

Package ‘GIMP’

November 17, 2024

Title Genomic Imprinting Methylation Package

Version 0.1.0

Author Francesco Cecere <francesco.cecerengs@gmail.com>

Maintainer Francesco Cecere <francesco.cecerengs@gmail.com>

Description A package for analyzing DNA methylation at imprinted loci

Depends R (>= 3.5.0)

License MIT

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Imports tidyverse, valr, reshape2, ggplot2, pheatmap, viridisLite,
ggplotify, readr, shiny, BiocManager

Contents

bed450k	2
bedEPICv1	2
bedEPICv2	3
create_bedmeth	3
DMRs.hg19	4
DMRs.hg38	4
iDMR_heatmap	5
make_cpgs	6
make_ICRs	6
plot_CpG_coverage	7
plot_line_region	8
Index	9

bed450k	<i>BED 450K probes</i>
---------	------------------------

Description

This dataset contains the 450K array probes coordinates.

Usage

```
data(bed450k)
```

Format

BED file.

Examples

```
data(bed450k)  
head(bed450k)
```

bedEPICv1	<i>BED EPICv1 probes</i>
-----------	--------------------------

Description

This dataset contains the EPICv1 probes coordinates.

Usage

```
data(bedEPICv1)
```

Format

BED file.

Examples

```
data(bedEPICv1)  
head(bedEPICv1)
```

bedEPICv2	<i>BED EPICv2 probes</i>
-----------	--------------------------

Description

This dataset contains the EPICv2 probes coordinates.

Usage

```
data(bedEPICv2)
```

Format

BED file.

Examples

```
data(bedEPICv2)
head(bedEPICv2)
```

create_bedmeth	<i>Create BED File Data from Methylation Array Annotations</i>
----------------	--

Description

This function generates a BED-format data frame from Illumina Human Methylation annotation files. The BED data includes chromosome, position, and probe ID information, and supports multiple annotation versions.

Usage

```
create_bedmeth(version = "v1")
```

Arguments

version	A character string specifying the annotation version to use. Options include "v1" for the EPIC version1 and "v2" for EPIC version2. Default is "v1".
---------	--

Value

A data frame in BED format containing columns:

chr	Chromosome name.
pos	Position on the chromosome.
probeID	Unique identifier for each probe.
end	End position, which is the same as 'pos' in this output.

Examples

```
# Create BED-format data with the default version (EPIC v1)
bed_data <- create_bedmeth()
head(bed_data) # View the first few rows

# Use a different annotation version if available
bed_data_v2 <- create_bedmeth(version = "v2")
```

DMRs.hg19*Imprinted Regions*

Description

This dataset contains the Human Imprinted regions coordinates in hg19.

Usage

```
data(DMRs.hg19)
```

Format

A data frame with iDMRs coordinates.

Examples

```
data(DMRs.hg19)
head(DMRs.hg19)
```

DMRs.hg38*Imprinted Regions*

Description

This dataset contains the Human Imprinted regions coordinates in hg38.

Usage

```
data(DMRs.hg38)
```

Format

A data frame with iDMRs coordinates.

Examples

```
data(DMRs.hg38)
head(DMRs.hg38)
```

iDMR_heatmap

*Generate Heatmap of Imprinted DMRs Methylation***Description**

This function generates a heatmap for visualizing methylation data of Imprinted Differentially Methylated Regions (DMRs).

Usage

```
iDMR_heatmap(
  df_ICR,
  sampleInfo,
  control_label = "Control",
  case_label = "Case",
  bedmeth = "v1",
  order_by = "cord",
  annotation_col = NULL,
  plot_type = "beta",
  sd_threshold = 3
)
```

Arguments

df_ICR	A data frame or matrix containing methylation beta values for Imprinted DMRs.
sampleInfo	A vector indicating the group labels (e.g., "Control" and "Case") for each sample in 'df_ICR'. Each element in 'sampleInfo' should correspond to a sample in 'df_ICR'.
control_label	A character string specifying the label for the control group in 'sampleInfo'. Default is "Control".
case_label	A character string specifying the label for the case group in 'sampleInfo'. Default is "Case".
bedmeth	A character string specifying the BED data version for DMR coordinates. Options are "v1", "v2", or "450k". Default is "v1".
order_by	A character string specifying the ordering rows in the heatmap. Options are "cord" for coordinates or "meth" for methylation values. Default is "cord".
annotation_col	A named list of colors for each unique value in 'sampleInfo'. If 'NULL', default colors are assigned using the "viridis" palette. Default is 'NULL'.
plot_type	A character string specifying the type of heatmap to generate. Options are "beta" for beta values, "delta" for values normalized against controls, and "defect" for defect matrix based on standard deviations. Default is "beta".
sd_threshold	A numeric value specifying the standard deviation threshold for detecting defects in the defect matrix. Only used if 'plot_type' is "defect". Default is '3'.

Value

A heatmap plot visualizing methylation of Imprinted DMRs.

Examples

```
# Example sampleInfo with "Case" and "Control" labels for each sample
sampleInfo <- c(rep("Case", 10), rep("Control", 10))
DMR_heatmap(df_ICR = my_ICR_data, sampleInfo = sampleInfo, annotation_col = list(Sample = c("darkgreen", "darkred")))
```

make_cpgs	<i>Create ICR CpG Matrix</i>
-----------	------------------------------

Description

This function generates a CpG matrix for Imprinted Control Regions (ICR) using methylation data. The CpG matrix is constructed based on the provided BED data version.

Usage

```
make_cpgs(Bmatrix, bedmeth = "v1")
```

Arguments

Bmatrix	A data frame or matrix containing methylation beta values. Rows typically represent individual probes or CpGs, and columns represent samples.
bedmeth	A character string specifying the BED data version to use for CpG mapping. Options are "v1" (EPIC v1), "v2" (EPIC v2), or "450k" (450k array). Default is "v1".

Value

A data frame representing the ICR CpG matrix, with rows as CpG probes and columns as samples.

Examples

```
# Generate the ICR CpG matrix with default BED version (EPIC v1)
ICRcpg <- make_cpgs(Bmatrix = df, bedmeth = "v1")

# Use a different BED version, such as EPIC v2
ICRcpg_v2 <- make_cpgs(Bmatrix = df, bedmeth = "v2")
```

make_ICRs	<i>Create the ICR Matrix</i>
-----------	------------------------------

Description

This function generates an ICR (Imprinted Control Region) matrix from a given beta matrix, using specified BED data for CpG mapping. The ICR matrix provides data organized by CpG probes and samples. The coordinates of the Human Imprinted regions are taken from <https://doi.org/10.1080/15592294.2016.1264561>

Usage

```
make_ICRs(Bmatrix, bedmeth = "v1")
```

Arguments

Bmatrix	A data frame or matrix containing methylation beta values. Rows should represent CpG probes, and columns represent samples.
bedmeth	A character string indicating the BED data version to use for CpG mapping. Options are "v1" (EPIC v1), "v2" (EPIC v2), or "450k" (450k array). Default is "v1".

Value

A data frame representing the ICR matrix, structured by CpG probes and samples.

Examples

```
ICRmatrix <- make_ICRs(Bmatrix = df, bedmeth = "v1")
```

plot_CpG_coverage	<i>Plot ICR CpG Matrix with Counts and Percentage Coverage</i>
-------------------	--

Description

This function plots the CpG coverage for Imprinted Control Regions (ICRs) using the provided data frame of CpG counts. It compares CpG counts in the specified BED data version for visual analysis and includes an additional plot for percentage coverage.

Usage

```
plot_CpG_coverage(df_ICR_cpg, bedmeth = "v1")
```

Arguments

df_ICR_cpg	A data frame containing CpG counts for ICR regions. Each row represents a different CpG probe, and columns contain sample-related information.
bedmeth	A character string specifying the BED data version to use for mapping CpG coverage. Options are "v1" (EPIC v1), "v2" (EPIC v2), or "450k" (450k array). Default is "v1".

Value

A list containing two plots (counts and percentage coverage) and the data frame with CpG counts and coverage information.

Examples

```
plot_CpG_coverage(df_ICR_cpg_counts, bedmeth = "v1")
```

plot_line_region	<i>Plot Line Plot for Imprinted DMR Methylations</i>
------------------	--

Description

This function generates a line plot to visualize methylation values across a specified Imprinted Differentially Methylated Region (ICR). Users can choose between a static ‘ggplot2’ plot or an interactive ‘plotly’ plot.

Usage

```
plot_line_region(significantDMPs, ICRcpg, ICR, sampleInfo, interactive = TRUE)
```

Arguments

significantDMPs	A data frame containing information about significant DMPs. Must include columns ‘ICR’, ‘start’, and ‘end’.
ICRcpg	A data frame or matrix containing CpG methylation data. Includes CpG coordinates (‘cstart’) and methylation values.
ICR	A character string specifying the name of the ICR region to be plotted.
sampleInfo	A character vector providing group labels (e.g., “Control” or “Case”) for each sample in the methylation data.
interactive	A logical value indicating whether to return an interactive ‘plotly’ plot (‘TRUE’) or a static ‘ggplot2’ plot (‘FALSE’). Default is ‘TRUE’.

Value

A plot representing the line plot of methylation values across the specified ICR region, highlighting significant DMPs. The plot is either a ‘ggplot2’ object or a ‘plotly’ object, depending on the value of ‘interactive’.

Examples

```
# Example data for significantDMPs
significantDMPs <- data.frame(
  ICR = c("ICR1", "ICR1", "ICR2"),
  start = c(100, 200, 150),
  end = c(150, 250, 200)
)

# Example data for ICRcpg
ICRcpg <- data.frame(
  CpG1 = runif(10, 0, 1), CpG2 = runif(10, 0, 1),
  cstart = seq(100, 190, by = 10)
)

# Example sample info
sampleInfo <- c("Control", "Case")

# Plot methylation for a specific ICR region
plot <- plot_line_region(significantDMPs, ICRcpg, ICR = "ICR1", sampleInfo = sampleInfo, interactive = TRUE)
plot
```


Index

* datasets

- bed450k, [2](#)
- bedEPICv1, [2](#)
- bedEPICv2, [3](#)
- DMRs.hg19, [4](#)
- DMRs.hg38, [4](#)

- bed450k, [2](#)
- bedEPICv1, [2](#)
- bedEPICv2, [3](#)

- create_bedmeth, [3](#)

- DMRs.hg19, [4](#)
- DMRs.hg38, [4](#)

- iDMR_heatmap, [5](#)

- make_cpgs, [6](#)
- make_ICRs, [6](#)

- plot_CpG_coverage, [7](#)
- plot_line_region, [8](#)