Package 'miRLAB'

October 18, 2022

Type Package

Title Dry lab for exploring miRNA-mRNA relationships

Version 1.26.0

Date 2021.10.07

Author Thuc Duy Le, Junpeng Zhang, Mo Chen, Vu Viet Hoang Pham

Maintainer Thuc Duy Le <Thuc.Le@unisa.edu.au>

Imports methods, stats, utils, RCurl, httr, stringr, Hmisc, energy, entropy, gplots, glmnet, impute, limma, pcalg,TCGAbiolinks,dplyr,SummarizedExperiment, ctc, InvariantCausalPrediction, Category, GOstats, org.Hs.eg.db

Description
Provide tools exploring miRNA-mRNA relationships, including popular miRNA target prediction methods, ensemble methods that integrate individual methods, functions to get data from online resources, functions to validate the results, and functions to conduct enrichment analyses.

License GPL (>=2)

Suggests knitr,BiocGenerics, AnnotationDbi,RUnit,rmarkdown

VignetteBuilder knitr

biocViews miRNA, GeneExpression, NetworkInference, Network

RoxygenNote 7.1.1

NeedsCompilation no

URL https://github.com/pvvhoang/miRLAB

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

git_branch RELEASE_3_15

git_last_commit 8276c9b

git_last_commit_date 2022-04-26

Date/Publication 2022-10-18

R topics documented:

`miRLAB-package` *A dry lab for exploring miRNA-mRNA relationships*

Description

Provide tools exploring miRNA-mRNA relationships, including popular miRNA target prediction methods using expression data, ensemble methods that integrate individual methods, functions to get data from online resources, functions to validate the results, and functions to conduct enrichment analyses.

Details

Package:	miRLAB
Type:	Package
Version:	0.99
Date:	2015-04-23
License:	GPL(>=2)

Author(s)

Thuc Duy Le, Junpeng Zhang

Maintainer: Thuc Duy Le <Thuc.Le@unisa.edu.au>

References

miRLAB: An R based dry lab for exploring miRNA-mRNA relationships

---

`Borda` *Ensemble method for miRNA target prediction using Borda count election*

Description

Use the Borda count election method to integrate the rankings from different miRNA target prediction methods

Usage

```
Borda(listCEmatrices)
```

Arguments

`listCEmatrices`	a list of matrices that include the correlation coefficients/causal effects/scores resulting from different target prediction methods

Value

a matrix of ranking scores (averaging all the rankings from different methods). Columns are miRNAs and rows are mRNAs

References

1. Le, T.D., Zhang, J., Liu, L., and Li, J. (2015) Ensemble Methods for miRNA Target Prediction from Expression Data, Plos ONE.

2. Marbach, D., Costello, J.C., Kuffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R. and DREAM5 Consortium (2012). Wisdom of crowds for robust gene network inference. Nat. Methods, 9, 796-804.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
ps=Pearson(dataset, cause=1:3, effect=4:18)
ida=IDA(dataset, cause=1:3, effect=4:18)
borda=Borda(list(ps, ida))
```

BordaTopk *Ensemble method for miRNA target prediction using Borda count election with topk targets*

Description

Use the Borda count election method to integrate the rankings from different miRNA target prediction methods, but only topk targets of each miRNA are included in the calculation. The targets outside the topk will be assigned a large and fixed rank, e.g. number of genes in the dataset.

Usage

```
BordaTopk(listCEmatrices, topk)
```

Arguments

`listCEmatrices`	a list of matrices that include the correlation/causal effects/scores resulting from a target prediction method
`topk`	number of targets of a miRNA to be included in the calculation (Borda count election)

Value

a matrix of ranking scores (averaging all the rankings from different methods). Columns are miRNAs and rows are mRNAs

References

Le, T.D., Zhang, J., Liu, L., and Li, J. (2015) Ensemble Methods for miRNA Target Prediction from Expression Data, Plos ONE.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
ps=Pearson(dataset, cause=1:3, effect=4:18)
ida=IDA(dataset, cause=1:3, effect=4:18)
borda=BordaTopk(list(ps, ida), topk=10)
```

---

bRank *Extract topk predicted targets of a miRNA Rank all the targets of a miRNA and extract the topk targets*

Description

Extract topk predicted targets of a miRNA Rank all the targets of a miRNA and extract the topk targets

Usage

```
bRank(CEmatrix, causeIndex, topk, downreg = TRUE)
```

Arguments

CEmatrix	the matrix of correlation/causal effect/score results with columns are miRNAs and rows are mRNAs
causeIndex	the column index of the miRNA that we would like to extract
topk	the number of targets being extracted
downreg	if TRUE the negative correlation/causal effect/score will be on the top of the ranking. This is to favour the negative regulations.

Value

a matrix with 3 columns, where the first column contains the miRNA, the second column contains the mRNAs and the last column contains the correlations/causal effects/scores

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
ps=Pearson(dataset, cause=1:3, effect=4:18)
miR200aTop10 = bRank(ps, 3, 10, TRUE)
```

convert *Convert miRNA symbols from a miRBase version to another*

Description

This function convert the miRNAs in the input file from the "source" miRBase version to the "Target" version. If users do not know the miRBase version of the input file, please set the source version to 0. The function will match the miRNAs in the input file to all miRBase versions to find the most likely miRBase version. Currently, we have versions 16-21.

Usage

```
convert(miRNAListFile, sourceV, targetV)
```

Arguments

miRNAListFile	the input file containing a list of miRNA symbols in csv format
sourceV	the miRBase version of the input miRNAs, e.g. 16. If users do not know the version, use 0.
targetV	the miRBase version that we want to convert into, e.g. 21.

Value

A csv file in the working directory containing the converted miRNA symbols.

Examples

```
miRs=system.file("extdata", "ToymiRs.csv", package="miRLAB")
convert(miRs, 17, 21)
```

Dcov *miRNA target prediction with the Distance correlation method*

Description

Calculate the Distance correlation of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Dcov(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Distance correlation values. Columns are miRNAs, rows are mRNAs.

References

Szekely, G., Rizzo, M. and Bakirov, N. (2007) Measuring and testing independence by correlation of distances. Ann. Stat., 35, 2769 - 94.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Dcov(dataset, 1:3, 4:18)
```

DiffExpAnalysis *Differentially expressed analysis*

Description

Find the top miRNAs and mRNAs that are differently expression between different conditions, e.g. cancer vs normal

Usage

```
DiffExpAnalysis(miR1, miR2, mR1, mR2, topkmiR, topkmR, p.miR, p.mR)
```

Arguments

miR1	the miRNA dataset for condition 1, e.g. cancer
miR2	the miRNA dataset for condition 1, e.g. normal
mR1	the mRNA dataset for condition 1, e.g. cancer
mR2	the mRNA dataset for condition 2, e.g. normal
topkmiR	the maximum number of miRNAs that we would like to extract, e.g. top 50 miRNAs.
topkmR	the maximum number of mRNAs that we would like to extract, e.g. top 2000 mRNAs.

p.miR	cutoff value for adjusted p-values when conducting differentially expressed analysis for miRNAs.
p.mR	cutoff value for adjusted p-values when conducting differentially expressed analysis for mRNAs.

Value

the dataset that includes differentially expressed miRNAs and mRNAs. columns are miRNAs and mRNAs and rows are samples

References

Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and computational biology solutions using R and Bioconductor (pp. 397-420). Springer New York.

Elastic *miRNA target prediction with the Elastic-net regression coefficient method*

Description

Calculate the Elastic-net regression coefficient of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Elastic(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Elastic-net regression coefficients. Columns are miRNAs, rows are mRNAs.

References

1. Le, T.D., Zhang, J., Liu, L., and Li, J. (2015) Ensemble Methods for miRNA Target Prediction from Expression Data, under review.

2. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. J. R. Stat. Soc. Series B Stat. Methodol., 67, 301-320.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Elastic(dataset, 1:3, 4:18)
```

`experiment` *Function for validate the results from all 12 methods.*

Description

Function for validate the results from all 12 methods.

Usage

```
experiment(allmethods, topk, Expgroundtruth, LFC, downreg)
```

Arguments

`allmethods`	A list of results (matrix with columns are miRNA and rows are mRNAs).
`topk`	Top k targets of each miRNA that will be extracted for validation
`Expgroundtruth`	The ground truth in .csv file for validation
`LFC`	log fold-change for validating the results using transfection experiments
`downreg`	If set to TRUE the negative effects will have higher ranks than the positives.

Value

The validation results for all 12 methods

`Extopk` *Extract top k miRNA-mRNA interactions*

Description

Rank the miRNA-mRNA interactions based on absolute values of the correlations/scores/causal effects, and return the topk interactions.

Usage

```
Extopk(cormat, topk)
```

Arguments

`cormat`	the correlation matrix that need to be extracted with columns are miRNAs and rows are mRNAs
`topk`	the number of interactions that need to be extracted.

Value

topk interactions

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
EMTresults=Pearson(dataset, 1:3, 4:18)
top10=Extopk(EMTresults, 10)
```

`filterAndCompare` *Filter and compare the validation results from 12 methods Keep the miRNAs that have at least noVal confirmed targets and compare the validation results from all methods.*

Description

Filter and compare the validation results from 12 methods Keep the miRNAs that have at least noVal confirmed targets and compare the validation results from all methods.

Usage

```
filterAndCompare(allresults, noVal)
```

Arguments

`allresults`	the results from all methods generated from experiment function. This is a list.
`noVal`	Number of confirmed targets in each method (threshold) to filter. Records (miRNA) with less than this will be removed

Value

the validation results of all methods

Examples

```
print("result=filterAndCompare(allresults, 2)")
```

---

getData *getData from GDC*

Description

getData from GDC

Usage

```
getData(cancerName)
```

Arguments

cancerName	The name of cancer in string format

Value

dataset in matrix format

GOBPenrichment *Functional enrichment analysis*

Description

GO BP enrichment analysis for a gene list

Usage

```
GOBPenrichment(Genes, Cutoff)
```

Arguments

Genes	a list of gene symbols
Cutoff	the significant level, e.g. 0.05

Value

a list of GO terms for the genes

References

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. and Sherlock, G. (2000) Gene Ontology: tool for the unification of biology. Nat. Genet., 25, 25-29.

Examples

```
print("result = GOBPenrichment(genelist, 0.05)")
```

Hoeffding *miRNA target prediction with the Hoeffding correlation coefficient method*

Description

Calculate the Hoeffding correlation coefficient of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Hoeffding(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Hoeffding correlation coefficients. Columns are miRNAs, rows are mRNAs.

References

Hoeffding, W. (1948) A non-parametric test of independence. Ann. Math. Stat., 19, 546 - 57.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Hoeffding(dataset, 1:3, 4:18)
```

`ICPPam50` *Identify miRNA targets by ICP and PAM50*

Description

This function identifies miRNA targets by ICP and PAM50.

Usage

```
ICPPam50(d, nmiR, nmR, fiftymRNAsData)
```

Arguments

`d`	A matrix of expression of miRNAs and mRNAs with columns being miRNA or mRNA names and rows being samples
`nmiR`	Number of miRNAs
`nmR`	Number of mRNAs
`fiftymRNAsData`	A matrix of expression of 50 mRNAs in PAM50 with columns being mRNA names and rows being samples

Value

The matrix of causal effects of miRNAs and mRNAs with columns being miRNAs and rows being mRNAs

References

1. Parker, J. S., et al. (2009). "Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes." Journal of Clinical Oncology 27(8): 1160-1167.

`IDA` *miRNA target prediction with the IDA method*

Description

Calculate the causal effect of each pair of miRNA-mRNA,and return a matrix of causal effects with columns are miRNAs and rows are mRNAs.

Usage

```
IDA(
  datacsv,
  cause,
  effect,
  pcmethod = "original",
  alpha = 0.05,
  targetbinding = NA
)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
pcmethod	choose different versons of the PC algorithm, including "original" (default) "stable", and "stable.fast"
alpha	significance level for the conditional independence test, e.g. 0.05.
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the causal effects. Columns are miRNAs, rows are mRNAs.

References

1. Le, T.D., Liu, L., Tsykin, A., Goodall, G.J., Liu, B., Sun, B.Y. and Li, J. (2013) Inferring microRNA-mRNA causal regulatory relationships from expression data. Bioinformatics, 29, 765-71.

2. Zhang, J., Le, T.D., Liu, L., Liu, B., He, J., Goodall, G.J. and Li, J. (2014) Identifying direct miRNA-mRNA causal regulatory relationships in heterogeneous data. J. Biomed. Inform., 52, 438-47.

3. Maathuis, H.M., Colombo, D., Kalisch, M. and Buhlmann, P. (2010) Predicting causal effects in large-scale systems from observational data. Nat. Methods, 7, 247-249.

4. Maathuis, H.M., Kalisch, M. and Buhlmann, P. (2009) Estimating high-dimensional intervention effects from observational data. Ann. Stat., 37, 3133-3164.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=IDA(dataset, 1:3, 4:18)
```

`identifymiRTargetsByEnsemble` *Identify the top miRNA targets by an ensemble method with ICP-PAM50, Pearson and Lasso*

Description

This function identifies the top miRNA targets by an ensemble method with ICP-PAM50, Pearson and Lasso.

Usage

```
identifymiRTargetsByEnsemble(d, nmiR, nmR, fiftymRNAsData, top = 1, topk = 500)
```

Arguments

`d`	A matrix of expression of miRNAs and mRNAs with columns being miRNA or mRNA names and rows being samples
`nmiR`	Number of miRNAs
`nmR`	Number of mRNAs
`fiftymRNAsData`	A matrix of expression of 50 mRNAs in PAM50 with columns being mRNA names and rows being samples
`top`	1 if getting the top of all miRNAs and 2 if getting the top of each miRNA
`topk`	Number of the top to get

Value

The top k miRNA targets

References

1. Parker, J. S., et al. (2009). "Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes." Journal of Clinical Oncology 27(8): 1160-1167.

`identifymiRTargetsByICPPam50` *Identify the top miRNA targets by ICP and PAM50*

Description

This function identifies the top miRNA targets by ICP and PAM50.

Usage

```
identifymiRTargetsByICPPam50(d, nmiR, nmR, fiftymRNAsData, top = 1, topk = 500)
```

Arguments

d	A matrix of expression of miRNAs and mRNAs with columns being miRNA or mRNA names and rows being samples
nmiR	Number of miRNAs
nmR	Number of mRNAs
fiftymRNAsData	A matrix of expression of 50 mRNAs in PAM50 with columns being mRNA names and rows being samples
top	1 if getting the top of all miRNAs and 2 if getting the top of each miRNA
topk	Number of the top to get

Value

The top k miRNA targets

References

1. Parker, J. S., et al. (2009). "Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes." Journal of Clinical Oncology 27(8): 1160-1167.

ImputeNormData *Filter, impute, and normalise data.*

Description

Remove the genes (rows) that have more than r% of missing data; use the impute package to fill in missing data, and finally normalise the data.

Usage

```
ImputeNormData(dataset, r)
```

Arguments

dataset	The input dataset in csv format. e.g. "EMT.csv"
r	The rate threshold to filter the records (genes). Genes with more than r% missing data will be removed.

Value

The processed dataset.

References

1. Hastie T, Tibshirani R, Narasimhan B and Chu G. impute: Imputation for microarray data. R package version 1.42.0.

2. Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and computational biology solutions using R and Bioconductor (pp. 397-420). Springer New York.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
impdata=ImputeNormData(dataset, 0.1)
```

`KEGGenrichment` *Functional enrichment analysis KEGG enrichment analysis for a gene list*

Description

Functional enrichment analysis KEGG enrichment analysis for a gene list

Usage

```
KEGGenrichment(Genes, Cutoff)
```

Arguments

`Genes`	a list of gene symbols
`Cutoff`	the significant level, e.g. 0.05

Value

a list of pathways for the genes

References

Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res., 28, 27-30.

Examples

```
print("result = KEGGenrichment(genelist, 0.05)")
```

`Kendall` *miRNA target prediction with the Kendall correlation coefficient method*

Description

Calculate the Kendall correlation coefficient of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Kendall(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Kendall correlation coefficients. Columns are miRNAs, rows are mRNAs.

References

Kendall, M. (1938) A new measure of rank correlation. Biometrika, 30, 81 - 9.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Kendall(dataset, 1:3, 4:18)
```

Lasso *miRNA target prediction with the Lasso method*

Description

Calculate the Lasso regression coefficient of each pair of miRNA-mRNA, and return a matrix of coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Lasso(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Lasso regression coefficients. Columns are miRNAs, rows are mRNAs.

References

1. Le, T.D., Zhang, J., Liu, L., and Li, J. (2015) Ensemble Methods for miRNA Target Prediction from Expression Data, submitted.

2. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Series B Stat. Methodol., 267-288.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Lasso(dataset, 1:3, 4:18)
```

`MI` *miRNA target prediction with mutual information method*

Description

Calculate the mutual information of each pair of miRNA-mRNA,and return a matrix of mutual information values with columns are miRNAs and rows are mRNAs.

Usage

```
MI(datacsv, cause, effect, targetbinding = NA)
```

Arguments

`datacsv`	the input dataset in csv format
`cause`	the column range that specifies the causes (miRNAs), e.g. 1:35
`effect`	the column range that specifies the effects (mRNAs), e.g. 36:2000
`targetbinding`	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the mutual information values. Columns are miRNAs, rows are mRNAs.

References

Moon, Y.I., Balaji, R., and Lall, U. (1995) Estimation of mutual information using kernel density estimators. Phys. Rev. E, 52, 2318 - 21.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=MI(dataset, 1:3, 4:18)
```

`Pearson` *miRNA target prediction with the Pearson correlation coefficient method*

Description

Calculate the Pearson correlation coefficient of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Pearson(datacsv, cause, effect, targetbinding = NA)
```

Arguments

`datacsv`	the input dataset in csv format
`cause`	the column range that specifies the causes (miRNAs), e.g. 1:35
`effect`	the column range that specifies the effects (mRNAs), e.g. 36:2000
`targetbinding`	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Pearson correlation coefficients. Columns are miRNAs, rows are mRNAs.

References

Pearson, K. (1920) Notes on the history of correlation. Biometrika, 13, 25 - 45.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Pearson(dataset, 1:3, 4:18)
```

RDC *miRNA target prediction with the Randomized Dependence Coefficient method*

Description

Calculate the Randomized Dependence coefficient of each pair of miRNA-mRNA,and return a matrix of coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
RDC(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the correlation coefficients. Columns are miRNAs, rows are mRNAs.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=RDC(dataset, 1:3, 4:18)
```

Read *Read dataset from csv file*

Description

Read dataset from csv file

Usage

```
Read(dataset)
```

Arguments

dataset	The input dataset in csv format

Value

dataset in matrix format

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
data=Read(dataset)
```

`ReadExtResult` *Read results from other methods*

Description

Read the results predicted by external methods (methods that are not in this package and may not be implemented in R). Consequently, we can compare the results predicted by the external methods and results predicted by the methods in the miRLAB package.

Usage

```
ReadExtResult(datacsv, cause, effect, ExtCEcsv)
```

Arguments

`datacsv`	the input dataset in csv format
`cause`	the column range that specifies the causes (miRNAs), e.g. 1:35
`effect`	the column range that specifies the effects (mRNAs), e.g. 36:2000
`ExtCEcsv`	score matrix predicted by an external matrix with columns are miRNAs and rows are mRNAs.

Value

a matrix of scores predicted by an external matrix and ready for further validation and comparison tasks.

Examples

```
print("GenemiR=ReadExtResult(dataset, cause=1:3, effect=4:18, 'genemirresults.csv')")
```

readHeader *Read the header of the dataset*

Description

Read the header of the dataset

Usage

```
readHeader(dataset)
```

Arguments

dataset	the character string of the names of the dataset in csv format, e.g. "ToyEMT.csv"

Value

the header of the dataset

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
header=readHeader(dataset)
```

Spearman *miRNA target prediction with the Spearman correlation coefficient method*

Description

Calculate the Spearman correlation coefficient of each pair of miRNA-mRNA,and return a matrix of correlation coefficients with columns are miRNAs and rows are mRNAs.

Usage

```
Spearman(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Spearman correlation coefficients. Columns are miRNAs, rows are mRNAs.

References

Spearman, C. (1904) General intelligence, objectively determined and measured. Am. J. Psychol., 15, 201 - 92.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Spearman(dataset, 1:3, 4:18)
```

Standardise

Stardarsise the dataset Stadardise the dataset to have mean=0 and std=1 in each column.

Description

Stardarsise the dataset Stadardise the dataset to have mean=0 and std=1 in each column.

Usage

```
Standardise(dataset)
```

Arguments

dataset	The input dataset in csv format. e.g. "ToyEMT.csv". The first column is the sample name.

Value

The standardised dataset.

Examples

```
## Not run:
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
stdata=Standardise(dataset)

## End(Not run)
```

`ValidateAll` *Validate the targets of all miRNA using both experimentally confirmed and transfection data*

Description

Given the predicted target of all miRNA, the function returns a list of targets of each miRNA that are confirmed based on the experimentally validated interactions or curated transfection data. Users need to download the file logFC.imputed.rda from nugget.unisa.edu.au/Thuc/miRLAB/ and place it in the working directory (this file is obtained from the TargetScoreData package)

Usage

```
ValidateAll(CEmatrix, topk, groundtruth, LFC, downreg = TRUE)
```

Arguments

`CEmatrix`	the matrix of correlation/causal effects/scores with columns are miRNAs and rows are mRNAs
`topk`	the number of targets of each miRNA that are being validated.
`groundtruth`	the csv file containing the ground truth.
`LFC`	the log fold change threshold for the transfection data. The targets that have the absolute value of log fold change greater than the LFC will be regarded as the confirmed targets.
`downreg`	if TRUE the negative correlation/causal effect/score values will be ranked on the top of the ranking. This is to favour the down regulations.

Value

a list of matrices that contains the confirmed interactions by both provided ground truth and built-in transfection data.

Examples

```
print("ps=Pearson(dataset, cause=1:3, effect=4:18)")
print("results=ValidateAll(ps, 10, groundtruth, LFC=0.5, downreg=TRUE)")
```

`Validation` *Validate the targets of a miRNA*

Description

Given the predicted target of a miRNA, the function returns a list of targets that are experimentally confirmed based on the provided ground truth. Users can provide their own ground truth or use the built-in ground truth which is the union of Tarbase, miRTarbase, miRecords, and miRWalk.

Usage

```
Validation(topkList, datacsv)
```

Arguments

`topkList`	a matrix with 3 columns. The first column is the miRNA name, the second contains the target mRNAs, and the third contains the correlation values/ causal effects/ scores
`datacsv`	the ground truth for the validation. The ground truth is a matrix with 2 columns, where the first column is the miRNA and the second is the mRNA.

Value

a matrix in the same format of the input matrix put only contains the confirmed interactions.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
ps=Pearson(dataset, cause=1:3, effect=4:18)
miR200aTop10=bRank(ps, 3, 10, TRUE)
groundtruth=system.file("extdata", "Toygroundtruth.csv", package="miRLAB")
miR200aTop10Confirmed = Validation(miR200aTop10, groundtruth)
```

`ValidationT` *Validate the targets of a miRNA using transfection data*

Description

Given the predicted target of a miRNA, the function returns a list of targets that are confirmed based on the curated transfection data. Users need to download the file logFC.imputed.rda from nugget.unisa.edu.au/Thuc/miRLAB/ and place it in the working directory (this file is obtained from the TargetScoreData package)

Usage

```
ValidationT(topkList, LFC)
```

Arguments

topkList	a matrix with 3 columns. The first column is the miRNA name, the second contains the target mRNAs, and the third contains the correlation values/ causal effects/ scores
LFC	the log fold change threshold. The targets that have the absolute value of log fold change greater than the LFC will be regarded as the confirmed targets.

Value

a matrix in the same format of the input matrix put only contains the confirmed interactions.

References

1. Le, T.D., Zhang, J., Liu, L., and Li, J. (2015) Ensemble Methods for miRNA Target Prediction from Expression Data, under review.

2. Li Y, Goldenberg A, Wong K and Zhang Z (2014). A probabilistic approach to explore human microRNA targetome using microRNA-overexpression data and sequence information. Bioinformatics, 30(5), pp. 621-628. http://dx.doi.org/10.1093/bioinformatics/btt599.

Examples

```
print("ps=Pearson(dataset, cause=1:35, effect=36:1189)")
print("miR200aTop100=bRank(ps, 11, 100, TRUE)")
print("miR200aTop100Confirmed = ValidationT(miR200aTop100, 1.0)")
```

Zscore *miRNA target prediction with the Z-score method*

Description

Calculate the Z-score value of each pair of miRNA-mRNA, and return a matrix of values with columns are miRNAs and rows are mRNAs.

Usage

```
Zscore(datacsv, cause, effect, targetbinding = NA)
```

Arguments

datacsv	the input dataset in csv format
cause	the column range that specifies the causes (miRNAs), e.g. 1:35
effect	the column range that specifies the effects (mRNAs), e.g. 36:2000
targetbinding	the putative target, e.g. "TargetScan.csv". If targetbinding is not specified, only expression data is used. If targetbinding is specified, the prediction results using expression data with be intersected with the interactions in the target binding file.

Value

A matrix that includes the Z-score values. Columns are miRNAs, rows are mRNAs.

References

Prill, R.J., Marbach, D., Saez-Rodriguez, J., Sorger, P.K., Alexopoulos, L.G., Xue, X., Clarke, N.D., Altan-Bonnet, G. and Stolovitzky, G. (2010) Towards a rigorous assessment of systems biology models: the DREAM3 challenges. PLoS One, 5, e9202.

Examples

```
dataset=system.file("extdata", "ToyEMT.csv", package="miRLAB")
results=Zscore(dataset, 1:3, 4:18)
```

Index