Package 'BUS'

September 17, 2025

Type Package

Title Gene network reconstruction

Depends R (>= 2.3.0), minet

Imports stats, infotheo

Version 1.64.0

Date 2010-04-19

Author Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

Maintainer Yuanhua Liu <liuyuanhua@picb.ac.cn>

Description This package can be used to compute associations among genes (gene-networks) or between genes and some external traits (i.e. clinical).

License GPL-3

biocViews Preprocessing

LazyLoad yes

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

git_branch RELEASE_3_21

git_last_commit bdf79c3

git_last_commit_date 2025-04-15

Repository Bioconductor 3.21

Date/Publication 2025-09-17

Contents

BUS-package 2
BUS 2
copasi 4
gene.pvalue 4
gene.similarity 6
gene.trait.pvalue 7
gene.trait.similarity 8
pred.network 9
tumors.miRNA 10
tumors.mRNA 10

Index 12

BUS-package *For network reconstruction.*

Description

This package can be used to compute associations among genes (gene-networks) or between genes and some external traits (i.e. clinical). [Function: BUS]

Both associations can be computed via correlation or mutual information (MI). [Functions: gene.similarity (gene-gene associations) and gene.trait.similarity (gene-trait associations)]

Statistical significance of the association is computed for single and multiple hypotheses testing, using random permutations method [Functions: gene.pvalue, gene.trait.pvalue]

The package can handle data with missing values using bootstrapping methods to fill NAs. [Arguments: na.replica]

Details

Package:	BUS
Type:	Package
Version:	1.0.2
Date:	2009-10-31
License:	GPL-3

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

Maintainer: Yuanhua Liu<liuyuanhua@picb.ac.cn.com>

BUS *A wrapper function for matrices of p-value and predicted network*

Description

A wrapper function to calculate the computation of two types of similarities (correlation and mutual information) with two different goals: (i) identification of the statistically significant similarities among the activity of molecules sampled across different experiments (option Unsupervised, U), (ii) identification of the statistically significant similarities between such molecules and other types of information (clinical etc., option supervised, S) .

Usage

```
BUS(EXP, trait = NULL, measure, method.permut = 2, n.replica = 400, net.trim = NULL, thresh = NULL, nflag
```

Arguments

EXP	Gene expression data in form of a matrix. Row stands for genes and column for experiments.
trait	Trait data in form of a matrix. The row stands for traits and column for experiments.
measure	Metric used to calculate similarity: "corr" for correlation, "MI" for mutual information.
method.permut	A flag to indicate which method is used to correct permutation p-values, default as 2. See gene.pvalue for details.
n.replica	Number of permutations used for the correction of multiple hypothesis testing; default value is 400.
net.trim	Method used to trim the network: "mrnet", "clr", "aracne" and "none" . "mrnet" infers a network using the maximum relevance/minimum redundancy feature selection method; "clr" use the CLR algorithm; "aracne" applies the data processing inequality to all triplets of nodes in order to remove the least significant edge in each triplet. These options come from the package *minet*, and they are used only for mutual information. "none" indicates no trim operation. It should be chosen when correlation is considered.
thresh	Threshold for significance of the corrected p-value. It is used, in the Unsupervised case, to trim the adjacency matrix (contains the results of the gene-gene association based on the chosen metric) and obtain a predicted gene interaction network. In the Supervised case, since no network is predicted, it is set as NULL.
nflag	A flag to indicate a gene-gene interaction case (Unsupervised) or a gene-trait interaction case (Supervised); 1 for Unsupervised and 2 for Supervised.

Value

similarity	A matrix of similarity, which could be correlation or mutual information
single.perm.p.value	A matrix of single p-values
multi.perm.p.value	A matrix of corrected p-values
net.pred.permut	Predicted network obtained trimming non-significant values

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

See Also

gene.pvalue,gene.trait.pvalue,pred.network

Examples

```
data(copasi)
mat<-as.matrix(copasi[1:10,])
rownames(mat)<-paste("G",1:nrow(mat), sep="")
BUS(EXP=mat,measure="corr",net.trim="none",thresh=0.05,nflag=1)
```

`copasi` *copasi data*

Description

This dataset is taken from Copasi2 (Complex Pathway Simulator), a software for simulation and analysis of biochemical networks. The system generates random artificial gene networks according to well-defined topological and kinetic properties. These are used to run in silico experiments simulating real laboratory micro-array experiments. Noise with controlled properties is added to the simulation results several times emulating measurement replicates, before expression ratios are calculated. This series consists of 150 artificial gene networks. Each network consists of 100 genes with a total of 200 gene interactions (on average each gene has 2 modulators).

Format

A data frame is size of 100x100, the 100 rows represent 100 genes and 100 columns for 100 experiments.

References

See http://www.comp-sys-bio.org/AGN/data.html for detailed information.

`gene.pvalue` *Calculates p-value for gene-gene interaction*

Description

To calculate p-value for the null hypothesis that there is no gene-gene interaction. For gene expression data with M genes, a p-value matrix under MxM single null hypotheses (each two genes have no interaction) is computed; besides, matrices with correct p-values are output: corrected permutation method using a distribution of MxMxP (P number of permutations) null hypotheses tests (multi.perm.p.value). p-values are calculated based on the adjacency matrix for gene-gene interaction computed by function *gene.similarity*.

Usage

```
gene.pvalue(EXP, measure, net.trim, n.replica = 400)
```

Arguments

EXP	Gene expression data in form of a matrix. Row stands for genes and column for experiments.
measure	Metric used to calculate similarity between genes: "corr" for correlation, "MI" for mutual information.
net.trim	Method used to trim the network: "mrnet", "clr", "aracne" and "none" . "mrnet" infers a network using the maximum relevance/minimum redundancy feature selection method; "clr" use the CLR algorithm; "aracne" applies the data processing inequality to all triplets of nodes in order to remove the least significant edge in each triplet. These options come from the package *minet*, and they are used only for mutual information. "none" indicates no trim operation. It should be chosen when correlation is considered.
n.replica	Number of permutations used for the correction of multiple hypothesis testing; default value is 400.

Details

Normally, in a permutation method, we use the empirical distribution of some statistics to estimate the p-value. To get a simple p-value for no interaction between gene i and j, empirical distribution of a vector with length of P (number of replicates) is used; to correct for multiple hypothesis with permutations, an empirical distribution of a vector with length of PxM (M being the number of hypotheses tested) is used.

Value

`single.perm.p.value`
: A matrix of single p-values obtained with permutation method + beta distribution for extreme values (for MI) or obtained with the exact distribution computed directly by *cor.test* (for correlation)

`multi.perm.p.value`
: A matrix of corrected p-values obtained with permutation method

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

See Also

gene.similarity

Examples

```
data(copasi)
mat=as.matrix(copasi)[1:10,]
rownames(mat)<-paste("G",1:nrow(mat), sep="")
gene.pvalue(mat,measure="MI",net.trim="mrnet")
```

`gene.similarity` *Calculate adjacency matrix for gene-gene interaction*

Description

To calculate an adjacency matrix for gene-gene interaction (using correlation/mutual information metric). For gene expression data with M genes and N experiments, the adjacency matrix is in size of MxM. It is optional to get a trimmed adjacency matrix according to the argument *net.trim*, i.e. *mrnet*, *clr* and*aracne* (from the package *minet*).

Usage

```
gene.similarity(EXP, measure, net.trim, na.replica = 50)
```

Arguments

`EXP`	Gene expression data in form of a matrix. Row stands for genes and column for experiments.
`measure`	Metric used to calculate similarity between genes: "corr" for correlation, "MI" for mutual information.
`net.trim`	Method used to trim the adjacency matrix: "mrnet", "clr", "aracne" and "none". "mrnet" infers a network using the maximum relevance/minimum redundancy feature selection method; "clr" use the CLR algorithm; "aracne" applies the data processing inequality to all triplets of nodes in order to remove the least significant edge in each triplet. These options come from the package *minet*, and they are used only for mutual information. "none" indicates no trim operation. It should be chosen when correlation is considered.
`na.replica`	Times of replication for filling NANs in the impute method; default value is 50. The (smooth) bootstrapping approach is used to give an estimation to missing value in the data.

Value

An adjacency matrix in size of MxM with rows and columns both standing for genes. Element in row i and column j indicates the similarity between gene i and gene j.

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

Examples

```
data(copasi)
mat=as.matrix(copasi)[1:10,]
rownames(mat)<-paste("G",1:nrow(mat), sep="")
res<-gene.similarity(mat,measure="corr",net.trim="none")
```

gene.trait.pvalue *Calculate p-value for gene-trait interaction*

Description

To calculate p-value for null hypothesis that there is no interaction between gene and trait. There are MxT interactions between M genes and T traits. Results are given with 3 possibilities 1 for single p-value, and 3 for different types of correction. p-values are calculated based on the adjacency matrix for gene-gene interaction computed by function *gene.trait.similarity*.

Usage

```
gene.trait.pvalue(EXP, trait, measure, method.permut = 2, n.replica = 400)
```

Arguments

EXP	Gene expression data in form of a matrix. Row stands for genes and column for experiments.
trait	Trait data in form of matrix. Row stands for traits and column for experiments.
measure	Metric used to calculate similarity: "corr" for correlation, "MI" for mutual information.
method.permut	A flag to indicate correction style when multiple hypotheses testing is considered. 1 for multiple traits correction, 2 for multiple genes and 3 for both genes and traits correction. The default value is 2.
n.replica	Number of permutations for the correction of multiple hypothesis testing; default value is 400.

Details

According to a permutation method, we use the empirical distribution of some statistics to estimate the p-value. For single p-value the empirical distribution is a vector of P (number of random replicates for each test) test values. It is then possible to correct p-value in different ways: method.permut = 1, it is the empirical distribution of a vector with length of TxP, corrects for the multiple traits tested; method.permut = 2, it is the empirical distribution of a vector with length of MxP, corrects for the multiple genes tested; method.permut = 3, it is empirical distribution of a vector with length of MxTxP, corrects for the multiple traits and genes tested.

Value

single.perm.p.value
: A matrix of single p-values obtained with permutation method + beta distribution for extreme values (for MI) or obtained with the exact distribution computed directly by *cor.test* (for correlation)

multi.perm.p.value
: A matrix of corrected p-values obtained with permutation method

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

See Also

gene.trait.similarity

Examples

```
data(tumors.mRNA)
data(tumors.miRNA)
exp<-tumors.mRNA
trait<-tumors.miRNA
gene.trait.pvalue(EXP=exp[1:10,],trait=trait[1:5,],measure="MI")
```

---

gene.trait.similarity *Calculate similarity for gene-trait interaction*

Description

To calculate similarity for gene-trait interaction (using correlation/mutual information metric).

Usage

```
gene.trait.similarity(EXP, trait, measure, na.replica = 50)
```

Arguments

EXP	Gene expression data in form of a matrix. Row stands for genes and column for experiments.
trait	Trait data in form of matrix. Row stands for traits and column for experiments.
measure	Metric used to calculate similarity: "corr" for correlation, "MI" for mutual information.
na.replica	Times of replicates for filling NANs in impute method; default value is 50. The (smooth) bootstrapping approach is used to give an estimation to missing value in the data.

Value

A matrix, row stands for gene and column for trait. Element in row i and column j stands for the association between the gene i and trait j.

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

Examples

```
data(tumors.mRNA)
data(tumors.miRNA)
exp<-tumors.mRNA
trait<-tumors.miRNA
gene.trait.similarity(EXP= exp[1:10, ],trait= trait[1:5, ],measure="MI")
```

---

pred.network *Predict the network*

Description

To predict the matrix of gene network, based on the similarity matrix and filtered according to a corrected p-value matrix.

Usage

```
pred.network(pM,similarity,thresh)
```

Arguments

pM	A corrected p-value matrix, a MxM matrix for significance of similarity among M genes.
similarity	A MxM matrix for similarity between genes.
thresh	Threshold for significance of the p-value.

Value

A MxM matrix of the predicted network, where cell emphij infers a link between gene i and j and set 0 when the p-value is not significant (no link).

Author(s)

Yin Jin, Hesen Peng, Lei Wang, Raffaele Fronza, Yuanhua Liu and Christine Nardini

Examples

```
data(copasi)
mat<-as.matrix(copasi[1:10,])
rownames(mat)<-paste("G",1:nrow(mat), sep="")
similarity=gene.similarity(mat,measure="MI",net.trim="mrnet")
pM=gene.pvalue(mat,measure="MI",net.trim="mrnet")$single.perm.p.value
pred.network(pM,similarity,thresh=0.05)
```

tumors.miRNA *miRNA data from Human brain tumors*

Description

MiRNA data obtained by RT-PCR from human brain tumors. 12 brain tumors at different levels are analyzed for both mRNA and miRNA levels to study the correlation of any mRNA-miRNA pair in the *reference* .

Usage

```
data(tumors.miRNA)
```

Format

`tumors.miRNA` is a matrix with miRNA as rows and tumor type as columns.

References

Liu T, Papagiannakopoulos T, Puskar K, Qi S, Santiago F, Clay W, Lao K, Lee Y, Nelson SF, Kornblum HI, Doyle F, Petzold L, Shraiman B, Kosik KS. Detection of a microRNA signal in an in vivo expression set of mRNAs. Plos One. 2007; 2(8):e804.

Examples

```
data(tumors.miRNA)
tumors.miRNA[1:10,]
```

tumors.mRNA *Gene expression data from Human brain tumors*

Description

Gene expression data obtained by microarray from human brain tumors. 12 brain tumors at different levels are analyzed for both mRNA and miRNA levels to study the correlation of any mRNA-miRNA pair in the *reference* .

Usage

```
data(tumors.mRNA)
```

Format

`tumors.mRNA` is a matrix with mRNA probe IDs as rows and tumor type as columns.

References

Liu T, Papagiannakopoulos T, Puskar K, Qi S, Santiago F, Clay W, Lao K, Lee Y, Nelson SF, Kornblum HI, Doyle F, Petzold L, Shraiman B, Kosik KS. Detection of a microRNA signal in an in vivo expression set of mRNAs. Plos One. 2007; 2(8):e804.

Examples

```
data(tumors.mRNA)
tumors.mRNA[1:10,]
```

Index

∗ **arith**
gene.similarity, 6
gene.trait.similarity, 8
pred.network, 9
∗ **datasets**
copasi, 4
tumors.miRNA, 10
tumors.mRNA, 10
∗ **htest**
BUS, 2
gene.pvalue, 4
gene.trait.pvalue, 7
∗ **multivariate**
BUS, 2
BUS-package, 2

BUS, 2
BUS-package, 2

copasi, 4

gene.pvalue, *3*, 4
gene.similarity, *5*, 6
gene.trait.pvalue, *3*, 7
gene.trait.similarity, *8*, 8

pred.network, *3*, 9

tumors.miRNA, 10
tumors.mRNA, 10