Package 'Anaquin'

April 14, 2017

Type Package

Title Statistical analysis of sequins

Version 1.0.0

Date 2016-10-05

Author Ted Wong

Maintainer Ted Wong `<t.wong@garvan.org.au>`

Description The project is intended to support the use of sequins (synthetic sequencing spike-in controls) owned and made available by the Garvan Institute of Medical Research. The goal is to provide a standard open source library for quantitative analysis, modelling and visualization of spike-in controls.

License BSD_3_clause + file LICENSE

VignetteBuilder knitr

URL www.sequin.xyz

Depends R (>= 3.3)

Imports ggplot2, ROCR, knitr, qvalue, locfit, methods, stats, utils, plyr

Suggests RUnit, rmarkdown

BugReports https://github.com/student-t/RAnaquin/issues

LazyData true

biocViews DifferentialExpression, Preprocessing, RNASeq, GeneExpression, Software

NeedsCompilation no

R topics documented:

---

analysis   *Return type of analysis*

Description

Return type of analysis from data set created by `AnaquinData`

Usage

```
analysis(object, ...)

## S4 method for signature 'AnaquinData'
analysis(object)
```

Arguments

`...`	Not used
`object`	S4 object created by `AnaquinData`

Details

This function takes an `AnaquinData` object and return the type of analysis.

Value

Type of analysis, eg: `PlotLinear`

Examples

```
# Sequin names
names <- row.names(UserGuideData_5.4.6.3)

#Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKM
measured <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                            seqs=names,
                          input=input,
                       measured=measured)
analysis(anaquin)
```

AnaquinData	*Create Anaquin dataset*

Description

Create an Anaquin dataset. The resulting S4 object is required for every Anaquin analysis.

Usage

```
AnaquinData(analysis, ...)
```

Arguments

analysis   Type of analysis to be performed (eg. `'PlotLinear'`, `'PlotLogistic'`, `'PlotROC'`, `'PlotLODR'`.)

...   Analysis specific data inputs.

The function requires the following mandatory data inputs:

analysis	Type of analysis
seqs	Sequin names
...	Analysis specific data inputs

`analysis` indicates the type of the analysis. The current release supports the following analysis:

PlotLinear	Linear model with sequins
PlotLogistic	GLM logistic model with sequins
PlotROC	ROC analysis with sequins
PlotLODR	LODR (LOESS) analysis with sequins

`seqs` gives the sequin names. The function will give an error message unless both `analysis` and `seqs` are given.

The function accepts the following optional data inputs:

std	Standard deviation
pval	P-value probability
qval	Q-value probability
ratio	Expected sequin ratio
input	Input concentration (attomol/ul)
measured	Measured variable (eg: FPKM)
label	Classified labels (eg: 'TP', 'FP')
score	Value used for ranking sequins

Details

Create an Anaquin dataset. The resulting R-object is required for every Anaquin analysis.

Value

An S4 object of class `AnaquinData`.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
#
# Example 1: Create Anaquin dataset for PlotLogistic
#

data(UserGuideData_5.4.5.1)

title <- 'Assembly Plot'
xlab  <- 'Input Concentration (log2)'
ylab  <- 'Sensitivity'

# Sequin names
seqs <- row.names(UserGuideData_5.4.5.1)

# Input concentration
input <- log2(UserGuideData_5.4.5.1$InputConcent)

# Measured sensitivity
measured <- UserGuideData_5.4.5.1$Sn

anaquin <- AnaquinData(analysis='PlotLogistic',
                           seqs=seqs,
                         input=input,
                    measured=measured)

plotLogistic(anaquin, title=title, xlab=xlab, ylab=ylab, showLOA=TRUE)

#
# Example 2: Create Anaquin dataset for PlotLinear
#

data(UserGuideData_5.4.6.3)
```

```
title <- 'Gene Expression'
xlab  <- 'Input Concentration (log2)'
ylab  <- 'FPKM (log2)'

# Sequin names
seqs <- row.names(UserGuideData_5.4.6.3)

# Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKKM
measured <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                             seqs=seqs,
                           input=input,
                       measured=measured)

plotLinear(anaquin, title=title, xlab=xlab, ylab=ylab, showLOQ=TRUE)

#
# Example 3: Create Anaquin dataset for plotROC
#

data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# How ROC points are ranked (scoring function)
score <- 1-UserGuideData_5.6.3$Pval

# Classified labels (TP/FP)
label <- UserGuideData_5.6.3$Label

anaquin <- AnaquinData(analysis='PlotROC',
                             seqs=seqs,
                           ratio=ratio,
                           score=score,
                           label=label)

plotROC(anaquin, title='ROC Plot', refRats=0)

#
# Example 4: Create Anaquin dataset for plotLODR
#

data(UserGuideData_5.6.3)

xlab  <- 'Average Counts'
ylab  <- 'P-value'
title <- 'LODR Curves'

# Sequin names
```

```
seqs <- row.names(UserGuideData_5.6.3)

# Measured average mean
measured <- UserGuideData_5.6.3$Mean

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# P-value
pval <- UserGuideData_5.6.3$Pval

# Q-value
qval <- UserGuideData_5.6.3$Qval

anaquin <- AnaquinData(analysis='PlotLODR',
                           seqs=seqs,
                       measured=measured,
                          ratio=ratio,
                           pval=pval,
                           qval=qval)

plotLODR(anaquin, xlab=xlab, ylab=ylab, title=title, FDR=0.1)
```

---

input   *Return input concentration*

Description

Return input concentarion from data set created by `AnaquinData`.

Usage

```
input(object, ...)
## S4 method for signature 'AnaquinData'
input(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return the input concentration in the data set.

Value

List of input concentration for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.4.6.3)

# Sequin names
names <- row.names(UserGuideData_5.4.6.3)

# Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKM
measured <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                              seqs=names,
                            input=input,
                        measured=measured)
input(anaquin)
```

---

label   *Return classified labels*

Description

Return classified labels from data set created by `AnaquinData`.

Usage

```
label(object, ...)
## S4 method for signature 'AnaquinData'
label(object)
```

Arguments

object   S4 object created by `AnaquinData`

...   Not used

Details

This function takes an `AnaquinData` object and return the classified labels.

Value

List of classified labels for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# How the ROC curves are ranked
score <- 1-UserGuideData_5.6.3$Pval

# Classified labels (TP/FP)
label <- UserGuideData_5.6.3$Label

anaquin <- AnaquinData(analysis='PlotROC',
                              seqs=seqs,
                          ratio=ratio,
                          score=score,
                          label=label)
label(anaquin)
```

---

measured   *Return measured abundance*

Description

Return measured abundance from data set created by `AnaquinData`.

Usage

```
measured(object, ...)
## S4 method for signature 'AnaquinData'
measured(object)
```

Arguments

object   S4 object created by `AnaquinData`

...   Not used

Details

This function takes an `AnaquinData` object and return the measured abundance.

Value

List of measured abundance for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.4.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.4.6.3)

# Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKM
FPKM <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                             seqs=seqs,
                           input=input,
                        measured=FPKM)
measured(anaquin)
```

---

mixtureA   *RnaQuin mixture A*

Description

RnaQuin sequins are combined together across a range of concentration to formulate a mixture for emulating expression level. The concentration of each sequin within a mixture is specificed in a CSV mixture file.

This is a staggered mixture (A) for RnaQuin, user can also download the file directly from s3.amazonaws.com/sequins/

Usage

```
data(mixtureA)
```

Format

Data frame for each RnaQuin sequin:

- Length: Length in nucleotide bases
- MXA: Input concentration in attomol/ul

Value

A data frame with columns defined in Format.

---

mixtureB   *RnaQuin mixture B*

Description

RnaQuin sequins are combined together across a range of concentration to formulate a mixture for emulating expression level. The concentration of each sequin within a mixture is specificed in a CSV mixture file.

This is a staggered mixture (B) for RnaQuin, user can also download the file directly from `s3.amazonaws.com/sequins/`

Usage

```
data(mixtureB)
```

Format

Data frame for each RnaQuin sequin:

- Length: Length in nucleotide bases
- MXB: Input concentration in attomol/ul

Value

A data frame with columns defined in `Format`.

---

mixtureF   *RnaQuin mixture F*

Description

RnaQuin sequins are combined together across a range of concentration to formulate a mixture for emulating expression level. The concentration of each sequin within a mixture is specificed in a CSV mixture file.

This is a flat mixture (F) for RnaQuin, user can also download the file directly from `s3.amazonaws.com/sequins/mixtur`

Usage

```
data(mixtureF)
```

Format

Data frame for each RnaQuin sequin:

- Length: Length in nucleotide bases
- MXF: Input concentration in attomol/ul

Value

A data frame with columns defined in `Format`.

---

plotLinear   *Plot linear model for sequins*

Description

Create a scatter plot with expected abundance on the x-axis, and measured abundance on the y-axis.

Usage

```
plotLinear(data, title, xlab, ylab, showSD, showLOQ, xBreaks, yBreaks, showAxis, ...)
```

Arguments

data	Anaquin dataset created by `AnaquinData`. It needs to define information in **Details**.
title	Label of the plot.
xlab	Label for the x-axis
ylab	Label for the y-axis
xBreaks	Breaks for the x-axis
yBreaks	Breaks for the y-axis
showSD	Show standard deviation bars vertically? Default to TRUE.
showLOQ	Show limit-of-quantification? Default to TRUE.
showAxis	Show x-axis and y-axis? Default to TRUE.
...	Reserved for internal testing

Details

plotLinear requires the following data inputs from AnaquinData.

seqs	List of sequin identifiers (eg. R2_11_2).
input	Input concentration of sequins in attomol/ul.
measured	Measured abundance of sequins (typically FPKM).

The plotLinear function plots a scatter plot with expected abundance on the x-axis, and measured abundance on the y-axis. The expected abundance is typically the input concentration of sequins in the mixture, although other measures (such as expected allele frequency) are also possible. The function builds a linear regression between the two variables, and reports associated statistics (R2, correlation and regression parameters) on the plot.

The function also estimates limit-of-quantification (LOQ) breakpoint, and reports it on the plot if found. LOQ is defined as the lowest empirical detection limit, a threshold value beyond which stochastic behavior occur. LOQ is estimated by fitting segmented linear regression with two segments on the entire data set, while minimizing the total sum of squares of the differences between the variables.

Value

The functions does not return anything but it prints a scatter plot.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
library(Anaquin)

#
# Data set generated by Cufflinks and Anaquin. described in Section 5.4.6.3 of
# the user guide.
#
data(UserGuideData_5.4.6.3)

title <- 'Gene Expression'
xlab  <- 'Input Concentration (log2)'
ylab  <- 'FPKM (log2)'

# Sequin names
seqs <- row.names(UserGuideData_5.4.6.3)

# Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKM
measured <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                             seqs=seqs,
                           input=input,
                        measured=measured)

plotLinear(anaquin, title=title, xlab=xlab, ylab=ylab, showLOQ=TRUE)
```

---

plotLODR   *Create Limit-of-Detection Ratio (LODR) plot*

Description

Create a Limit-of-Detection Ratio (LODR) plot between measured abundance (x-axis) and p-value probability (y-axis).

Usage

```
plotLODR(data, FDR, title, xlab, ylab, legTitle, showConf, ...)
```

Arguments

data	Anaquin dataset created by AnaquinData. It needs to define information in **Details**.
FDR	Chosen false-discovery-rate. Default to NULL.
title	Label of the plot. Default to NULL.
xlab	Label for the x-axis. Default to NULL.

ylab	Label for the y-axis. Default to NULL.
legTitle	Title for the legend. Default to 'Ratio'.
showConf	Show confidence interval? Default to FALSE.
...	Reserved for internal testing

Details

plotLODR requires the following data inputs from AnaquinData.

seqs	List of sequin identifiers (eg. R2_11_2).
measured	Measured abundance (eg: average counts, DP field in a VCF file etc)
ratio	Expected ratio; eg: expected log-fold ratio or expected allele frequency etc
pval	P-value probability

Create a Limit-of-Detection Ratio (LODR) plot between measured abundance (x-axis) and p-value probability (y-axis).

The LODR plot indicates the confidence in measurement relative to the magnitude of the measurement. For example, p-value should converge to zero as the sequencing depth increases.

The function also fits non-parametric curves for each sequin ratio group. The curves are modelled with local regression analysis, and are colored by the sequin group.

plotLODR is an amendment from the LODR code in the ERCC dashboard R-package. Further details on the statistical algorithm is available in the ERCC documentation at https://bioconductor.org/packages/release/bioc

Value

The functions does not return anything but it prints a LODR plot.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
library(Anaquin)

#
# Data set generated by DESeq2 and Anaquin. described in Section 5.6.3.3 of
# the user guide.
#
data(UserGuideData_5.6.3)

xlab  <- 'Average Counts'
ylab  <- 'P-value'
title <- 'LODR Curves'

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# Measured average abundance
```

```
measured <- UserGuideData_5.6.3$Mean

# P-value
pval <- UserGuideData_5.6.3$Pval

# Q-value
qval <- UserGuideData_5.6.3$Qval

anaquin <- AnaquinData(analysis='PlotLODR',
                          seqs=seqs,
                     measured=measured,
                        ratio=ratio,
                         pval=pval,
                         qval=qval)

plotLODR(anaquin, xlab=xlab, ylab=ylab, title=title, FDR=0.1)
```

plotLogistic	*Plot logistic model for sequins*

Description

Create a scatter plot with input concentration on the x-axis, and measured proportion on the y-axis.

Usage

```
plotLogistic(data, title, xlab, ylab, showLOA, threshold, ...)
```

Arguments

data	Anaquin dataset created by `AnaquinData`. It needs to define information in **Details**.
title	Title of the plot. (Default to NULL).
xlab	Label for the x-axis. (Default to NULL).
ylab	Label for the y-axis. (Default to NULL).
showLOA	Display limit-of-assembly. (Default to TRUE).
threshold	Threshold required for limit-of-assembly (LOA)
...	Reserved for internal testing

Details

`plotLogistic` requires the following data inputs from `AnaquinData`.

seqs	List of sequin identifiers (eg. R2_11_2).
input	Input concentration of sequins in attomol/ul.
measured	Measured variable (eg: sensitivity).

The `plotLogistic` function creates a scatter plot with input concentration on the x-axis, and measured proportion on the y-axis. Common measured statistics include p-value, percentage and sensi-

tivity. The plot builds a GLM logistic regression model between the two variables.

The function also estimates limit-of-assembly (LOA) breakpoint, and reports it on the plot if found. The LOA breakpoint is an empirical detection limit, and also the abundance whereby the fitted logistic curve exceeds a user-defined threshold.

Value

The functions does not return anything but it prints a scatter plot with a fitted logistic function.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
library(Anaquin)

#
# Data set generated by Cufflinks and Anaquin. described in Section 5.4.5.1 of
# the user guide.
#
data(UserGuideData_5.4.5.1)

title <- 'Assembly Plot'
xlab  <- 'Input Concentration (log2)'
ylab  <- 'Sensitivity'

# Sequin names
seqs <- row.names(UserGuideData_5.4.5.1)

# Input concentration
input <- log2(UserGuideData_5.4.5.1$InputConcent)

# Measured sensitivity
measured <- UserGuideData_5.4.5.1$Sn

anaquin <- AnaquinData(analysis='PlotLogistic',
                          seqs=seqs,
                         input=input,
                      measured=measured)

plotLogistic(anaquin, title=title, xlab=xlab, ylab=ylab, showLOA=TRUE)
```

---

plotROC   *Create ROC plot*

Description

Create a receiver operating characteristic (ROC) plot at various threshold settings.

Usage

```
plotROC(data, refRats, title, legTitle, ...)
```

Arguments

`data`	Anaquin dataset created by `AnaquinData`. It needs to define information in **Details**.
`refRats`	Reference ratio groups
`title`	Label of the plot. Default to `NULL`.
`legTitle`	Title of the legend. Default to `Ratio`.
`...`	Reserved for internal testing

Details

`plotROC` requires the following data inputs from `AnaquinData`.

`seqs`	List of sequin identifiers (eg. R2_11_2)
`label`	Classified labels ('TP' or 'FP')
`score`	How the ROC points should be ranked
`ratio`	Expected ratio; eg: expected log-fold ratio

Create a receiver operating characteristic (ROC) plot at various threshold settings. The true positive rate (TPR) is plotted on the x-axis and false positive rate (FPR) is plotted on the y-axis.

The function requires a scoring threshold function, and illustrates the performance of the data as the threshold is varied. Common scoring threshold include p-value, sequencing depth and allele frequency, etc.

ROC plot is a useful diagnostic performance tool; it provides tools to select possibly optimal models and to discard suboptimal ones. In particularly, the AUC statistics indicate the performance of the model relatively to a random experiment (AUC 0.5).

Value

The functions does not return anything but it prints a ROC plot and it's AUC statistics.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
library(Anaquin)

#
# Data set generated by DESeq2 and Anaquin. described in Section 5.6.3.3 of
# the user guide.
#
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# How the ROC curves are ranked
```

```
score <- 1-UserGuideData_5.6.3$Pval

# Classified labels (TP/FP)
label <- UserGuideData_5.6.3$Label

anaquin <- AnaquinData(analysis='PlotROC',
                             seqs=seqs,
                          ratio=ratio,
                          score=score,
                          label=label)

plotROC(anaquin, title='ROC Plot', refRats=0)
```

---

pval   *Return p-value probability*

Description

Return p-value probability from data set created by `AnaquinData`.

Usage

```
pval(object, ...)
## S4 method for signature 'AnaquinData'
pval(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return it's p-value probability.

Value

List of p-value proability for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC
```

```
# Measured average abundance
measured <- UserGuideData_5.6.3$Mean

# P-value
pv <- UserGuideData_5.6.3$Pval

# Q-value
qv <- UserGuideData_5.6.3$Qval

anaquin <- AnaquinData(analysis='PlotLODR',
                               seqs=seqs,
                         measured=measured,
                            ratio=ratio,
                             pval=pv,
                             qval=qv)
pval(anaquin)
```

---

qval   *Return q-value probability*

Description

Return q-value probability from data set created by `AnaquinData`.

Usage

```
qval(object, ...)
## S4 method for signature 'AnaquinData'
qval(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return it's q-value probability.

Value

List of q-value proability for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# Measured average abundance
measured <- UserGuideData_5.6.3$Mean

# P-value
pv <- UserGuideData_5.6.3$Pval

# Q-value
qv <- UserGuideData_5.6.3$Qval

anaquin <- AnaquinData(analysis='PlotLODR',
                             seqs=seqs,
                       measured=measured,
                          ratio=ratio,
                           pval=pv,
                           qval=qv)
qval(anaquin)
```

---

ratio	*Return ratio*

Description

Return ratio from data set created by `AnaquinData`.

Usage

```
ratio(object, ...)
## S4 method for signature 'AnaquinData'
ratio(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` data set object and return its sequin ratio.

Value

List of sequin ratio for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# How the ROC curves are ranked
score <- 1-UserGuideData_5.6.3$Pval

# Classified labels (TP/FP)
label <- UserGuideData_5.6.3$Label

anaquin <- AnaquinData(analysis='PlotROC',
                        seqs=seqs,
                       ratio=ratio,
                       score=score,
                       label=label)
ratio(anaquin)
```

RnaQuin.gen2iso   *Convert sequin isoforms to genes*

Description

Convert RnaQuin sequin isoforms to genes.

Usage

```
RnaQuin.gen2iso(names)
```

Arguments

names   Sequin isoform names

Details

For example:

R1_11_1 to R1_11 R1_11_2 to R1_11 R1_12_2 to R1_12

Value

A list of RnaQuin sequin genes.

Author(s)

Ted Wong <t.wong@garvan.org.au>

Examples

```
data(UserGuideData_5.4.5.1)
RnaQuin.gen2iso(row.names(UserGuideData_5.4.5.1))
```

RnaQuin.genes   *Return RnaQuin genes*

Description

Return RnaQuin genes from data set created by `AnaquinData`.

Usage

```
RnaQuin.genes(object, ...)
## S4 method for signature 'AnaquinData'
RnaQuin.genes(object)
```

Arguments

object   S4 object created by `AnaquinData`

...   Not used

Details

This function takes an `AnaquinData` object and return the RnaQuin genes identifiers.

Value

Data frame with the genes and their input concentration. Return `NULL` if unavailable.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
library(Anaquin)

#
# Data set generated by Cufflinks and Anaquin. described in Section 5.4.5.1 of
# the user guide.
#
data(UserGuideData_5.4.5.1)

data <- UserGuideData_5.4.5.1

# Sequin names
seqs <- row.names(data)

# Input concentration
input <- log2(data$InputConcent)

# Measured sensitivity
```

```
measured <- data$Sn

anaquin <- AnaquinData(analysis='PlotLogistic',
                                seqs=seqs,
                              input=input,
                          measured=measured)

RnaQuin.genes(anaquin)
```

RnaQuin.isoforms   *Return RnaQuin isoforms*

Description

Return RnaQuin isoforms from data set created by `AnaquinData`.

Usage

```
RnaQuin.isoforms(object, ...)
## S4 method for signature 'AnaquinData'
RnaQuin.isoforms(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return the RnaQuin isoform identifiers.

Value

Data frame with the isoforms and their input concentration. Return `NULL` if unavailable.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
library(Anaquin)

#
# Data set generated by Cufflinks and Anaquin. described in Section 5.4.5.1 of
# the user guide.
#
data(UserGuideData_5.4.5.1)

data <- UserGuideData_5.4.5.1

# Sequin names
seqs <- row.names(data)
```

```
# Input concentration
input <- log2(data$InputConcent)

anaquin <- new("AnaquinData", analysis = 'Mixture',
                              seqs = row.names(data),
                             input = log2(data$InputConcent))
RnaQuin.isoforms(anaquin)
```

---

score   *Return scores*

Description

Return scores from data set created by `AnaquinData`.

Usage

```
score(object, ...)
## S4 method for signature 'AnaquinData'
score(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return the ROC ranking scores.

Value

List of scores for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
seqs <- row.names(UserGuideData_5.6.3)

# Expected log-fold
ratio <- UserGuideData_5.6.3$ExpLFC

# How the ROC curves are ranked
score <- 1-UserGuideData_5.6.3$Pval

# Classified labels (TP/FP)
label <- UserGuideData_5.6.3$Label
```

```
anaquin <- AnaquinData(analysis='PlotROC',
                              seqs=seqs,
                            ratio=ratio,
                            score=score,
                            label=label)
score(anaquin)
```

---

seqs   *Return sequin names*

Description

Return sequin names from `AnaquinData` data set.

Usage

```
seqs(object, ...)
## S4 method for signature 'AnaquinData'
seqs(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return the sequin names in the data set.

Value

List of sequin names.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
data(UserGuideData_5.4.6.3)

# Sequin names
names <- row.names(UserGuideData_5.4.6.3)

# Input concentration
input <- log2(UserGuideData_5.4.6.3$InputConcent)

# Measured FPKM
measured <- log2(UserGuideData_5.4.6.3[,2:4])

anaquin <- AnaquinData(analysis='PlotLinear',
                              seqs=names,
```

```
                              input=input,
                        measured=measured)
    seqs(anaquin)
```

---

std   *Return standard deviation*

Description

Return standard deviation from `AnaquinData` data set.

Usage

```
std(object, ...)
## S4 method for signature 'AnaquinData'
std(object)
```

Arguments

object	S4 object created by `AnaquinData`
...	Not used

Details

This function takes an `AnaquinData` object and return the standard deviation in the data set.

Value

List of standard deviation for each sequin. Return `NULL` if unavailable.

Author(s)

Ted Wong `<t.wong@garvan.org.au>`

Examples

```
data(UserGuideData_5.6.3)

# Sequin names
names <- row.names(UserGuideData_5.6.3)

# Expected LFC
input <- log2(UserGuideData_5.6.3$ExpLFC)

# Measured LFC
measured <- log2(UserGuideData_5.6.3$ObsLFC)

# Standard deviation
sd <- UserGuideData_5.6.3$SD

anaquin <- AnaquinData(analysis='PlotLinear',
                            seqs=names,
                          input=input,
                            std=sd,
```

```
                                      measured=measured)
std(anaquin)
```

UserGuideData_5.4.5.1   *Section 5.4.5.1 Assembly Dataset*

Description

Assembly sensitivity estimated by Cuffcompare. `Section 5.4.5.1` of the Anaquin user guide has details on the data set.

Usage

```
data(UserGuideData_5.4.5.1)
```

Format

Data frame:

- InputConcent: Input concentration in attomol/ul
- Sn: Measured sensitivity

Value

Data frame with columns defined in `Format`.

Source

S.A Hardwick. Spliced synthetic genes as internal controls in RNA sequencing experiments. `Nature Methods`, 2016.

UserGuideData_5.4.6.3   *Gene expression (RnaQuin)*

Description

Gene expression estimated by Cufflinks. `Section 5.4.6.3` of the Anaquin user guide has details on the data set.

Usage

```
data(UserGuideData_5.4.6.3)
```

Format

Data frame:

- InputConcent: Input concentration in attomol/ul
- Observed1: Measured FPKM for the first replicate
- Observed2: Measured FPKM for the second replicate
- Observed3: Measured FPKM for the third replicate

Value

Data frame with columns defined in `Format`.

Source

S.A Hardwick. Spliced synthetic genes as internal controls in RNA sequencing experiments. `Nature Methods`, 2016.

`UserGuideData_5.6.3`   *Differential expression (RnaQuin)*

Description

Differential gene expression estimated by DESeq2. `Section 5.6.3` has details on the data set.

Usage

```
data(UserGuideData_5.6.3)
```

Format

Data frame:

- ExpLFC: Expected log-fold change
- ObsLFC: Observed log-fold change
- SD: Standard deviation of the measurment
- Pval: P-value probability
- Qval: Q-value probability
- Mean: Average counts across the samples
- Label: Average counts across the samples

Value

Data frame with columns defined in `Format`.

Source

S.A Hardwick. Spliced synthetic genes as internal controls in RNA sequencing experiments. `Nature Methods`, 2016.

Index