Package 'BisRNA'

October 12, 2022

Type Package

Title Analysis of RNA Cytosine-5 Methylation

Version 0.2.2

Author C Legrand

Maintainer Carine Legrand <c.legrand@dkfz.de>

Description Bisulfite-treated RNA non-conversion in a set of samples is
analysed as follows : each sample's non-conversion distribution is
identified to a Poisson distribution. P-values adjusted for multiple
testing are calculated in each sample. Combined non-conversion P-values
and standard errors are calculated on the intersection of the set of
samples. For further details, see C Legrand, F Tuorto, M Hartmann,
R Liebers, D Jakob, M Helm and F Lyko (2017) <doi:10.1101/gr.210666.116>.

Date 2017-12-15

License GPL (>= 2)

LazyData TRUE

Depends R (>= 3.3.2)

Imports MASS, graphics, stats, utils

Suggests IHW, testthat

NeedsCompilation no

Repository CRAN

Date/Publication 2017-12-15 13:33:54 UTC

R topics documented:

BisRNA-package   *Analysis of RNA Cytosine-5 Methylation from bisulfite sequencing experiments*

Description

BisRNA analyses non-conversion of bisulfite-treated RNA of a set of samples, in order to distinguish actual cytosine-5 methylation marks from artifacts. Calculations are done as follows:

- each sample's ratio (Poisson parameter / coverage) is determined, assuming a null Poisson distribution (`RNAmeth.poisson.par`),

- non-conversion p-values are calculated for each sample (`RNAmeth.poisson.test`),

- combined non-conversion p-values and standard error of the non-conversion ratios are calculated on the intersection of the set of samples `samples.combine`.

A small non-conversion p-value points to methylation of a RNA cytosine, or another event blocking bisulfite conversion.

Examples

```
## Load data
data(Bisdata,package="BisRNA")

## Ratio (Poisson parameter / coverage), derived p-values,
#  adjusted for multiple testing using either BH (here), or IHW if available.
lambda1 <- RNAmeth.poisson.par(Bisdata1)$estimate
BisXP1  <- RNAmeth.poisson.test(Bisdata1,lambda1,method="BH")
#
lambda2 <- RNAmeth.poisson.par(Bisdata2)$estimate
BisXP2  <- RNAmeth.poisson.test(Bisdata2,lambda2,method="BH")
#
lambda3 <- RNAmeth.poisson.par(Bisdata3)$estimate
BisXP3  <- RNAmeth.poisson.test(Bisdata3,lambda3,method="BH")

## Combined p-values ; median and standard error of
## bisulfite non-conversion ratio
BisXP.combined <- samples.combine(BisXP1,BisXP2,BisXP3)
```

Bisdata	*Three synthetic samples of RNA bisulfite sequencing for examples and tests*

Description

A dataset containing the RNA names, C position, coverage and non-conversion ratio for three synthetic samples of a RNA bisulfite sequencing experiment.

Usage

```
data(Bisdata)
```

Format

Three data frames with 330 to 345 rows and the 4 following variables:

RNA RNA name or identifier

Cpos Cytosine position in RNA

coverage Number of available reads at Cpos

ncratio Bisulfite non-conversion ratio at Cpos

Examples

```
data(Bisdata)
ls(pattern="Bisdata")
head(Bisdata1)
```

Bisdata1	*Synthetic sample 1 of RNA bisulfite sequencing*

Description

A dataset containing the RNA names, C position, coverage and non-conversion ratio for synthetic sample 1 of a RNA bisulfite sequencing experiment.

Usage

```
data(Bisdata)
```

Format

Data frame with 330 to 345 rows and the 4 following variables:

RNA RNA name or identifier

Cpos Cytosine position in RNA

coverage Number of available reads at Cpos

ncratio Bisulfite non-conversion ratio at Cpos

Examples

```
data(Bisdata)
ls(pattern="Bisdata")
head(Bisdata1)
```

Bisdata2   *Synthetic sample 2 of RNA bisulfite sequencing*

Description

A dataset containing the RNA names, C position, coverage and non-conversion ratio for synthetic sample 2 of a RNA bisulfite sequencing experiment.

Usage

```
data(Bisdata)
```

Format

Data frame with 330 to 345 rows and the 4 following variables:

RNA RNA name or identifier

Cpos Cytosine position in RNA

coverage Number of available reads at Cpos

ncratio Bisulfite non-conversion ratio at Cpos

Examples

```
data(Bisdata)
ls(pattern="Bisdata")
head(Bisdata2)
```

Bisdata3 *Synthetic sample 3 of RNA bisulfite sequencing*

Description

A dataset containing the RNA names, C position, coverage and non-conversion ratio for synthetic sample 3 of a RNA bisulfite sequencing experiment.

Usage

```
data(Bisdata)
```

Format

Data frame with 330 to 345 rows and the 4 following variables:

RNA RNA name or identifier

Cpos Cytosine position in RNA

coverage Number of available reads at Cpos

ncratio Bisulfite non-conversion ratio at Cpos

Examples

```
data(Bisdata)
ls(pattern="Bisdata")
head(Bisdata3)
```

---

class.BisXP *Cast bisulfite experiment data into a BisXP object*

Description

`class.BisXP` creates a BisXP object from a table containing RNA name, C position, bisulfite non-conversion ratio and adjusted p-value.

Usage

```
class.BisXP(BisData)
```

Arguments

BisData
A data frame with 4 columns:
- RNA name
- C position (integer, in [1,+Inf])
- bisulfite non-conversion ratio (numeric, in [0,1])
- pvalue.adj (numeric, in [0,1])

Details

This function takes bisulfite experiment data as an input table, performs checks, and casts the table into a BisXP object, with rows labelled after a RNA_C.position pattern.

Value

If the input table is correct, then the output will be a BisXP object corresponding to a consolidated data frame with rows labelled after a RNA_C.position pattern.

Examples

```
RNA     <- c("NM_00001","NM_00001","NM_00002")
Cpos    <- as.integer(c(1,5,1))
ncratio <- c(0.1,0.5,0.3)
pv.adj  <- c(0.001,0.1,0.3)
BSdata  <- data.frame(RNA, Cpos, ncratio, pv.adj, stringsAsFactors = FALSE)
bsXP    <- class.BisXP(BSdata)
```

---

fisher.method	*Function implementing Fisher's method to combine independent p-values*

Description

`fisher.method` takes a list of independent p-values and combines them using Fisher's method.

Usage

```
fisher.method(pvalues)
```

Arguments

pvalues   A list of p-values

Details

This function takes a list of p-values as input, determines the number of degrees of freedom (2 * number of p-values), combines the p-values using Fisher's method and returns the combined p-value in output.

Value

The p-value combined using Fisher's method.

References

Fisher RA (1925) Statistical Methods for Research Workers. Edinburg: Oliver and Boyd.

Fisher RA (1948) Questions and Answers #14. In: Mosteller F, Fisher RA (1948) The American Statistician, 2:30-31 http://www.jstor.org/stable/2681650

Examples

```
list.pv    <- c(0.0001,0.0142,0.0150)
p.combined <- fisher.method(list.pv)
```

---

intersectMatrix   *Take intersection of 2 tables*

Description

intersectMatrix takes 2 matrices and outputs their intersection based on common row.names.

Usage

```
intersectMatrix(Tab1, Tab2)
```

Arguments

Tab1	A matrix or data frame with defined row.names
Tab2	A matrix or data frame with defined row.names

Details

This function takes 2 matrices as input, determines the intersection of their row names, and returns a single matrix containing the rows in the intersection and concatenated columns of the initial matrices.

Value

A matrix with rows common to both Tab1 and Tab2, and concatenated columns.

---

read.BisXP   *Read RNA bisulfite experiment data and cast it into a BisXP object*

Description

read.BisXP reads a table containing RNA name, C position, bisulfite non-conversion ratio and adjusted p-value from a file and casts it into a BisXP object.

Usage

```
read.BisXP(filename)
```

Arguments

filename   Address of the file containing data from a bisulfite experiment in 4 columns
                  separated by a tabulation, with header on the first line:

- RNA name
- C position (integer, in [1,+Inf])
- bisulfite non-conversion ratio (numeric, in [0,1])
- pvalue.adj (numeric, in [0,1])

Details

This function takes a file name as input, reads the bisulfite data table contained in this file, performs
checks, and casts the data into a BisXP object, which contain input data with rows labelled after a
RNA_C.position pattern.

Value

If the input table is correct, then the output will be a BisXP object corresponding to a consolidated
data frame with rows labelled after a RNA_C.position pattern.

Examples

```
RNAs    <- c("NM_00001","NM_00001","NM_00002")
Cpos    <- c(1,5,1)
ncratio <- c(0.1,0.5,0.3)
pv.adj  <- c(0.001,0.1,0.3)
BSdata <- data.frame(RNAs, Cpos, ncratio, pv.adj)
```

RNAmeth.poisson.par   *Find ratio (Poisson parameter / coverage) for one sample of bisulfite-
                            converted RNA*

Description

RNAmeth.poisson.par determines the ratio (Poisson parameter / coverage), based on the assump-
tion that bisulfite-treated RNA's non-conversion ratios follow a Poisson distribution.

Usage

```
RNAmeth.poisson.par(BSrna)
```

Arguments

BSrna   A dataframe containing RNA name, C position, coverage and non-conversion
                  ratio, in this order.

Details

This function takes RNA bisulfite sequencing data from one sample as input, restricts the dataset to RNAs for which coverage is at least 10, and divides data into coverage bins. Then, the Poisson parameter is estimated at each coverage bin using fitdistr from package MASS. The ratio (Poisson parameter / coverage) (median and 95% confidence interval) is returned in output.

Value

This function returns a list containing:

estimate The ratio (Poisson parameter / coverage) (median taken over coverage bins)

bca.ci A confidence interval for the ratio (Poisson parameter / coverage) (bootstrap confidence interval of the median, type "bca")

Examples

```
## Load data, find out ratio (Poisson parameter / coverage).
data(Bisdata,package="BisRNA")
lambda1 <- RNAmeth.poisson.par(Bisdata1)$estimate

## P-values adjusted for multiple testing, using either BH (here) or IHW method.
BisXP1  <- RNAmeth.poisson.test(Bisdata1,lambda1,method="BH")

## Print BisXP1
BisXP1

## Display BisXP1 as a data frame
BisXP1.df <- data.frame(BisXP1$nonconv.ratio, BisXP1$pv.adj, row.names=BisXP1$RNA.pos)
BisXP1.df
```

RNAmeth.poisson.test   *Test RNA Cytosine methylation based on Poisson model distribution*

Description

RNAmeth.poisson.test tests RNA bisulfite sequencing non-conversion based on a sample and the ratio (Poisson parameter / coverage) for this sample.

Usage

```
RNAmeth.poisson.test(BisRNA, lambda, method="BH")
```

Arguments

`BisRNA`	A dataframe containing RNA name, C position, coverage and non-conversion ratio
`lambda`	Ratio (Poisson parameter / coverage)
`method`	Adjustment method for multiple testing, either "BH" (Benjamini-Hochberg) or "IHW" (Independent Hypothesis Weighting, from R package IHW)

Details

This function takes RNA bisulfite sequencing data from one sample and the ratio (Poisson parameter / coverage) as inputs. Then, the dataset is restricted to those C positions where non-conversion ratio is larger than (Poisson parameter / coverage). Finally, it carries out poisson.test and a correction for multiple testing. The output of function RNAmeth.poisson.test is a BisXP object which contains RNA names, C positions, non-conversion ratios and adjusted p-values. The formatting into a BisXP object guarantees that some consistency checks are satisfied.

Value

This function returns a BisXP object whose elements correspond to RNAs where non-conversion ratio is higher than lambda = ratio (Poisson parameter / coverage). The variables contained in this object are the non-conversion ratio and the adjusted p-value.

Examples

```
## Load data, find out ratio Poison parameter / coverage,
## and produce corresponding adjusted p-values.
data(Bisdata,package="BisRNA")
lambda1 <- RNAmeth.poisson.par(Bisdata1)$estimate
BisXP1  <- RNAmeth.poisson.test(Bisdata1,lambda1,method="BH")

## Print BisXP1
BisXP1

## Display BisXP1 as a data frame
BisXP1.df <- data.frame(BisXP1$nonconv.ratio, BisXP1$pv.adj, row.names=BisXP1$RNA.pos)
BisXP1.df
```

---

samples.combine	*Combine samples p-values and ratios*

Description

`samples.combine` takes RNA bisulfite sequencing samples as input, combines their p-values using Fisher's method, and produces median and standard error of the bisulfite non-conversion ratios.

Usage

```
samples.combine(BisXP1, ...)
```

Arguments

BisXP1	A BisXP object containing non-conversion ratio and p-value
...	One or more additional samples, in the form of BisXP objects

Details

This function takes several bisulfite sequencing samples, in form of BisXP objects, as inputs. It is recommended to provide at least 3 samples and in any case all available, relevant samples. Using RNA and C positions present in all samples, the adjusted p-values of each sample are combined using Fisher's method. Median and standard error of the non-conversion ratio are also given in output.

Value

This function returns a data frame whose row names correspond to the RNA and C position which are present in all samples, and the following variables:

p.adj.combined p-value adjusted (done in the preparation of the BisXP object) and combined (done here)

nonconv.ratio.med Median of bisulfite non-conversion ratio for a specific RNA and C positions

nonconv.ratio.se Standard error of bisulfite non-conversion ratio for a specific RNA and C positions

References

Fisher RA (1925) Statistical Methods for Research Workers. Edinburg: Oliver and Boyd.

Fisher RA (1948) Questions and Answers #14. In: Mosteller F, Fisher RA (1948) The American Statistician, 2:30-31 http://www.jstor.org/stable/2681650

Examples

```
## Load data
data(Bisdata,package="BisRNA")

## Obtain the ratio (Poisson parameter / coverage), and p-values
## adjusted for multiple testing using BH (here) or IHW method.
lambda1 <- RNAmeth.poisson.par(Bisdata1)$estimate
BisXP1  <- RNAmeth.poisson.test(Bisdata1,lambda1,method="BH")
lambda2 <- RNAmeth.poisson.par(Bisdata2)$estimate
BisXP2  <- RNAmeth.poisson.test(Bisdata2,lambda2,method="BH")
lambda3 <- RNAmeth.poisson.par(Bisdata3)$estimate
BisXP3  <- RNAmeth.poisson.test(Bisdata3,lambda3,method="BH")

## Combine samples
```

```
BisXP.combined <- samples.combine(BisXP1,BisXP2,BisXP3)
```

---

testMeth   *Apply poisson.test to BS coverage and non-conversion ratio.*

Description

testMeth applies poisson.test to a sample from bisulfite-treated RNA experiment taking into account the ratio (Poisson parameter / coverage) characteristic of this sample.

Usage

```
testMeth(X, lambda)
```

Arguments

X   A list containing coverage as 1st element and non-conversion ratio as 2nd element, for one C position.

lambda   ratio (Poisson parameter / coverage)

Details

This function takes RNA bisulfite sequencing coverage and non-conversion ratio, applies poisson.test and outputs the p-value.

Value

This function returns the p-value from poisson.test.

Index