BicARE

October 25, 2011

`BicARE-package` *BicARE*

Description

Biclustering Analysis and Results Exploration

Details

Package:	BicARE
Version:	0.1.0
Date:	2008-06-05
Depends:	R (>= 1.8.0), Biobase, multtest, GSEABase
License:	GPL
biocViews:	Microarray, Transcription, Statistics, Clustering
URL:	http://bioinfo.curie.fr
Packaged:	Tue Aug 19 14:21:38 2008; pgestraud
Built:	R 2.8.0; i686-pc-linux-gnu; 2008-08-19 14:39:57; unix

Index:

Further information is available in the following vignettes:

`BicARE`	BicARE (source, pdf)

Author(s)

Pierre Gestraud

Maintainer: Pierre Gestraud , <pierre.gestraud@curie.fr>

`FLOC` *Performs the FLOC algorithm*

Description

Find a given number of biclusters using the a modified version of the FLOC algorithm.

Usage

```
FLOC(Data, k = 20, pGene = 0.5, pSample=pGene, r = NULL, N = 8, M
= 6, t = 500, blocGene = NULL, blocSample = NULL)
```

Arguments

`Data`	an `ExpressionSet` or a matrix (with genes on rows and conditions on columns)
`k`	the number of biclusters searched
`pGene`	genes initial probability of membership to the biclusters
`pSample`	samples initial probability of membership to the biclusters
`r`	the residue threshold
`N`	minimal number of gene per bicluster
`M`	minimal number of conditions per bicluster
`t`	number of iterations
`blocGene`	a matrix indicating the directed initialisation for the genes (see details)
`blocSample`	a matrix indicating the directed initialisation for the conditions (see details)

Details

This biclustering algorithm is based on the FLOC algorithm (FLexible Overlapped biClustering) defined by Yang et al. (see references). It can discover a set of `k`, possibly overlapping, biclusters. If `r` is set to NULL, the residue threshold used in the analysis is the residue of `Data` divided by 10.

`blocGene` and `blocSample` are matrix of 0 and 1 with the rows representing the features (gene or samples) and the columns the biclusters. A 1 on line i and column j indicates that the feature i (gene or sample) will be include in the bicluster j during the initialisation step and will not be removed from it during the analysis. If the number of columns in these matrices is different from the number of bicluster searched, `k` is set to the maximal value of these two.

See `bicluster` to extract a bicluster from the biclustering result.

Value

Returns an object of class 'biclustering', a list containing at least :

`Call`	the matched call.
`ExpressionSet`	the data used
`param`	a data.frame with the algorithm parameters
`bicRow`	a matrix of boolean indicating the belonging of the genes to the biclusters
`bicCol`	the same as for bicRow but for the conditions
`mat.resvol.bic`	a matrix describing the biclusters

Author(s)

Pierre Gestraud (<pierre.gestraud@curie.fr>)

References

J. Yang, H. Wang, W. Wang, and P.S. Yu. An improved biclustering method for analyzing gene expression. International Journal on Artificial Intelligence Tools, 14(5):771-789, 2005

Examples

```
data(sample.bicData)     ## subset of sample.ExpressionSet from Biobase
residue(sample.bicData)  ##  0.3401921
resBic <- FLOC(sample.bicData, k=10, pGene=0.5,r=0.05,N=8,M=10,t=500)
resBic

## initialising samples of 2 biclusters
iniSample <- matrix(0, ncol=2, nrow=26)
## first bicluster initialised around Female cases
iniSample[pData(sample.bicData)$sex=="Female",1] <- 1
## second bicluster initialised around control cases
iniSample[pData(sample.bicData)$type=="Control",2] <- 1
resBic <- FLOC(sample.bicData, k=10, pGene=0.5, r=0.05, N=8, M=10, t=500, blocSample=iniS
resBic
```

`bicluster` *Extract a bicluster*

Description

Extract a bicluster from an object of class biclustering

Usage

```
bicluster(biclustering, k, graph=TRUE)
```

Arguments

biclustering	an object of class "biclustering" created by function FLOC
k	the number of the bicluster considered in the "biclustering" object
graph	boolean, indicating whether the graph should be plotted or not

Value

Returns the bicluster as a matrix with the genes on rows and the samples on columns. Result matrix is of class "bicluster". The "graph" option allows to plot the expression profiles of the genes across the conditions in the bicluster.

Author(s)

Pierre Gestraud

Examples

```
### extract the first bicluster
data(sample.biclustering)
sample.biclustering
bic <- bicluster(sample.biclustering, 1, graph=TRUE)
plot(bic)
```

---

makeReport *Export the results as html files*

Description

Creates a directory with html files containing the biclustering results.

Usage

```
makeReport(dirPath, dirName, resBic, browse=TRUE)
```

Arguments

dirPath	path to the directory
dirName	the name of the directory where the report will be created
resBic	a biclustering result
browse	logical. If TRUE the web browser will be opened

Details

makeReport produces a html report of biclustering results in a new directory named dirName. If the browse argument is set to TRUE the web browser will be opened on the "home.html" file.

Make sure to have rights to create the result directory.

Author(s)

Pierre Gestraud <pierre.gestraud@curie.fr>

Examples

```
data(sample.biclustering)
dirPath <- getwd() ## report created in the current working directory
dirName <- "test"
makeReport(dirPath, dirName, sample.biclustering, browse=FALSE)
```

`residue` *Residue of a matrix*

Description

Returns the residue of a matrix.

Usage

```
residue(Data)
```

Arguments

`Data`	an `ExpressionSet-class` or a matrix

Details

This function computes the residue of a matrix as defined by Yang et al (see references).

Author(s)

Pierre Gestraud

References

J. Yang, H. Wang, W. Wang, and P.S. Yu. An improved biclustering method for analyzing gene expression. International Journal on Artificial Intelligence Tools, 14(5):771-789, 2005

See Also

`FLOC`

Examples

```
data(sample.bicData)
residue(sample.bicData)
```

sample.bicData *Example data set for BicARE*

Description

A subset of sample.ExpressionSet from package Biobase. The data for 26 cases, labeled A to Z and 350 genes. Each case has three covariates: sex (male/female), type (case/control) and score (testing score).

Usage

```
sample.bicData
```

Format

An ExpressionSet

sample.biclustering *Example biclustering object*

Description

A biclustering object created by the FLOC function on the sample.bicData with the following options : k=10, pGene = 0.3, pSample = 0.5, r = 0.025, N = 8, M = 8, t = 1000.

Usage

```
sample.biclustering
```

Format

a biclustering object

testAnnot *Find samples annotations over-represented covariates in biclusters*

Description

Characterisation of the biclusters in term of over-representation of sample covariates.

Usage

```
testAnnot(resBic, annot=NULL, covariates="all")
```

Arguments

resBic	a biclustering result from FLOC
annot	annotation matrix, default value is set to NULL, then phenoData of the ExpressionSet is used
covariates	the names of the covariates that should be tested, default value is set to "all"

Details

For each bicluster and each covariate a chi-squarred test is performed to test the adequation between the distribution of the levels of the covariates in the bicluster and in the original dataset.

Multiple testing correction is performed by the Benjamini-Yekutieli procedure. The residuals of the tests indicate if the level is over or down represented in the bicluster.

Due to the amount of results it is advised to use the makeReport function to get a html report.

Value

A biclustering object containing resBic and updated with the results of the tests in resBic$covar.

The results are presented as a list with :

covar	the samples covariates tested
pvalues	a matrix with the p-values of the tests
adjpvalues	a matrix with the p-values adjusted by the Benjamini Yekutieli procedure
index	a list of matrices with the numbers of each level in each bicluster
residuals	a list of matrices with the residuals of the tests for each modality in each bicluster

Author(s)

Pierre Gestraud

Examples

```
data(sample.biclustering)
resBic <- testAnnot(sample.biclustering, annot=NULL, covariates=c("sex", "type"))
```

---

testSet *Find gene sets that are enriched in a bicluster*

Description

Test of the over-representation of gene sets in the biclusters

Usage

```
testSet(resBic, geneSetCol)
```

Arguments

resBic	a biclustering object created by FLOC
geneSetCol	a GeneSetCollection-class

Details

The over-representation of a gene set in a bicluster is evaluated by an hypergeometric test.

The genes identifiers of the gene sets will automatically be mapped to the same as those used in the data.

Due to the amount of results it is advised to use the `makeReport` function to get a html report.

Value

A `biclustering` object containing `resBic` and updated with the results of the tests in `resBic$geneSet`.

The results are presented as a list with :

`GeneSetCollection`	the GeneSetCollection used
`pvalues`	a matrix containing the pvalues of the tests for each geneSet and each bicluster
`adjpvalue`	a matrix containing the p-values adjusted by the Benjamini Yekutieli procedure

Author(s)

Pierre Gestraud <pierre.gestraud@curie.fr>

Examples

```
data(sample.biclustering)
gss <- GeneSetCollection(sample.biclustering$ExpressionSet[1:50,], setType=GOCollection()
resBic <- testSet(sample.biclustering, gss)
```

Index