Package 'methylKit'

November 29, 2011

Type Package

Title DNA methylation analysis from high-throughput bisulfate sequencing

Version 0.2

Date 2011-11-10

Author Altuna Akalin

Maintainer ME <yourfault@somewhere.net>

Description DNA methylation analysis RRBS,ERRBS,BS-seq

License Artistic License, Version 2.0

LazyLoad yes

Depends R (>= 2.13.0), methods, GenomicRanges, IRanges

Imports methods, GenomicRanges, IRanges

Collate 'backbone.R' 'diffMeth.R' 'annotate.R' 'clusterSamples.R'

R topics documented:

| clusterSamples | *CpG Dinucleotide Methylation Hierarchical Cluster Analysis* |

Description

CpG Dinucleotide Methylation Hierarchical Cluster Analysis

Arguments

.Object   a `methylBase` object

dist   the distance measure to be used. This must be one of `"correlation"`, `"euclidean"`, `"maximum"`, `"manhattan"`, `"canberra"`, `"binary"` or `"minkowski"`. Any unambiguous substring can be given. (default:`"correlation"`)

method   the agglomeration method to be used. This should be (an unambiguous abbreviation of) one of `"ward"`, `"single"`, `"complete"`, `"average"`, `"mcquitty"`, `"median"` or `"centroid"`. (default:`"ward"`)

plot   a logical value indicating whether to plot hierarchical clustering. (default:TRUE)

Value

a `tree` object of a hierarchical cluster analysis using a set of dissimilarities for the n objects being clustered.

| getAssembly | *get assembly of the genome* |

Description

get assembly of the genome

Arguments

x   a methylBase object

Value

the assembly string for the object

---

getCorrelation	*get correlation between samples in methylBase object*

Description

get correlation between samples in methylBase object

Arguments

`.Object`	a methylBase object
`plot`	scatterPlot if TRUE (deafult:False)

Value

a correlation matrix object and plot scatterPlot

---

getCoverageStats	*get coverage stats from methylRaw object*

Description

get coverage stats from methylRaw object

Arguments

`.Object`	a `methylRaw` object
`plot`	plot a histogram of coverage if TRUE (deafult:FALSE)
`both.strands`	do stats and plot for both strands if TRUE (deafult:FALSE)

Value

a summary of coverage statistics or plot a histogram of coverage

---

getData	*gets the data slot from the methylBase object*

Description

The data retrived from this function is of a `data.frame`. This is basically containing all relevant methylation information per region

Arguments

`x`	a methylBase object

Value

data.frame for methylation events

---

getMethylationStats   *get Methylation stats from methylRaw object*

Description

get Methylation stats from methylRaw object

Arguments

.Object   a methylRaw object

plot   plot a histogram of Methylation if TRUE (deafult:FALSE)

both.strands   do plots and stats for both strands seperately if TRUE (deafult:FALSE)

Value

a summary of Methylation statistics or plot a histogram of coverage

---

methylBase-class   *An S4 class that holds base-pair resolution methylation information for multiple experiments, only bases that are covered in all experiments are held in this class*

Description

extends data.frame

Slots

sample.ids character vector for ids of samples in the object

assembly name of the genome assembly

treatment treatment vector denoting which samples are test and control

coverage.index vector denoting which columns in the data correspons to coverage values

numCs.index vector denoting which columns in the data correspons to number of methylatedCs values

numTs.index vector denoting which columns in the data correspons to number of unmethylated Cs values

---

methylRaw-class   *An S4 class for holding raw methylation data from alignment pipeline.*

Description

This object stores the raw mehylation data that is read in through read function and extends data.frame

Slots

sample.id: string for an identifier of the sample

assembly: string for genome assembly, ex: hg18,hg19,mm9

---

methylRawList-class   *An S4 class for holding a list of methylRaw objects.*

Description

This object stores the list of raw mehylation data that is read in through read function and extends data.frame

Slots

treatment: numeric vector denoting control and test samples

PCASamples   *CpG Dinucleotide Methylation Principal Components Analysis*

Description

CpG Dinucleotide Methylation Principal Components Analysis

Arguments

.Object	a methylBase object
cor	a logical value indicating whether the calculation should use the correlation matrix or the covariance matrix. (default: TRUE)
screeplot	a logical value indicating whether to plot the variances against the number of the principal component. (default: FALSE)

Value

The form of the value returned by PCASamples is the summary of principal component analysis by princomp.

---

read   *read file(s) to a methylrawList or methylraw object*

Description

read a list of locations or one location and create a methylrawList or methylraw object

Arguments

file	location(s)
sample.id	sample.id(s)
assembly	a string that defines the genome assembly such as hg18, mm9
header	if the input file has a header or not (default: TRUE)
pipeline	name of the alignment pipeline, currently only supports AMP (default: AMP)
treatment	a vector contatining 0 and 1 denoting which samples are control which samples are test

---

unite   *unites methylRawList to a single table*

Description

This functions unites `methylRawList` object that only bases with coverage from all samples are retained. The resulting object is a class of `methylBase`

Arguments

destrand   if TRUE, reads covering both strands of a CpG dinucleotide will be merged, do not set to TRUE if not only interested in CpGs (default: FALSE)

Value

a methylBase object

Index