Package 'MiRSEA'

January 20, 2025

Type Package

Version 1.1.1

Date 2015-06-30

Title 'MicroRNA' Set Enrichment Analysis

Author Junwei Han, Siyao Liu

Maintainer Junwei Han <hanjunwei1981@163.com>

Description The tools for 'MicroRNA Set Enrichment Analysis' can identify risk pathways(or prior gene sets) regulated by microRNA set in the context of microRNA expression data. (1) This package constructs a correlation profile of microRNA and pathways by the hypergeometric statistic test. The gene sets of pathways derived from the three public databases (Kyoto Encyclopedia of Genes and Genomes ('KEGG'); 'Reactome'; 'Biocarta') and the target gene sets of microRNA are provided by four databases('TarBaseV6.0'; 'mir2Disease'; 'miRecords'; 'miRTarBase';). (2) This package can quantify the change of correlation between microRNA for each pathway(or prior gene set) based on a microRNA expression data with cases and controls. (3) This package uses the weighted Kolmogorov-Smirnov statistic to calculate an enrichment score (ES) of a microRNA set that coregulate to a pathway , which reflects the degree to which a given pathway is associated with the specific phenotype. (4) This package can provide the visualization of the results.

Collate Corrp2miRfile.R EnrichmentScore.R S2N.R MirSEA.R MsReport.R GetExampleData.R getEnvironmentData.R GetMiRTargetData.R GetPathwayData.R PlotCorrelation.R PlotHeatMap.R PlotRunEnrichment.R

LazyData Yes

Depends R (>= 2.15.1)

Suggests stats,graphics,grDevices

License GPL (>= 2)

biocViews Statistics, Pathways, microRNA, enrichment analysis

NeedsCompilation no

Repository CRAN

Date/Publication 2024-01-24 09:49:16 UTC

Contents

`MiRSEA-package` *MicroRNA Set Enrichment Analysis*

Description

This package can identify dysregulated pathways(or prior gene sets) regulated by microRNAs set in the context of miRNA expression data.

Introduction

The package can identify dysregulated pathways(or prior gene sets) regulated by microRNAs set in the context of miRNA expression data. (1) The MiRSEA package constructs a correlation profile of miRNAs and pathways by hypergeometric.The gene sets of pathways derived from the three public databases(KEGG;Reactome;Biocarta;).The target gene sets of miRNAs are provided by four databases(TarBaseV6.0; mir2Disease; miRecords; miRTarBase;). (2) The MiRSEA package can quantify the change of correlation between miRNAs for each pathway (or prior gene set) based on miRNA expression data with cases and controls. (3) The MiRSEA package uses the weighted Kolmogorov-Smirnov statistic to calculate an enrichment score(ES)of a miRNA set that co-regulate to a pathway , which reflects the degree to which a given pathway is associated with the specific phenotype. (4) The MiRSEA package can provide the visualization of the results.

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

Corrp2miRfile *create a pathway-microRNA(miRNA) profile and a p value matrix*

Description

The function Corrp2miRfile create a p value matrix and a pathway-miRNA correlation profile

Usage

```
Corrp2miRfile(pathway="kegg", species = "hsa")
```

Arguments

pathway	choose database of pathway,"kegg","biocarta" or"reactome"
species	Species of miRNAs(default: hsa)

Details

When users input interesting species and pathway, the function can calculate the p value between pathway and miRNA using hypergeometric.The p value can quantify the strength of the pathway regulated by each miRNA.The smaller p value is represent the bigger strength of regulate.Then p2m can get miRNA set(pmSET)for each pathway,which is a co-regulated miRNA set of this pathway(w>0).

Value

p	A p value weighted matrix (rows are pathway ,cols are miRNAs)
p2miR	pathway-miRNA correlation(pmSET) profile

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

References

Rivals I, Personnaz L, Taing L, & Potier MC (2007) Enrichment or depletion of a GO category within a class of genes: which test? (Translated from eng) Bioinformatics 23(4):401-407 (in eng).

See Also

MirSEA,MsReport

Examples

```
## Not run:
p2m<-Corrp2miRfile(pathway="kegg", species = "example")

p2m$p[1,1:10]
p2m$p2miR[1,1:5]

## End(Not run)
```

EnrichmentScore *Computes the enrichment score*

Description

Computes the enrichment score of a microRNA(miRNA) set in a ordered miRNA list.

Usage

```
EnrichmentScore(miR.list, miR.set, weighted.score.type = 1, correl.vector = NULL)
```

Arguments

miR.list	The ordered miRNA list ,integers indicating the original position in the input dataset.
miR.set	A miRNA set ,integers indicating the location of those miRNAs in the input dataset.
weighted.score.type	Type of score,weight=0,ES reduces to the standard Kolmogorov-Smirnov statistic,when weight=1, we are weighting the miRNAs by their tw-score normalized by the sum of the tw-scores over all of the miRNAs in the miRNA set.
correl.vector	A vector with the correlations(tw-scores) corresponding to the miRNAs in the miRNA list

Details

The function can computes the enrichment score of a miRNA set in a miRNA list. The weighted score type is the exponent of the correlation(e.g.tw-score) (1) Rank order the miRNAs in a miRNA set to form a list according to the correlation(e.g.tw-score) of their expression profiles and regulated pathway (2) Evaluate the fraction of miRNAs in the miRNA set(hits) weighted by their correlation and the fraction of miRNAs not in the miRNA set(misses)present up to a given position i in the miRNA list.The ES is the maximum deviation from zero of 'P(hit)-P(miss)'. For a randomly distributed miRNA set,The enrichment score will be relatively small, but if it is concentrated at the top or bottom of the list,or otherwise nonrandomly distributed, then the Enrichment score will be correspondingly high.

Value

ES	Enrichment score.
arg.ES	Location in the miRNA list where the peak running enrichment occurs.
RES	Numerical vector containing the running enrichment score for all locations in the miRNA list.
tag.indicator	Binary vector indicating the location of the miRNA sets in the miRNA list.

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

EnrichmentScore2

Examples

```
#Computes the enrichment score of a miRNA set in a ordered miRNA list.
E1<-EnrichmentScore(miR.list=sample(1:1000),miR.set=c(39,281,37,381,39,11,3,34),
correl.vector=rep(0.3,1000))
#show results
#EnrichmentScore of this set
E1$ES
#peak running enrichment
E1$arg.ES
#running enrichment score of top ten miRNAs
E1$RES[1:10]
#Binary vector indicating the location of top ten miRNA in the miRNA list
E1$tag.indicator[1:10]
```

EnrichmentScore2 *Computes the enrichment score faster*

Description

Computes the enrichment score of a microRNA(miRNA) set in miRNA list.

Usage

```
EnrichmentScore2(miR.list, miR.set, weighted.score.type = 1, correl.vector = NULL)
```

Arguments

miR.list	The ordered miRNA list,integers indicating the original position in the input dataset
miR.set	A miRNA set,integers indicating the location of those miRNAs in the input dataset
weighted.score.type	Type of score, weight=0 (unweighted = Kolmogorov-Smirnov), 1 (weighted), and 2 (over-weighted)
correl.vector	A vector with the correlations (e.g.tw-scores) corresponding to the miRNAs in the miRNA list

Details

Computes the weighted enrichment score of a miRNA set in miRNA list. It is the same calculation as in EnrichmentScore but faster without producing the RES, arg.RES and tag.indicator outputs. This call is intended to be used to asses the enrichment of random permutations rather than the observed one.The weighted score type is the exponent of the correlation.

Value

ES	Enrichment score (real number between -1 and +1)

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

EnrichmentScore

Examples

```
#Computes the enrichment score of a miRNA set in miRNA list
R2<-EnrichmentScore2(miR.list=sample(1:100),miR.set=c(39,28,37,38,11,3,34),
correl.vector=rep(0.04,100))
#show the result
R2$ES
```

---

envData *The variables in the environment include predefine pathway, target information of miRNAs,an expression profile and a example result of miRNA list*

Description

The pathway information is download on the GSEA website,concluding three pathway database (KEGG,Biocarta,Reactome).We arranged the data for miRNAs and their target genes,which is according to four database including miRTarBase,TarBaseV6.0,miRecords and mir2Disease. example.GCT is an interesting miRNA expression data and example.cls is the vector of binary labels(class.labels) p is a p value weighted matrix (rows are pathway ,cols are miRNAs).p2miR is a correlation profile between kegg pathways and each human miRNA. miRList is a list of drawing parameters of KEGG ERBB signaling Pathway.

Format

An environment variable

Details

The environment variable includes the variable `pathway`, `mfile`,`example.cls`,`example.gct`,`p`,`p2miR`,`miRList`

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

`GetExampleData` *Get the example data*

Description

Get the example data.

Usage

```
GetExampleData(exampleData)
```

Arguments

`exampleData`	A character string, must be one of "dataset", "class.labels" , "miRList","p_value"and "p2miR".

Details

The function GetExampleData(exampleData="dataset") obtains miRNA expression dataset from the environment variable `envData`.

The function GetExampleData(exampleData="class.labels") obtains class labels from the environment variable `envData`.

The function GetExampleData(exampleData="miRList") obtains the drawing parameters of a miRNA List from the environment variable `envData`.

The function GetExampleData(exampleData="p_value") obtains the weighting matrix from the environment variable envData.

The function GetExampleData(exampleData="p2miR") obtains the correlation profile between kegg pathways and each human miRNA from the environment variable envData.

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

Examples

```
## Not run:

#obtain the gene expression dataset.
dataset<-GetExampleData(exampleData="dataset")


#obtain the class labels.
class.labels<-GetExampleData(exampleData="class.labels")

#obtain the drawing parameters of a miRNA List
miRList<-GetExampleData(exampleData="miRList")


#obtain the weighting matrix
p_value<-GetExampleData(exampleData="p_value")

#obtain the correlation profile
p2miR<-GetExampleData(exampleData="p2miR")



## End(Not run)
```

GetMiRTargetData *Get the data of microRNA(miRNA) and target genes*

Description

Get the data of miRNA and target genes

Usage

```
GetMiRTargetData()
```

Details

The data for target genes of miRNAs are obtained from the environment variable envData,which is obtained from four database(TarBaseV6.0,mir2Disease,miRecords,miRTarBase).

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

Examples

```
#Get the data for target genes of miRNAs
MiRTarget<-GetMiRTargetData()
```

`GetPathwayData` *Get the gene sets of pathways*

Description

Get the gene sets of pathways for the three pathway databases (KEGG; Biocarta; Reactome)

Usage

```
GetPathwayData(pathway)
```

Arguments

`pathway`	choose database of pathway,"kegg","biocarta" or"reactome"

Details

The gene sets of pathways for the three pathway database (KEGG; Biocarta; Reactome)are obtained from the environment variable `envData`.

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

Examples

```
## Not run:

#obtain the gene sets of  kegg pathways.
pathway<-GetPathwayData("kegg")



## End(Not run)
```

HeatMapPlot *Plot a heatmap*

Description

Plot a heatmap of a microRNA(miRNA) expression

Usage

```
HeatMapPlot(V, row.names = FALSE, col.labels, col.classes, col.names = FALSE,
main = " ", xlab = " ", ylab = " ")
```

Arguments

V	A miRNA expression matrix
row.names	A name list of row vector,default=FALSE
col.labels	Phenotype of class disticntion of interest. A vector of binary labels having first the 1's and then the 0's
col.classes	phenotype name
col.names	A name list of samples,default=FALSE
main	a main title for the heatmap
xlab	a label for the x axis, defaults to a description of x
ylab	a label for the y axis, defaults to a description of y

Details

Plots a heatmap of a miRNA expression matrix including phenotype vector and miRNA, sample and phenotype labels

Note

return a heatmap

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

References

Andy Liaw, original,R. Gentleman, M. Maechler, W. Huber.

See Also

PlotHeatMap

Examples

```
#example of expression profile
V<-matrix(runif(200),10,20)
#example of class.labels ("0"or "1")
 a1<-rep(0,20)
a1[sample(1:20,5)]=1
#plot heat map
 HeatMapPlot(V =V, row.names = FALSE, col.labels = a1, col.classes =c("a","b"), col.names =FALSE,
 main =" Heat Map for MiRs in MiR Set", xlab=" ", ylab=" ")
```

`MirSEA` *Identify dysregulated pathways based on microRNA (miRNA) set enrichment analysis*

Description

This function propose a novel method of miRNA set enrichment analysis(MiRSEA)to identify the dysregulated pathways by calculating the enrichment score of miRNA set which co-regulate a biological pathway(or prior gene set)

Usage

```
MirSEA(input.ds, input.cls, p_value,p2miR,
reshuffling.type = "miR.labels", nperm = 1000,
weighted.score.type = 1, ms.size.threshold.min = 10,
ms.size.threshold.max = 500)
```

Arguments

`input.ds`	Input miRNA expression Affymetrix dataset file in GCT format
`input.cls`	Input class vector (phenotype) file in CLS format
`p_value`	A weighting matrix of p value of the hypergeometric. (rows are pathway ,cols are microRNAs(miRNAs))
`p2miR`	pathway-miRNA correlation(pmSET) profile
`reshuffling.type`	Type of permutation reshuffling: "sample.labels" or "miR.labels" (default: "miR.labels")
`nperm`	Number of random permutations (default: 1000)
`weighted.score.type`	Enrichment correlation based weighting:When weight= 0, ES reduces to the standard Kolmogorov-Smirnov statistic,when weight=1, we are weighting the miRNAs by their dw-score normalized by the sum of the dw-scores over all of the miRNAs in the miRNA set,when weight=2,it represent over weight (default: 1)
`ms.size.threshold.min`	Minimum size (in miRNAs) for database miRNA sets to be considered (default: 10)

`ms.size.threshold.max`
: Maximum size (in miRNAs) for database miRNA sets to be considered (default: 500)

Details

MiRSEA integrates pathway (e.g.the strength of the pathway regulated by miRNAs.) and differential expression among miRNAs in identifying dysregulated pathways.MiRSEA can order pathway by the enrichment score of miRNA set, which is co-regulated by a miRNA set.

Value

`report.phen1`
: It is the summary of the result of the up regulated pathway

`report.phen2`
: It is the summary of the result of the down regulated pathway.Each rows of the dataframe represents a pathway. Its columns include "Pathway Name", "SIZE", "Pathway Source", "Pathway Enrichment Score", "NOM p-val", "FDR q-val", "Tag percentage"(Percent of miRNA set before running enrichment peak),"MiR percentage"(Percent of miRNA list before running enrichment peak),"Signal strength" (enrichment signal strength).

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

References

Subramanian A, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 102(43):15545-15550.

Lu M, Shi B, Wang J, Cao Q, & Cui Q (2010) TAM: a method for enrichment and depletion analysis of a microRNA category in a list of microRNAs. BMC bioinformatics 11:419.

See Also

`EnrichmentScore`, `EnrichmentScore2`,`S2N`,`Corrp2miRfile`

Examples

```
## Not run:
#get example of expression data
#input.ds <- readLines("F:/lsy/xin data/GSE36915.gct")
#input.cls <- readLines("F:/lsy/xin data/GSE36915.cls")
input.ds <- GetExampleData("dataset")
input.cls <- GetExampleData("class.labels")

#get example of p value matrix
p_value <- GetExampleData("p_value")
#get example of correlation profile
p2miR <- GetExampleData("p2miR")
```

```
#identify dysregulated pathways by using the function MirSEA
MirSEAresult <- MirSEA(input.ds,input.cls,p_value,p2miR,
reshuffling.type = "miR.labels", nperm = 1000,
weighted.score.type = 1, ms.size.threshold.min = 10,
ms.size.threshold.max = 500)
#print the summary results of pathways to screen

summaryResult1 <- MirSEAresult$report.phen1
summaryResult1[1:5,]
summaryResult2 <- MirSEAresult$report.phen2
summaryResult2[1:5,]

#write the summary results of pathways to tab delimited file.
write.table(summaryResult1,file="summaryResult1.txt",sep="\t",row.names=FALSE)

write.table(summaryResult2,file="summaryResult2.txt",sep="\t",row.names=FALSE)

## End(Not run)
```

`MsReport` *Produce report for a microRNA (abbreviated miRNA) set*

Description

The miR.report includes miRNA names, locstion, S2N, RES and whether is core-enrichment miRNA,

Usage

```
MsReport(MsNAME = "", input.ds, input.cls, p_value,
p2miR,weighted.score.type = 1)
```

Arguments

`MsNAME`	An interesting pathway name
`input.ds`	Input miRNA expression Affymetrix dataset file in RES or GCT format
`input.cls`	Input class vector (phenotype) file in CLS format
`p_value`	A weighting matrix of p value of the hypergeometric. (rows are pathway ,cols are microRNAs(miRNAs))
`p2miR`	pathway-miRNA correlation(pmSET) profile
`weighted.score.type`	Enrichment correlation-based weighting: 0=no weight (KS), 1=standard weigth, 2 = over-weigth (default: 1)

Details

When users input a interesting pathway,the function MsReport can create a report for a miRNA set that coordinated regulate this pathway. MiR : the name of miRNAs. For example the probe accession number, miRNA symbol or the miRNA identifier in the dataset. LIST LOC : location of the miRNA in the sorted miRNA list. S2N : correlation(tw-score) of the miRNA in the miRNA list. RES : value of the running enrichment score at the miRNA location. CORE_ENRICHMENT:whether this miRNA is the "core enrichment" section of the list, Yes or No variable specifying in the miRNA location is before (positive ES) or after (negative ES) the running enrichment peak.

Value

A list. It includes two elements: Msreport and miRList.

Msreport is matrix of input pathway which present the detail results . Its columns include "miRNA name", "location of the miRNA in the sorted miRNA list", "tw-scoe of miRNA", "Running enrichment score", "Property of contribution".

miRList is a list of drawing parameters for function PlotHeatMap,PlotCorrelation and PlotRunEnrichment.

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

MirSEA,S2N,EnrichmentScore,PlotHeatMap,PlotCorrelation,PlotRunEnrichment

Examples

```
## Not run:
#get example data
#input.ds <- readLines("F:/lsy/xin data/GSE36915.gct")
#input.cls <- readLines("F:/lsy/xin data/GSE36915.cls")
input.ds <- GetExampleData("dataset")
input.cls <- GetExampleData("class.labels")

#get example of p value matrix
p_value <- GetExampleData("p_value")
#get example of correlation profile
p2miR <- GetExampleData("p2miR")

#get a miRNA.SET report for KEGG ERBB PATHWAY
Results<-MsReport(MsNAME = "KEGG_ERBB_SIGNALING_PATHWAY", input.ds, input.cls,p_value,p2miR)
# show the report of top five miRNA in the pathway
Results[[1]][1:5,]

## End(Not run)
```

PlotCorrelation *Plot global microRNA(miRNA) correlation profile*

Description

plot global miRNA correlation profile for differential weighted scores(dw-score) of miRNAs

Usage

```
PlotCorrelation(miRlist)
```

Arguments

miRlist	A list of miRNA LIST result obtained from the MsReport function

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

MsReport

Examples

```
## Not run:
#get a list of miRNA list result
miRlist<-GetExampleData("miRList")
#Plot global miRNA correlation profile
PlotCorrelation(miRlist)


## End(Not run)
```

PlotHeatMap *Plot a heat map*

Description

Plot a heat map for a microRNA(miRNA) set which co-regulate pathway

Usage

```
PlotHeatMap(miRlist,input.ds,input.cls)
```

Arguments

miRlist	A list of miRNA LIST result obtained from the MsReport function
input.ds	Input miRNA expression Affymetrix dataset file in GCT format
input.cls	Input class vector (phenotype) file in CLS format

Details

Plots a heatmap of a miRNA set in the expression matrix including phenotype vector and miRNA, sample and phenotype labels

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

PlotHeatMap, MsReport

Examples

```
## Not run:
#get example data
#input.ds <- readLines("F:/lsy/xin data/GSE36915.gct")
#input.cls <- readLines("F:/lsy/xin data/GSE36915.cls")
input.ds <- GetExampleData("dataset")
input.cls <- GetExampleData("class.labels")

#get a list of miRNA list result
miRlist<-GetExampleData("miRList")
#Plot a heat map
PlotHeatMap(miRlist,input.ds,input.cls)

## End(Not run)
```

PlotRunEnrichment *Plot running microRNAs(miRNAs) enrichment score*

Description

Plot running miRNAs enrichment score for the input pathway

Usage

```
PlotRunEnrichment(miRlist)
```

Arguments

miRlist	A list of miRNA LIST result obtained from the MsReport function

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

See Also

MsReport

Examples

```
## Not run:
#get a list of miRNA list result
miRlist<-GetExampleData("miRList")
#Plot running miRNA enrichment score
PlotRunEnrichment(miRlist)

## End(Not run)
```

S2N *calculate signal to noise ratio for microRNAs(miRNAs)*

Description

This function calculate the signal to noise ratio for miRNAs for the actual phenotype and also random permutations

Usage

```
S2N(A, class.labels, miR.labels, nperm )
```

Arguments

A	Matrix of miRNAs expression values (rows are miRNAs, columns are samples)
class.labels	Phenotype of class disticntion of interest. A vector of binary labels having first the 1's and then the 0's
miR.labels	miRNA labels,Vector of probe ids or accession numbers for the rows of the expression matrix
nperm	Number of random permutations to perform

Details

The function uses matrix operations to implement the signal to noise calculation in stages and achieves fast execution speed.

Value

s2n.matrix	Matrix with random permuted or bootstraps signal to noise ratios (rows are miRNAs, columns are permutations or bootstrap subsamplings
obs.s2n.matrix	Matrix with observed signal to noise ratios (rows are miRNAs, columns are boostraps subsamplings. If fraction is set to 1.0 then all the columns have the same values

Author(s)

Junwei Han<hanjunwei1981@163.com>,Siyao Liu <liusiyao29@163.com>

References

Subramanian A, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 102(43):15545-15550.

See Also

MirSEA

Examples

```
##Matrix of miRNAs expression values
A<-matrix(runif(200),10,20)
##class.labels("0" or "1")
a1<-rep(0,20)
a1[sample(1:20,5)]=1
a1<-sort(a1,decreasing=FALSE)
#calculate signal to noise ratio for example data
M1<-S2N(A, class.labels=a1, miR.labels=seq(1,10), nperm=100)
#show actual results for top five in the matrix
M1$obs.s2n.matrix[1:5,1]
#show permutation results
M1$s2n.matrix[1:5,1:5]
```

Index