug1",
                pathway.name = "pathway5", show.plot = TRUE)
```

---

drug.ident	*Dr. Insight Drug Identification*

Description

This function allows user to use the differential expression data of their interested disease phenotype or drug profile to query against CMap reference drug expression profiles. This function returns an obejct of three elements: a table of drug p-values reflecting the connectivity between query data and CMap drugs; a table contains all the meta information of CMap drugs; and an object of p-values of identified CEGs for each drug instance.

Usage

```
drug.ident(
  query.data = NULL,
  cmap.ref.profiles = NULL,
  repurposing.unit = "treatment",
  CEG.threshold = 0.05,
  connectivity = "negative"
)
```

Arguments

query.data	User provided differential expression analysis results(e.g. t-test statistic scores) of querying data (either disease phenotype data, or drug expression data). The column names for gene symboles and statistic scores must be "geneSymbol" and "score".
cmap.ref.profiles	
	Dr. Insight provided CMap drug rank matrix containing all 6100 instances of CMap data set.The example of reference data can be loaded with data("example.profiles"). The real cmap data can be loaded with the function 'get.cmap.ref' (See instructions in vignette).
repurposing.unit	
	The parameter of either "treatment" or "drug", which indicates if user want the algorithm to test drug repurposing p value at treatment level or drug level. The default is "treatment", which treats the drug data from different cell lines separately.
CEG.threshold	The p value threshold to select the consistently differential expressed genes (CEGs).
connectivity	The type of connectivity, either "negative" or "positive". Negative connectivity is used when the query data is the differential scores from disease data, and Dr. Insight will repurpose drugs that can potentially reverse the query disease phenotype. Positive connectivity is used when the query data is from a drug profile, and Dr. Insight will return the drugs that are similar to the query drug.

Examples

```
## load in the example query disease data
data("example.disease")
data("example.drug.profiles")

## get the Dr. Insight drug identification results
drug.ident.res = drug.ident(query.data = example.disease, cmap.ref.profiles = example.drug.profiles,
                  repurposing.unit = "treatment", connectivity = "negative")

drug.pvals = drug.ident.res$drug.pvals
```

| drug.info | *The matrix containing the drug instances information of CMap drug data* |

Description

This matrix gives the drug instance information of CMap drug data, including drug names, instance numbers, cell line, dosage, etc.

Usage

```
drug.info
```

Format

A matrix of 6100 rows and 7 columns.

Source

CMap

| example.disease | *An example gene differential expression analysis test statistic score data* |

Description

This data gives an example of gene differential expression analysis test statistic score as the inpur for DrInsight analysis.

Usage

```
example.disease
```

Format

A matrix containing the simulated scores of 1000 genes

Source

simulation

---

example.drug.profiles   *An example cmap drug reference profiles*

Description

This data gives an simulated example of cmap drug perturbed reference profiles.

Usage

```
example.drug.profiles
```

Format

An object containing the simulated reference drug rank matrix and drug information table

Source

simulation

---

example.pathway   *example.pathway*

Description

This data set gives the geneset list of a simulated pathways.

Usage

```
example.pathway
```

Format

A list containing 10 pathways.

Source

simulation

---

get.cmap.ref *Read In and Process CMap Reference Profiles*

Description

This function allows user to load in the CMap drug rank matrix user downloaded from CMap website: https://portals.broadinstitute.org/cmap/ (data matrix in the "downloads" section).

Usage

```
get.cmap.ref(cmap.data.path = NULL, probe.to.genes = NULL, drug.info = NULL)
```

Arguments

cmap.data.path   The local path and the name of the downloaded data matrix file. The data matrix file should be in txt format.

probe.to.genes   The ID converter between Affymetrix probe IDs (the IDs used in CMap data matrix) and official gene symbol. The packge comes with an embeded probe.to.genes file that can be directly used. User can use their own converter file with two columns named "ID", and "Gene.Symbol'.

drug.info   The drug instance information from CMap data. The deafault drug.info file comes with DrInsight package.

---

make.cytoscape.network

Drug and Pathway connection output files for Cytoscape visulization

---

Description

This fucniton allows user to get the two files needed for Cytoscape to visulize the drug-pathway network

Usage

```
make.cytoscape.network(
  path.analysis.res = path.analysis.res,
  pathway.FDR.cutoff = 0.1
)
```

Arguments

path.analysis.res

The pathway analysis results. Output of pathway.analysis().

pathway.FDR.cutoff

The FDR threshold to select significant drug specific pathways and the default is 0.1.

Examples

```
## get the Dr. Insight drug identification results
drug.ident.res = drug.ident(query.data = example.disease, cmap.ref.profiles = example.drug.profiles,
                  repurposing.unit = "treatment", connectivity = "negative")

## load in example pathway data
data("example.pathway")

## Performe pathway analysis (for the drugs that are identified by ident.drug())
path.analysis.res = pathway.analysis(drug.ident.res = drug.ident.res,
                      pathway.list = example.pathway, drug.FDR.cutoff = 0.5)

## get the pathway analysis ouput that can be loaded into Cytoscape for visulization
network.cytoscape = make.cytoscape.network(path.analysis.res = path.analysis.res,
pathway.FDR.cutoff = 0.5)
```

---

network.graph   *Plot Drug and Pathway Interaction Network*

Description

This function allows user to plot the drug-pathway interaction network and returns the drug-pathway interaction contigency table.

Usage

```
network.graph(
  path.analysis.res = path.analysis.res,
  pathway.FDR.cutoff = 0.1,
  pathway.label = FALSE,
  drug.label.size = 1.2,
  path.label.size = 0.8,
  return.adj.table = TRUE
)
```

Arguments

path.analysis.res

>   The pathway analysis results. Output of function pathway.analysis().

pathway.FDR.cutoff

>   The FDR threshold to select significant drug specific pathways and the default
>   is 0.1.

pathway.label   True or False specofying if show pathway labels in the graph.

drug.label.size

>   The number indicating the size of drug labels in the graph.

```
path.label.size
```
                    The number indicating the size of pathway labels in the graph.
```
return.adj.table
```
                    True or False specifying if return the resulted drug-pathway contigency table.
                    In the returned table, 1's represent that the pathway is up-regulated by the drug;
                    -1's reprensent down-regulation and 0's no-regulation.

Examples

```
## get the Dr. Insight drug identification results
drug.ident.res = drug.ident(query.data = example.disease, cmap.ref.profiles = example.drug.profiles,
                    repurposing.unit = "treatment", connectivity = "negative")

## load in example pathway data
data("example.pathway")

## Performe pathway analysis (for the drugs that are identified by ident.drug())
path.analysis.res = pathway.analysis(drug.ident.res = drug.ident.res,
                    pathway.list = example.pathway,drug.FDR.cutoff = 0.5)

drug.pathway.network = network.graph(path.analysis.res, pathway.FDR.cutoff = 0.5,
                        return.adj.table = TRUE, pathway.label = TRUE)
```

---

pathway.analysis	*Drug Mechanism of Action Pathway Analysis*

Description

This function allows user to run pathway analysis on the identified significat drugs, therefore to
detect drug mechanism related pathways, and the CEGs in the pathways.

Usage

```
pathway.analysis(
  drug.ident.res = drug.ident.res,
  pathway.list = NULL,
  pathway.list.path = NULL,
  drug.FDR.cutoff = 0.1,
  CEG.threshold = 0.05
)
```

Arguments

```
drug.ident.res
```
Dr. Insight drug identification analysis results. Output of function drug.ident().

```
pathway.list
```
The pathways used to analyze drug mechanism. Should be a list of pathways
where the names of lists are pathway names. Dr.Insight provides NIH pathway
interaction database (PID) pathways, which can be called by data("pathway.PID").
If you want to analyze with other pathways (from GSEA MsigDB), please use
the parameter: pathway.list.path.

pathway.list.path

> The local path and file name of the pathways downloaded from GSEA MsigDB
> http://software.broadinstitute.org/gsea/msigdb/collections.jsp, if pathway.list is
> NULL. Should be in full directory and file name (See instructions in vignette).

drug.FDR.cutoff

> The FDR threshold to select significant repurposed drugs by Dr. Insight for
> pathway analysis.

CEG.threshold   The p value threhold for selecting significant consistently differential expressed
                 genes (CEGs).

Examples

```
## get the Dr. Insight drug identification results
drug.ident.res = drug.ident(query.data = example.disease, cmap.ref.profiles = example.drug.profiles,
                 repurposing.unit = "treatment", connectivity = "negative")

## load in example pathway data
data("example.pathway")

## Performe pathway analysis (for the drugs that are identified by ident.drug())
path.analysis.res = pathway.analysis(drug.ident.res = drug.ident.res,
                     pathway.list = example.pathway, drug.FDR.cutoff = 0.5)
```

---

pathway.PID   *pathway.PID*

Description

This data set gives the geneset list of PID pathways

Usage

```
pathway.PID
```

Format

A large list containing 222 pathways.

Source

PID genesets from Pahtway Commons

probe.to.genes	*The matrix of Affymetrix probe IDs and gene symbols relationship*

Description

This matrix gives the resource to convert the Affymetrix probe IDs in CMap data into gene symbols.

Usage

```
probe.to.genes
```

Format

A matrix of two columns and 12994 rows

Source

Affymetrix

Index