

# Package ‘BFWGCNA’

February 15, 2016

**Type** Package

**Title** Differential Co-Expression Network Analysis

**Version** 1.0

**Date** 2015-10-18

**Author** Duolin Wang

**Maintainer** Duolin Wang <deepduoduo@gmail.com>

**Description** A Bayesian Approach to Detect Differential Co-Expression Gene Modules under case/control conditions.

**License** GPL (>= 2)

**Imports** Rcpp (>= 0.12.0)

**LinkingTo** Rcpp

**Depends** R (>= 3.0), dynamicTreeCut (>= 1.62), fastcluster, WGCNA , igraph

**NeedsCompilation** yes

## R topics documented:

BF_output_networks . . . . .	2
BF_pagerank . . . . .	4
BF_similarity . . . . .	5
BF_WGCNA . . . . .	5
Compute_bf . . . . .	7
SelectPower . . . . .	8
SimulationSmall . . . . .	9
<b>Index</b>	<b>10</b>

---

BF\_output\_networks *Output results*


---

**Description**

Output differential co-expression modules and information of gene-gene interactions.

**Usage**

```
BF_output_networks (
  dataExp,
  class,
  BFobt,
  showMST2=TRUE,
  mst2file="MST2.txt",
  bfthr=6,
  pthres=0.05,
  corthres=0.3)
```

**Arguments**

dataExp	data frame or matrix containing expression data. Columns correspond to genes and rows to samples. The same as the dataExp in function <a href="#">Compute_bf</a> .
class	vector containing class information. It is sorted according to the rows of dataExp. Must be the same as the class in function <a href="#">Compute_bf</a> .
BFobt	A BFobt object generated from function <a href="#">BF_WGCNA</a> .
showMST2	logical: generating First and Second Minimal Spanning Tree (MST2) or not? If TRUE, generate MST2 skeletons for each differential co-expression module, and output these skeletons to a file specified in argument mst2file below. If FALSE, leave out this procedure.
mst2file	If showMST2 is TRUE, specify a file to output mst2 skeletons. The default output is a file called "MST2.txt".
bfthr	thresholding bayes factor.
pthres	thresholding p-value.
corthres	thresholding Pearson Correlation Coefficient of gene-gene interactions for classing edges into up, down and unchange classes.

**Details**

The function output resulting differential co-expression modules and gene-gene interaction from modules.

**Value**

A list containing `genegroups` and `network`.

`genegroups` containing following 4 columns:

1. gene: A character vector giving gene ids.

2. colors: A character vector giving color labels of modules, with "grey" meaning unassigned.
3. labels: Numerical labels of modules, with 0 meaning unassigned.
4. gene\_weight: Numerical vector giving weight for each gene, with "NA" for unassigned genes.

network containing following 16 columns:

1. geneid1: A character vector giving ids for gene1.
2. geneid2: A character vector giving ids for gene2.
3. groupid: Numerical labels of modules for gene1 and gene2, connected genes belong to same modules.
4. gene1\_weight: Numerical vector giving weight for gene1.
5. gene2\_weight: Numerical vector giving weight for gene2.
6. bf.value: Numerical vector giving pair-wise bayes factor values for gene1 and gene2.
7. p.value: Numerical vector giving p-values of the pair-wise bayes factor values.
8. pc1: Numerical vector giving Pearson Correlation Coefficient of gene1 and gene2 under condition1.
9. pc2: Numerical vector giving Pearson Correlation Coefficient of gene1 and gene2 under condition2.
10. correlation: A character vector indicate the direction of correlation changing in terms of Pearson Correlation Coefficient. Compare the difference of pc1 and pc2 with corthres, it gives values from ("down", "up", "unchange").
11. mean\_g1\_c1: Numerical vector giving the mean expression for gene1 under condition1.
12. mean\_g1\_c2: Numerical vector giving the mean expression for gene1 under condition2.
13. mean\_foldchange\_gene1: Numerical vector giving the fold change of mean expression for gene1 between condition1 and condition2.
14. mean\_g2\_c1: Numerical vector giving the mean expression for gene2 under condition1.
15. mean\_g2\_c2: Numerical vector giving the mean expression for gene2 under condition2.
16. mean\_foldchange\_gene2: Numerical vector giving the fold change of mean expression for gene2 between condition1 and condition2.

### Note

only output the items which can pass the `bfthr` thresholding and `pthres` thresholding.

### Author(s)

Duolin Wang

### References

~~ Literature or other references for background information ~~

BF\_pagerank

*Estimate gene importance by PageRank algorithm.***Description**

Run PageRank algorithm on pair-wise bayes factor to estimate global gene importance.

**Usage**

```
BF_pagerank(bfmatrix,
            bfthr=6,
            pthres=0.05,
            d=0.85)
```

**Arguments**

<code>bfmatrix</code>	A <code>bfmatrix</code> object containing information of pair-wise bayes factors generated from function <code>Compute_bf</code> . Details in <code>Compute_bf</code> .
<code>bfthr</code>	thresholding bayes factor.
<code>pthres</code>	thresholding p-value.
<code>d</code>	the damping factor in Pagerank algorithm. see " <a href="http://www.sciencedirect.com/science/article/pii/S016975529800110X">http://www.sciencedirect.com/science/article/pii/S016975529800110X</a> ".

**Details**

This function constructs a weighted undirected graph by using pair-wise bayes factor as weight and applies a PageRank algorithm on this weighted undirected graphs, giving each gene a PageRank value and rank sorted by PageRank value in descending order.

**Value**

A data frame containing the following elements:

<code>geneid</code>	A character vector giving gene ids.
<code>pagerank.value</code>	Numerical vector giving PageRank values.
<code>pagerank.rank</code>	Numerical vector giving gene ranks sorted by <code>pagerank.value</code> in descending order.

**Author(s)**

Duolin Wang

**References**

"<http://www.sciencedirect.com/science/article/pii/S016975529800110X>"  
 "<https://web.eecs.umich.edu/~mihalcea/papers/mihalcea.emnlp04.pdf>"

---

BF_similarity	<i>Convert pair-wise bayes factors into network adjacency</i>
---------------	---

---

### Description

Convert pair-wise bayes factors into network adjacency which can be used as an input for other methods.

### Usage

```
BF_similarity(bfmatrix,bfthr=6,pthres=0.05)
```

### Arguments

<code>bfmatrix</code>	data frame containing information of pair-wise bayes factors generated from function <a href="#">Compute_bf</a> . Details in <a href="#">Compute_bf</a> .
<code>bfthr</code>	thresholding bayes factor.
<code>pthres</code>	thresholding p-value.

### Details

The function generating similarity matrix from pair-wise differential co-expression under two different conditions measured by pair-wise bayes factors. It transfer pair-wise bayes factor matrix (`bfmatrix`) into an adjacency matrix which can be used as an input for other methods.

### Value

Adjacency matrix. If `bfmatrix` contains  $n$  different genes, adjacency matrix is of dimensions  $n$  times  $n$ .

### Author(s)

Duolin Wang

### References

~~ Literature or other references for background information ~~

---

BF_WGCNA	<i>Identify differential co-expression modules</i>
----------	--

---

### Description

Generate differential co-expression modules through pair-wise bayes factors.

## Usage

```
BF_WGCNA(bfmatrix,
         softPower=6,
         bfthr=6,
         pthres=0.05,
         plotTree=TRUE,
         plotfile="Gene_dendrogram_and_module_colors.pdf",
         ...)
```

## Arguments

<code>bfmatrix</code>	A <code>bfmatrix</code> object containing information of pair-wise bayes factors generated from function <code>Compute_bf</code> . Details in <code>Compute_bf</code> .
<code>softPower</code>	soft thresholding power. Same as the power in package WGCNA
<code>bfthr</code>	thresholding bayes factor.
<code>pthres</code>	thresholding p-value.
<code>plotTree</code>	if TRUE, plots a hierarchical clustering dendrogram and color annotation(s) of module(s) in file specified by argument <code>plotfile</code> below.
<code>plotfile</code>	if <code>plotTree</code> is TRUE, specify a file (in pdf format) to output the dendrogram plot. The default output is a file called "Gene_dendrogram_and_module_colors.pdf".
<code>...</code>	Arguments to be passed to method <code>cutreeDynamic</code> in package <code>dynamicTreeCut</code> . see " <a href="https://CRAN.R-project.org/package=dynamicTreeCut">https://CRAN.R-project.org/package=dynamicTreeCut</a> " Arguments, like <code>minClusterSize</code> , <code>cutHeight</code> , <code>deepSplit</code> and method are very crucial for clustering.
<code>minClusterSize</code>	Minimum cluster size.
<code>cutHeight</code>	<code>cutHeight</code> Maximum joining heights that will be considered. see " <a href="https://CRAN.R-project.org/package=dynamicTreeCut">https://CRAN.R-project.org/package=dynamicTreeCut</a> "
<code>deepSplit</code>	provides a rough control over sensitivity to cluster splitting. see " <a href="https://CRAN.R-project.org/package=dynamicTreeCut">https://CRAN.R-project.org/package=dynamicTreeCut</a> "
<code>method</code>	Chooses the method to use. Recognized values are "hybrid" and "tree".

## Details

This function generates differential co-expression modules from pair-wise differential co-expression under two different conditions measured by pair-wise bayes factors. It transfer pair-wise bayes factor matrix into adjacency matrix first, then transfer this adjacency matrix into topological overlap matrix by TOMdist function from package WGCNA, then do hierarchical clustering using `cutreeDynamic` function from package `dynamicTreeCut`. Only pair-wise gene-gene interactions which satisfy the bayes factor thresholding and permutation p-value thresholding can be considered in this procedure.

## Value

A `BFobt` object containing gene modules, adjacency matrix, bayes factor matrix, and p-value matrix.

## Author(s)

Duolin Wang

## References

~~ Literature or other references for background information ~~

---

Compute_bf	<i>Estimation the strength of pair-wise differential co-expression by bayes factor</i>
------------	--

---

## Description

Calculation of the pair-wise bayes factors from a given expression matrix.

## Usage

```
Compute_bf(dataExp, class, permutation=TRUE, nperm=20, bfthr=6)
```

## Arguments

dataExp	data frame or matrix containing expression data. Columns correspond to genes and rows to samples.
class	vector containing class information. It is sorted according to the rows of data-Exp. By far package only accept binary conditions, use different value for each condition.
permutaion	logical: permutation or not? If TRUE, do permutation. If FALSE, don't do permutaion.
nperm	integer: times of permutaion. It is valid only if permutation is set as TRUE.
bfthr	thresholding bayes factor.

## Details

The function estimate the stength of pair-wise differential co-expression by computing pair-wise bayes factors from a given expression matrix.

## Value

A `bfmatrix` object (data frame) containing the information of pair-wise bayes factors. It contains the following elements:

geneid1	A character vector giving gene ids for gene1?
geneid2	A character vector giving gene ids for gene2?
bf.value	Numerical vector giving pair-wise bayes factor values for gene1 and gene2.
p.value	Numerical vector giving p-values of the pair-wise bayes factor values.

## Note

only output the items which can pass the `bfthr` thresholding.

## Author(s)

Duolin Wang

## References

~~ Literature or other references for background information ~~

---

SelectPower	<i>Choosing the soft-thresholding power by analysing of network topology</i>
-------------	--

---

## Description

choose the soft thresholding power based on the criterion of approximate scale-free topology.

## Usage

```
SelectPower(bfsimilarity, powers)
```

## Arguments

`bfsimilarity` Adjacency matrix based on pair-wise bayes factor generated from [BF\\_similarity](#).  
`powers` A vector of candidate soft-thresholding powers provided by customer, soft-thresholding power will be choosed from it.

## Details

This function plots a network topology under a set of soft-thresholding powers (a set of candidate `softpower` parameters for [BF\\_WGCNA](#)) provided by customer. It constructs two panels in the plot, the left panel shows the scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis). The right panel shows the mean connectivity (degree, y-axis) as a function of the soft-thresholding power (x-axis). According to WGCNA (<https://labs.genetics.ucla.edu/horvath/htdocs/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/FemaleLiver-02-pdf>), they proposed to choose the lowest soft thresholding power and achieve an approximate scale-free topology (scale-free topology fit index reaches 0.9).

## Value

None.

## Author(s)

Duolin Wang

## References

B. Zhang and S. Horvath. A general framework for weighted gene co-expression network analysis. Statistical Applications in Genetics and Molecular Biology, 4(1):Article 17, 2005.



---

SimulationSmall	<i>small simulation data for test.</i>
-----------------	--

---

## Description

This matrix gives class information and expression data generated by a simulation procedure.

## Usage

```
data(SimulationSmall)
```

## Format

A 200 x 241 matrix. The first column (class) lists class labels, 1 indicate condition1 and 2 indicate condition2. Other columns correspond to genes. The rows are correspond to samples. There are 7 co-expression modules in total, each contains 20 genes, and 100 randomly generated genes. Generated by

$$dataCond1 = ModuleMaineneC1 * \sqrt{varMainC1} + RandomMatrix * \sqrt{varRdmC1} + SubModuleMatrix * \sqrt{varSubC1}$$

$$dataCond2 = ModuleMaingeneC2 * \sqrt{varMainC2} + RandomMatrix * \sqrt{varRdmC2} + SubModuleMatrix * \sqrt{varSubC2}$$

module1-module2 have the same distribution in both condition1 and condition2, so they are stable modules.

module3 and module4: are differential co-expression modules. Genes in these modules have different co-expression patterns. For module3,  $varMainC1=0.5$  and  $varMainC2=0$ ; For module4,  $varMainC1=0.2$  and  $varMainC2=0.8$ . ModuleMaingeneC1 and ModuleMaingeneC2 are draw from  $norm(0,1)$ .

module5 is differential co-expression module. Genes in this module have different mean values for different conditions, ModuleMaingeneC1 is draw from  $norm(0,1)$  but ModuleMaingeneC2 is draw from  $norm(4,1)$ ;

module6 and module7 have Module-to-Module differential co-expression but no within-module differential co-expression. module6 is generated as module1 and module2. ModuleMaingeneC1 of module6 and ModuleMaingeneC1 of module7 have positive correlation but ModuleMaingeneC2 of module6 and ModuleMaingeneC2 of module7 have negative correlation.

## Examples

```
data(SimulationSmall)
head(SimulationSmall)
```

# Index

## \*Topic **datasets**

SimulationSmall, [9](#)

## \*Topic **package**

BF\_output\_networks, [2](#)

BF\_pagerank, [4](#)

BF\_similarity, [5](#)

BF\_WGCNA, [5](#)

Compute\_bf, [7](#)

SelectPower, [8](#)

BF\_output\_networks, [2](#)

BF\_pagerank, [4](#)

BF\_similarity, [5](#), [8](#)

BF\_WGCNA, [2](#), [5](#), [8](#)

Compute\_bf, [2](#), [4-6](#), [7](#)

cutreeDynamic, [6](#)

SelectPower, [8](#)

SimulationSmall, [9](#)