Package 'SCATEData'

October 18, 2022

Version 1.6.0

Title Data for SCATE (Single-cell ATAC-seq Signal Extraction and Enhancement)

Description SCATEData is an ExperimentHub package for SCATE which is a software tool for extracting and enhancing the sparse and discrete Single-cell ATAC-seq Signal.

BugReports <https://github.com/Winnie09/SCATEData/issues>

License MIT + file LICENSE

VignetteBuilder knitr

Encoding UTF-8

biocViews ExperimentHub, ExperimentData

Depends ExperimentHub, GenomicAlignments, GenomicRanges, base

Imports utils

Suggests knitr, rmarkdown

git_url https://git.bioconductor.org/packages/SCATEData

git_branch RELEASE_3_15

git_last_commit f906eb9

git_last_commit_date 2022-04-26

Date/Publication 2022-10-18

Author Zhicheng Ji [aut],
Weiqiang Zhou [aut],
Wenpin Hou [cre, aut] (<<https://orcid.org/0000-0003-0972-2192>>),
Hongkai Ji [aut]

Maintainer Wenpin Hou <wp.hou3@gmail.com>

R topics documented:

GSM1596831.bam   *'GSM1596831.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596831.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

`GSM1596840.bam`	*'GSM1596840.bam.Rd' data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596840.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

`GSM1596874.bam`	*'GSM1596874.bam.Rd' data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596874.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

GSM1596881.bam	*'GSM1596881.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596881.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

GSM1596940.bam	*'GSM1596940.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596940.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

GSM1596942.bam	*'GSM1596942.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596942.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

GSM1596944.bam	*'GSM1596944.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596944.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

| GSM1596961.bam | *'GSM1596961.bam.Rd'data* |

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1596961.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

| GSM1597041.bam | *'GSM1597041.bam.Rd'data* |

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1597041.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

GSM1597096.bam	*'GSM1597096.bam.Rd' data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "GSM1597096.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

hg19.rds	*'hg19' dataset*

Description

hg19 annotation file for SCATE

Details

This is a hg19 annotation file for SCATE package.

File sizes: 66 MB.

Source: https://github.com/zji90/SCATE/tree/master/inst/extdata

Examples

```
fl <- system.file("extdata", "hg19.rds", package = "SCATEData")
a <- readRDS(fl)
```

| mm10.rds | *'mm10' dataset* |

Description

mm10 annotation file for SCATE

Details

This is a mm10 annotation file for SCATE package.

File sizes: 58 MB.

Source: https://github.com/zji90/SCATE/tree/master/inst/extdata

Examples

```
fl <- system.file("extdata", "mm10.rds", package = "SCATEData")
a <- readRDS(fl)
```

| SCATEData | *Data package for SCATE to extract and enhance the sparse and discrete Single-cell ATAC-seq Signal* |

Description

Data package for SCATE which is a software for extracting and enhancing the sparse and discrete Single-cell ATAC-seq Signal. This package contains data in the GenomicRanges format that SCATE needs for compiling examples or running the algorithm. It is hosted on Bioconductor ExperimentHub.

Details

Overview

The 'SCATEData' is an extensive data resources containing the location of genomic regions bins, the mean and standard deviation of normalized bulk DNase-seq signals across ENCODE samples, and the clustering of the genomic bins based on bulk DNase-seq signals. These data have been formatted as 'GenomicRanges' Bioconductor objects and hosted on Bioconductor's 'ExperimentHub' platform.

'SCATEData' provides data resources for the package 'SCATE' to extract and enhance the sparse and discrete Single-cell ATAC-seq Signal.

Datasets

The package contains the following datasets, which can be used by 'SCATE' to extract and enhance the sparse Single-cell ATAC-seq signals. They are grouped into datasets useful for (i) running 'SCATE', and (ii) providing examples.

- Loaded by 'SCATE':

- hg19.rds
- mm10.rds

- scATAC-seq example files:

- GSM1596831.bam
- GSM1596940.bam
- GSM1597041.bam
- SRR1779856.bam
- SRR1779973.bam
- GSM1596840.bam
- GSM1596942.bam
- GSM1597096.bam
- SRR1779874.bam
- SRR1780018.bam
- GSM1596874.bam
- GSM1596944.bam
- SRR1779746.bam
- SRR1779956.bam
- SRR1780020.bam
- GSM1596881.bam
- GSM1596961.bam
- SRR1779805.bam
- SRR1779959.bam
- SRR1780054.bam
- GSM1596831.bam.bai
- GSM1596940.bam.bai
- GSM1597041.bam.bai
- SRR1779856.bam.bai
- SRR1779973.bam.bai
- GSM1596840.bam.bai
- GSM1596942.bam.bai
- GSM1597096.bam.bai
- SRR1779874.bam.bai
- SRR1780018.bam.bai
- GSM1596874.bam.bai
- GSM1596944.bam.bai
- SRR1779746.bam.bai
- SRR1779956.bam.bai

- `SRR1780020.bam.bai`

- `GSM1596881.bam.bai`

- `GSM1596961.bam.bai`

- `SRR1779805.bam.bai`

- `SRR1779959.bam.bai`

- `SRR1780054.bam.bai`

These '.bam' files are Single-cell ATAC-seq reads and the corresponding '.bam.bai' files are for indexing and searching for the reads in these files. For more extensive documentation, please refer to the metadata or help files for the objects.

How to load data

Please refer to ['SCATE' tutorial](https://github.com/Winnie09/SCATE) for the loading and use of the data.

Citation

If the 'SCATEData' package is useful in your work, please cite the following paper: - Zhicheng Ji, Weiqiang Zhou, Wenpin Hou, Hongkai Ji, [*Single-cell ATAC-seq Signal Extraction and Enhancement with SCATE*](https://doi.org/10.1186/s13059-020-02075-3), Genome Biol 21, 161 (2020).

SRR1779746.bam	*'SRR1779746.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779746.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779805.bam	*'SRR1779805.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779805.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779856.bam	*'SRR1779856.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779856.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779874.bam	'*SRR1779874.bam.Rd*'data

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779874.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779956.bam	'*SRR1779956.bam.Rd*'data

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779956.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779959.bam	*'SRR1779959.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779959.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1779973.bam	*'SRR1779973.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1779973.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1780018.bam	*'SRR1780018.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1780018.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1780020.bam	*'SRR1780020.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1780020.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

SRR1780054.bam	*'SRR1780054.bam.Rd'data*

Description

Bam file for single-cell ATAC-seq data (GEO Accession GSE65360).

Details

Original .fastq files were downloaded and then alignments were performed to obtained these files

Source: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65360

Examples

```
library(GenomicAlignments)
fl <- system.file("extdata", "SRR1780054.bam", package = "SCATEData")
a <- readGAlignments(fl)
```

Index