Package 'EBSEA'

October 12, 2016

Type Package

Title Exon Based Strategy for Expression Analysis of genes

Version 1.0.0

Date 2015-12-15

Author Arfa Mehmood, Asta Laiho, Laura L. Elo

Maintainer Arfa Mehmood <arfa.mehmood@utu.fi>

Description Calculates differential expression of genes based on exon counts of genes obtained from RNA-seq sequencing data.

License GPL-2

biocViews Software, DifferentialExpression, GeneExpression, Sequencing

Imports edgeR, limma, gtools, graphics, stats

NeedsCompilation no

R topics documented:

EBSEA *Exon Based Startegy for Expression Analysis of genes*

Description

EBSEA takes as input unnormalized counts of exons, normalizes them and then performs a two group comparison of the samples to detect differentially expressed between the groups. Both paired or unpaired comparison are supported. It calculates fold changes, p-values and false discovery rate of the genes between the groups.

Usage

```
EBSEA(countData, group, paired = FALSE, effects = NULL, plot = FALSE)
```

Arguments

countData	A dataframe of exon count data
group	A vector indicating the sample groups in the experiment
paired	A logical indicating whether the samples are paired or unpaired. Default: FALSE
effects	A vector indicating the paired samples.
plot	A logical indicating whether a volcano plot is visualized. Default: FALSE

Value

EBSEA returns a list of two dataframes. ExonTable is a dataframe that contains exon statistics including log fold change, p-values, adjusted p-values, average expression and fold change. GeneTable is a dataframe that contains the corresponding fold change, log fold change, p-values and false discovery rate.

References

Laiho, A., & Elo, L. L. (2014). A note on an exon-based strategy to identify differentially expressed genes in RNA-seq experiments. PloS One, 9(12), e115964.

See Also

visualizeGenes

Examples

```
data(origCounts)
group <- c('Group1', 'Group1', 'Group1', 'Group2', 'Group2', 'Group2', 'Group2')
result <- EBSEA(origCounts, group)
```

---

filterCounts *Filter Count Data*

Description

The exons are filtered based on their expression levels so that each exon has a cpm (count per million) of more than 1 in user defined percent of the samples.

Usage

```
filterCounts(x, noOfSamples)
```

Arguments

x	A numeric dataframe of counts in the sample with gene and exon number as the row names and samples as the column names
noOfSamples	Percentage of the number of samples that should have cpm greater than 1.

Value

A dataframe of filtered counts of exons

See Also

EBSEA

Examples

```
data(origCounts)
res <- filterCounts(origCounts, 20)
```

---

filterGenes *Filter Gene List*

Description

The differentially expressed genes are filtered based on the FC and FDR provided by the user.The default thresholds are FC => 1.25 and fdr <= 0.01

Usage

```
filterGenes(x, fc = 1.25, fdr = 0.01)
```

Arguments

x	A dataframe containing the gene statistics returned by EBSEA.
fc	A fold change threshold for the genes to be filtered. Default: 1.25.
fdr	A FDR threshold for the genes to be filtered. Default: 0.01.

Value

A list containing upregulated and downregulated genes.

See Also

EBSEA

Examples

```
data(origCounts)
group <- c('Group1', 'Group1', 'Group1', 'Group2', 'Group2', 'Group2', 'Group2')
result <- EBSEA(origCounts, group)
filteredGenes <- filterGenes(result$GeneTable)
```

`origCounts` *Subset of Pasilla Dataset*

Description

origCounts consists of a subset of the exon counts from the pasilla dataset.

Usage

```
data("origCounts")
```

Format

A data frame with 1000 observations on the following 7 variables.

`treated1fb` a numeric vector

`treated2fb` a numeric vector

`treated3fb` a numeric vector

`untreated1fb` a numeric vector

`untreated2fb` a numeric vector

`untreated3fb` a numeric vector

`untreated4fb` a numeric vector

Value

Dataset

See Also

EBSEA

Examples

```
data(origCounts)
```

visualizeGenes *Visualize Gene*

Description

Plots for each exon of the gene entered by the user, the mean of the counts and the fold changes.

Usage

```
visualizeGenes(gene, group, countData, result)
```

Arguments

gene	Gene Name. The gene name should be the from the genes in count data.
group	A vector indicating the sample group in the experiment.
countData	A dataframe of the original exon count data.
result	Results returned by EBSEA.

Value

A plot of mean counts and fold changes of exons of a gene.

See Also

EBSEA

Examples

```
data(origCounts)
group <- c('Group1', 'Group1', 'Group1', 'Group2', 'Group2', 'Group2', 'Group2')
result <- EBSEA(origCounts, group)
visualizeGenes('FBgn0000017', group, origCounts, result)
```

Index