Package ‘msgsbsR’

April 26, 2017

Type Package

Title msgbsR: methylation sensitive genotyping by sequencing (MS-GBS)
R functions

Version 1.0.0

Date 2017-04-24

Author Benjamin Mayne

Maintainer Benjamin Mayne <benjamin.mayne@adelaide.edu.au>

Depends R (>= 3.4), GenomicRanges, methods
Imports BSgenome, easyRNASeq, edgeR, GenomicAlignments,
    GenomicFeatures, GenomeInfoDb, ggbio, ggplot2, IRanges,
    parallel, plyr, Rsamtools, R.utils, stats,
    SummarizedExperiment, S4Vectors, utils
Suggests roxygen2, BSgenome.Rnorvegicus.UCSC.rm6

biocViews DifferentialMethylation, DataImport, Epigenetics, MethylSeq
Description Pipeline for the anaysis of a MS-GBS experiment.
License GPL-2

LazyLoad yes

Collate 'msgsbsR.R' 'rawCounts.R' 'checkCuts.R' 'plotCounts.R'
    'diffMeth.R' 'plotCircos.R'

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

    checkCuts ......................................................... 2
    cuts .......................................................... 3
    diffMeth ......................................................... 3
    msgbsR ........................................................ 4
    plotCircos ...................................................... 4
    plotCounts ...................................................... 5
    ratdata ......................................................... 5
    ratdata2 ....................................................... 6
    rawCounts ...................................................... 7

Index 8
checkCuts

**Description**
Determines the sequence around a cut site using a fasta file or BSgenome

**Usage**
```
checkCuts(cutSites, genome, fasta = FALSE, seq)
```

**Arguments**
- `cutSites` A GRanges object containing the locations of the cut sites to be checked for sequence match. The names of the correct cut sites will be returned as a GRanges object.
- `genome` The path to a fasta file or a BSgenome object to check for genomic sequences.
- `fasta` TRUE if a fasta file has been supplied. Default = FALSE
- `seq` The desired recognition sequence that the enzyme should have cut.

**Value**
A GRanges object containing the names of the sites that had the correct sequence.

**Author(s)**
Benjamin Mayne

**Examples**
```r
library(GenomicRanges)
library(SummarizedExperiment)
library(BSgenome.Rnorvegicus.UCSC.rn6)
# Load the positions of possible MspI cut sites
data(ratdata)
# Adjust the cut sites to overlap recognition site on each strand
start(cutSites) <- ifelse(test = strand(cutSites) == '+',
yes = start(cutSites) - 1, no = start(cutSites) - 2)
end(cutSites) <- ifelse(test = strand(cutSites) == '+',
yes = end(cutSites) + 2, no = end(cutSites) + 1)
correctCuts <- checkCuts(cutSites = cutSites, genome = "rn6", seq = "CCGG")
```
A GRanges object of differentially methylated MspI cut sites on chromosome 20 in Rat from a MS-GBS experiment.

Description
The GRanges object was created from a list of differentially methylated cut sites from a MS-GBS experiment between two groups of rats that were fed either a control diet or a high fat diet.

Usage
data(cuts)

Format
A GRanges object of length 10.

Details
• Positions of MspI cut sites differentially methylated in the prostate on chromosome 20 in Rats.
The data set contains 10 differentially methylated sites in the prostate between rats fed a control or high fat diet.

Value
A GRanges object of length 10.

diffMeth
diffMeth

Description
Determines differential methylated sites from a RangedSummarizedExperiment

Usage
diffMeth(se, cateogory, condition1, condition2, block = NULL, cpmThreshold, thresholdSamples)

Arguments
se A RangedSummarizedExperiment containing meta data of the samples.
cateogory The heading name in the sample data to be tested for differential methylation.
condition1 The reference group within the category.
condition2 The experimental group within the category.
block The heading name in the sample data if differential methylation is to be tested with a blocking factor. Default is NULL.
cpmThreshold Counts per million threshold of read counts to be filtered out of the analysis.
thresholdSamples Minimum number of samples to contain the counts per million threshold.
Value
A data frame containing which cut sites that are differentially methylated.

Author(s)
Benjamin Mayne

Examples
# Load data
data(ratdata2)
top <- diffMeth(se = ratdata2, category = "Group",
condition1 = "Control", condition2 = "Experimental",
cpmThreshold = 1, thresholdSamples = 1)

Description
msgbsR
plotCircos

Usage
plotCircos(cutSites, seqlengths, cutSite.colour, seqlengths.colour)

Arguments
cutSites A GRanges object containing the locations of the cut sites to be plotted.
seqlengths An integer with the lengths of the chromosomes.
cutSite.colour The colour of the cut sites.
seqlengths.colour The colour of the chromosomes

Value
A circos plot showing the locations of the cut sites.

Author(s)
Benjamin Mayne
plotCounts

Examples

# load example cut site positions
data(cuts)
# Obtain the length of chromosome 20 in rn6
library(BSgenome.Rnorvegicus.ucsc.rn6)
chr20 <- seqlengths(BSgenome.Rnorvegicus.ucsc.rn6)["chr20"]
plotCircos(cutSites = cuts, seqlengths = chr20,
cutSite.colour = "red", seqlengths.colour = "blue")

plotCounts

Description

Plots the total number of reads vs total number of cut sites per sample

Usage

plotCounts(se, category)

Arguments

se
A RangedSummarizedExperiment containing meta data of the samples.

category
The heading name in the sample data to distinguish groups.

Value

Produces a plot showing the total number reads vs total number of cut sites per sample.

Author(s)

Benjamin Mayne

Examples

data(ratdata2)
plotCounts(se = ratdata2, category = "Group")

ratdata

Read counts of potential MspI cut sites from a MS-GBS experiment of prostates from rats

Description

A RangedSummarizedExperiment containing read counts generated from a MS-GBS experiment using the restriction enzyme MspI, focusing on chromosome 20 of Rat.

Usage

data(ratdata)
ratdata2

Format

RangedSummarizedExperiment

Details

- ratdata A RangedSummarizedExperiment with 16047 potential MspI cut sites on chromosome 20 in Rat and six samples (3 Control and 3 Experimental).

This dataset contains six prostate samples from rats: 3 control and 3 experimental high fat diet.

Value

RangedSummarizedExperiment

ratdata2

Read counts of correct MspI cut sites from a MS-GBS experiment of prostates from rats

Description

A RangedSummarizedExperiment containing read counts generated from a MS-GBS experiment using the restriction enzyme MspI, focusing on chromosome 20 of Rat. The sites have been checked for the correct recognition site.

Usage

data(ratdata2)

Format

RangedSummarizedExperiment

Details

- ratdata2 A RangedSummarizedExperiment containing data for 13983 MspI cut sites on chromosome 20 in Rat and six samples (3 Control and 3 Experimental).

This dataset contains six prostate samples from rats: 3 control and 3 experimental high fat diet. The data can be used for differential methylation analyses.

Value

RangedSummarizedExperiment
**rawCounts**

**Description**
Imports the raw read counts from sorted and indexed bam file(s)

**Usage**

```r
rawCounts(bamFilepath, threads = 1)
```

**Arguments**

- `bamFilepath` The path to the location of the bam file(s).
- `threads` The total number of usable threads to be used. Default is 1.

**Value**

Produces a RangedSummarizedExperiment. Columns are samples and the rows are cut sites. The cut site IDs are in the format chr:position-position:strand.

**Author(s)**
Benjamin Mayne, Sam Buckberry

**Examples**

```r
my_path <- system.file("extdata", package = "msgbsR")
my_data <- rawCounts(bamFilepath = my_path)
```
Index

*Topic datasets
  cuts, 3
  ratdata, 5
  ratdata2, 6

checkCuts, 2
cuts, 3

diffMeth, 3

msgbsR, 4
msgbsR-package (msgbsR), 4

plotCircos, 4
plotCounts, 5

ratdata, 5
ratdata2, 6
rawCounts, 7