Package 'SeqMADE'

January 20, 2025

Type Package

Title Network Module-Based Model in the Differential Expression
Analysis for RNA-Seq

Version 1.0

Date 2016-06-27

Author Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

Maintainer Mingli Lei<leimingli2013@sjtu.edu.cn>

Description A network module-
based generalized linear model for differential expression analysis with the count-
based sequence data from RNA-Seq.

License GPL (>= 2)

Depends R (>= 2.1.0),MASS

Repository CRAN

LazyData TRUE

NeedsCompilation no

Date/Publication 2016-06-27 06:59:54

Contents

SeqMADE–package	*Network Module-Based Model in the Differential Expression Analysis for RNA-Seq*

Description

A network module-based generalized linear model for differential expression analysis with the count-based sequence data from RNA-Seq.

Details

Package:	SeqMADE
Type:	Package
Version:	1.0
Date:	2016-06-27
License:	GPL (>2)
LazyLoad:	yes

The main functions in this package are [Factor](#) A function of constructing the Group variables, Direction variables, and the Count variables, [moduleMatrix](#) a function of constructing the modulematrix for all the modules, [nbGLM](#) Identify differential expression modules based on the GLM method using Group and Module variables, [nbGLMdir](#) Identify differential expression modules based on the Generalized Linear Model(GLM) using Group, Module and Direction variables, and [nbGLMdirperm](#) Identify differential expression modules based on the GLM method by shuffling the phenotypic variables.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li Maintainer: Mingli Lei<leimingli2013@sjtu.edu.cn>

References

Xu, J., Wang, L. and Li, J. (2014) Biological network module-based model for the analysis of differential expression in shotgun proteomics, J Proteome Res, 13, 5743-5750.

See Also

glm(),lm()

Examples

```
data(exprs)
data(networkModule)
case <- c("A1","A2","A3","A4","A5","A6","A7")
control <- c("B1","B2","B3","B4","B5","B6","B7")
factors <- Factor(exprs,case,control)
```

```
modulematrix <- moduleMatrix(exprs,networkModule)
Result1<- nbGLM(factors,14,networkModule,modulematrix,distribution="NB")
Result2<- nbGLMdir(factors,14,networkModule,modulematrix,distribution="NB")
Result3<- nbGLMdirperm(exprs,case,control,factors,networkModule,
                       modulematrix,10,distribution="NB")
```

---

exprs	*Gene Expression Dataset*

Description

Gene expression dataset, containing 100 genes and 14 samples(7 case and 7 control respectively).

Usage

```
data(exprs)
```

Details

In this dataset, there are 100 genes and 14 samples which consist of the expression dataset, in which 7 samples are in case groups and other 7 samples are in control groups.

Author(s)

Mingli Lei

Examples

```
data(exprs)
```

Factor	*Construction of Variable Factors*

Description

A function of constructing the Group variables, Direction variables, and the Count variables.

Usage

```
Factor(exprs, case, control)
```

Arguments

exprs	exprs is a data frame or matrix for two groups or conditions, with rows as variables (genes) and columns as samples.
case	case is the sample names in case groups.
control	control is the sample names in control groups.

Details

Two indicator variables Group and Direction corresponding to the different groups and the direction of the gene expression changes in the context of an RNA-Seq experiment, respectively. And in this part, 1 represents that a gene belongs to case group or up-regulated and 0 represents a gene belongs to control group or down-regulated. Besides, Count variables are the expression value in different samples for genes.

Value

Count	The gene expression count values.
Group	The indicator variables represent that whether a gene belongs to case group or not.
Direction	The indicator variables represent that a gene is up-regulated or down-regulated.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

Examples

```
data(exprs)
case <- c("A1","A2","A3","A4","A5","A6","A7")
control <- c("B1","B2","B3","B4","B5","B6","B7")
factors <- Factor(exprs, case, control)
```

---

moduleMatrix   *Modulematrix Construction*

Description

A function of constructing the modulematrix for the modules was used to indicate whether genes belong to a given module or not.

Usage

```
moduleMatrix(exprs, networkModule)
```

Arguments

exprs	exprs is a data frame or matrix for two groups or conditions, with rows as variables (genes) and columns as samples
networkModule	NetworkModule is the gene sets or modules in the biological network or metabolic pathway, with the 1th column as the module names and the 2th columnn as the gene symbols constituting the module

Details

Modulematrix is a matrix, in which the indicator variables 1 or 0 represent whether a gene belong to a given module or not.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

Examples

```
data(exprs)
data(networkModule)
modulematrix <- moduleMatrix(exprs,networkModule)
```

---

nbGLM	*Identify Differential Expression Modules Based on the Generalized Linear Model*

Description

The algorithm identify differential expression modules using Generalized Linear Model (GLM) for differential expression analysis in RNA-Seq data, and in the model two indicator variables Group and Module are adopted to fit the GLM.

Usage

```
nbGLM(factors, N, networkModule, modulematrix, distribution = c("poisson", "NB")[1])
```

Arguments

factors	Factors with three variables including Count, Group, Direction.
N	The total sample sizes.
networkModule	NetworkModule is the gene sets or modules in the biological network or metabolic pathway, with the 1th column as the module names and the 2th column as the gene symbol constituting the module.
modulematrix	Modulematrix is a matrix, in which the indicator variables 1 or 0 represent whether a gene belong to a given module or not.
distribution	a character string indicating the distribution of RNA-Seq count value, default is 'NB'.

Details

The GLM method was determined by the distribution of RNA-Seq count value including Poisson and Negative Binomial distribution, and there are two indicator variables Group and Module, Module=1 when a gene belongs to the module and Module= 0 otherwise; Group=1 for case values and Group=0 for control values. Group * Module represents the interaction effects between Group and Module, and the significance of a module is decided by the interaction and adjusted p-values are calculated to correct for multiple testing.

Value

The nominal pvalue and FDR for the significance of each gene set or module.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

See Also

glm()

Examples

```
data(exprs)
data(networkModule)
case <- c("A1","A2","A3","A4","A5","A6","A7")
control <- c("B1","B2","B3","B4","B5","B6","B7")
factors <- Factor(exprs, case, control)
modulematrix <- moduleMatrix(exprs,networkModule)
Result <- nbGLM(factors, 14, networkModule, modulematrix, distribution = "NB")
```

---

nbGLMdir | *Identify Differential Expression Modules Based on the GLM Model with Up or Down-regulated Change*

Description

The algorithm identify differential expression modules using Generalized Linear Model (GLM) for differential expression analysis in RNA-Seq data, and in the model three indicator variables Group, Module and Direction are adopted to fit the GLM.

Usage

```
nbGLMdir(factors, N, networkModule, modulematrix, distribution = c("poisson", "NB")[1])
```

Arguments

factors	Factors with three variables including Count, Group, Direction.
N	The total sample size.
networkModule	NetworkModule is the gene sets or modules in the biological network or metabolic pathway, with the 1th column as the module names and the 2th columnn as the gene symbol constituting the module.
modulematrix	Modulematrix is a matrix, in which the indicator variables 1 or 0 represent whether a gene belong to a given module or not.
distribution	a character string indicating the distribution of RNA-Seq count value, default is 'NB'.

Details

The GLM method was determined by the distribution of RNA-Seq count value, such as poisson or negative binomial, and there are three indicator variables Group, Module and Direction. Module=1 when a gene belongs to the module and Module= 0 otherwise; Group=1 for case values and Group=0 for control values; Direction=1 for up-regulated and Direction=-1 for down-regualted. Group * Module * Direction represents the interaction effects between Group, Module and Direction.

Value

The nominal pvalue and FDR for the significance of each gene set or module.

Author(s)

Mingli Lei, Li-Ching Huang

See Also

glm()

Examples

```
data(exprs)
data(networkModule)
case <- c("A1","A2","A3","A4","A5","A6","A7")
control <- c("B1","B2","B3","B4","B5","B6","B7")
factors <- Factor(exprs, case, control)
modulematrix <- moduleMatrix(exprs,networkModule)
Result <- nbGLMdir(factors, 14, networkModule, modulematrix,distribution="NB")
```

---

nbGLMdirperm | *Identify Differential Expression Modules Based on the GLM Method by Shuffling the Phenotypic Variables*

Description

Identify differential expression modules based on the Generalized Linear Model(GLM), including Group, Module and Direction variables, then generate the empirical null distribution for the statistic z-values and calculate a empirical estimate of p-value of each module in the permutation null distribution by shuffling the phenotypic variables.

Usage

```
nbGLMdirperm(exprs, case, control, factors,
             networkModule, modulematrix, N,
 distribution = c("poisson", "NB")[1])
```

Arguments

`exprs`	exprs is a data frame or matrix for two groups or conditions, with rows as variables (genes) and columns as samples.
`case`	case is the sample names in case groups.
`control`	control is the sample names in control groups.
`factors`	Factors with three variables including Count, Group, Direction.
`networkModule`	NetworkModule is the gene sets or modules in the biological network or metabolic pathway, with the 1th column as the module names and the 2th columnn as the gene symbol constituting the module.
`modulematrix`	Modulematrix is a matrix, in which the indicator variables 1 or 0 represent whether a gene belong to a given module or not.
`N`	permutation times. If N>0, the permutation step will be implemented. The default value for N is 0.
`distribution`	a character string indicating the distribution of RNA-Seq count value, default is 'NB'.

Details

The GLM method was determined by the distribution of RNA-Seq count value including poisson and Negative Binomial distribution, and there are three indicator variables Group, Module and Direction, in which Module=1 when a gene belongs to the module and Module= 0 otherwise; Group=1 for case values and Group=0 for control values;Direction=1 for up-regulated and Direction=-1 for down-regualted. We therefore construct the contrast vector to test the null hypothesis by fitting the GLM and then focus on the interaction term Group*Module*Direction. Then the samples between the two conditions will be disturbed and by shuffling the phenotypic variables, we can generate the empirical null distribution for each module. Repeat the above process for N times. Pool all the z score together to form a null distribution of z-value. The corresponding statistical significance (p-value) is estimated against null statistics.

Value

The matrix for the sigificance of each module in differential expression analysis.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

See Also

glm()

Examples

```
data(exprs)
data(networkModule)
case <- c("A1","A2","A3","A4","A5","A6","A7")
control <- c("B1","B2","B3","B4","B5","B6","B7")
```

```
factors <- Factor(exprs, case, control)
modulematrix <- moduleMatrix(exprs,networkModule)
result <- nbGLMdirperm(exprs,case,control,factors,
                       networkModule, modulematrix,
   5, distribution="NB")
```

---

networkModule   *NetworkModule*

Description

Different gene sets or modules in the biological network or metabolic pathway.

Usage

```
data(networkModule)
```

Details

In this networkModule, there are five modules consist of different genes.

Author(s)

Mingli Lei, Jia Xu, Li-Ching Huang, Lily Wang, Jing Li

Examples

```
data(networkModule)
```

Index