

Package ‘PeakError’

June 20, 2017

Maintainer Toby Dylan Hocking <toby.hocking@r-project.org>

Author Toby Dylan Hocking

Version 2017.06.19

License GPL-3

Title Compute the Annotation Error of Peak Calls

Description Chromatin immunoprecipitation DNA sequencing results in genomic tracks that show enriched regions or peaks where proteins are bound. This package implements fast C code that computes the true and false positives with respect to a database of annotated regions.

Suggests testthat, ggplot2

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-06-19 22:48:28 UTC

R topics documented:

checkChrom	1
checkPositions	2
PeakError	2
PeakErrorChrom	4
Peaks	4

Index	6
--------------	----------

checkChrom	<i>checkChrom</i>
------------	-------------------

Description

Check for a valid data.frame with chrom names.

Usage

```
checkChrom(df)
```

Arguments

```
df
```

Author(s)

Toby Dylan Hocking

checkPositions	<i>checkPositions</i>
----------------	-----------------------

Description

Check for a valid data.frame with chromStart, chromEnd.

Usage

```
checkPositions(df)
```

Arguments

```
df
```

Author(s)

Toby Dylan Hocking

PeakError	<i>PeakError</i>
-----------	------------------

Description

Compute true and false positive peak calls, with respect to a database of annotated regions.

Usage

```
PeakError(peaks, regions)
```

Arguments

peaks	data.frame with columns chrom, chromStart, chromEnd. NOTE: chromStart should be 0-based and chromEnd should be 1-based. EXAMPLE: the first 100 base of of a chromosome are chromStart=0, chromEnd=100. The second 100 bases are chromStart=100, chromEnd=200.
regions	data.frame with columns chrom, chromStart, chromEnd, annotation.

Value

data.frame for each region with additional counts of true positives (tp, possible.tp), false positives (fp, possible.fp, fp.status), and false negatives (fn, fn.status).

Author(s)

Toby Dylan Hocking

Examples

```
x <- seq(5, 85, by=5)
peaks <- rbind(Peaks("chr2", x, x+3),
              Peaks("chr3", c(25, 38, 57), c(33, 54, 75)),
              Peaks("chr4", c(5, 32, 38, 65), c(15, 35, 55, 85)),
              Peaks("chr5", c(12, 26, 56, 75), c(16, 54, 59, 85)))
regions <- NULL
for(chr in 1:5){
  regions <- rbind(regions, {
    data.frame(chrom=paste0("chr", chr),
              chromStart=c(10, 30, 50, 70),
              chromEnd=c(20, 40, 60, 80),
              annotation=c("noPeaks", "peakStart", "peakEnd", "peaks"))
  })
}
err <- PeakError(peaks, regions)
ann.colors <-
  c(noPeaks="#f6f4bf",
    peakStart="#ffafaf",
    peakEnd="#ff4c4c",
    peaks="#a445ee")
library(ggplot2)
ggplot()+
  geom_rect(aes(xmin=chromStart+1/2, xmax=chromEnd+1/2,
              ymin=-1, ymax=1,
              fill=annotation,
              linetype=fn.status,
              size=fp.status),
            data=err, color="black")+
  scale_y_continuous("", breaks=NULL)+
  scale_linetype_manual(values=c("false negative"="dotted", correct="solid"))+
  scale_size_manual(values=c("false positive"=3, correct=1))+
  scale_fill_manual(values=ann.colors, breaks=names(ann.colors))+
  facet_grid(chrom ~ .)+
  theme_bw()+
  guides(fill=guide_legend(order=1),
         linetype=guide_legend(order=2, override.aes=list(fill="white")),
         size=guide_legend(order=3, override.aes=list(fill="white")))+
  theme(panel.margin=grid::unit(0, "cm"))+
  geom_segment(aes(chromStart+1/2, 1/2, xend=chromEnd+1/2, yend=1/2),
              data=peaks, color="deepskyblue", size=2)+
  scale_x_continuous("position on chromosome",
                    breaks=seq(10, 90, by=10))+
```

```
geom_text(aes(base, -1/2, label="N"), data.frame(base=10:90),
          color="deepskyblue")
```

PeakErrorChrom	<i>PeakErrorChrom</i>
----------------	-----------------------

Description

Compute the PeakError assuming that peaks and regions are on the same chrom.

Usage

```
PeakErrorChrom(peaks, regions)
```

Arguments

peaks	data.frame with columns chromStart, chromEnd. NOTE: chromStart should be 0-based and chromEnd should be 1-based. EXAMPLE: the first 100 base of of a chromosome are chromStart=0, chromEnd=100. The second 100 bases are chromStart=100, chromEnd=200.
regions	data.frame with columns chromStart, chromEnd.

Value

data.frame with 1 row for each region and error columns.

Author(s)

Toby Dylan Hocking

Peaks	<i>Peaks</i>
-------	--------------

Description

Make a data.frame that represents a list of peaks.

Usage

```
Peaks(chrom = factor(), base.before = integer(), last.base = integer())
```

Arguments

chrom	character or factor with chrom name for example "chr22"
base.before	integer, base before peak.
last.base	integer, last base of peak.

Value

data.frame with columns chrom, chromStart, chromEnd.

Author(s)

Toby Dylan Hocking

Index

checkChrom, 1
checkPositions, 2

PeakError, 2
PeakErrorChrom, 4
Peaks, 4