

Package ‘RNAmoDR.AlkAnilineSeq’

March 30, 2021

Type Package

Title Detection of m7G, m3C and D modification by AlkAnilineSeq

Version 1.4.0

Date 2020-02-28

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 (>= 3.6), RNAmoDR

Imports methods, S4Vectors, IRanges, BiocGenerics, GenomicRanges, Gviz

Suggests BiocStyle, knitr, rmarkdown, testthat, rtracklayer,
Biostrings, RNAmoDR.Data

Collate 'RNAmoDR.AlkAnilineSeq.R' 'Modifier-AlkAnilineSeq-class.R'
'Modifier-AlkAnilineSeq-viz.R'

VignetteBuilder knitr

RoxygenNote 7.0.2

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_12

git_last_commit a2e2b37

git_last_commit_date 2020-10-27

Date/Publication 2021-03-29

Author Felix G.M. Ernst [aut, cre] (<<https://orcid.org/0000-0001-5064-0928>>),
Denis L.J. Lafontaine [ctb, fnd]

Maintainer Felix G.M. Ernst <felix.gm.ernst@outlook.com>

R topics documented:

ModAlkAnilineSeq	2
ModAlkAnilineSeq-functions	4
RNAmodR.AlkAnilineSeq	6
RNAmodR.AlkAnilineSeq-datasets	6

Index	7
--------------	----------

ModAlkAnilineSeq	<i>ModAlkAnilineSeq class to analyze AlkAnilineSeq data</i>
------------------	---

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 call AlkAnilineSeq and allows modified position be detected by an accumulation of 5'-ends at the N+1 position. Since the identify 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

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

```
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 parameter is only required if x is 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 parameter is only required if x is 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: `minScoreSR = 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

- Marchand V, Ayadi L, __Ernst FGM__, Hertler J, Bourguignon-Igel V, Galvanin A, Kotter A, Helm M, __Lafontaine DLJ__, Motorin Y (2018): "AlkAniline-Seq: Profiling of m7 G and m3 C RNA Modifications at Single Nucleotide Resolution." *Angewandte Chemie (International ed. in English)* 57 (51), P. 16785–16790. DOI: [10.1002/anie.201810946](https://doi.org/10.1002/anie.201810946).

Examples

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

ModAlkAnilineSeq-functions

Functions for ModAlkAnilineSeq

Description

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

Usage

```
## 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,
  ...
)

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

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

```
## 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="RNAmoR.AlkAnilineSeq")  
maas <- msaas[[1]]  
settings(maas)  
aggregate(maas)  
modify(maas)  
getDataTrack(maas, "1", mainScore(maas))
```

RNAmodR.AlkAnilineSeq *RNAmodR.AlkAnilineSeq*

Description

'RNAmodR.AlkAnilineSeq' implements the detection of 7-methyl guanosine, 3-methyl cytidine and dihydrouridine from AlkAnilineSeq data using the workflow and class the package 'RNAmodR' provides.

Author(s)

Felix G M Ernst [aut], Denis L J Lafontaine [fnd]

See Also

Further details are described in the man pages of the [Modifier](#) object and the vignettes.

RNAmodR.AlkAnilineSeq-datasets

Example data in the RNAmodR.AlkAnilineSeq package

Description

This contains an example ModifierSet object of type ModSetAlkAnilineSeq

Usage

```
data(msaas)
```

Format

a ModSetAlkAnilineSeq instance

Index

- * **datasets**
 - RNAmodR.AlkAnilineSeq-datasets, 6
- aggregate, 5
- aggregate (ModAlkAnilineSeq-functions), 4
- aggregateData, ModAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4
- AlkAnilineSeq (ModAlkAnilineSeq), 2

- DataTrack, 5

- End5SequenceData, 3

- findMod, ModAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4

- getDataTrack, ModAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4

- ModAlkAnilineSeq, 2
- ModAlkAnilineSeq-class (ModAlkAnilineSeq), 2
- ModAlkAnilineSeq-functions, 4
- Modifier, 2, 4–6
- ModifierSet, 2, 4, 5
- modify, 5
- modify (ModAlkAnilineSeq-functions), 4
- ModSetAlkAnilineSeq (ModAlkAnilineSeq), 2
- ModSetAlkAnilineSeq-class (ModAlkAnilineSeq), 2
- msaas (RNAmodR.AlkAnilineSeq-datasets), 6

- NormEnd5SequenceData, 2

- PileupSequenceData, 2
- plotData, 5
- plotData (ModAlkAnilineSeq-functions), 4
- plotData, ModAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4
- plotData, ModSetAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4
- plotDataByCoord, 5
- 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, 2
- settings, 5
- settings (ModAlkAnilineSeq-functions), 4
- settings<-, ModAlkAnilineSeq-method (ModAlkAnilineSeq-functions), 4