Package ‘motifStack’

March 26, 2013

Type Package

Version 1.0.4

Date 2012-10-11

Title Plot stacked logos for single or multiple DNA, RNA and amino acid sequence

Author Jianhong Ou, Michael Brodsky, Scot Wolfe and Lihua Julie Zhu

Maintainer Jianhong Ou <jianhong.ou@umassmed.edu>

Imports grImport, grid, XML

Depends R (>= 2.15.1), methods, grImport

Suggests RUnit, BiocGenerics, MotIV

biocViews SequenceMatching, GenomicsSequence, Visualization

Description The motifStack package is designed for graphic
representation of multiple motifs with different similarity
scores. It works with both DNA/RNA sequence motif and amino
acid sequence motif. In addition, it provides the flexibility
for users to customize the graphic parameters such as the font type and symbol colors.

License GPL (>= 2)

Lazyload yes

R topics documented:

colorset *retrieve color setting for logo*

Description

retrieve color setting for logo

Usage

```
colorset(alphabet="DNA", colorScheme='auto')
```

Arguments

alphabet	character, 'DNA', 'RNA' or 'AA'
colorScheme	'auto', 'charge', 'chemistry', 'classic' or 'hydrophobicity' for AA, 'auto' or 'basepairing' for DNA ro RNA

Value

A character vector of color scheme

Examples

```
col <- colorset("AA", "hydrophobicity")
```

DNAmotifAlignment *align DNA motifs*

Description

align DNA motifs for plotting motifs stack

Usage

```
DNAmotifAlignment(pfms, threshold=0.4)
```

Arguments

pfms	a list of position frequency matrices, pfms must be a list of class pfm
threshold	information content cutoff threshold for useful postions

Value

a list of aligned motifs

Examples

```
pcms<-readPCM(file.path(find.package("motifStack"), "extdata"),"pcm$")
pcms<-lapply(pcms,function(.ele){.ele<-.ele[,3:ncol(.ele)];rownames(.ele)<-c("A","C","G","T");.ele})
motifs<-lapply(pcms,pcm2pfm)
motifs<-lapply(names(motifs), function(.ele, motifs){new("pfm",mat=motifs[[.ele]], name=.ele)},motifs)
motifs<-DNAmotifAlignment(motifs)
```

pcm2pfm *convert pcm to pfm*

Description

convert position count matrix to position frequency matrix

Usage

```
pcm2pfm(pcm)
```

Arguments

pcm	position count matrix

Value

position weight matrix

Examples

```
pcms<-readPCM(file.path(find.package("motifStack"), "extdata"),"pcm$")
pcms<-lapply(pcms,function(.ele){.ele<-.ele[,3:ncol(.ele)];rownames(.ele)<-c("A","C","G","T");.ele})
motifs<-lapply(pcms,pcm2pfm)
```

pfm-class *Class* "pfm"

Description

An object of class "pfm" represents the position frequency matrix of a DNA/RNA/amino-acid sequence motif. The entry stores a matrix, which in row i, column j gives the frequency of observing nucleotide/or amino acid i in position j of the motif.

Objects from the Class

Objects can be created by calls of the form `new("pfm", mat, name, alphabet, color, background)`.

Slots

`mat` Object of class "matrix" The position frequency matrix

`name` code"character" The motif name

`alphabet` "character" The sequence alphabet. "DNA", "RNA", "AA" or "others".

`color` a "character" vector. The color setting for each symbol

`background` a "numeric" vector. The background frequency.

Methods

- **addBlank** signature(x="pfm", n="numeric", b="logical") add space into the position frequency matrix for alignment. b is a bool value, if TRUE, add space to the 3' end, else add space to the 5' end. n indicates how many spaces should be added.
- **coerce** signature(from = "pfm", to = "matrix"): convert object pfm to matrix
- **getIC** signature(x = "pfm",) Calculate information content profile for position frequency matrix.
- **getIC** signature(x = "matrix", p = "numeric") Calculate information content profile for matrix. p is the background frequency
- **matrixReverseComplement** signature(x = "pfm") get the reverse complement of position frequency matrix.
- **plot** signature(x = "pfm") Plots the sequence logo of the position frequency matrix.

Examples

```
pcm <- read.table(file.path(find.package("motifStack"), "extdata", "bin_SOLEXA.pcm"))
pcm <- pcm[,3:ncol(pcm)]
rownames(pcm) <- c("A","C","G","T")
motif <- pcm2pfm(pcm)
motif <- new("pfm", mat=motif, name="bin_SOLEXA")
plot(motif)
```

pfm-methods *"pfm" methods*

Description

methods for pfm objects.

Usage

```
## S4 method for signature 'pfm,numeric,logical'
addBlank(x,n,b)
## S4 method for signature 'pfm,ANY'
getIC(x,p="missing")
## S4 method for signature 'matrix,numeric'
getIC(x,p)
## S4 method for signature 'pfm'
matrixReverseComplement(x)
## S4 method for signature 'pfm'
plot(x,y="missing",...)
```

Arguments

x	An object of class pfm. For getIC, if parameter p is followed, x should be an object of matrix.
y	Not use.
p	p is the background frequency.
n	how many spaces should be added.
b	logical value to indicate where the space should be added.
...	Further potential arguments passed to plotMotifLogo.

Methods

- **addBlank** signature(x="pfm", n="numeric", b="logical") add space into the position frequency matrix for alignment. b is a bool value, if TRUE, add space to the 3' end, else add space to the 5' end. n indicates how many spaces should be added.
- **getIC** signature(x = "pfm",) Calculate information content profile for position frequency matrix.
- **getIC** signature(x = "matrix", p = "numeric") Calculate information content profile for matrix. p is the background frequency
- **matrixReverseComplement** signature(x = "pfm") change the position frequency matrix into reverse complement.
- **plot** signature(x = "pfm") Plots the sequence logo of the position frequency matrix.

Examples

```
pcm <- read.table(file.path(find.package("motifStack"), "extdata", "bin_SOLEXA.pcm"))
pcm <- pcm[,3:ncol(pcm)]
rownames(pcm) <- c("A","C","G","T")
motif <- pcm2pfm(pcm)
motif <- new("pfm", mat=motif, name="bin_SOLEXA")
getIC(motif)
matrixReverseComplement(motif)
addBlank(motif, 1, FALSE)
addBlank(motif, 3, TRUE)
as(motif,"matrix")
```

plotMotifLogo *plot sequence logo*

Description

plot amino acid or DNA sequence logo

Usage

```
plotMotifLogo(pfm, motifName, p=rep(0.25, 4), font="Helvetica-Bold",
colset=c("#00811B","#2000C7","#FFB32C","#D00001"),
xaxis=TRUE,yaxis=TRUE,xlab="position",ylab="bits",
xlcex=1.2, ylcex=1.2, ncex=1.2)
```

Arguments

pfm	a position frequency matrices
motifName	motif name
p	background possibility
font	font of logo
colset	color setting for each logo letter
xaxis	draw x-axis or not
yaxis	draw y-axis or not
xlab	x-label, do nothing if set xlab as NA

ylab	y-label, do nothing if set ylab as NA
xlcex	cex value for x-label
ylcex	cex value for y-label
ncex	cex value for motif name

Value

none

Examples

```
pcm<-matrix(runif(40,0,100),nrow=4,ncol=10)
pfm<-pcm2pfm(pcm)
rownames(pfm)<-c("A","C","G","T")
plotMotifLogo(pfm)
```

plotMotifLogoStack *plot sequence logos stack*

Description

plot sequence logos stack

Usage

```
plotMotifLogoStack(pfms, ...)
```

Arguments

pfms	a list of position frequency matrices, pfms must be a list of class pfm
...	other parameters can be passed to plotMotifLogo function

Value

none

Examples

```
pcm1<-matrix(c(0,50,0,50,
   100,0,0,0,
   0,100,0,0,
   0,0,100,0,
   0,0,0,100,
   50,50,0,0,
   0,0,50,50), nrow=4)
pcm2<-matrix(c(50,50,0,0,
   0,100,0,0,
   0,50,50,0,
   0,0,0,100,
   50,50,0,0,
   0,0,50,50), nrow=4)
rownames(pcm1)<-c("A","C","G","T")
rownames(pcm2)<-c("A","C","G","T")
pfms<-list(p1=new("pfm",mat=pcm2pfm(pcm1),name="m1"),
  p2=new("pfm",mat=pcm2pfm(pcm2),name="m2"))
pfms<-DNAmotifAlignment(pfms)
plotMotifLogoStack(pfms)
```

---

plotMotifLogoStackWithTree

plot sequence logos stack with hierarchical cluster tree

Description

plot sequence logos stack with hierarchical cluster tree

Usage

```
plotMotifLogoStackWithTree(pfms, hc, treewidth=1/8, trueDist=FALSE, ...)
```

Arguments

pfms	a list of position frequency matrices, pfms must be a list of class pfm
hc	an object of the type produced by stats::hclust
treewidth	the width to show tree
trueDist	logical flags to use hclust height or not.
...	other parameters can be passed to plotMotifLogo function

Value

none

Examples

```
pcm1<-matrix(c(0,50,0,50,
  100,0,0,0,
  0,100,0,0,
  0,0,100,0,
  0,0,0,100,
  50,50,0,0,
  0,0,50,50), nrow=4)
pcm2<-matrix(c(50,50,0,0,
  0,100,0,0,
  0,50,50,0,
  0,0,0,100,
  50,50,0,0,
  0,0,50,50), nrow=4)
pcm3<-matrix(c(50,50,0,50,
  0,100,0,0,
  0,50,50,0,
  0,0,0,100,
  50,50,0,0,
  0,0,50,50), nrow=4)
rownames(pcm1)<-c("A","C","G","T")
rownames(pcm2)<-c("A","C","G","T")
rownames(pcm3)<-c("A","C","G","T")
pfms<-list(p1=new("pfm",mat=pcm2pfm(pcm1),name="m1"),
  p2=new("pfm",mat=pcm2pfm(pcm2),name="m2"),
  p3=new("pfm",mat=pcm2pfm(pcm3),name="m3"))
pfms<-DNAmotifAlignment(pfms)
dis<-function(mj,nj){
sum((mj-nj)^2)
}
mdis<-function(m,n){
d<-0
for(j in 1:ncol(m)) d<-d+dis(m[,j],n[,j])
d
}
m<-matrix(nrow=3,ncol=3,dimnames=list(c("p1","p2","p3"),c("p1","p2","p3")))
for(i in 1:3)
for(j in 1:3)
m[i,j]<-mdis(pfms[[i]]@mat,pfms[[j]]@mat)
d<-dist(m)
hc<-hclust(d)
plotMotifLogoStackWithTree(pfms, hc=hc)
```

plotXaxis *plot x-axis*

Description

plot x-axis for the sequence logo

Usage

```
plotXaxis(pfm, p=rep(0.25, 4))
```

Arguments

pfm	position frequency matrices
p	background possibility

Value

none

plotYaxis *plot y-axis*

Description

plot y-axis for the sequence logo

Usage

```
plotYaxis(pfm)
```

Arguments

pfm	position frequency matrices

Value

none

---

readPCM *read pcm from a path*

Description

read position count matrix from a path

Usage

```
readPCM(path=".", pattern=NULL)
```

Arguments

path	a character vector of full path names
pattern	an optional regular expression

Value

A list of position count matrix

Examples

```
pcms<-readPCM(file.path(find.package("motifStack"), "extdata"),"pcm$")
```

Index