Package ‘nullrangesData’

April 4, 2024

Title  ExperimentHub datasets for the nullranges package

Version 1.8.0

Description  Provides datasets for the nullranges package vignette, in particular example datasets for DNase hypersensitivity sites (DHS), CTCF binding sites, and CTCF genomic interactions. These are used to demonstrate generation of null hypothesis feature sets, either through block bootstrapping or matching, in the nullranges vignette. For more details, see the data object man pages, and the R scripts for object construction provided within the package.

Imports  utils

Depends  R (>= 4.1.0), ExperimentHub, GenomicRanges, InteractionSet

Suggests  knitr, rmarkdown

biocViews  ExperimentHub, Homo_sapiens_Data, SequencingData, ChIPSeqData, ENCODE

License  GPL-3

Encoding  UTF-8

VignetteBuilder  knitr

Roxygen  list(markdown = TRUE)

RoxygenNote  7.2.1

LazyData  true

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

git_branch  RELEASE_3_18

git_last_commit  b4405be

git_last_commit_date  2023-10-24

Repository  Bioconductor 3.18

Date/Publication  2024-04-04

Author  Michael Love [aut, cre] (<https://orcid.org/0000-0001-8401-0545>), Wancen Mu [aut] (<https://orcid.org/0000-0002-5061-7581>), Eric Davis [aut] (<https://orcid.org/0000-0003-4051-3217>), Mikhail Dozmorov [aut]

Maintainer  Michael Love <michaelisaiahlove@gmail.com>
### DHSA549Hg38

DNA hypersensitivity (DHS) peaks in A549 cell example data, lifted to hg38

An example dataset containing narrowPeak file from ENCODE. Retrieve record with object[["AH22505"]]
on Annotation Hub. Construction script is in 'inst/script/DHSA549Hg38.R'. Function returns a GRanges object with metadata score, signal value, p/q value and peak.

```r
dhs <- DHSA549Hg38()
dhs
```

### hg19_10kb_bins

10Kb bins from hg19 with GM12878 metadata annotation features

10Kb bins were tiled across hg19 and annotated with CTCF and DNase site features from GM12878. Feature annotations for each bin include 1) the number of CTCF sites, 2) the CTCF signal strength (from peak calls), 3) the number of DNase sites, 4) the DNase signal strength (from signal tracks), and finally 5) the presence/absence of a loop to any other bin. Function returns a GRanges object with covariate metadata.

```r
bins <- hg19_10kb_bins()
bins
```
**hg19_10kb_ctcfBoundBinPairs**

*CTCF-bound 10Kb paired genomic interactions*

**Description**

10Kb bins were tiled across hg19 then subset by those which contained CTCF sites. All pairs of CTCF-bound 10Kb bins were generated and annotated with feature overlaps from GM12878. Feature annotations include 1) presence/absence of a loop between bin-pairs, 2) the total CTCF signal from both bin-pairs, 3) the number of CTCF sites from both bin-pairs, 4) the distance between bin-pairs, and finally 4) whether a convergent set of CTCF sites exists between bin-pairs. Function returns a GInteractions object with covariate metadata.

**Examples**

```r
binPairs <- hg19_10kb_ctcfBoundBinPairs()
binPairs
```

**nullrangesData**

*ExperimentHub datasets for the nullranges package*

**Description**

DNase hypersensitivity sites (DHS), CTCF binding sites, and CTCF genomic interactions for demonstration of functions in the nullranges package.

**Examples**

```r
suppressPackageStartupMessages(library(GenomicRanges))
dhs <- DHSA549Hg38()
dhs
```

**sc_promoter**

*Chromium Single Cell Multiome ATAC + Gene Expression assay*

**Description**

An example dataset of Single cell Gene and ATAC-seq data for Construction script is in 'inst/script/make-multiome-data.R'. Function returns a promoter GRanges object with pseudo bulk counts stored in metadata column "counts2" as NumericList.

**Examples**

```r
data("sc_promoter")
```
sc_rna

Chromium Single Cell Multiome ATAC + Gene Expression assay

Description

An example dataset of Single cell Gene and ATAC-seq data for Construction script is in `inst/script/make-multiome-data.R`. Function returns a gene expression GRanges object with pseudo bulk counts stored in metadata column "counts1" as NumericList.

Examples

data("sc_rna")
Index

DHSA549Hg38, 2

hg19_10kb_bins, 2
hg19_10kb_ctcfBoundBinPairs, 3

nullrangesData, 3

sc_promoter, 3
sc_rna, 4