

# Package ‘miRSM’

May 30, 2024

**Type** Package

**Title** Inferring miRNA sponge modules in heterogeneous data

**Version** 2.0.0

**Description** The package aims to identify miRNA sponge or ceRNA modules in heterogeneous data. It provides several functions to study miRNA sponge modules at single-sample and multi-sample levels, including popular methods for inferring gene modules (candidate miRNA sponge or ceRNA modules), and two functions to identify miRNA sponge modules at single-sample and multi-sample levels, as well as several functions to conduct modular analysis of miRNA sponge modules.

**Depends** R (>= 3.5.0)

**License** GPL-3

**URL** <https://github.com/zhangjunpeng411/miRSM>

**Encoding** UTF-8

**biocViews** GeneExpression, BiomedicalInformatics, Clustering, GeneSetEnrichment, Microarray, Software, GeneRegulation, GeneTarget

**RoxygenNote** 7.3.1

**Imports** WGCNA, flashClust, dynamicTreeCut, GFA, igraph, linkcomm, MCL, fabia, NMF, biclust, iBBiG, BicARE, isa2, s4vd, BiBitR, rqubic, Biobase, PMA, stats, dbscan, subspace, mclust, SOMbrero, ppclust, Repp, utils, SummarizedExperiment, GSEABase, org.Hs.eg.db, clusterProfiler, ReactomePA, DOSE, MatrixCorrelation, energy

**Suggests** BiocStyle, knitr, rmarkdown, testthat

**VignetteBuilder** knitr

**BugReports** <https://github.com/zhangjunpeng411/miRSM/issues>

**git\_url** <https://git.bioconductor.org/packages/miRSM>

**git\_branch** RELEASE\_3\_19

**git\_last\_commit** d3b3055

**git\_last\_commit\_date** 2024-04-30

**Repository** Bioconductor 3.19

**Date/Publication** 2024-05-29

**Author** Junpeng Zhang [aut, cre]

**Maintainer** Junpeng Zhang <zjp@dali.edu.cn>

## Contents

BRCA_genes . . . . .	2
ceRExp . . . . .	3
cor_binary . . . . .	4
diff_module . . . . .	5
miRExp . . . . .	6
miRSM . . . . .	6
miRSM_SS . . . . .	9
miRTarget . . . . .	10
module_biclust . . . . .	10
module_CEA . . . . .	12
module_clust . . . . .	13
module_Coexpress . . . . .	15
module_FA . . . . .	16
module_GFA . . . . .	18
module_group_sim . . . . .	19
module_igraph . . . . .	20
module_miRdistribute . . . . .	21
module_miRsponge . . . . .	22
module_miRtarget . . . . .	23
module_NMF . . . . .	24
module_ProNet . . . . .	25
module_Validate . . . . .	26
module_WGCNA . . . . .	27
mRExp . . . . .	28
share_miRs . . . . .	28
<b>Index</b>	<b>30</b>

---

BRCA\_genes

*BRCA genes*

---

## Description

BRCA genes

## Format

BRCA\_genes: A SummarizedExperiment object with 4819 BRCA related genes (including lncRNAs and mRNAs).

## Details

The BRCA related lncRNAs are from LncRNADisease v2.0, Lnc2Cancer v2.0 and MNDR v2.0. The BRCA related mRNAs are from DisGeNET v5.0 and COSMIC v86.

## References

- Bao Z, Yang Z, Huang Z, Zhou Y, Cui Q, Dong D. (2019) "LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases". *Nucleic Acids Res.*, 47(D1):D1034-D1037.
- Cui T, Zhang L, Huang Y, Yi Y, Tan P, Zhao Y, Hu Y, Xu L, Li E, Wang D. (2018) "MNDR v2.0: an updated resource of ncRNA-disease associations in mammals". *Nucleic Acids Res.*, 46, D371-D374.
- Gao Y, Wang P, Wang Y, Ma X, Zhi H, Zhou D, Li X, Fang Y, Shen W, Xu Y, Shang S, Wang L, Wang L, Ning S, Li X. (2019) "Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers". *Nucleic Acids Res.*, 47, D1028-D1033.
- Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T, Campbell PJ. (2017) "COSMIC: somatic cancer genetics at high-resolution". *Nucleic Acids Res.*, 45, D777-D783
- Pinero J, Bravo A, Queralt-Rosinach N, Gutierrez-Sacristan A, Deu-Pons J, Centeno E, Garcia-Garcia J, Sanz F, Furlong LI. (2017) "DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants". *Nucleic Acids Res.*, 45, D833-D839.

---

ceRExp

*ceRNA expression data*

---

## Description

ceRNA expression data

## Format

ceRExp: A SummarizedExperiment object with 72 BRCA and 72 normal samples (rows) and 305 lncRNAs (columns).

## Details

The matched breast invasive carcinoma (BRCA) miRNA, lncRNA and mRNA expression data is obtained from TCGA (<http://cancergenome.nih.gov/>). lncRNA expression data is regarded as ceRNA expression data. The data focuses on 72 individuals for which the complete sets of tumor and matched normal (i.e., normal tissue taken from the same patient) profiles are available. A lncRNA which has missing values in more than 10 are imputed using the k-nearest neighbours (KNN) algorithm from the impute R package. We use the limma R package to infer differentially expressed lncRNAs between tumour and normal samples. After the analysis, we select top 305 lncRNAs which are differentially expressed at a significant level (adjusted p-value < 1E-02, adjusted by Benjamini & Hochberg method).

---

`cor_binary`*cor\_binary*

---

**Description**

Generation of positively correlated binary matrix between ceRNAs, or ceRNAs and mRNAs

**Usage**

```
cor_binary(  
  ceExp,  
  mRExp = NULL,  
  cor.method = "pearson",  
  pos.p.value.cutoff = 0.01  
)
```

**Arguments**

<code>ceExp</code>	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
<code>mRExp</code>	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
<code>cor.method</code>	The method of calculating correlation selected, including 'pearson' (default), 'kendall', 'spearman'.
<code>pos.p.value.cutoff</code>	The significant p-value cutoff of positive correlation.

**Value**

A binary matrix.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008, 9:559.

**Examples**

```
data(BRCASampleData)  
cor_binary_matrix <- cor_binary(ceExp, mRExp)
```

---

diff_module	<i>diff_module</i>
-------------	--------------------

---

## Description

Inferring differential modules between two list of module groups

## Usage

```
diff_module(  
  Module.group1,  
  Module.group2,  
  sim.cutoff = 0.8,  
  sim.method = "Simpson"  
)
```

## Arguments

Module.group1	List object, the first list of module group.
Module.group2	List object, the second list of module group.
sim.cutoff	Similarity cutoff between modules, the interval is [0 1].
sim.method	Methods for calculating similarity between two modules, select one of three methods (Simpson, Jaccard and Lin). Default method is Simpson.

## Value

A list of differential modules

## Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## Examples

```
library(GSEABase)  
data(BRCASampleData)  
modulegenes_WGCNA_all <- module_WGCNA(ceRExp, mRExp)  
modulegenes_WGCNA_1 <- module_WGCNA(ceRExp[-1, ], mRExp[-1, ])  
Differential_module <- diff_module(geneIds(modulegenes_WGCNA_all), geneIds(modulegenes_WGCNA_1))
```

---

miRExp	<i>miRNA expression data</i>
--------	------------------------------

---

### Description

miRNA expression data

### Format

miRExp: A SummarizedExperiment object with 72 BRCA and 72 normal samples (rows) and 226 miRNAs (columns).

### Details

The matched breast invasive carcinoma (BRCA) miRNA, lncRNA and mRNA expression data is obtained from TCGA (<http://cancergenome.nih.gov/>). The data focuses on 72 individuals for which the complete sets of tumor and matched normal (i.e., normal tissue taken from the same patient) profiles are available. A miRNA which has missing values in more than 10 are imputed using the k-nearest neighbours (KNN) algorithm from the impute R package. We use the limma R package to infer differentially expressed miRNAs, ceRNAs and mRNAs between tumour and normal samples. After the analysis, we select top 226 miRNAs which are differentially expressed at a significant level (adjusted p-value < 1E-02, adjusted by Benjamini & Hochberg method).

---

miRSM	<i>miRSM</i>
-------	--------------

---

### Description

Identify miRNA sponge modules using sensitivity canonical correlation (SCC), sensitivity distance correlation (SDC), sensitivity RV coefficient (SRVC), sensitivity similarity index (SSI), sensitivity generalized coefficient of determination (SGCD), sensitivity Coxhead's or Rozeboom's coefficient (SCRC), and sponge module (SM) methods.

### Usage

```
miRSM(
  miRExp = NULL,
  ceRExp,
  mRExp = NULL,
  miRTarget,
  CandidateModulegenes,
  typex = "standard",
  typez = "standard",
  nperms = 100,
  method = c("SCC", "SDC", "SRVC", "SM", "SSI", "SGCD", "SCRC"),
```

```

num_shared_miRNAs = 3,
pvalue.cutoff = 0.05,
MC.cutoff = 0.8,
SMC.cutoff = 0.1,
RV_method = c("RV", "RV2", "RVadjMaye", "RVadjGhaziri"),
BCmethod = "BCPlaid",
CRC_method = c("Coxhead", "Rozeboom")
)

```

## Arguments

miRExp	NULL (default) or a SummarizedExperiment object. miRNA expression data: rows are samples and columns are miRNAs.
ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
miRTarget	A SummarizedExperiment object. Putative miRNA-target binding information.
CandidateModulegenes	List object: a list of candidate miRNA sponge modules. Only for the SCC, SDC, SRVC, SSI, SGCD and SCRC methods.
typex	The columns of x unordered (type='standard') or ordered (type='ordered'). Only for the SCC method.
typez	The columns of z unordered (type='standard') or ordered (type='ordered'). Only for the SCC method.
nperms	The number of permutations. Only for the SCC method.
method	The method selected to identify miRNA sponge modules, including 'SCC', 'SDC', 'SRVC', 'SM', 'SSI', 'SGCD' and 'SCRC'.
num_shared_miRNAs	The number of common miRNAs shared by a group of ceRNAs and mRNAs. Only for the SCC, SDC, SRVC, SSI, SGCD and SCRC methods.
pvalue.cutoff	The p-value cutoff of significant sharing of common miRNAs by a group of ceRNAs and mRNAs or significant correlation.
MC.cutoff	The cutoff of matrix correlation (canonical correlation, distance correlation and RV coefficient). Only for the SCC, SDC, SRVC, SSI, SGCD and SCRC methods.
SMC.cutoff	The cutoff of sensitivity matrix correlation (sensitivity canonical correlation, sensitivity distance correlation and sensitivity RV coefficient). Only for the SCC, SDC, SRVC, SSI, SGCD and SCRC methods when miRExp is not NULL.
RV_method	the method of calculating RV coefficients. Select one of 'RV', 'RV2', 'RVadj-Maye' and 'RVadjGhaziri' methods. Only for the SRVC method.
BCmethod	Specification of the biclustering method, including 'BCBimax', 'BCCC', 'BC-Plaid' (default), 'BCQuest', 'BCSpectral', 'BCXmotifs'. Only for the SM method.
CRC_method	the method of calculating matrix correlation. Select one of 'Coxhead' and 'Rozeboom' methods. Only for the SCRC method.

**Value**

List object: Group competition of miRNA sponge modules, and miRNA sponge modules.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

- Witten DM, Tibshirani R, Hastie T. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*. 2009, 10(3):515-34.
- Szekely GJ, Rizzo ML. Partial distance correlation with methods for dissimilarities. *Annals of Statistics*. 2014, 42(6):2382-2412.
- Szekely GJ, Rizzo ML, Bakirov NK. Measuring and Testing Dependence by Correlation of Distances. *Annals of Statistics*, 2007, 35(6):2769-2794.
- Robert P, Escoufier Y. A unifying tool for linear multivariate statistical methods: the RV-Coefficient. *Applied Statistics*, 1976, 25(3):257-265.
- Smilde AK, Kiers HA, Bijlsma S, Rubingh CM, van Erk MJ. Matrix correlations for high-dimensional data: the modified RV-coefficient. *Bioinformatics*, 2009, 25(3):401-405.
- Maye CD, Lorent J, Horgan GW. Exploratory analysis of multiple omics datasets using the adjusted RV coefficient". *Stat Appl Genet Mol Biol.*, 2011, 10, 14.
- EIGHaziri A, Qannari EM. Measures of association between two datasets; Application to sensory data, *Food Quality and Preference*, 2015, 40(A):116-124.
- Indahl UG, Næs T, Liland KH. A similarity index for comparing coupled matrices. *Journal of Chemometrics*. 2018; 32:e3049.
- Yanai H. Unification of various techniques of multivariate analysis by means of generalized coefficient of determination (GCD). *Journal of Behaviormetrics*, 1974, 1(1): 45-54.
- Coxhead P. Measuring the relationship between two sets of variables. *British journal of mathematical and statistical psychology*, 1974, 27(2): 205-212.
- Rozeboom WW. Linear correlations between sets of variables. *Psychometrika*, 1965, 30(1): 57-71.

**Examples**

```
data(BRCASampleData)
modulegenes_igraph <- module_igraph(ceRExp[, seq_len(10)],
  mRExp[, seq_len(10)])
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_igraph_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
  modulegenes_igraph, method = "SRVC",
  SMC.cutoff = 0.01, RV_method = "RV")
```



---

miRSM\_SS

*miRSM\_SS*


---

**Description**

Inferring sample-specific miRNA sponge modules

**Usage**

```
miRSM_SS(
  Modulelist.all,
  Modulelist.exceptk,
  sim.cutoff = 0.8,
  sim.method = "Simpson"
)
```

**Arguments**

`Modulelist.all` List object, modules using all of samples.  
`Modulelist.exceptk` List object, modules using all of samples excepting sample k.  
`sim.cutoff` Similarity cutoff between modules, the interval is [0 1].  
`sim.method` Methods for calculating similarity between two modules, select one of three methods (Simpson, Jaccard and Lin). Default method is Simpson.

**Value**

A list of sample-specific miRNA sponge modules

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**Examples**

```
data(BRCASampleData)
nsamples <- 3
modulegenes_all <- module_igraph(ceRExp[, 151:300], mRExp[, 151:300])
modulegenes_exceptk <- lapply(seq(nsamples), function(i)
  module_WGCNA(ceRExp[-i, seq(150)],
    mRExp[-i, seq(150)]))

miRSM_SRVC_all <- miRSM(miRExp, ceRExp[, 151:300], mRExp[, 151:300],
  miRTarget, modulegenes_all,
  method = "SRVC", SMC.cutoff = 0.01,
  RV_method = "RV")
miRSM_SRVC_exceptk <- lapply(seq(nsamples), function(i) miRSM(miRExp[-i, ],
  ceRExp[-i, seq(150)], mRExp[-i, seq(150)],
```

```

miRTarget, modulegenes_exceptk[[i]],
method = "SRVC",
SMC.cutoff = 0.01, RV_method = "RV"))

Modulegenes_all <- miRSM_SRVC_all[[2]]
Modulegenes_exceptk <- lapply(seq(nsamples), function(i)
  miRSM_SRVC_exceptk[[i]][[2]])

Modules_SS <- miRSM_SS(Modulegenes_all, Modulegenes_exceptk)

```

---

miRTarget

*miRNA-target interactions*


---

### Description

miRNA-target interactions

### Format

miRTarget: A SummarizedExperiment object with 29901 miRNA-target interactions.

### Details

The miRNA-target binding information is from miRTarBase v8.0 (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>), and LncBase v2.0 ([http://carolina.imis.athena-innovation.gr/diana\\_tools/web/index.php?r=lnccbasev2/index](http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lnccbasev2/index)). Among 226 miRNAs, 305 lncRNAs and 500 mRNAs which are differentially expressed, we obtain 29901 miRNA-target interactions (including miRNA-lncRNA and miRNA-mRNA interactions).

### References

Hastie T, Tibshirani R, Narasimhan B, Chu G. impute: Imputation for microarray data. R package version 1.54.0. doi: 10.18129/B9.bioc.impute.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015; 43(7):e47.

---

module\_biclust

*module\_biclust*


---

### Description

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using a series of biclustering packages, including biclust, iBBiG, fabia, BicARE, isa2, s4vd, BiBitR and rqubic

**Usage**

```

module_biclust(
  ceRExp,
  mRExp = NULL,
  BCmethod = "fabia",
  num.modules = 10,
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)

```

**Arguments**

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
BCmethod	Specification of the biclustering method, including 'BCBimax', 'BCCC', 'BC-Plaid' (default), 'BCQuest', 'BCSpectral', 'BCXmotifs', 'iBBiG', 'fabia', 'fabiap', 'fabias', 'mfsc', 'nmfdiv', 'nmfeu', 'nmfsc', 'FLOC', 'isa', 'BCs4vd', 'BC-ssvd', 'bibit' and 'quBicluster'.
num.modules	The number of modules to be identified. For the 'BCPlaid', 'BCSpectral', 'isa' and 'bibit' methods, no need to set the parameter. For the 'quBicluster' method, the parameter is used to set the number of biclusters that should be reported.
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

**Value**

GeneSetCollection object: a list of module genes.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

- Prelic A, Bleuler S, Zimmermann P, Wille A, Buhmann P, Gruissem W, Hennig L, Thiele L, Zitzler E. A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*. 2006, 22(9):1122-9.
- Cheng Y, Church GM. Biclustering of expression data. *Proc Int Conf Intell Syst Mol Biol*. 2000, 8:93-103.
- Turner H, Bailey T, Krzanowski W. Improved biclustering of microarray data demonstrated through systematic performance tests. *Comput Stat Data Anal*. 2003, 48(2): 235-254.
- Murali TM, Kasif S. Extracting conserved gene expression motifs from gene expression data. *Pac Symp Biocomput*. 2003:77-88.

- Kluger Y, Basri R, Chang JT, Gerstein M. Spectral biclustering of microarray data: coclustering genes and conditions. *Genome Res.* 2003, 13(4):703-16.
- Gusenleitner D, Howe EA, Bentink S, Quackenbush J, Culhane AC. iBBiG: iterative binary biclustering of gene sets. *Bioinformatics.* 2012, 28(19):2484-92.
- Hochreiter S, Bodenhofer U, Heusel M, Mayr A, Mitterecker A, Kasim A, Khamiakova T, Van Sanden S, Lin D, Talloen W, Bijmans L, G'ohlmann HW, Shkedy Z, Clevert DA. FABIA: factor analysis for bicluster acquisition. *Bioinformatics.* 2010, 26(12):1520-7.
- Yang J, Wang H, Wang W, Yu, PS. An improved biclustering method for analyzing gene expression. *Int J Artif Intell Tools.* 2005, 14(5): 771-789.
- Bergmann S, Ihmels J, Barkai N. Iterative signature algorithm for the analysis of large-scale gene expression data. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2003, 67(3 Pt 1):031902.
- Sill M, Kaiser S, Benner A, Kopp-Schneider A. Robust biclustering by sparse singular value decomposition incorporating stability selection. *Bioinformatics.* 2011, 27(15):2089-97.
- Lee M, Shen H, Huang JZ, Marron JS. Biclustering via sparse singular value decomposition. *Biometrics.* 2010, 66(4):1087-95.
- Rodriguez-Baena DS, Perez-Pulido AJ, Aguilar-Ruiz JS. A biclustering algorithm for extracting bit-patterns from binary datasets. *Bioinformatics.* 2011, 27(19):2738-45.
- Li G, Ma Q, Tang H, Paterson AH, Xu Y. QUBIC: a qualitative biclustering algorithm for analyses of gene expression data. *Nucleic Acids Res.* 2009, 37(15):e101.

## Examples

```
data(BRCASampleData)
modulegenes_biclust <- module_biclust(ceRExp[, seq_len(30)],
  mRExp[, seq_len(30)])
```

---

module\_CEA

*module\_CEA*

---

## Description

Cancer enrichment analysis of miRNA sponge modules using hypergeometric distribution test

## Usage

```
module_CEA(ceRExp, mRExp = NULL, Cancergenes, Modulelist)
```

## Arguments

- |             |  |
|-------------|--|
| ceRExp      | A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.                 |
| mRExp       | NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs. |
| Cancergenes | A SummarizedExperiment object: a list of cancer genes given.   |
| Modulelist  | List object: a list of the identified miRNA sponge modules.  |

**Value**

Cancer enrichment significance p-values of the identified miRNA sponge modules

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

Johnson NL, Kotz S, Kemp AW (1992) "Univariate Discrete Distributions", Second Edition. New York: Wiley.

**Examples**

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
                          modulegenes_WGCNA, method = "SRVC",
                          SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM.CEA.pvalue <- module_CEA(ceRExp, mRExp, BRCA_genes,
                              miRSM_WGCNA_SRVC_genes)
```

---

module\_clust

*module\_clust*

---

**Description**

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using a series of clustering packages, including stats, flashClust, dbscan, subspace, mclust, SOMbrero and ppclust packages.

**Usage**

```
module_clust(
  ceRExp,
  mRExp = NULL,
  cluster.method = "kmeans",
  num.modules = 10,
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)
```

**Arguments**

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
cluster.method	Specification of the clustering method, including 'kmeans'(default), 'hclust', 'dbscan', 'clique', 'gmm', 'som' and 'fcm'.
num.modules	Parameter of the number of modules to be identified for the 'kmeans', 'hclust', 'gmm' and 'fcm' methods. Parameter of the number of intervals for the 'clique' method. For the 'dbscan' and 'som' methods, no need to set the parameter.
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

**Value**

GeneSetCollection object: a list of module genes.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

- Forgy EW. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. *Biometrics*, 1965, 21:768-769.
- Hartigan JA, Wong MA. Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 1979, 28:100-108.
- Lloyd SP. Least squares quantization in PCM. Technical Note, Bell Laboratories. Published in 1982 in *IEEE Transactions on Information Theory*, 1982, 28:128-137.
- MacQueen J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, eds L. M. Le Cam & J. Neyman, 1967, 1, pp.281-297. Berkeley, CA: University of California Press.
- Langfelder P, Horvath S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *Journal of Statistical Software*. 2012, 46(11):1-17.
- Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise, *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 1996, 96(34): 226-231.
- Campello RJGB, Moulavi D, Sander J. Density-based clustering based on hierarchical density estimates, *Pacific-Asia conference on knowledge discovery and data mining*. Springer, Berlin, Heidelberg, 2013: 160-172.
- Agrawal R, Gehrke J, Gunopulos D, Raghavan P. Automatic subspace clustering of high dimensional data for data mining applications. In *Proc. ACM SIGMOD*, 1998.
- Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models *The R Journal* 8/1, 2016, pp. 205-233.

- Kohonen T. Self-Organizing Maps. Berlin/Heidelberg: Springer-Verlag, 3rd edition, 2001.
- Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 1973, 3(3):32-57.
- Bezdek JC. Cluster validity with fuzzy sets. *Journal of Cybernetics*, 1974, 3: 58-73.
- Bezdek JC. Pattern recognition with fuzzy objective function algorithms. Plenum, NY, 1981.

### Examples

```
data(BRCASampleData)
modulegenes_clust <- module_clust(ceRExp[, seq_len(30)],
  mRExp[, seq_len(30)])
```

---

module_Coexpress	<i>module_Coexpress</i>
------------------	-------------------------

---

### Description

Co-expression analysis of each miRNA sponge module and its corresponding random miRNA sponge modules

### Usage

```
module_Coexpress(
  ceRExp,
  mRExp = NULL,
  Modulelist,
  resample = 1000,
  method = c("mean", "median"),
  test.method = c("t.test", "wilcox.test")
)
```

### Arguments

- |            |   |
|------------|---|
| ceRExp     | A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.  |
| mRExp      | NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.  |
| Modulelist | List object: a list of the identified miRNA sponge modules.   |
| resample   | The number of random miRNA sponge modules generated, and 1000 times in default.   |
| method     | The method used to evaluate the co-expression level of each miRNA sponge module. Users can select "mean" or "median" to calculate co-expression value of each miRNA sponge module and its corresponding random miRNA sponge module. |

`test.method` The method used to evaluate statistical significance p-value of co-expression level higher than random miRNA sponge modules. Users can select "t.test" or "wilcox.test" to calculate statistical significance p-value of co-expression level higher than random miRNA sponge modules.

### Value

List object: co-expression values of miRNA sponge modules and their corresponding random miRNA sponge modules, and statistical significance p-value of co-expression level higher than random miRNA sponge modules.

### Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

### Examples

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
                        modulegenes_WGCNA, method = "SRVC",
                        SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM_WGCNA_Coexpress <- module_Coexpress(ceRExp, mRExp,
                                         miRSM_WGCNA_SRVC_genes,
                                         resample = 10, method = "mean",
                                         test.method = "t.test")
```

---

module\_FA

*module\_FA*

---

### Description

Functional analysis of miRNA sponge modules, including functional enrichment and disease enrichment analysis

### Usage

```
module_FA(
  Modulelist,
  GOont = "BP",
  Diseaseont = "DO",
  KEGGorganism = "hsa",
  Reactomeorganism = "human",
  OrgDb = "org.Hs.eg.db",
  padjustvaluecutoff = 0.05,
  padjustedmethod = "BH",
```





```

miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM_WGCNA_SRVC_FEA <- module_FA(miRSM_WGCNA_SRVC_genes, Analysis.type = 'FEA')
miRSM_WGCNA_SRVC_DEA <- module_FA(miRSM_WGCNA_SRVC_genes, Analysis.type = 'DEA')

## End(Not run)

```

---

 module\_GFA

*module\_GFA*


---

### Description

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using GFA package

### Usage

```

module_GFA(
  ceRExp,
  mRExp = NULL,
  StrengthCut = 0.9,
  iter.max = 5000,
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)

```

### Arguments

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
StrengthCut	Desired minimum strength (absolute value of association with interval [0 1]) for each bicluster.
iter.max	The total number of Gibbs sampling steps (default 1000).
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

### Value

GeneSetCollection object: a list of module genes.

### Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## References

Bunte K, Leppäaho E, Saarinen I, Kaski S. Sparse group factor analysis for biclustering of multiple data sources. *Bioinformatics*. 2016, 32(16):2457-63.

Leppäaho E, Ammad-ud-din M, Kaski S. GFA: exploratory analysis of multiple data sources with group factor analysis. *J Mach Learn Res*. 2017, 18(39):1-5.

## Examples

```
data(BRCASampleData)
modulegenes_GFA <- module_GFA(ceRExp[seq_len(20), seq_len(15)],
  mRExp[seq_len(20), seq_len(15)], iter.max = 3000)
```

---

module\_group\_sim      *module\_group\_sim*

---

## Description

Calculating similarity between two list of module groups

## Usage

```
module_group_sim(Module.group1, Module.group2, sim.method = "Simpson")
```

## Arguments

Module.group1    List object, the first list of module group.  
Module.group2    List object, the second list of module group.  
sim.method        Methods for calculating similarity between two modules, select one of three methods (Simpson, Jaccard and Lin). Default method is Simpson.

## Value

Similarity between two list of module groups

## Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## References

Simpson E H. Measurement of diversity. *Nature*, 1949, 163(4148): 688-688.

Jaccard P. The distribution of the flora in the alpine zone. 1. *New phytologist*, 1912, 11(2): 37-50.

Lin D. An information-theoretic definition of similarity. in: *Icml*. 1998, 98(1998): 296-304.

**Examples**

```
library(GSEABase)
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
modulegenes_igraph <- module_igraph (ceRExp, mRExp)
Sim <- module_group_sim(geneIds(modulegenes_WGCNA), geneIds(modulegenes_igraph))
```

---

module_igraph	<i>module_igraph</i>
---------------	----------------------

---

**Description**

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using igraph package

**Usage**

```
module_igraph(
  ceRExp,
  mRExp = NULL,
  cor.method = "pearson",
  pos.p.value.cutoff = 0.01,
  cluster.method = "greedy",
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)
```

**Arguments**

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
cor.method	The method of calculating correlation selected, including 'pearson' (default), 'kendall', 'spearman'.
pos.p.value.cutoff	The significant p-value cutoff of positive correlation.
cluster.method	The clustering method selected in <b>igraph</b> package, including 'betweenness', 'greedy' (default), 'infomap', 'prop', 'eigen', 'louvain', 'walktrap'.
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

**Value**

GeneSetCollection object: a list of module genes.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

Csardi G, Nepusz T. The igraph software package for complex network research, InterJournal, Complex Systems. 2006:1695.

**Examples**

```
data(BRCASampleData)
modulegenes_igraph <- module_igraph(ceRExp[, seq_len(10)],
  mRExp[, seq_len(10)])
```

---

module\_miRdistribute    *module\_miRdistribute*

---

**Description**

miRNA distribution analysis of sharing miRNAs by the identified miRNA sponge modules

**Usage**

```
module_miRdistribute(share_miRs)
```

**Arguments**

share\_miRs      List object: a list of common miRNAs of each miRNA sponge module generated by share\_miRs function.

**Value**

Matrix object: miRNA distribution in each miRNA sponge module.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**Examples**

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
  modulegenes_WGCNA, method = "SRVC",
  SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM_WGCNA_share_miRs <- share_miRs(miRExp, miRTarget, miRSM_WGCNA_SRVC_genes)
```

```
miRSM_WGCNA_miRdistribute <- module_miRdistribute(miRSM_WGCNA_share_miRs)
```

---

module_miRsponge	<i>module_miRsponge</i>
------------------	-------------------------

---

## Description

Extract miRNA sponge interactions of each miRNA sponge module

## Usage

```
module_miRsponge(ModuleList)
```

## Arguments

`ModuleList` List object: a list of the identified miRNA sponge modules.

## Value

List object: miRNA sponge interactions of each miRNA sponge module.

## Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## Examples

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
                        modulegenes_WGCNA, method = "SRVC",
                        SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM_WGCNA_miRsponge <- module_miRsponge(miRSM_WGCNA_SRVC_genes)
```



---

`module_NMF`*module\_NMF*

---

## Description

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using NMF package

## Usage

```
module_NMF(  
  ceExp,  
  mRExp = NULL,  
  NMF.algorithm = "brunet",  
  num.modules = 10,  
  num.ModuleceRs = 2,  
  num.ModulemRs = 2  
)
```

## Arguments

<code>ceExp</code>	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
<code>mRExp</code>	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
<code>NMF.algorithm</code>	Specification of the NMF algorithm, including 'brunet' (default), 'Frobenius', 'KL', 'lee', 'nsNMF', 'offset', 'siNMF', 'snmf/l', 'snmf/r'.
<code>num.modules</code>	The number of modules to be identified.
<code>num.ModuleceRs</code>	The minimum number of ceRNAs in each module.
<code>num.ModulemRs</code>	The minimum number of mRNAs in each module.

## Value

GeneSetCollection object: a list of module genes.

## Author(s)

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## References

Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010, 11:367.



**Examples**

```

data(BRCASampleData)
# Reimport NMF package to avoid conflicts with DelayedArray package
library(NMF)
modulegenes_NMF <- module_NMF(ceRExp[, seq_len(10)],
  mRExp[, seq_len(10)])

```

---

module\_ProNet

*module\_ProNet*


---

**Description**

Identification of gene modules from matched ceRNA and mRNA expression data or single gene expression data using ProNet package

**Usage**

```

module_ProNet(
  ceRExp,
  mRExp = NULL,
  cor.method = "pearson",
  pos.p.value.cutoff = 0.01,
  cluster.method = "MCL",
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)

```

**Arguments**

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
cor.method	The method of calculating correlation selected, including 'pearson' (default), 'kendall', 'spearman'.
pos.p.value.cutoff	The significant p-value cutoff of positive correlation
cluster.method	The clustering method selected in <b>ProNet</b> package, including 'FN', 'MCL' (default), 'LINKCOMM', 'MCODE'.
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

**Value**

GeneSetCollection object: a list of module genes.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**References**

Clauset A, Newman ME, Moore C. Finding community structure in very large networks. Phys Rev E Stat Nonlin Soft Matter Phys., 2004, 70(6 Pt 2):066111.

Enright AJ, Van Dongen S, Ouzounis CA. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res., 2002, 30(7):1575-84.

Kalinka AT, Tomancak P. linkcomm: an R package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. Bioinformatics, 2011, 27(14):2011-2.

Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 2003, 4:2.

**Examples**

```
data(BRCASampleData)
modulegenes_ProNet <- module_ProNet(ceRExp[, seq_len(10)],
  mRExp[, seq_len(10)])
```

---

module_Validate	<i>module_Validate</i>
-----------------	------------------------

---

**Description**

Validation of miRNA sponge interactions in each miRNA sponge module

**Usage**

```
module_Validate(Modulelist, Groundtruth)
```

**Arguments**

Modulelist	List object: a list of the identified miRNA sponge modules.
Groundtruth	Matrix object: a list of experimentally validated miRNA sponge interactions.

**Value**

List object: a list of validated miRNA sponge interactions in each miRNA sponge module

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**Examples**

```

data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
                          modulegenes_WGCNA, method = "SRVC",
                          SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
Groundtruthcsv <- system.file("extdata", "Groundtruth_high.csv", package="miRSM")
Groundtruth <- read.csv(Groundtruthcsv, header=TRUE, sep=",")
miRSM.Validate <- module_Validate(miRSM_WGCNA_SRVC_genes, Groundtruth)

```

---

module\_WGCNA

*module\_WGCNA*


---

**Description**

Identification of co-expressed gene modules from matched ceRNA and mRNA expression data or single gene expression data using WGCNA package

**Usage**

```

module_WGCNA(
  ceRExp,
  mRExp = NULL,
  RsquaredCut = 0.9,
  num.ModuleceRs = 2,
  num.ModulemRs = 2
)

```

**Arguments**

ceRExp	A SummarizedExperiment object. ceRNA expression data: rows are samples and columns are ceRNAs.
mRExp	NULL (default) or a SummarizedExperiment object. mRNA expression data: rows are samples and columns are mRNAs.
RsquaredCut	Desired minimum scale free topology fitting index $R^2$ with interval [0 1].
num.ModuleceRs	The minimum number of ceRNAs in each module.
num.ModulemRs	The minimum number of mRNAs in each module.

**Value**

GeneSetCollection object: a list of module genes.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

## References

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008, 9:559.#'

## Examples

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp[, seq_len(80)],
  mRExp[, seq_len(80)])
```

---

mRExp	<i>mRNA expression data</i>
-------	-----------------------------

---

## Description

mRNA expression data

## Format

mRExp: A SummarizedExperiment object with 72 BRCA and 72 normal samples (rows) and 226 miRNAs (columns).

## Details

The matched breast invasive carcinoma (BRCA) miRNA, lncRNA and mRNA expression data is obtained from TCGA (<http://cancergenome.nih.gov/>). The data focuses on 72 individuals for which the complete sets of tumor and matched normal (i.e., normal tissue taken from the same patient) profiles are available. A mRNA which has missing values in more than 10 are imputed using the k-nearest neighbours (KNN) algorithm from the impute R package. We use the limma R package to infer differentially expressed mRNAs between tumour and normal samples. After the analysis, we select top 500 mRNAs which are differentially expressed at a significant level (adjusted p-value < 1E-02, adjusted by Benjamini & Hochberg method).

---

share_miRs	<i>share_miRs</i>
------------	-------------------

---

## Description

Extract common miRNAs of each miRNA sponge module

## Usage

```
share_miRs(miRExp = NULL, miRTarget, Modulelist)
```

**Arguments**

miRExp	NULL (default) or a SummarizedExperiment object. miRNA expression data: rows are samples and columns are miRNAs.
miRTarget	A SummarizedExperiment object. Putative miRNA-target binding information.
Modulelist	List object: a list of the identified miRNA sponge modules.

**Value**

List object: a list of common miRNAs of each miRNA sponge module.

**Author(s)**

Junpeng Zhang (<https://www.researchgate.net/profile/Junpeng-Zhang-2>)

**Examples**

```
data(BRCASampleData)
modulegenes_WGCNA <- module_WGCNA(ceRExp, mRExp)
# Identify miRNA sponge modules using sensitivity RV coefficient (SRVC)
miRSM_WGCNA_SRVC <- miRSM(miRExp, ceRExp, mRExp, miRTarget,
                          modulegenes_WGCNA, method = "SRVC",
                          SMC.cutoff = 0.01, RV_method = "RV")
miRSM_WGCNA_SRVC_genes <- miRSM_WGCNA_SRVC[[2]]
miRSM_WGCNA_share_miRs <- share_miRs(miRExp, miRTarget, miRSM_WGCNA_SRVC_genes)
```

# Index

BRCA\_genes, 2

ceRExp, 3

cor\_binary, 4

diff\_module, 5

miRExp, 6

miRSM, 6

miRSM\_SS, 9

miRTarget, 10

module\_biclust, 10

module\_CEA, 12

module\_clust, 13

module\_Coexpress, 15

module\_FA, 16

module\_GFA, 18

module\_group\_sim, 19

module\_igraph, 20

module\_miRdistribute, 21

module\_miR sponge, 22

module\_miRtarget, 23

module\_NMF, 24

module\_ProNet, 25

module\_Validate, 26

module\_WGCNA, 27

mRExp, 28

share\_miRs, 28