Package 'PICS'

March 26, 2013

Type Package

Version 2.2.0

Index

Title Probabilistic inference of ChIP-seq

Author Xuekui Zhang <xzhang@stat.ubc.ca>, Raphael Gottardo <rgottard@fhcrc.org></rgottard@fhcrc.org></xzhang@stat.ubc.ca>
Maintainer Renan Sauteraud <rsautera@fhcrc.org></rsautera@fhcrc.org>
Depends R (>= 2.14.0), BiocGenerics (>= 0.1.3)
Imports methods, stats4, IRanges, GenomicRanges, graphics, grDevices, stats
Suggests ShortRead, rtracklayer, parallel
Description Probabilistic inference of ChIP-Seq using an empirical Bayes mixture model approach.
biocViews Clustering, Visualization, Sequencing, ChIPseq
License Artistic-2.0
makeRangedDataOutput
and Decide DE

 summary
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...

19

make Ranged Data Output

Create a RangedData object from a PICS output

Description

Create a list of 'RangedData' objects from a 'pics' object. The resulting RangedData object can then be analyzed with the 'IRanges' packages and/or exported to bed/wig files with the 'rtracklayer' package.

Usage

makeRangedDataOutput(obj, type="fixed", filter=list(delta=c(0,Inf), se=c(0,Inf), sigmaSqF=c(0,Inf), sigmaS

Arguments

obj	An object of class 'picsList' as returned by 'PICS' when running it on the IP/Control data.
type	The type of intervals to be created. The different types are 'bed', 'wig', 'ci' and 'fixed'. See details for more info.
filter	A list of filters to be used before computing the FDR. By default all regions are included, see details for more info on how to specify the filters.
length	The length to be used for the fixed type 'RangedData', see details.

Details

'bed' will generate intervals from the forward peak max to the reverse peak max. 'wig' will generate a density profile for the forward and reverse reads. 'bed' and 'wig' types should be used to be exported to wig/bed files to be used with the UCSC genome browser. 'ci' corresponds to the binding site estimates +/-3*se, while 'fixed' corresponds to the binding site estimates +/-3*length. 'bed' and 'wig' files can be exported using the 'export' function fo the 'rtracklayer' package.

Value

An object of type 'RangedData'.

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

See Also

export

pics 3

Examples

```
\label{eq:proposed_proposed_proposed} $\# $\operatorname{Not} \ run: $ rdBed<-makeRangedDataOutput(pics,type="bed",filter=list(delta=c(50,Inf),se=c(0,50),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sigmaSqF=c(0,22500),sig
```

pics

Estimation of binding site positions

Description

This object contains Estimation of binding site positions and has the following slots: segReadsList, dataType.

Usage

PICS(segReadsList,dataType=NULL, paraEM=NULL, paraPrior=NULL)

Arguments

segReadsList This object contains segmentation of Genome

dataType The type of data you are processing: specified 'TF' for transcription factor.

paraEM A list of parameters for the EM algorithm as returned by the setParaEm func-

tion. The default parameters should be good enough for most usages.

minK: an integer, default=1. The minimum number of binding events per region. If the value is 0, the minimum number is automatically calculated.

maxK: an integer, default=15. The maximum number of binding events per region. If the value is 0, the maximum number is automatically calculated.

tol: a numeric, default=1e-4. The tolerance for the EM algorithm.

B: an integer, default=100. The maximum number of iterations to be used.

mSelect: a character string specifying the information criteria to be used when selecting the number of binding events. Default="BIC"

mergePeaks: a logical stating whether overlapping binding events should be picked. Default=TRUE

mapCorrect: a logical stating whether mappability profiles should be incorporated in the estimation, i.e. missing reads estimated. Default=TRUE

paraPrior

A list of parameters for the prior distribution as returned by the setParaPrior function. The default parameters should be good enough for most usages.

xi: an integer, default=200. The average DNA fragment size.

rho: an integer, default=1. A variance parameter for the average DNA fragment size distribution.

alpha: an integer, default=20. First hyperparameter of the inverse Gamma distribution for sigma^2 in the PICS model

4 pics-class

beta: an integer, default=40000. Second hyperparameter of the inverse Gamma distribution for sigma^2 in the PING model

lambda: an integer, default=0. The precision of the prior for mu used for histone data.

dMu: an integer, default=0. Our best guess for the distance between two neighboring nucleosomes.

Methods

code signature(x = ``pics''): return the error code for each list element (i.e. candidate region) of a PICS object. If the string is empty, there were no errors.

plot $\operatorname{signature}(x = \text{``pics''})$: Plot all regions in the PICS object. This might be long, and should only be used to plot a few regions, so subset the object before plotting.

sigmaSqR signature(x = "pics"): return the variance parameter of the reverse (R) distribution for each binding event.

sigmaSqF signature(x = "pics"): return the variance parameter of the forward (F) distribution for each binding event.

score $\operatorname{signature}(x = \text{``pics''})$: return the score for each binding event.

scoreF signature(x = ``pics''): return the score of the forward (F) for each binding event.

scoreR signature(x = ``pics''): return the score of the forward (R) for each binding event.

maxRange signature(x = "pics"): return the range maximum.

minRange signature(x = "pics"): return the range minimal.

K signature(x = "pics"): subset PICS object.

wigDensity signature(x = "pics"): return the density for each binding event.

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

See Also

pics

pics-class

The pics class

Description

This object is used to gather all parameters from fitting PICS to a single candidate region. The objet contains the following slots: 'estimates', 'infMat', 'Nmerged', 'converge', 'chr'. 'estimates' is a list containing all parameters estimates as well as standard errors. 'infMat' is the Cholesky decomposition of the information matrix, 'converge' is a logical value indicating whether the EM algorithm has converged, while 'chr' is a character string corresponding to a candidate region's chromosome. 'Nmerged' gives the number of binding events that were merged; binding events that overlap are merged (see the cited paper below for details).

pics-class 5

Accessors

The PICS package provide accessors to directly access to most of the parameters/standard errors and chromosome. In the code snippets below, 'x' is a 'pics' object.

'chromosome(x)' Gets the chromosome name of the candidate region.

'mu(x)' Gets the position estimates of all binding sites identified in the region.

'delta(x)' Gets the average fragment lengths of all binding sites identified in the region.

'sigmaSqF(x)' Gets the F peak variances of all binding sites identified in the region.

'sigmaSqR(x)' Gets the R peak variances of all binding sites identified in the region.

'se $\mathbf{F}(\mathbf{x})$ ' Gets the standard errors of all binding site position estimates identified in the region.

'se $\mathbf{F}(\mathbf{x})$ ' Gets the standard errors of all F peak modes identified in the region.

'seR(x)' Gets the standard errors of all R peak modes identified in the region.

score $\operatorname{signature}(x = "\operatorname{pics}")$: return the score for each binding event.

scoreF signature(x = "pics"): return the score of the forward (F) for each binding event.

scoreR signature(x = "pics"): return the score of the forward (R) for each binding event.

Constructor

newPics(w,mu,delta,sigmaSqF,sigmaSqR,seMu,seMuF,seMuR,score,Nmerged,converge,infMat,chr) construct a new 'pics' object with the following arguments:

w The mixture weights (a vector)

mu The binding site positions (a vector)

delta The DNA fragment lengths (a vector)

sigmaSqF The variance parameters for the forward distribution (vector)

sigmaSqR The variance parameters for the forward distribution (vector)

seMu The standard errors for mu (vector)

seMuF The standard errors for muF (vector)

seMuR The standard errors for muR (vector)

seMuR The standard errors for muR (vector)

score The scores for each binding event (vector)

Nmerged The number of peaks that got merged (integer)

converge A logical value, TRUE, if the EM as converged

infMat The information matrix

chr The chromosome for the region

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

6 picsError-class

See Also

```
pics picsError
```

Examples

```
# Here is an example of how to construct such a region.
# Typically, you would not do this manually, you would use the pics function to return a 'picsList' that contains a list of 'r
mu<-10000
delta < -150
sigmaSqF < -5000
sigmaSqR < -5000
seMu < -10
seMuF < -10
seMuR < -10
score < -5
Nmerged < -0
converge<-TRUE
chr < -"chr1"
range<-c(1000,2000)
# Contructor
#myPICS<-newPics(w,mu,delta,sigmaSqF,sigmaSqR,seMu,seMuF,seMuR,score,Nmerged,as.integer(range),chr)
```

picsError-class

The pics class

Description

This object is used to return an error code when the PICS function failed to return a valid set of estimates for a candidate regions. This could be due to non-convergence of the EM algorithm, a singular information matrix, or a number of reads below the limit specified by the user. All of these are typically due to too few reads in the region and do not affect the rest of the analysis, as such regions would most likely be labelled as false positives.

Accessors

All of the accessors defined for a 'pics' object still work for a 'picsError' object but will simply return a NULL pointer.

Constructor

newPicsError(string) where 'string' is the error code.

Constructor

```
newPicsError<-function(string)
string The mixture weights (a vector)</pre>
```

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

picsFDR 7

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

See Also

pics

Examples

```
\# Here is an example on how to construct such a picsError object
```

Typically, you would not do this manually, you would use the pics function to return a 'picsList' that contains a list of 'p

Contructor

myPicsError < -newPicsError ("Singular information matrix")

Accessors

Get the standard error of Mu

se(myPicsError)

Get the standard error of MuF

seF(myPicsError)

Get the scores

score(myPicsError)

picsFDR

Estimate the FDR.

Description

Estimate the false detection rate for an object of class pics or picsList.

Usage

```
picsFDR(picsIP,picsCont,filter=list(delta=c(0,Inf),se=c(0,Inf),sigmaSqF=c(0,Inf),sigmaSqR=c(0,Inf)))\\
```

Arguments

picsIP An object of class pics or picsList containing the informations for the IP reads.

picsCont An object of class pics or picsList containing the informations for the control

reads.

filter filterA list of ranges for filtering regions based on PICS parameters. By default

filter is set to 'NULL' and all regions are used.

delta Length of the binding sites.

se Standard error.

sigmaSqF Forward peak variance
sigmaSqR Reverse peak variance

Value

A 3 columns data.frame with the following columns: FDR, score, N.

8 picsList-class

Author(s)

Xuekui Zhang

See Also

picsList pics

picsList-class

The pics class

Description

This object is used to gather all parameters from fitting PICS to multiple candidate regions (as returned by the 'segmentReads' function). The objet contains the following slots: 'List', 'paraPrior', 'paraEM', 'minReads', 'N', 'Nc'. 'List' is a list of 'pics' or 'picsError' objects. 'paraPrior' is a list containing the hyperparameters used for the prior, 'paraEM' is a list of convergence parameters for the EM, 'minReads' is a list containing the minimum number of reads used to fit a region with 'PICS', 'N' is the total number of reads in the ChIP samples while 'Nc' is the total number of reads in the control sample.

Arguments

object

An object of class pics.

Accessors

The PICS package provide accessors to directly access to most of the parameters/standard errors and chromosomes. In the code snippets below, 'x' is a 'picsList' object. For all accessors, the 'picsError' objects are omitted, so that the accessors only return values for the 'pics' objects (i.e. all valid binding events).

'chromosome(x)' Gets the chromosome names of all candidate regions.

'mu(x)' Gets the position estimates of all binding sites identified in all candidate regions.

'delta(x)' Gets the average fragment lengths of all binding sites identified in all candidate regions.

'sigmaSqF(x)' Gets the F peak variances of all binding sites identified in all candidate regions.

'sigmaSqR(x)' Gets the R peak variances of all binding sites identified in all candidate regions.

'seF(x)' Gets the standard errors of all binding site position estimates identified in all candidate regions.

'se $\mathbf{F}(\mathbf{x})$ ' Gets the standard errors of all F peak modes identified in all candidate regions.

'seR(x)' Gets the standard errors of all R peak modes identified in all candidate regions.

'score(x)' Gets the scores of all binding events identified in all candidate regions.

Constructor

newPicsList(List, paraEM, paraPrior, minReads, N, Nc)

List The mixture weights (a vector)

paraEM The binding site positions (a vector)

paraPrior The DNA fragment lengths (a vector)

N The variance parameters for the forward distribution (vector)

Nc The variance parameters for the forward distribution (vector)

picsList-class 9

Methods

```
[ signature(x = "pics"): subset PICS object.
```

Methods

```
length signature(x = "pics"): subset PICS object.
```

Constructor

newPicsList<-function(List, paraEM, paraPrior, minReads, N, Nc) constructs a new 'picsList' object with the following arguments.

newPicsList

```
w The mixture weights (a vector)
```

mu The binding site positions (a vector)

delta The DNA fragment lengths (a vector)

sigmaSqF The variance parameters for the forward distribution (vector)

sigmaSqR The variance parameters for the reverse distribution (vector)

seMu The standard errors for mu (vector)

seMuF The standard errors for muF (vector)

seMuR The standard errors for muR (vector)

seMuR The standard errors for muR (vector)

score The scores for each binding event (vector)

Nmerged The number of peaks that were merged (integer)

converge A logical value, TRUE, if the EM as converged

infMat The information matrix

chr The chromosome for the region

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

```
X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.
```

See Also

pics

 $\begin{array}{l} sigmaSqF\!<\!\!-5000 \\ sigmaSqR\!<\!\!-5000 \end{array}$

Examples

```
\# Here is an example of how to construct such a region \# Typically, you would not do this manually, you would use the pics function to return a 'picsList' that contains a list of 'pw<-1 mu<-10000 delta<-150
```

10 plot-FDR

```
seMu < -10
seMuF{<}\text{-}10
seMuR < -10
score{<}\text{-}5
Nmerged < -0
converge {<-} TRUE
\inf Mat < -matrix(0)
chr < -"chr1"
range < -c(1000, 2000)
# Contructor
#myPICS1<-newPics(w,mu,delta,sigmaSqF,sigmaSqR,seMu,seMuF,seMuR,score,Nmerged,converge,infMat,as.integer(
#myPICS2<-newPics(w,mu+1000,delta,sigmaSqF,sigmaSqR,seMu,seMuF,seMuR,score,Nmerged,converge,infMat,as.in
#minReads<-list(perPeak=2,perRegion=5)
\#paraPrior<-list(xi=200,rho=1,alpha=20,beta=40000)
\#paraEM<-list(minK=1,maxK=15,tol=10e-6,B=100)
#N<-100
#Nc<-200
#mynewPicsList<-newPicsList(list(myPICS1,myPICS2), paraEM, paraPrior, minReads, as.integer(100), as.integer(200
# Accessors
# Get the standard error of Mu
#se(mynewPicsList)
\# Get the standard error of MuF
#seF(mynewPicsList)
# Get the scores
#score(mynewPicsList)
```

plot-FDR

FDR plot for PICS

Description

This method plots an FDR curve showing the FDR as a function of the PICS scores.

Usage

```
## S4 method for signature 'picsList,picsList' plot(x, y, filter=NULL, h=.1, ...)
```

Arguments

X	A picsList object as returned by the function PICS run on the treatment data.
У	A picsList object as returned by the function PICS run on the control data.
filter	A list of ranges for filtering regions based on PICS parameters. By default filter is set to 'NULL' and all regions are used.
h	A value between 0 and 1, representing the desired FDR. This simply draws a horizontal line at the given value.
	Further graphical parameters passed to the generic function plot.

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

segChrRead 11

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq," Biometrics, iss. In press, 2010.

See Also

PICS

segChrRead Segmentation of paired-end sequencing data

Description

These two functions are part of the segmentation step for paired-end sequencing data and are exported to be used in PING package.

segmentPICS	Segment the genome into candidate regions	
-------------	-------------------------------------------	--

Description

Pre-process bidirectional aligned reads data from a single ChIP-Seq experiment to detect candidate regions with a minimum number of forward and reverse reads. These candidate regions will then be processed by PICS.

Usage

```
segmentPICS(data, dataC=NULL, map=NULL, minReads=2, minReadsInRegion=3, jitter=FALSE, dataType="TF", maxLregion=0, minLregion=100)
```

Arguments

dataType

maxLregion

minLregion

Type of experiment.

The maximum length.

The minimum length.

guments		
data	A link S4class{GRanges} object containing the IP reads. See details for more information on how to set up the data.	
dataC	A link S4class{GRanges} object containing the control reads. Set to NULL by default, i.e. no control.	
map	A 'RangedData' object containing the mappability profiles. Set to NULL by default, i.e. no profiles.	
minReads	The minimum number of F/R reads to be present in the sliding window.	
minReadsInRegion		
	The minimum number of F/R reads to be present in the region.	
jitter	A logical value stating whether some noise should be added to the read locations. This is recommended if the read positions have lots of duplicates.	

12 segReads

Value

An object of class segReadsList containing the results for all regions pre-processed.

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009.

See Also

segReadsList

Examples

```
# Read data path<-system.file("extdata",package="PICS")
## Note that the col name for the chromosome needs to be space and not chr dataIP<-read.table(file.path(path,"Treatment_tags_chr21_sort.bed"),header=TRUE,colClasses=c("factor","integer" dataIP<-as(dataIP,"GRanges")

dataCont<-read.table(file.path(path,"Input_tags_chr21_sort.bed"),header=TRUE,colClasses=c("factor","integer","idataCont<-as(dataCont,"GRanges")

map<-read.table(file.path(path,"mapProfileShort"),header=TRUE,colClasses=c("factor","integer","integer","NULL")
map<-as(map,"GRanges")
seg<-segmentPICS(dataIP, dataC=dataCont, map=map, minReads=1)
```

segReads

Segment the genome into candidate regions

Description

Pre-process bidirectional aligned reads data from a single ChIP-Seq experiment to detect candidate regions with a minimum number of forward and reverse reads. These candidate regions will then be processed by PICS.

Methods

```
map signature(x = "pics"): subset PICS object.
```

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

segReadsList 13

See Also

pics

segReadsList

Segment the genome into candidate regions

Description

Pre-process bidirectional aligned reads data from a single ChIP-Seq experiment to detect candidate regions with a minimum number of forward and reverse reads. These candidate regions will then be processed by PICS.

Methods

```
[ signature(x = "pics"): subset gadem object.
[[ signature(x = "pics"): subset gadem object.
```

Methods

```
length signature(x = "pics"): subset PICS object.
```

Author(s)

Xuekui Zhang, Arnaud Droit <<arnaud.droit@crchuq.ualaval.ca>> and Raphael Gottardo <<rgottard@fhcrc.org>>

References

X. Zhang, G. Robertson, M. Krzywinski, K. Ning, A. Droit, S. Jones, and R. Gottardo, "PICS: Probabilistic Inference for ChIP-seq" arXiv, 0903.3206, 2009. To appear in Biometrics.

See Also

pics

 ${\bf segReadsListPE}$

List of segReadsPE objects

Description

A list of segReadsPE. The class also store information related to the segmentation process, keeping a trace of the parameters used and the proportion of forward and reverse reads for the input and the control.

Methods

```
[ signature(x = "pics"): subset gadem object.
[[ signature(x = "pics"): subset gadem object.
```

14 setParaEM

Extends

Class segReadsList, directly.

Author(s)

Xuekui Zhang

See Also

 ${\tt segReadsPE\ segReadsList}$

segReadsPE

Class to store post-segmentation result

Description

This class stores the information of the segmentation performed by segmentPING. It is used as the input of the PING function.

Extends

Class segReadsList, directly.

Author(s)

Xuekui Zhang

See Also

 ${\tt segReadsListPE}$

setParaEM

Function that returns a list of parameters for the EM algorithm that can be used as an argument of PICS.

Description

This function takes from 0 to 7 EM algorithm parameters as argument, check if they are valid and returns a list to be used in a call to PICS.

Usage

set ParaEM (minK=1, maxK=15, tol=1e-4, B=100, mSelect="BIC", mergePeaks=TRUE, mapCorrect=TRUE, data to the context of the co

setParaPrior 15

Arguments

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 string specifying the information criteria to be used when selecting

the number of binding events.

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.

dataType A character. If a dataType is set, the algorithm will use the default parameters

for this type of data (all the previous arguments will be ignored).

Value

Returns a list of parameters to be used in PICS.

Author(s)

Renan Sauteraud

See Also

PICS

setParaPrior	Function that returns a list of parameters that can be used as an argu-
	ment of PICS.

Description

This function takes from 0 to 6 parameters as argument, check if they are valid and returns a list to be used in a call to PICS.

Usage

```
set Para Prior(xi=200, rho=1, alpha=20, beta=40000, lambda=0, dMu=0, data Type=NULL, PExi=0)
```

Arguments

xi An integer. The average DNA fragment size.

rho An integer. A variance parameter for the average DNA fragment size distribu-

tion.

alpha An integer. First hyperparameter of the inverse Gamma distribution for sigma^2

in the PICS model

16 show

beta An integer. Second hyperparameter of the inverse Gamma distribution for

sigma^2 in the PICS model

lambda An integer. The precision of the prior for mu used for histone data.

dMu An integer. Our best guess for the distance between two neighboring nucleo-

somes.

dataType A character string. If a valid dataType is specified, use our suggested parameters.

"MNase" or "sonicated"

PExi A numeric. With paired end data, 'xi' can be calculated directly from the reads.

If PExi is set, it will overwrite the xi determined by the dataType.

Value

Returns a list of 6 parameters to be used in PICS.

Author(s)

Renan Sauteraud

See Also

PICS

Examples

```
\# set prior for PICS data para
Prior<-setParaPrior() \# set prior for sonicated data using our selected default parameters para
Prior<-setParaPrior(dataType="sonicated")
```

show show PICS

Description

This methods show the objects of PICS

Usage

```
## S4 method for signature 'pics' show(object)
## S4 method for signature 'picsError' show(object)
## S4 method for signature 'picsList' show(object)
## S4 method for signature 'segReads' show(object)
## S4 method for signature 'segReadsList' show(object)
```

summary 17

Arguments

object Object returned from pics.

Details

List of the slots include in the object

Author(s)

 $Xuekui\ Zhang,\ Arnaud\ Droit << {\rm arnaud.droit@crchuq.ualaval.ca}>> \ and\ Raphael\ Gottardo << rgottard@fhcrc.org>> \ and\ Raphael\ Gottardo << rgottard.group <> \ and\ Raphael\ Gottardo << rgottard.gr$

See Also

summary

summary

summary PICS

Description

This methods summarized 'pics', 'picsList', 'seg' or 'segList' objects.

Usage

```
## S4 method for signature 'pics' summary(object)
## S4 method for signature 'picsList' summary(object)
## S4 method for signature 'segReads' summary(object)
## S4 method for signature 'segReadsList' summary(object)
```

Arguments

object

Object returned from pics.

Author(s)

 $Xuekui\ Zhang,\ Arnaud\ Droit << arnaud.droit@crchuq.ualaval.ca>> \ and\ Raphael\ Gottardo << rgottard@fhcrc.org>> \ arnaud.droit@crchuq.ualaval.ca>> \ ar$

See Also

show

18 summarySeg

summarySeg

Summarize a segReadsList object.

Description

Returns info about a segReadsList object in a data.frame containing the following informations: chr: chromosome id NF: number of forward reads NR: number of reverse reads L: length of segment min: start location of segments max: location of segments

Usage

```
summarySeg(seg)
```

Arguments

seg

An object of class segReadsList

Value

A six columns data.frame.

Author(s)

Xuekui Zhang

See Also

 ${\bf segReadsList}$

Index

*Topic classes	as.list, AlignedRead-method (pics), 3
segReadsListPE, 13	
segReadsPE, 14	candidate.region (segChrRead), 11
*Topic data	chromosome (pics-class), 4
makeRangedDataOutput, 2	chromosome,data.frame-method
*Topic functions	(pics-class), 4
picsFDR, 7	chromosome, pics-method (pics-class), 4
segmentPICS, 11	chromosome,picsError-method
setParaEM, 14	(picsError-class), 6
setParaPrior, 15	chromosome, picsList-method
summarySeg, 18	(picsList-class), 8
*Topic graphs	code (pics), 3
plot-FDR, 10	code,pics-method (pics), 3
*Topic models	code,picsError-method (pics), 3
pics, 3	code,picsList-method (pics), 3
pics-class, 4	coerce (pics), 3
picsError-class, 6	coerce, Aligned Read, Genome Data-method
picsList-class, 8	(pics), 3
segReads, 12	coerce,data.frame,GenomeData-method
segReadsList, 13	(pics), 3
*Topic print	coerce, data.frame, GRanges-method (pics),
show, 16	3
summary, 17	coerce, data.frame, picsList-method (pics),
*Topic process	coerce, picsList, data.frame-method (pics),
makeRangedDataOutput, 2	coerce, picsList, RangedData-method (pics)
[,picsList,ANY,ANY-method	3
(picsList-class), 8	coerce, Ranged Data, Genome Data-method
[,picsList-method (picsList-class), 8	(pics), 3
[,segReadsList,ANY,ANY-method	delta (pics-class), 4
(segReadsList), 13	delta,data.frame-method (pics-class), 4
[,segReadsList-method (segReadsList), 13	delta, pics-method (pics-class), 4
[,segReadsListPE-method	delta,picsError-method (picsError-class), 6
(segReadsListPE), 13	delta,picsList-method (picsList-class), 8
[],picsList,ANY,ANY-method	density, picsList-method (picsList-class), 8
(picsList-class), 8	V/1 (1 //
[[,picsList-etass), 8	FDR (plot-FDR), 10
[[,segReadsList,ANY,ANY-method	
(segReadsList), 13	K (pics), 3
[[,segReadsList-method (segReadsList), 13	K,pics-method (pics), 3
[],segReadsListPE,ANY,ANY-method	K,picsError-method (pics), 3
(segReadsListPE), 13	K,picsList-method (pics), 3
%in%,ANY,segReads-method (segReads),	length (picsList-class), 8
12	length,picsList-method (picsList-class), 8
12	iongui,prosensu-mounou (prosensu-ciass), o

20 INDEX

length,segReadsList-method	score, picsList-method (picsList-class), 8
(segReadsList), 13	scoreForward (pics), 3
	scoreForward,data.frame-method (pics), 3
makeRangedDataOutput, 2	scoreForward, pics-method (pics), 3
map (segReads), 12	scoreForward, picsError-method (pics), 3
map,segReads-method (segReads), 12	scoreForward,picsList-method (pics), 3
map,segReadsList-method (segReads), 12	scoreReverse (pics), 3
maxRange (pics), 3	scoreReverse,data.frame-method (pics), 3
maxRange,pics-method (pics), 3	scoreReverse, pics-method (pics), 3
maxRange,picsError-method (pics), 3	scoreReverse, pics-method (pics), 3 scoreReverse, picsError-method (pics), 3
maxRange,picsList-method (pics), 3	·-
minRange (pics), 3	scoreReverse,picsList-method (pics), 3
minRange,pics-method (pics), 3	se (pics-class), 4
	se,data.frame-method (pics-class), 4
minRange, picsError-method (pics), 3	se,pics-method (pics-class), 4
minRange,picsList-method (pics), 3	se,picsError-method (picsError-class), 6
mu (pics-class), 4	se,picsList-method (picsList-class), 8
mu,data.frame-method (pics-class), 4	seF (pics-class), 4
mu,pics-method (pics-class), 4	seF,data.frame-method (pics-class), 4
mu,picsError-method (picsError-class), 6	seF,pics-method (pics-class), 4
mu,picsList-method (picsList-class), 8	seF,picsError-method (picsError-class), 6
	seF,picsList-method (picsList-class), 8
newPics (pics-class), 4	segChrRead, 11
newPics,pics-method (pics-class), 4	segmentPICS, 11
newPicsError (picsError-class), 6	segReads, 12, <i>14</i>
newPicsError,picsError-method	
(picsError-class), 6	segReads-class (segReads), 12
newPicsList (picsList-class), 8	segReadsGeneric (segmentPICS), 11
newPicsList,picsList-method	segReadsList, 12, 13, 14, 18
(picsList-class), 8	segReadsList-class (segReadsList), 13
//	segReadsListPE, 13, 14
PICS, 11	segReadsListPE-class (segReadsListPE), 13
PICS (pics), 3	segReadsPE, 14, 14
pics, 3, 4, 6–9, 13, 17	segReadsPE-class (segReadsPE), 14
pics-class, 4	seR (pics-class), 4
pics-class (pics), 3	seR,data.frame-method (pics-class), 4
picsError, 6	seR,pics-method (pics-class), 4
picsError (picsError-class), 6	seR,picsError-method (picsError-class), 6
	seR,picsList-method (picsList-class), 8
picsError-class, 6	setParaEM, 14
picsFDR, 7	setParaPrior, 15
picsList, 8	show, 16, 17
picsList (picsList-class), 8	show, pics-method (show), 16
picsList-class, 8	
plot,pics,segReads-method (pics), 3	show,picsError-method (show), 16
plot,picsError,segReads-method (pics), 3	show,picsList-method (show), 16
plot,picsList,picsList-method (plot-FDR),	show,segReads-method (show), 16
10	show,segReadsList-method (show), 16
plot,picsList,segReadsList-method (pics), 3	show-methods (show), 16
plot-FDR, 10	sigmaSqF (pics), 3
	sigmaSqF, data.frame-method (pics), 3
score (pics), 3	sigmaSqF,pics-method (pics), 3
score,data.frame-method (pics-class), 4	sigmaSqF, picsError-method (pics), 3
score,pics-method (pics-class), 4	sigmaSqF,picsList-method (pics), 3
score,picsError-method (picsError-class), 6	sigmaSqR (pics), 3

INDEX 21

```
sigmaSqR,data.frame-method (pics), 3
sigmaSqR,pics-method (pics), 3
sigmaSqR, picsError-method\ (pics),\ 3
sigmaSqR,picsList-method (pics), 3
summary, 17, 17
summary, pics-method (summary), 17
summary, picsList-method (summary), 17
summary,segReads-method (summary), 17
summary,segReadsList-method (summary),
summary-methods (summary), 17
summarySeg, 18
w (pics-class), 4
w,data.frame-method (pics-class), 4
w,pics-method (pics-class), 4
w,picsError-method (picsError-class), 6
w,picsList-method (picsList-class), 8
wigDensity (pics), 3
wigDensity, pics-method (pics), 3
wigDensity,picsError-method (pics), 3
wigDensity,picsList-method (pics), 3
```