Package ‘RNAmodR.AlkAnilineSeq’

May 4, 2024

Type Package
Title Detection of m7G, m3C and D modification by AlkAnilineSeq
Version 1.18.0
Date 2021-01-12
Description RNAmodR.AlkAnilineSeq implements the detection of m7G, m3C and D modifications on RNA from experimental data generated with the AlkAnilineSeq protocol. The package builds on the core functionality of the RNAmodR package to detect specific patterns of the modifications in high throughput sequencing data.
biocViews Software, WorkflowStep, Visualization, Sequencing
License Artistic-2.0
Encoding UTF-8
LazyData false
Depends R (>= 4.0), RNAmodR (>= 1.5.3)
Imports methods, S4Vectors, IRanges, BiocGenerics, GenomicRanges, Gviz
Suggests BiocStyle, knitr, rmarkdown, testthat, rtracklayer,
Biostings, RNAmodR.Data
‘Modifier-AlkAnilineSeq-viz.R’
VignetteBuilder knitr
RoxygenNote 7.1.1
BugReports https://github.com/FelixErnst/RNAmodR.AlkAnilineSeq/issues
URL https://github.com/FelixErnst/RNAmodR.AlkAnilineSeq
git_url https://git.bioconductor.org/packages/RNAmodR.AlkAnilineSeq
git_branch RELEASE_3_19
git_last_commit 24499e8
git_last_commit_date 2024-04-30
Repository Bioconductor 3.19
Date/Publication 2024-05-03
Description

7-methyl guanosine (m7G), 3-methyl cytidine (m3C) and Dihydrouridine (D) are commonly found in rRNA and tRNA and can be detected classically by primer extension analysis. However, since the modifications do not interfere with Watson-Crick base pairing, a specific chemical treatment is employed to cause strand breaks specifically at the modified positions.

This classical protocol was converted to a high throughput sequencing method called AlkAniline-Seq and allows modified position to be detected by an accumulation of 5'-ends at the N+1 position. Since the identity of the unmodified nucleotide is different for the three modified nucleotides, they modification can be detected at the same time from the same samples.

dataType is c("NormEnd3SequenceData","PileupSequenceData"):

The `ModAlkAnilineSeq` class uses the `NormEnd5SequenceData` class to store and aggregate data along the transcripts. This includes normalized values against the whole transcript (normalized cleavage) and normalized values against the overlapping reads (stop ratio), which are used to score for modified positions.

In addition the `PileupSequenceData` class is used as well, to check, whether the base is can be called according to the expected sequence identity.

Only samples named treated are used for this analysis. Normalization to untreated samples is currently not used.

Usage

```r
ModAlkAnilineSeq(x, annotation = NA, sequences = NA, seqinfo = NA, ...)
```

```r
ModSetAlkAnilineSeq(x, annotation = NA, sequences = NA, seqinfo = NA, ...)
```
Arguments

- **x**
  the input which can be of the different types depending on whether a `ModRibomethSeq` or a `ModSetRibomethSeq` object is to be constructed. For more information have a look at the documentation of the `Modifier` and `ModifierSet` classes.

- **annotation**
  annotation data, which must match the information contained in the BAM files. This is parameter is only required if x if not a `Modifier` object.

- **sequences**
  sequences matching the target sequences the reads were mapped onto. This must match the information contained in the BAM files. This is parameter is only required if x if not a `Modifier` object.

- **seqinfo**
  An optional `Seqinfo` argument or character vector, which can be coerced to one, to subset the sequences to be analyzed on a per chromosome basis.

... Optional arguments overwriting default values, which are

- **minLength**
  The minimal read length to be used for the analysis (default: `minLength = 9L`).

- **minSignal**
  The minimal signal at the position as integer value (default: `minSignal = 10L`). If the reaction is very specific a lower value may need to be used.

- **minScoreNC**
  minimum for score (normalized cleavage) to identify m7G, m3C and D positions de novo (default: `minScoreNC = 50L`)

- **minScoreSR**
  minimum for score (stop ration) to identify m7G, m3C and D positions de novo (default: `minScoreSR = 0.5`)

- **minScoreBaseScore**
  minimum score for base calling (0.0-1.0) (default: `minScoreBaseScore = 0.9`)

- **scoreOperator**
  how the minimal score should be used as logical operator. "&" requires all minimal values to be exceeded, whereas "|" detects positions, if at least one minimal values is exceeded (default: `scoreOperator = "&"`).

- other arguments which are passed on to `End5SequenceData`

Value

- a `ModAlkAnilineSeq` or `ModSetAlkAnilineSeq` object

Author(s)

Felix G.M. Ernst [aut]

References

Examples

```r
library(RNAmodR.Data)
library(rtracklayer)
annotation <- GFF3File(RNAmodR.Data.example.AAS.gff3())
sequences <- RNAmodR.Data.example.AAS.fasta()
files <- list("wt" = c(treated = RNAmodR.Data.example.wt.1()),
               "Bud23del" = c(treated = RNAmodR.Data.example.bud23.1()),
               "Trm8del" = c(treated = RNAmodR.Data.example.trm8.1()))

# Creating a Modifier object of type ModRiboMethSeq
maas <- ModAlkAnilineSeq(files[[1]], annotation = annotation,
                          sequences = sequences)

# Creating a ModifierSet object of type ModSetRiboMethSeq
msaas <- ModSetAlkAnilineSeq(files, annotation = annotation,
                             sequences = sequences)
```

Description

All of the functions of `Modifier` and the `ModifierSet` classes are inherited by the `ModAlkAnilineSeq` and `ModSetAlkAnilineSeq` classes.

Usage

```r
## S4 replacement method for signature 'ModAlkAnilineSeq'
settings(x) <- value

## S4 method for signature 'ModAlkAnilineSeq'
aggregateData(x)

## S4 method for signature 'ModAlkAnilineSeq'
findMod(x)

## S4 method for signature 'ModAlkAnilineSeq'
getDataTrack(x, name, type, ...)

## S4 method for signature 'ModAlkAnilineSeq,GRanges'
plotDataByCoord(
    x,
    coord,
    type = c("ends", "scoreNC", "scoreSR"),
    window.size = 15L,
    ...
)
```
ModAlkAnilineSeq-functions

```r
## S4 method for signature 'ModAlkAnilineSeq'
plotData(
  x,
  name,
  from = 1L,
  to = 30L,
  type = c("ends", "scoreNC", "scoreSR"),
  ...
)
```

```r
## S4 method for signature 'ModSetAlkAnilineSeq,GRanges'
plotDataByCoord(
  x,
  coord,
  type = c("scoreNC", "scoreSR", "ends"),
  window.size = 15L,
  ...
)
```

```r
## S4 method for signature 'ModSetAlkAnilineSeq'
plotData(
  x,
  name,
  from = 1L,
  to = 30L,
  type = c("scoreNC", "scoreSR", "ends"),
  ...
)
```

**Arguments**

- `x` a `Modifier` or a `ModifierSet` object. For more details see also the man pages for the functions mentioned below.

- `value` See `settings`

- `coord, name, from, to, type, window.size,...` See `plotData`.

**Details**

`ModAlkAnilineSeq` specific arguments for `plotData`:

- `colour` - a named character vector of length = 4 for the colours of the individual histograms. The names are expected to be c("scoreNC","scoreSR")

**Value**

- `settings` See `settings`.
- `aggregate` See `aggregate`.
• modify See modify.
• getDataTrack a list of DataTrack object.
• plotData See plotDataByCoord.
• plotDataByCoord See plotDataByCoord.

Examples

data(msaas, package="RNAmodR.AlkAnilineSeq")
maas <- msaas[[1]]
settings(maas)
aggregate(maas)
modify(maas)
getDataTrack(maas, "1", mainScore(maas))
Index

* datasets
  RNAmodR.AlkAnilineSeq-datasets, 6
  aggregate, 5
  aggregate (ModAlkAnilineSeq-functions), 4
  aggregateData, ModAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4
  AlkAnilineSeq (ModAlkAnilineSeq), 2
  DataTrack, 6
  End5SequenceData, 3
  findMod, ModAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4
  getDataTrack, ModAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4
  ModAlkAnilineSeq, 2
  ModAlkAnilineSeq-class
    (ModAlkAnilineSeq), 2
  ModAlkAnilineSeq-functions, 4
  Modifier, 3–6
  ModifierSet, 3–5
  modify, 6
  modify (ModAlkAnilineSeq-functions), 4
  ModSetAlkAnilineSeq (ModAlkAnilineSeq), 2
  ModSetAlkAnilineSeq-class
    (ModAlkAnilineSeq), 2
  msaas (RNAmodR.AlkAnilineSeq-datasets), 6
  NormEnd5SequenceData, 2
  PileupSequenceData, 2
  plotData, ModAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4
  plotData, ModSetAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4
  plotDataByCoord, 6
  plotDataByCoord
    (ModAlkAnilineSeq-functions), 4
  plotDataByCoord, ModAlkAnilineSeq, GRanges-method
    (ModAlkAnilineSeq-functions), 4
  plotDataByCoord, ModSetAlkAnilineSeq, GRanges-method
    (ModAlkAnilineSeq-functions), 4
  RNAmodR.AlkAnilineSeq, 6
  RNAmodR.AlkAnilineSeq-datasets, 6
  Seqinfo, 3
  settings, 5
  settings (ModAlkAnilineSeq-functions), 4
  settings<-, ModAlkAnilineSeq-method
    (ModAlkAnilineSeq-functions), 4

7