

Package ‘Anaquin’

October 17, 2017

Type Package

Title Statistical analysis of sequins

Version 1.2.0

Date 2016-11-24

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), ggplot2 (>= 2.2.0)

Imports ggplot2, ROCR, knitr, qvalue, locfit, methods, stats, utils,
plyr, DESeq2

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	2
AnaquinData	3
input	6
label	7
measured	8
plotLinear	9
plotLODR	11
plotLogistic	12
plotMA	14

plotROC	15
pval	16
qval	17
ratio	18
RnaQuin.aggregate	19
RnaQuin.genes	20
RnaQuin.iso2gen	21
RnaQuin.isoforms	22
RnaQuinMixtureA	22
RnaQuinMixtureB	23
score	24
seqs	25
std	26
UserGuideData_5.4.5.1	27
UserGuideData_5.4.6.3	27
UserGuideData_5.6.3	28

Index 29

analysis	<i>Return type of analysis</i>
----------	--------------------------------

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 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)

#
# 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=ylob, title=title, FDR=0.1)
```

input

Return input concentration

Description

Return input concentration 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	<i>Return classified labels</i>
-------	---------------------------------

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)

```

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, errors,
           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.
errors	How errors bar should be calculated. SD or Range
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 (R^2 , 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 sensitivity. 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=ylob, showLOA=TRUE)
```

plotMA

Create MA plot

Description

Create an MA plot between the M (log ratios) and A (mean average) for endogenous features and sequins.

Usage

```
plotMA(data, ...)
```

Arguments

data	Anaquin dataset created by AnaquinData. It needs to define information in Details .
...	Reserved for internal testing

Details

plotMA requires the following data inputs from AnaquinData.

seqs	List of sequin identifiers (eg. R2_11_2).
mean	Log mean average
lfc	Log-fold ratio

Create an MA plot between the M (log ratios) and A (mean average) for endogenous features and sequins.

This function can used for differential isoform/gene analysis. The distribution of the sequins are compared with the samples.

Value

The functions returns a 'ggplot2 object for the plot.

Author(s)

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

Examples

```
library(Anaquin)

data('UserGuideData_5.6.3')
data <- AnaquinData(analysis='PlotMA',
                    seqs=row.names(UserGuideData_5.6.3),
                    mean=log2(UserGuideData_5.6.3$Mean),
                    lfc=UserGuideData_5.6.3$ObsLFC)

plotMA(data)
```

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 particular, 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 probability 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 probability 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.aggregate

Aggregate isoform concentration

Description

Aggregate isoform concentration for genes

Usage

```
RnaQuin.aggregate(isos, concent)
```

Arguments

```
isos          List of RnaQuin isoform names
concent       List of input concentration for each RnaQuin isoform
```

Details

This function provides a convenient mechanism to add up the RnaQuin input concentration for isoforms to the gene level. For example, if we have the following isoforms: R1_101_1: 10 attomol/ul and R1_101_2: 15 attomol/ul.

R1_101_1 and R1_101_2 forms the splice variants of a gene. Their input concentration would add up to 25 attomol/ul for R1_101.

Value

Data frame with two columns. The first column gives the RnaQuin gene names. The second column gives the aggregated gene level input concentration attomol/ul.

Author(s)

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

Examples

```
data('RnaQuinMixtureA')

# Aggregate input concentration for isoforms into genes
RnaQuin.aggregate(RnaQuinMixtureA$ID, RnaQuinMixtureA$MixA)
```

RnaQuin.genes	<i>Filter RnaQuin genes</i>
---------------	-----------------------------

Description

Filter RnaQuin genes from a list of genes names.

Usage

```
RnaQuin.genes(genes)
```

Arguments

```
genes          List of character strings for the gene names
```

Details

This function provides a convenient mechanism to filter RnaQuin genes from a list of gene names. A most common usage is for filtering a gene expression table.

Value

List of RnaQuin genes from the input. If no sequin gene is found, the function returns an empty list.

Author(s)

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

Examples

```
# This could also come from a count matrix table
x <- c('R1_11', 'R1_12')

RnaQuin.genes(x)
```

RnaQuin.iso2gen

Convert RnaQuin isoforms to genes

Description

Convert RnaQuin isoform names to gene names.

Usage

```
RnaQuin.iso2gen(isos)
```

Arguments

`isos` List of character strings for the isoform names

Details

This function provides a convenient mechanism to convert RnaQuin isoform names to gene names.

Value

List of RnaQuin gene names for each input isoform.

Author(s)

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

Examples

```
# This could also come from a count matrix table
x <- c('R1_1_1', 'R1_1_2', 'R1_1_3', 'R1_2_1')

# Convert RnaQuin isoforms to genes
RnaQuin.iso2gen(x)
```

RnaQuin.isoforms *Filter RnaQuin isoforms*

Description

Filter RnaQuin isoforms from a list of isoforms names.

Usage

```
RnaQuin.isoforms(isos)
```

Arguments

isos List of character strings for the isoform names

Details

This function provides a convenient mechanism to filter RnaQuin isoforms from a list of isoform names. A most common usage is for filtering a isoform expression table.

Value

List of RnaQuin isoforms from the input. If no sequin isoforms is found, the function returns an empty list.

Author(s)

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

Examples

```
# This could also come from a count matrix table
x <- c('R1_11_1', 'This_Is_Not_Sequin', 'R1_12_2')

RnaQuin.isoforms(x)
```

RnaQuinMixtureA *RnaQuin mixture A*

Description

Individual sequins are combined across a range of precise concentrations to formulate mixtures. By modulating the concentration at which each sequin is present in the mixture, we can emulate quantitative features of genome biology.

This is the mixture A in RnaQuin.

Usage

```
data(RnaQuinMixtureA)
```

Format

Data frame:

- ID: Sequin name
- Length: Sequin length
- MixA: Input concentration

Value

Data frame with columns defined in Format.

RnaQuinMixtureB	<i>RnaQuin mixture B</i>
-----------------	--------------------------

Description

Individual sequins are combined across a range of precise concentrations to formulate mixtures. By modulating the concentration at which each sequin is present in the mixture, we can emulate quantitative features of genome biology.

This is the mixture B in RnaQuin.

Usage

```
data(RnaQuinMixtureB)
```

Format

Data frame:

- ID: Sequin name
- Length: Sequin length
- MixA: Input concentration

Value

Data frame with columns defined in Format.

seqs	<i>Return sequin names</i>
------	----------------------------

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	<i>Return standard deviation</i>
-----	----------------------------------

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 measurement
- 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

analysis, [2](#)
analysis, AnaquinData-method (analysis),
[2](#)
AnaquinData, [3](#)

input, [6](#)
input, AnaquinData-method (input), [6](#)

label, [7](#)
label, AnaquinData-method (label), [7](#)

measured, [8](#)
measured, AnaquinData-method (measured),
[8](#)

plotLinear, [9](#)
plotLODR, [11](#)
plotLogistic, [12](#)
plotMA, [14](#)
plotROC, [15](#)
pval, [16](#)
pval, AnaquinData-method (pval), [16](#)

qval, [17](#)
qval, AnaquinData-method (qval), [17](#)

ratio, [18](#)
ratio, AnaquinData-method (ratio), [18](#)
RnaQuin.aggregate, [19](#)
RnaQuin.genes, [20](#)
RnaQuin.iso2gen, [21](#)
RnaQuin.isoforms, [22](#)
RnaQuinMixtureA, [22](#)
RnaQuinMixtureB, [23](#)

score, [24](#)
score, AnaquinData-method (score), [24](#)
seqs, [25](#)
seqs, AnaquinData-method (seqs), [25](#)
std, [26](#)
std, AnaquinData-method (std), [26](#)

UserGuideData_5.4.5.1, [27](#)
UserGuideData_5.4.6.3, [27](#)
UserGuideData_5.6.3, [28](#)