

Package ‘PING’

May 9, 2024

Type Package

Title Probabilistic inference for Nucleosome Positioning with
MNase-based or Sonicated Short-read Data

Description Probabilistic inference of CHIP-Seq using an empirical
Bayes mixture model approach.

Version 2.49.0

Author Xuekui Zhang <xuezhang@jhsph.edu>, Raphael Gottardo
<rgottard@fredhutch.org>, Sangsoo Woo <swoo@fhcrc.org>

Maintainer Renan Sauteraud <renan.sauteraud@gmail.com>

Depends R(>= 3.5.0)

Imports methods, PICS, graphics, grDevices, stats, Gviz, fda,
BSgenome, stats4, BiocGenerics, IRanges, GenomicRanges,
S4Vectors

Suggests parallel, ShortRead, rtracklayer

Collate setClasses.R setMethods.R PING.R postPING.R segmentPING.R

License Artistic-2.0

biocViews Clustering, StatisticalMethod, Visualization, Sequencing

RoxygenNote 7.0.1

git_url <https://git.bioconductor.org/packages/PING>

git_branch devel

git_last_commit 4eff825

git_last_commit_date 2024-04-30

Repository Bioconductor 3.20

Date/Publication 2024-05-08

Contents

| | |
|--------------------|----------|
| postPING | 2 |
| Index | 5 |

 postPING

Post process Estimation of binding site positions obtained from PING

Description

Post process Estimation of binding site positions obtained from PING. Refit mixture models with stronger prior in candidate regions contain potential problems, and then convert final result into dataframe.

Usage

```
postPING(
  ping,
  seg,
  rho2 = NULL,
  sigmaB2 = NULL,
  alpha2 = NULL,
  beta2 = NULL,
  min.dist = 100,
  paraEM = NULL,
  paraPrior = NULL,
  score = 0.05,
  dataType = "MNase",
  nCores = 1,
  makePlot = FALSE,
  FragmentLength = 100,
  mart = NULL,
  seg.boundary = NULL,
  DupBound = NULL,
  IP = NULL,
  datname = ""
)
```

Arguments

| | |
|------------------------------|--|
| ping | A pingList object containing estimation of nucleosome positions as returned by the PING function. |
| seg | An object of class segmentReadsList containing the results for all pre-processed regions as returned by segmentReads. |
| rho2, sigmaB2, alpha2, beta2 | Integer values, the parameters in the prior of mixture models to be re-fitted. |
| min.dist | The minimum distance of two adjacent nucleosomes predicted from different candidate regions, smaller than that will be treated as duplicated predictions for the same nucleosomes. |
| paraEM | A list of parameters for the EM algorithm. The default parameters should be good enough for most usages. |

| | |
|---------------------------------------|---|
| paraPrior | A list of parameters for the prior distribution. The default parameters should be good enough for most usages. |
| score | A numeric. The score threshold used when calling FilterPING. |
| dataType | A character that can be set to use selected default parameters for the algorithm. |
| nCores | An integer. The number of cores that should be used in parallel by the function. |
| makePlot | A logical. Plot a summary of the output. |
| FragmentLength | An integer. The length of XSET profile extension |
| mart, seg.boundary, DupBound, datname | Plotting parameters and options. |
| IP | A GRanges object. The reads used in segmentation process. |
| minK | An integer. The minimum number of binding events per region. If the value is 0, the minimum number is automatically calculated. |
| maxK | An integer. The maximum number of binding events per region. If the value is 0, the maximum number is automatically calculated. |
| tol | A numeric. The tolerance for the EM algorithm. |
| B | An integer. The maximum number of iterations to be used. |
| mSelect | A character specifying the information criteria to be used when selecting the number of binding events. Default="AIC3" |
| mergePeaks | A logical stating whether overlapping binding events should be picked. |
| mapCorrect | A logical stating whether mappability profiles should be incorporated in the estimation, i.e: missing reads estimated. |
| xi | An integer. The average DNA fragment size. |
| rho | An integer. A variance parameter for the average DNA fragment size distribution. |
| alpha | An integer. First hyperparameter of the inverse Gamma distribution for σ^2 in the PICS model |
| beta | An integer. Second hyperparameter of the inverse Gamma distribution for σ^2 in the PING model |
| lambda | An integer. The lambda control Gaussian Markov Random Field prior on the distance of adjacent nucleosomes, we do not recommend user change the default value. |
| dMu | An integer. Our best guess for the distance between two neighboring nucleosomes. |

Value

A data.frame containing the estimated binding site positions

Note

Based on our experient on a few real data sets, we suggestion to use following values of parameters. For sonication data we use $\rho_1=1.2$; $\sigma_B^2=6400$; $\rho=15$; $\alpha_1=10$; $\alpha_2=98$; $\beta_2=200000$. For MNase data we use $\rho_1=3$; $\sigma_B^2=4900$; $\rho=8$; $\alpha_1=20$; $\alpha_2=100$; $\beta_2=100000$. The value of ξ depends on specs of sample, since that affect the length of linker-DNA. For example, we use $\xi=160$ for yeast and $\xi=200$ for mouse.

See Also

PING, plotSummary

Index

postPING, [2](#)