Package 'jmosaics'

October 9, 2013

Type Package

Title Joint analysis of multiple ChIP-Seq data sets

Version 1.0.0

Date 2013-01-29

Author Xin Zeng

Maintainer Xin Zeng <xinzeng@stat.wisc.edu>

Depends R (>= 2.15.2), mosaics

Description jmosaics detects enriched regions of ChIP-seq data sets jointly.

License GPL (>= 2)

LazyLoad yes

biocViews ChIPseq, Sequencing, Transcription, Genetics, Bioinformatics

R topics documented:

`jmosaics-package` *Joint analysis of multiple ChIP-seq data sets*

Description

Joint analysis of multiple ChIP-seq data sets

Details

Package:	jmosaics
Type:	Package
Version:	1.0
Date:	2012-06-24
License:	GPL(>=2)
LazyLoad:	yes

`readBinsMultiple` match coordinates for multiple datasets,`jmosaicsPattern` call E_LAYER and B_LAYER peaks

Author(s)

Xin Zeng, Sunduz Keles

Maintainer: Xin Zeng <xinzeng@stat.wisc.edu>

References

jMOSAiCS: Joint Analysis of Multiple ChIP-seq Datasets

Examples

```
## step by step not run
## Not run:
bin1 <- readBins(type = c("chip", "M", "GC", "N","input"),
 fileName = c("h3k27me3_chip_chr10.txt",
 "./M_chr10.txt", "./GC_chr10.txt", "./N_chr10.txt",
"h3k27me3_input_chr10.txt"))
bin2 <- readBins(type = c("chip", "M", "GC", "N","input"),
 fileName = c("h3k4me1_chip_chr10.txt",
 "./M_chr1.txt", "./GC_chr10.txt", "./N_chr10.txt",
 "h3k4me1_input_chr10.txt"))
origin_bin <- list(bin1, bin2)

## End(Not run)

data("jmosaics_example_data")
bin <- readBinsMultiple(origin_bin)
fit1 <- mosaicsFit(bin[[1]], analysisType = "IO")
fit2 <- mosaicsFit(bin[[2]], analysisType = "IO")
fit <- list(fit1,fit2)
result <- jmosaicsPattern(fit, region_length=1, FDR=0.01, thres=c(10,10), type=c('B','E','Pattern'), patternInfo=
```

jmosaicsPattern *Call peaks and obtain combinatorial enrichment patterns*

Description

Call peaks and obtain combinatorial enrichment patterns

Usage

```
jmosaicsPattern(fit_all,region_length,FDR,thres=NULL,type=c('B','E','Pattern'),patternInfo='FALSE')
```

Arguments

fit_all	A list of fitted MOSAiCS models obtained using function mosaicsFit.
region_length	Region length, the number of bins covered by each region.
FDR	False discovery rate.
thres	A peak within initial peak is removed if its ChIP tag counts are less than thres. thres could be NULL or a numeric vector, corresponding to each dataset. If NULL, no threshold for average ChIP tag counts for all data sets would be used. Default is NULL.
type	a vector of characters to decide which objects would be returned. Possible values are 'E': return E_LAYER result; 'B': return B_LAYER result and 'Pattern': return Pattern (enrichment pattern).
patternInfo	Parameter for returning information on regions. Possible values are TRUE (return information on regions) or FALSE (do not return information on regions). Default is FALSE.

Details

The function returns objects based on 'type'. 'B_LAYER' object is a list of regions which are enriched in at least one dataset. Peak information can be accessed by 'chrID', 'PeakStart', 'PeakStop', 'Postprob'(P(B=0jdata information)), 'aveChipCount E_*', 'aveInputCount E_*'. * indicates the index of the datasets, for example: aveChipCount E_1 is the average tagCount for the first dataset. Each list of 'E_LAYER' object reports enriched regions for each dataset which can be accessed by 'chrID', 'PeakStart', 'PeakStop', 'Postprob'(P(E=0j data sets)), 'aveChipCount', 'maxChipCount', 'aveInputCount', ''aveInputCountScaled', 'aveLog2Ratio'. If region length is 1, it can be accessed by 'chrID', 'PeakStart', 'PeakStop', 'Postprob', 'ChipCount', 'InputCount', 'Input- CountScaled', 'Log2Ratio'. In the object of Pattern, it reports the enrichment patterns to the regions which cover the whole genome. When the region covering more than one bin, to get the average ChIP and input tagCount would be time consuming, the argument of 'patternInfo' let users decide whether to report the average ChIP and input tagCount.

Value

A list with following components:

E_LAYER	lists of enriched E regions, each list includes enriched regions for each data set.
B_LAYER	list of enriched B regions.
Pattern	list of regions annotated with patterns.

Author(s)

Xin Zeng

Examples

```
data("jmosaics_example_data")
bin <- readBinsMultiple(origin_bin)
fit1 <- mosaicsFit(bin[[1]], analysisType = "IO")
fit2 <- mosaicsFit(bin[[2]], analysisType = "IO")
fit <- list(fit1,fit2)
result <- jmosaicsPattern(fit, region_length=1, FDR=0.01, thres=c(10,10), type=c('B','E','Pattern'), patternInfo=
```

---

jmosaics_example_data *ChIP-seq data of Histone H3K4me1 and H3K27me3.*

Description

This is an example chr10 ChIP-seq data of Histone H3K4me1 and H3K27me3.

Usage

```
data("jmosaics_example_data")
```

Format

jmosaics_example_data names of included datasets.

Examples

```
data(jmosaics_example_data)
jmosaics_example_data
```

origin_bin *ChIP-seq data of Histone H3K4me1 and H3K27me3.*

Description

This is an example chr10 ChIP-seq data of Histone H3K4me1 and H3K27me3.

Usage

```
data("jmosaics_example_data")
```

Format

`origin_bin` a list of bin-level data sets.

Examples

```
data(jmosaics_example_data)
str(origin_bin)
```

readBinsMultiple *Match coordinates for multiple data sets*

Description

Match coordinates for multiple data sets

Usage

```
readBinsMultiple(dataset)
```

Arguments

`dataset`	A list of bin-level data sets.

Details

Bin-level data can be generated by readBins from mosaics package for each data set. This function is used to match the coordinates of multiple data sets to get the shared regions.

Value

List of Bin-level data sets.

Author(s)

Xin Zeng

Examples

```
data("jmosaics_example_data")
bin<- readBinsMultiple(origin_bin)
```

Index