Package 'DIGREsyn'

April 4, 2018

Type Package

Title Drug Induced Genomic Residual Effect (DIGRE) model for predicting drug pair synergistic score

Version 0.2.0

Date 2016-10-7

Author Jicheng Yang, Sangin Lee, Minzhe Zhang

Maintainer Minzhe Zhang <zenroute.mzhang@gmail.com>

Description DIGREsyn is developed to predict drug pair synergism/antagonism using Drug In-
duced Genomic Residual Effect (DIGRE) model. DIGRE utilize drug perturbated cell gene ex-
pression profile (GEP), drug dose response curve (DRC) and gene regulatory network (GRN) in-
formation (KEGG pathway database or self-provided) to do the prediction, and out-
put the rank of drug pairs from most synergistic to most antagnoistic.

Imports preprocessCore, KEGGgraph, org.Hs.eg.db, AnnotationDbi, BiocGenerics, ggplot2

License GPL-3

LazyData TRUE

RoxygenNote 6.0.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

CGP.mat *Cancer Growth Pathway (CGP)*

Description

A matrix contains the Cancer Growth Pathway (CGP) gene-gene interaction network information

Usage

```
data(CGP.mat)
```

Format

An object of class `matrix` with 932 rows and 932 columns.

Details

The CGP was built using pathways empirically selected from KEGG pathway database based on our knowledge firstly, and then refined by a small drug combination training dataset. In particular, we selected 12 pathways which were highly related to cell growth. Then we remove one of the 12 pathways each time and merged the remaining to build a CGP with which we applied our approach to the external training dataset. Based on the results, we screened out 8 pathways that contributed mostly to the performance and merged them to build final CGP. The 8 KEGG pathways are: aminoacyl-tRNA biosynthesis, MAPK signaling pathway, NF-kappa B signaling pathway, Cell Cycle, p53 signaling pathway, Apoptosis, TGF-beta signaling pathway, Cancer pathway.

---

constGeneNet *Construct Gene Network Matrix*

Description

This function is to convert a gene-gene interaction table to gene network matrix, which is an input of DIGREscore function.

Usage

```
constGeneNet(geneNet)
```

Arguments

geneNet   a data frame contains gene-gene interaction. See 'geneNetLymph' for example.

Details

Input gene-gene interaction table should have two columns with gene SYMBOL names. Each raw represents two connected genes. The interaction is regarded as undirected.

Value

a matrix of gene connectivity matrix

Author(s)

Jichen Yang, Sangin Lee, Minzhe Zhang(<zenroute.mzhang@gmail.com>)

Examples

```
geneNetLymph.mat <- constGeneNet(geneNet = geneNetLymph)
```

DIGREscore   *Drug Induced Genomic Residual Effect (DIGRE) model*

Description

This function estimate the compound pair synergism/antagonism. It takes drug treated gene expression data and drug dose response curve data as input, and caculate the synergistic score using Drug Induced Genomic Residual Effect (DIGRE) model.

Usage

```
DIGREscore(geneExpDiff, doseRes, pathway = "KEGG", geneNet, fold = 0.6)
```

Arguments

geneExpDiff	a matrix of drug treated gene expression data. Each column represent one drug, each row represent one gene. The value represent the fold change of the expression of a particular gene after drug treated compared to negative control.
doseRes	a matrix contains drug dose response data. Each column represent one drug, two rows of two different drug dose response curve. See 'doseRes.demo' for example.
pathway	pathway information used in DIGRE model. User would specify either "KEGG" to use the KEGG pathway information or "GeneNet" to use the gene network information.
geneNet	optional parameter. If pathway parameter is "GeneNet", then specify this parameter to use your own gene network data. If pathway parameter is "KEGG", then do not set this parameter.
fold	a value between 0 and 1. Gene expression fold change above this value would be considered as upregulated, below the opposite number would be considered as downregulated, otherwise would be considered as no effect. The default value is 0.6.

Details

This function takes drug treated gene expression data, dose-response curve data and pathway information as inputs, and calculate pair synergistic score of all the possible combination of the compound you provided, and their rank from the most synergistic to the most antagonist. Larger score indicates high possibility of the pair to have synergistic effect, and vice versa. (Notice that this algorithm focus more on predicting the relative rank of your compound pairs not the exact synergistic strength. If you want to do that, maybe you should involve positive control in your experiment. And also the score calculated by two pathway information is not comparable.)

Value

a list contains two matrices. One gives the drug pair synergistic score and their rank, the other contains the raw data to calculate the score.

Author(s)

Jichen Yang, Sangin Lee, Minzhe Zhang(<zenroute.mzhang@gmail.com>)

See Also

Bansal M, Yang J, Karan C, et al. A community computational challenge to predict the activity of pairs of compounds[J]. Nature biotechnology, 2014, 32(12): 1213-1222.

Yang J, Tang H, Li Y, et al. DIGRE: Drug Induced Genomic Residual Effect Model for Successful Prediction of Multidrug Effects[J]. CPT: pharmacometrics & systems pharmacology, 2015, 4(2): 91-97.

Examples

```
### profile gene expression data
geneExpDiff <- profileGeneExp(geneExp = geneExp.demo)
### DIGRE prediction
res.KEGG <- DIGREscore(geneExpDiff = geneExpDiff, doseRes = doseRes.demo, pathway = "KEGG", fold = 0.6) # KEGG p
res.geneNet <- DIGREscore(geneExpDiff = geneExpDiff, doseRes = doseRes.demo, pathway = "GeneNet", geneNet = gen
```

DIGREvis	*Visualization of compound pairs synergistic score*

Description

This function plot heatmap and barplot of compound pairs synergistic score.

Usage

```
DIGREvis(pred.pair, type)
```

Arguments

pred.pair	data frame of score rank table generated by DIGREscore function.
type	specify "heat" or "bar" to plot whether heatmap or barplot

Details

Please directly apply this funtion to the result got from DIGREscore funtion.

Value

ggplot object of the heatmap or barplot of compound pairs synergistic score.

Author(s)

Jichen Yang, Sangin Lee, Minzhe Zhang(<zenroute.mzhang@gmail.com>)

Examples

```
vis.heat <- DIGREvis(pred.pair = res.KEGG$scoreRank, type = "heat")
vis.bar <- DIGREvis(pred.pair = res.KEGG$scoreRank, type = "bar")
plot(vis.heat)
plot(vis.bar)
```

---

doseRes.demo *Dose response data*

Description

Examplary drug dose response data of 14 drugs in NCI-DREAM challenge.

Usage

```
data(doseRes.demo)
```

Format

An object of class data.frame with 2 rows and 14 columns.

Details

This demo data contains dose response data of 14 drugs from NCI-DREAM challenge. It contain the cell viability reduction when treated with drug in two different dose. One dose is IC20 of the drug, therefore the cell viability reduction is always 0.2 for all drugs. The other dose is double dose of IC20, this value is infered from the dose response curve of each drug.

See Also

http://www.nature.com/nbt/journal/v32/n12/full/nbt.3052.html

---

geneExp.demo *Drug treated gene expression data*

Description

Examplary gene expression data of DLBCL cell after treated with 14 different drugs.

Usage

```
data(geneExp.demo)
```

Format

An object of class data.frame with 48704 rows and 51 columns.

Details

This data is from NCI-DREAM challenge competition for predicting drug pairs synergy. OCI-LY3 human diffuse large B-cell lymphoma (DLBCL) cell line was treated by 14 different drugs in its dose of IC20. 24 hours after Perturbation, gene expression level was measured.

See Also

http://www.nature.com/nbt/journal/v32/n12/full/nbt.3052.html

---

geneNetLymph   *Lymphoma Specific Gene Network*

Description

Gene-gene interaction information refined from lymphoma patients gene expression data.

Usage

    data(geneNetLymph)

Format

An object of class matrix with 13248 rows and 2 columns.

Details

We use gene expression data from lymphoma patients (GSE10846) to construct Lymphoma-specific gene network. The statistical algorithm sparse partial correlation estimation (SPACE) is used to infer the network structure from the expression data.

See Also

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10846 https://cran.r-project.org/web/packages/space/index.html

---

KEGGnet.mat   *Universal KEGG pathway Gene Network*

Description

A matrix contains the global pathway (GP) gene-gene interaction network merged from KEGG pathway database.

Usage

    data(KEGGnet.mat)

Format

An object of class matrix with 2709 rows and 2709 columns.

Details

We selected KEGG pathways belonging to genetic information processing, environmental information processing, cellular processing, and cancer disease. From this set of selected pathways we removed any pathway with fewer than 10 edges. Finally we merged the remaining 32 KEGG pathways into a global pathway (GP) which included 11642 interactions among 2322 genes.

---

profileGeneExp   *Profiling Drug Treated Gene Expression Data*

Description

This function profile the drug treated gene expression data to prepare the input for DIGREscore function.

Usage

```
profileGeneExp(geneExp)
```

Arguments

geneExp   a data frame contains the drug treated gene expression data with each column representing one drug, and each row representing one gene. See 'geneExp.demo' for example.

Details

Gene expression is measured after cell is treated by a single compound (or negative control). Raw data (micro array or RNA-Seq) should already be log transformed to have proper scale. This function will average duplicated data with same drug name, and collapse multiple probes to gene level.

Value

a matrix of processed gene expression data.

Author(s)

Jichen Yang, Sangin Lee, Minzhe Zhang(<zenroute.mzhang@gmail.com>)

Examples

```
geneExpDiff <- profileGeneExp(geneExp = geneExp.demo)
```

Index