

Package ‘lncDIFF’

January 13, 2020

Title Long Non-Coding RNA Differential Expression Analysis

Version 1.0.0

Description We developed an approach to detect differential expression features in long non-coding RNA low counts, using generalized linear model with zero-inflated exponential quasi likelihood ratio test. Methods implemented in this package are described in Li (2019) <doi:10.1186/s12864-019-5926-4>.

Depends R (>= 3.5.0)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.0.0

NeedsCompilation no

Author Qian Li [aut, cre]

Maintainer Qian Li <qian.li10000@gmail.com>

Repository CRAN

Date/Publication 2020-01-13 22:20:08 UTC

R topics documented:

cov	2
design	2
hnc.edata	2
lncDIFF	3
LRT	4
tissue	5
ZIQML.fit	5
Index	7

cov	<i>Batch information for samples in hnsk.edata.</i>
-----	---

Description

Batch information for samples in hnsk.edata.

Usage

cov

Format

A matrix of covariate(s) in columns.

design	<i>Design matrix for samples in hnsk.edata.</i>
--------	---

Description

Design matrix for samples in hnsk.edata.

Usage

design

Format

A model matrix with 80 rows (i.e. samples) and 3 columns of tissue type and batch.

hnsk.edata	<i>lncRNA Fragments Per Killobase per Million (FPKM) in a head and neck squamous cell carcinomas (hnsk) study.</i>
------------	--

Description

lncRNA Fragments Per Killobase per Million (FPKM) in a head and neck squamous cell carcinomas (hnsk) study.

Usage

hnsk.edata

Format

A data frame of lncRNA FPKM with 1000 rows (i.e. genes) and 80 columns (i.e. samples).

lncDIFF

*lncRNA Differential Expression (DE) analysis***Description**

lncDIFF returns DE analysis results based on lncRNA counts and grouping variables.

Usage

```
lncDIFF(
  edata,
  group,
  covariate = NULL,
  link.function = "log",
  CompareGroups = NULL,
  simulated.pvalue = FALSE,
  permutation = 100
)
```

Arguments

edata	Normalized counts matrix with genes in rows and samples in columns.
group	Primary factor of interest in DE analysis, e.g., treatment groups, tissue types, other phenotypes.
covariate	Other variables (or covariates) associated with expression level. Input must be a matrix or data frame with each column being a covariate matching to group
link.function	Link function for the generalized linear model, either 'log' or 'identity', default as 'log'.
CompareGroups	Labels of treatment groups or phenotypes of interest to be compared in DE analysis. Input must be a vector of group labels without duplicates.
simulated.pvalue	If empirical p-values are computed, simulated.pvalue=TRUE. The default is FALSE.
permutation	The number of permutations used in simulating pvalues. The default value is 100.

Value

DE.results	Likelihood ratio test results with test statistics, p-value, FDR, DE genes, group-wise mean expression, fold change (if two groups are compared). If simulated.pvalue=TRUE, test.results also includes simulated p-value and FDR.
full.model.fit	Generalized linear model with zero-inflated Exponential likelihood function, estimating group effect compared to a reference group.

References

Li, Q., Yu, X., Chaudhary, R. et al. 'lncDIFF: a novel quasi-likelihood method for differential expression analysis of non-coding RNA'. *BMC Genomics* (2019) 20: 539.

Examples

```
data('hncsc.edata', 'tissue', 'cov')

# DE analysis comparing two groups (normal vs tumor) for 100 genes
result=lncDIFF(edata=hncsc.edata[1:100,], group=tissue, covariate=cov)

# Recommend at least 50 permutations if simulated.pvalue=TRUE
```

LRT

Likelihood ratio test based on ZIQML.fit()

Description

ZIQML.LRT returns the likelihood ratio test statistics and p-value based on the object returned by ZIQML.fit().

Usage

```
LRT(ZIQML.fit, coef = NULL)
```

Arguments

ZIQML.fit	Object returned by ZIQML.fit()
coef	An integer or vector indicating the coefficient(s) in design matrix to be tested. coef=1 is the intercept (i.e. baseline group effect), and should not be tested.

Value

LRT.stat	Likelihood ratio test statistics.
LRT.pvalue	Likelihood ratio test p-value.

Examples

```

data('hnscc.edata','design')
# 'hnscc.edata' contains FPKM of 1132 lncRNA genes and 80 samples.
# 'design' is the design matrix of tissue type (tumor vs normal).

# Fit GLM by ZIQLM.fit for the first 100 genes
fit.log=ZIQLM.fit(edata=hnscc.edata[1:100,],design.matrix=design)

# Likelihood ratio test to compare tumor vs normal in gene expression level.
LRT.results=LRT(fit.log,coef=2)

```

tissue	<i>Tissue type for samples in hnscc.edata.</i>
--------	--

Description

Tissue type for samples in hnscc.edata.

Usage

```
tissue
```

Format

A character vector of tissue type.

ZIQLM.fit	<i>Group and covariate effects on lncRNA counts by Generalized Linear Model</i>
-----------	---

Description

ZIQLM.fit estimates the group effect on gene expression using zero-inflated exponential quasi likelihood.

Usage

```
ZIQLM.fit(edata, design.matrix, link = "log")
```

Arguments

edata	Normalized counts matrix with genes in rows and samples in columns.
design.matrix	Design matrix for groups and covariates, generated by model.matrix().
link	Link function for the generalized linear model and likelihood function, either 'log' or 'identity'. The default is 'log'.

Value

Estimates	Estimated group effect on gene expression by zero-inflated exponential quasi maximum likelihood (ZIQML) estimator.
logLikelihood	The value of zero-inflated quasi likelihood.
edata	lncRNA counts or expression matrix.
design.matrix	The design matrix of groups and covariates.
link	The specified link function.

Examples

```
data('hnc.edata','design')
# 'hnc.edata' contains FPKM of 1000 lncRNA genes and 80 samples
# 'design' is the design matrix for tissue and batch.

# For the first 100 genes
# Fit GLM by ZIQML with logarithmic link function
fit.log=ZIQML.fit(edata=hnc.edata[1:100,],design.matrix=design,link='log')

# Fit GLM by ZIQML with identity link function
fit.identity=ZIQML.fit(edata=hnc.edata[1:100,],design.matrix=design,link='identity')
```

Index

*Topic **datasets**

cov, [2](#)

design, [2](#)

hncsc.edata, [2](#)

tissue, [5](#)

cov, [2](#)

design, [2](#)

hncsc.edata, [2](#)

lncDIFF, [3](#)

LRT, [4](#)

tissue, [5](#)

ZIQML.fit, [5](#)