Package 'DeepTarget'

November 14, 2024

Type Package

Title Deep characterization of cancer drugs

Version 1.0.0

Description This package predicts a drug's primary target(s) or secondary target(s) by integrating large-scale genetic and drug screens from the Cancer Dependency Map project run by the Broad Institute. It further investigates whether the drug specifically targets the wild-type or mutated target forms. To show how to use this package in practice, we provided sample data along with step-by-step example.

License GPL-2

Encoding UTF-8

biocViews GeneTarget, GenePrediction,Pathways, GeneExpression, RNASeq, ImmunoOncology,DifferentialExpression, GeneSetEnrichment, ReportWriting,CRISPR

RoxygenNote 7.2.3

VignetteBuilder knitr

Suggests BiocStyle, knitr, rmarkdown

Imports fgsea, ggplot2, stringr, ggpubr, BiocParallel, pROC, stats, grDevices, graphics, depmap, readr, dplyr

Depends R (>= 4.2.0)

git_url https://git.bioconductor.org/packages/DeepTarget

git_branch RELEASE_3_20

git_last_commit 5e2d027

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2024-11-13

Author Sanju Sinha [aut], Trinh Nguyen [aut, cre] (<https://orcid.org/0000-0002-6606-6948>), Ying Hu [aut]

Maintainer Trinh Nguyen <tinh.nguyen@nih.gov>

Contents

`computeCor` *Compute a correlation between the every gene vs each drug response*

Description

Compute correlations between the viability of cell lines after CRISPR Knock Out of each gene and of the same cell lines after drug treatment.

Usage

```
computeCor(DrugName,DRS,GES)
```

Arguments

`DrugName`	Drug Name
`DRS`	Drug's response scores
`GES`	Gene effect scores from Knock-out method such as CRISPR.

Value

a list of matrices for the interesting drugs, where each matrix containing gene names with the correlation values and P values associated with response scores from a given drug ID.

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
```

```
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out ) <- S.Drugs
head(sim.out)
```

`Depmap2DeepTarget` *Retrieval and preparation of input data required from Depmap to Deeptarget package.*

Description

Retrieve gene expression, Cripr, mutation data from KO method, and drug matrix and then preperation the matrix compatible as input for Deeptarget.

Usage

```
Depmap2DeepTarget(FileN,version)
```

Arguments

`FileN`	File Named used as input for DeepTarget: "CCLE_expression.csv", "CRISPRGeneEffect.csv" ,"OmicsSomaticMutations.csv", or "secondary-screen-dose-response-curve-parameters.csv"
`version`	Version of data

Value

a data frame for each required input data

Author(s)

Trinh Nguyen, Ying Hu, and sanju

Examples

```
library(readr)
library(depmap)
# expresion
CCLE.exp <- Depmap2DeepTarget("CCLE_expression.csv","19Q4")
```

`DMB` *Predicting Drug Mutant Binding for mutant or non-mutant form*

Description

Predicting whether the drug is likely bind to mutant or non-mutant form and also generates the plot for visualization.

Usage

```
DMB(DrugName,GOI,Pred,Mutant,DRS,GES,plot=TRUE)
```

Arguments

DrugName	Drug of interest
GOI	Gene of interest
Pred	Prediction object resulting from both PredTarget and PredMaxSim functions to predict whether it is a primary target or secondary target
Mutant	Mutant matrix
DRS	Drug response matrix
GES	Gene Effect Scores
plot	Default is TRUE for plotting

Value

The plot of viability after KO as the X-axis vs drug response in a mutant target as the Y-axis.

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
S.Drugs <- c('K70301465','K09951645')
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
d.mt <- OntargetM$mutations_mat
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim.out,Meta.data)
Drug.Gene.max.sim <- PredMaxSim(sim.out,Meta.data)
identical ( DrugTargetSim[,1],Drug.Gene.max.sim[,1])
Pred.d <-cbind (DrugTargetSim,Drug.Gene.max.sim)
DOI = 'dabrafenib'
GOI = 'BRAF'
DMB (DOI,GOI,Pred.d,d.mt,sec.prism,KO.GES)
```

DoInteractExp *Compute the interaction between the drug and KO expression*

Description

Computes interaction between the drug and KO expression in term of lower vs higher expression using linear model.

Usage

```
DoInteractExp(Predtargets,Exp,DRS, GES,CutOff=3)
```

Arguments

Predtargets	a dataframe of drugs information and their most targeted gene with stats of correlation
Exp	Expression matrix
DRS	Drug scores matrix
GES	Gene effect scores matrix from KO method
CutOff	desired cut-off for low expression

Value

A list of drug names with their interaction values from two groups low and high expression based on the desired cut-off.

drug1	interaction with estimate and P vals from the linear model
drug2	interaction with estimate and P vals from the linear model
drugN	interaction with estimate and P vals from the linear model

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim.out,D.M = Meta.data)
d.expr <- OntargetM$expression_20Q4
ExpInteract <- DoInteractExp (DrugTargetSim,d.expr,sec.prism,KO.GES,CutOff = 2)
```

DoInteractMutant *Compute interaction between the drug and KO expression in term of mutant vs non-mutant*

Description

Compute interaction between the drug and KO expression in term of mutant vs non-mutant

Usage

```
DoInteractMutant(Predtargets,Mutant,DRS,GES)
```

Arguments

Predtargets	a dataframe of drugs information and their most targeted gene with stats of correlation
Mutant	Mutant matrix
DRS	Drug scores matrix
GES	Gene effect scores matrix from KO method

Value

A list of drug names with their interaction values from two groups mutant and non-mutant

drug1	interaction with estimate and P vals from the linear model
drug2	interaction with estimate and P vals from the linear model
drugN	interaction with estimate and P vals from the linear model

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
sim <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim,Meta.data)
d.mt <- OntargetM$mutations_mat
MutantInteract <- DoInteractMutant (DrugTargetSim,d.mt,sec.prism,KO.GES)
```

DoPWY *Provide a probability score for each pathway for the primary of mechanism of action (MOA) of a drug*

Description

Predicts a Primary Target at a pathway Level. It next finds the pathways that are most enriched in the genes with high DKS scores. It does this by performing a pathway enrichment test on the ranked gene list by DKS score. The output is a data frame of pathway-level probabilities for each drug to be the primary of mechanism of action.

Usage

```
DoPWY(Sim.GES.DRS,D.M)
```

Arguments

Sim.GES.DRS	The list of result from "GetSim" function.
D.M	meta data from drug

Value

a list of drugs, where each of them is data frame containing the pathway level probability to be a primary of mechanism of action.

drug1	a dataframe contain the pathway level probability to be a primary MOA
drug2	a dataframe contain the pathway level probability to be a primary MOA
drugN	a dataframe contain the pathway level probability to be a primary MOA

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
sim <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
Pwy.Enr <- DoPWY(sim,Meta.data)
```

DTR *Predicting Drug Target Response (DTR) for primary or secondary targets*

Description

Predicting whether the drug is likely response to primary or secondary targets and also generates the plot for visualization.

Usage

```
DTR(DN,GN,Pred,Exp,DRS,GES,CutOff= 3,plot = TRUE)
```

Arguments

DN	Drug of interest
GN	Gene of interest
Pred	Prediction object, an output result from prediction whether it is a primary target or secondary target
Exp	Expression matrix

DRS	Drug response matrix
GES	Gene Effect Scores
plot	whether users want to plot, default is true
CutOff	cutoff value for gene expression of gene of interest high or low

Value

vialbility after KO vs drug response of gene of interest low vs high cut-off values set by users

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
S.Drugs <- c('K70301465','K09951645')
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
d.expr <- OntargetM$expression_20Q4
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim.out,Meta.data)
Drug.Gene.max.sim <- PredMaxSim(sim.out,Meta.data)
identical ( DrugTargetSim[,1],Drug.Gene.max.sim[,1] )
Pred.d <-cbind (DrugTargetSim,Drug.Gene.max.sim )
DOI = 'ibrutinib'
GOI ='BTK'
DTR(DOI,GOI,Pred.d,d.expr,sec.prism,KO.GES,CutOff= 2)
```

OntargetM *An object containing a small part of the data from the Cancer Dependency Map (depmap.org) to demonstrate in DeepTarget pipeline*

Description

An object containing Viability matrix after CRISPR-KO; Viability after Drug Treatment; Drug metadata from Broad, mutation matrix, and expression matrix with common cell-lines and common drugs. This is a subset of the total data due to memory constraints, full data can be downloaded from depmap.org/portal.

Usage

```
data("OntargetM")
```

Format

A list of one dataframe and 4 matrices

`DrugMetadata` a dataframe containing 11 unique drugs as rownames with their associated information: broad_id_trimmed as ID of the drug, name, target, drug_category, and moa as columns

`secondary_prism` a viability scores matrix (after Drug Treatment) with 16 drugs as row names across 392 unique celllines as column names

`avana_CRISPR` a Gene effect scores (after CRISPR-KO) matrix for 487 genes as row names across 392 unique celllines as column names

`mutations_mat` Mutation binary matrix for 476 genes as row names across 392 unique cell lines as column names; 0 is WT; 1 is mutated

`expression_20Q4` Expression matrix for 550 genes as row names across 392 unique celllines as column names

Details

For a full list data used in the paper, please use the link below to download data

Source

DrugMetadata: Please download full data from this link https://depmap.org/repurposing/#:~:text=Corsello_supplemental_tables.xlsx

Secondary prism: please download full data from this link https://depmap.org/portal/download/all/?releasename=PRISM+Repurposing+19Q4&filename=secondary-screen-dose-response-curve-parameters.csv

avana_CRISPR: please download full data from this link https://depmap.org/portal/download/all/?releasename=DepMap+Public+22Q4&filename=CRISPRGeneEffect.csv

mutations_mat: Please download full data from this link https://depmap.org/portal/download/all/?releasename=DepMap+Public+22Q4&filename=OmicsSomaticMutations.csv

expression_20Q4: Please download full data of file named "CCLE_expression.csv" from this link https://depmap.org/portal/download/all/

Examples

```
data(OntargetM)
```

---

plotCor *Plot the correlation*

Description

Plot the correlation of a predicted target

Usage

```
plotCor(DN,GN,Pred,DRS,GES,plot=TRUE)
```

Arguments

DN	Drug Name
GN	Gene Name
Pred	Output from prediction object
DRS	Drug response score
GES	Gene Effect scores
plot	default is plot=TRUE

Value

Correlation plot

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
S.Drugs <- c('K70301465','K09951645')
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
d.expr <- OntargetM$expression_20Q4
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim.out,Meta.data)
Drug.Gene.max.sim <- PredMaxSim(sim.out,Meta.data)
identical ( DrugTargetSim[,1],Drug.Gene.max.sim[,1] )
Pred.d <-cbind (DrugTargetSim,Drug.Gene.max.sim )
DOI = 'ibrutinib'
GOI ='BTK'
plotCor (DOI,GOI,Pred.d,sec.prism,KO.GES)
```

---

plotSim *Plot the similarty between corelation values and P vals for all genes. The top 5 genes are labeled.*

Description

Plot the similarty between corelation values and P val;

Usage

```
plotSim(dx,dy,clr=NULL, plot=TRUE)
```

Arguments

dx	a matrix of p vals
dy	a matrix of correlation vals
clr	Desired range of color
plot	default plot =TRUE

Value

a plot of similarity

Author(s)

Ying Hu,Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
Sample.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
sim.out <- bplapply(Sample.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- Sample.Drugs
P.Values=vapply(sim.out, function(x) x[,1],FUN.VALUE=numeric(nrow(sim.out[[1]])))
estimate.cor.values=vapply(sim.out, function(x) x[,2],FUN.VALUE=numeric(nrow(sim.out[[1]])))
par(mar=c(4,4,5,2), xpd=TRUE, mfrow=c(3,3));
plotSim(dx=P.Values,dy=estimate.cor.values);
```

PredMaxSim *Predict the most similar gene to the drug response*

Description

Predicts the gene that has the most similarity associated with drug's response scores from the set of all genes.

Usage

```
PredMaxSim (Sim.GES.DRS,D.M)
```

Arguments

Sim.GES.DRS	similarity between Drug's response scores and Gene effect scores from Knock-out method such as CRISPR
D.M	Drug Metadata

Value

a dataframe of drug(s) information with the most predicted gene(s) with the max corelation value(s), P value(s), and FDR value(s).

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data (OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
Drug.Gene.max.sim <- PredMaxSim(sim.out,Meta.data)
```

PredTarget *Prediction of the most similar known targeted gene.*

Description

Predicts the gene that has the most similarity to a drug's response scores. This is done based on selecting a gene that has the most correlation across the known targeted genes by their drug.

Usage

```
PredTarget(Sim.GES.DRS,D.M)
```

Arguments

Sim.GES.DRS	similarity between Drug's response scores and Gene effect scores from Knock-out method such as CRISPR.
D.M	Drug Metadata

Value

a dataframe of drug(s) information with the most known predicted gene(s) with the max corelation value(s), P value(s), and FDR value(s).

Author(s)

sanjusinha7, Trinh Nguyen

Examples

```
library(BiocParallel)
data(OntargetM)
set.seed (12345)
All.Drugs <- OntargetM$DrugMetadata[,"broad_id_trimmed"]
S.Drugs <- sample(All.Drugs, 5)
KO.GES <- OntargetM$avana_CRISPR
sec.prism <- OntargetM$secondary_prism
```

```
sim.out <- bplapply(S.Drugs,function(x) computeCor(x,sec.prism,KO.GES))
names(sim.out) <- S.Drugs
Meta.data <- OntargetM$DrugMetadata
DrugTargetSim <- PredTarget(sim.out,Meta.data)
```

Index