

Package ‘Rqc’

October 9, 2015

Type Package

Title Quality Control Tool for High-Throughput Sequencing Data

Version 1.2.0

Author Welliton Souza, Benilton Carvalho <beniltoncarvalho@gmail.com>

Maintainer Welliton Souza <well1309@gmail.com>

Description Rqc is an optimised tool designed for quality control and assessment of high-throughput sequencing data. It performs parallel processing of entire files and produces a report which contains a set of high-resolution graphics.

License GPL (>= 2)

Depends BiocParallel, ShortRead, ggplot2

Imports BiocGenerics, Biostrings, IRanges, methods, S4Vectors, knitr (>= 1.7), BiocStyle, plyr, markdown, grid, reshape2

VignetteBuilder knitr

biocViews Sequencing, QualityControl, DataImport

URL <https://github.com/labbc/Rqc>

NeedsCompilation no

R topics documented:

| | |
|-------------------------------|----|
| Rqc-package | 2 |
| perCycleBasecall | 3 |
| perCycleQuality | 3 |
| perFileInformation | 4 |
| perReadQuality | 5 |
| perReadWidth | 6 |
| rqc | 6 |
| rqcCycleAverageQualityCalc | 8 |
| rqcCycleAverageQualityPcaCalc | 9 |
| rqcCycleAverageQualityPcaPlot | 9 |
| rqcCycleAverageQualityPlot | 10 |

| | |
|-------------------------------------|----|
| rqcCycleBaseCallsCalc | 11 |
| rqcCycleBaseCallsLinePlot | 12 |
| rqcCycleBaseCallsPlot | 12 |
| rqcCycleGCCalc | 13 |
| rqcCycleGCPlot | 14 |
| rqcCycleQualityBoxCalc | 15 |
| rqcCycleQualityBoxPlot | 16 |
| rqcCycleQualityCalc | 16 |
| rqcCycleQualityPlot | 17 |
| rqcQA | 18 |
| rqcReadQualityBoxCalc | 19 |
| rqcReadQualityBoxPlot | 20 |
| rqcReadQualityCalc | 21 |
| rqcReadQualityPlot | 21 |
| rqcReadWidthCalc | 22 |
| rqcReadWidthPlot | 23 |
| rqcReport | 24 |
| RqcResultSet-class | 25 |
| stats4trim | 25 |

| | |
|--------------|-----------|
| Index | 27 |
|--------------|-----------|

Rqc-package

Quality Control Tool for High-Throughput Sequencing Data

Description

Rqc is an optimized tool designed for quality assessment of high-throughput sequencing data. It performs parallel processing of entire files and produces a report, which contains a set of high-resolution images that can be directly used on publications.

Author(s)

Welliton Souza, Benilton Carvalho

Maintainer: Welliton Souza <well309@gmail.com>

Examples

```
options(device.ask.default = FALSE)
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
rqc(folder, ".fastq.gz")
```

| | |
|------------------|-----------------------------------------------------------|
| perCycleBasecall | <i>Frequency distribution of cycle-specific base call</i> |
|------------------|-----------------------------------------------------------|

Description

This function combines frequency tables of all RqcResultSet objects in one table and also add file name information.

Usage

```
perCycleBasecall(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by `rqc` and `rqcQA` functions.

Value

A data.frame with 'cycle', 'base', 'count' and 'filename' columns. 'count' variable represents how many a 'base' was identified on determined 'cycle'.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- perCycleBasecall(rqcResultSet)
head(df)
qplot(x=base, y=count, fill=base, data=df, geom="bar", stat="identity",
      facets=~filename)
```

| | |
|-----------------|---------------------------------------------------------|
| perCycleQuality | <i>Frequency distribution of cycle-specific quality</i> |
|-----------------|---------------------------------------------------------|

Description

This function combines frequency tables of all RqcResultSet objects in one table and also add file name information.

Usage

```
perCycleQuality(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by `rqc` and `rqcQA` functions.

Value

A data.frame with 'cycle', 'quality', 'score', 'count' and 'filename' columns. 'count' variable represents how many a 'quality/score' was identified on determined 'cycle'.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- perCycleQuality(rqcResultSet)
head(df)
qplot(x=quality, y=count, data=df, geom="bar", stat="identity",
      facets=~filename)
```

perFileInformation *File information*

Description

This function combines file information (such as name and path) of all RqcResultSet objects in one data.frame.

Usage

```
perFileInformation(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by `rqc` and `rqcQA` functions.

Value

A data.frame containing information about files.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
perFileInformation(rqcResultSet)
```

| | |
|----------------|--------------------------------------------------------|
| perReadQuality | <i>Frequency distribution of per read mean quality</i> |
|----------------|--------------------------------------------------------|

Description

This function returns a frequency distribution of per read mean quality from all RqcResultSet objects included on input list.

Usage

```
perReadQuality(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

A data.frame with 'average', 'count' and 'filename' columns. 'count' variable represents how many the 'average' values are found.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- perReadQuality(rqcResultSet)
head(df)
# create a histogram
qplot(x=average, weight=count, data=df, geom="histogram", binwidth=1,
      facets=~filename)
```

| | |
|--------------|---------------------------------------------|
| perReadWidth | <i>Frequency distribution of read width</i> |
|--------------|---------------------------------------------|

Description

This function combines frequency table of all RqcResultSet objects in one table and also add file name information.

Usage

```
perReadWidth(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by `rqc` and `rqcQA` functions.

Value

A data.frame with 'width', 'count' and 'filename' columns. 'count' variable represents how many reads have the same 'width'.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- perReadWidth(rqcResultSet)
head(df)
```

| | |
|-----|--------------------------|
| rqc | <i>Main Rqc function</i> |
|-----|--------------------------|

Description

Rqc is an optimized tool designed for quality assessment of high-throughput sequencing data. It performs parallel processing of entire files and produces an HTML report, which contains a set of high-resolution images that can be directly used on publications.

Usage

```
rqc(path=".", pattern,  
     sample=TRUE, n=1e6,  
     groupFactor = rep("None", length(files)),  
     outdir=tempdir(), file="rqc_report",  
     openBrowser=TRUE,  
     BPPARAM=bpparam())
```

Arguments

| | |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| path | directory path that contains input files. |
| pattern | a regex expression that matches to input file names |
| sample | it reads a random sample from files if this parameter is TRUE. |
| n | number of sequences to read from each input file. This represents sample size if 'sample' parameter is TRUE, if not represents the chunk size to read on each iteration. By default, it reads a sample of one million sequences from each input file. |
| groupFactor | group name for each input file. |
| outdir | output directory path. Is created a temporary directory by default. |
| file | output file name. |
| openBrowser | if TRUE opens report file on default Internet Browser. |
| BPPARAM | An instance of a BiocParallelParam. See bpparam . |

Value

A invisible named list of RqcResultSet objects, each one represents a file.

Author(s)

Welliton Souza

See Also

[rqcQA](#)

Examples

```
options(device.ask.default = FALSE)  
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")  
rqc(folder, ".fastq.gz", openBrowser=FALSE)
```

rqcCycleAverageQualityCalc

Per cycle average quality calculation

Description

Calculates average quality per cycle.

Usage

```
rqcCycleAverageQualityCalc(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Data frame ready for plot.

Author(s)

Welliton Souza

See Also

[rqcCycleAverageQualityPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcCycleAverageQualityCalc(rqcResultSet)
df$cycle <- as.numeric(levels(df$cycle))[df$cycle]
qplot(x=cycle, y=quality, colour=filename, data=df, geom="line")
```

`rqcCycleAverageQualityPcaCalc`*PCA calculation of per cycle read average quality*

Description

This function calculates Principal Components of per cycle read average quality.

Usage

```
rqcCycleAverageQualityPcaCalc(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

An object of class `prcomp`.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
fit <- rqcCycleAverageQualityPcaCalc(rqcResultSet)
biplot(fit)
```

`rqcCycleAverageQualityPcaPlot`*Biplot of PCA of per cycle read average quality*

Description

This function creates a Biplot of PCA of per cycle read average quality

Usage

```
rqcCycleAverageQualityPcaPlot(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleAverageQualityPcaCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleAverageQualityPcaPlot(rqcResultSet)
```

`rqcCycleAverageQualityPlot`

Per cycle average quality plot

Description

Creates a line graph of per cycle average quality.

Usage

```
rqcCycleAverageQualityPlot(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleAverageQualityCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleAverageQualityPlot(rqcResultSet)
```

`rqcCycleBaseCallsCalc` *Per cycle base call calculation*

Description

Calculates average base call per cycle.

Usage

```
rqcCycleBaseCallsCalc(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

List of data frames ready for plot.

Author(s)

Welliton Souza

See Also

[rqcCycleBaseCallsPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcCycleBaseCallsCalc(rqcResultSet)
qplot(x=cycle, y=value, colour=variable, group=variable, data=df,
      geom="line", facets=~filename)
```

rqcCycleBaseCallsLinePlot

Per cycle base calls line plot

Description

Creates a line graph of per cycle base calls.

Usage

```
rqcCycleBaseCallsLinePlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleBaseCallsCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleBaseCallsLinePlot(rqcResultSet)
```

rqcCycleBaseCallsPlot *Per cycle base calls plot*

Description

Creates a bar graph of per cycle base calls.

Usage

```
rqcCycleBaseCallsPlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleBaseCallsCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleBaseCallsPlot(rqcResultSet)
```

| | |
|----------------|---------------------------------|
| rqcCycleGCCalc | <i>Per cycle percentual GC.</i> |
|----------------|---------------------------------|

Description

Calculates per cycle percentual GC.

Usage

```
rqcCycleGCCalc(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Data frame ready for plot.

Author(s)

Welliton Souza

See Also

[rqcCycleGCPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcCycleGCCalc(rqcResultSet)
qplot(x=cycle, y=gc, colour=filename, data=df, geom="line")
```

rqcCycleGCPlot

Per cycle percentual GC plot

Description

Creates a line graph of per cycle percentual GC.

Usage

```
rqcCycleGCPlot(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

Plot object from `ggplot` function.

Author(s)

Welliton Souza

See Also

[rqcCycleGCCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleGCPlot(rqcResultSet)
```

`rqcCycleQualityBoxCalc`*Per cycle quality percentiles calculation for boxplot*

Description

Calculates per cycle quality percentiles to create boxplot.

Usage

```
rqcCycleQualityBoxCalc(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

Data frame ready to plot.

Author(s)

Welliton Souza

See Also

[rqcCycleQualityBoxPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcCycleQualityBoxCalc(rqcResultSet)
qplot(x=cycle, ymin=ymin, lower=lower, middle=middle, upper=upper, ymax=ymax,
      data=df, geom="boxplot", stat="identity", facets=~filename)
```

rqcCycleQualityBoxPlot

Per cycle quality box plot

Description

Plots per cycle quality box plot.

Usage

```
rqcCycleQualityBoxPlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleQualityBoxCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleQualityBoxPlot(rqcResultSet)
```

rqcCycleQualityCalc *Per cycle quality calculation*

Description

Calculates per cycle quality percentiles.

Usage

```
rqcCycleQualityCalc(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Data frame ready for plot.

Author(s)

Welliton Souza

See Also

[rqcCycleQualityPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcCycleQualityCalc(rqcResultSet)
qplot(x=cycle, y=percentiles, fill=value, data=df, geom="bar",
      stat="identity", facets=~filename)
```

`rqcCycleQualityPlot` *Per cycle quality plot*

Description

Creates a graph of per cycle quality.

Usage

```
rqcCycleQualityPlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcCycleQualityCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcCycleQualityPlot(rqcResultSet)
```

rqcQA

Quality Assessment Rqc function

Description

Process a set of files and returns a list of quality control data. Files must be FASTQ format, compressed or not.

Usage

```
rqcQA(files, sample = TRUE, n = 1e6,
       groupFactor = rep("None", length(files)), BPPARAM = bpparam())
```

Arguments

| | |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| files | A vector of file paths. |
| sample | It reads a random sample from files if this parameter is TRUE. |
| n | Number of sequences to read from each input file. This represents sample size if 'sample' parameter is TRUE, if not represents the chunk size to read on each iteration. Default is read a sample of one million sequences from each input file. |
| groupFactor | group name for each input file. |
| BPPARAM | An instance of a BiocParallelParam. See bpparam . |

Details

Input files are read using `FastStreamer` and `FastSampler` classes of [ShortRead](#) package. Process multiple files in parallel using `bplapply` function of [BiocParallel](#) package.

Value

A named list of `RqcResultSet` objects, each one represents a file.

Author(s)

Welliton Souza

See Also

[rqc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(full.names=TRUE, path=folder)
rqcResultSet <- rqcQA(files)
rqcReadQualityPlot(rqcResultSet)
```

rqcReadQualityBoxCalc *Per read mean quality percentiles calculation for boxplot*

Description

Calculates per read mean quality percentiles to create boxplot.

Usage

```
rqcReadQualityBoxCalc(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by [rqc](#) and [rqcQA](#) functions.

Value

A data frame ready to plot, it contains the following columns:

- filename: name of input files
- min: minimum value
- ymin: 10% of data
- lower: 25% of data
- middle: 50% of data, median
- upper: 75% of data
- ymax: 90% of data
- max: maximum value

Author(s)

Welliton Souza

See Also

[rqcReadQualityBoxPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcReadQualityBoxCalc(rqcResultSet)
p <- qplot(x=filename, ymin=ymin, lower=lower, middle=middle, upper=upper,
          ymax=ymax, data=df, geom="boxplot", stat="identity") +
  coord_flip()

p
# add minimum and maximum values
p + geom_point(aes(y=min)) + geom_point(aes(y=max))
```

`rqcReadQualityBoxPlot` *Per read mean quality box plot*

Description

This function creates a graphic charts with box plots describing per read mean quality distribution for each input file

Usage

```
rqcReadQualityBoxPlot(rqcResultSet)
```

Arguments

`rqcResultSet` list of `RqcResultSet` objects created by `rqc` and `rqcQA` functions.

Value

Plot object from `ggplot` function.

Author(s)

Welliton Souza

See Also

[rqcReadQualityBoxCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcReadQualityBoxPlot(rqcResultSet)
```

rqcReadQualityCalc *Per read quality calculation*

Description

Calculates per read quality

Usage

```
rqcReadQualityCalc(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Data frame ready for plot.

Author(s)

Welliton Souza

See Also

[rqcReadQualityPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcReadQualityCalc(rqcResultSet)
qplot(x=quantile, y=value, colour=filename, data=df, geom="line")
```

rqcReadQualityPlot *Per read quality plot*

Description

Plots the quality of all the files by read.

Usage

```
rqcReadQualityPlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcReadQualityCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcReadQualityPlot(rqcResultSet)
```

`rqcReadWidthCalc` *Per read width calculation*

Description

Calculates amount of per read width

Usage

```
rqcReadWidthCalc(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Data frame ready for plot.

Author(s)

Welliton Souza

See Also

[rqcReadWidthPlot](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
df <- rqcReadWidthCalc(rqcResultSet)
qplot(x=width, y=count, data=df, geom="bar", stat="identity", facets=~filename)
```

| | |
|------------------|----------------------------|
| rqcReadWidthPlot | <i>Per read width plot</i> |
|------------------|----------------------------|

Description

Creates bar graph of per read width from all elements of input list.

Usage

```
rqcReadWidthPlot(rqcResultSet)
```

Arguments

rqcResultSet list of RqcResultSet objects created by [rqc](#) and [rqcQA](#) functions.

Value

Plot object from [ggplot](#) function.

Author(s)

Welliton Souza

See Also

[rqcReadWidthCalc](#)

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
rqcReadWidthPlot(rqcResultSet)
```

`rqcReport`*Quality Control HTML Report*

Description

Generates an HTML report file.

Also creates a directory called "figure" in outdir path.

Usage

```
rqcReport(rqcResultSet, outdir=tempdir(), file="rqc_report", keepMD=FALSE)
```

Arguments

| | |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>rqcResultSet</code> | list of RqcResultSet objects created by rqc and rqcQA functions. |
| <code>outdir</code> | output directory path. It is created a temporary directory by default. |
| <code>file</code> | output file name. |
| <code>keepMD</code> | If true Rqc does not delete markdown file. knit function takes RMarkdown template file (within package) and generates a temporary Markdown file. Next markdownToHTML function takes this markdown file and creates final HTML file. |

Value

Report file path.

Author(s)

Welliton Souza

See Also

[rqc](#)

[rqcQA](#)

Examples

```
options(device.ask.default = FALSE)
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(full.names=TRUE, path=folder)
rqcResultSet <- rqcQA(files)
reportFile <- rqcReport(rqcResultSet)
browseURL(reportFile)
```

| | |
|--------------------|-----------------------------|
| RqcResultSet-class | <i>Quality control data</i> |
|--------------------|-----------------------------|

Description

RqcResultSet class represents quality assessment data created by `rqc` and `rqcQA` functions. This class extends `.QA` class from `ShortRead` package.

See Also

`perReadQuality` returns read mean quality frequency table.
`perReadWidth` returns read width frequency table.
`perCycleQuality` returns cycle-specific quality frequency table.
`perCycleBasecall` returns cycle-specific base call frequency table.
`perFileInformation` returns informations of the files.

| | |
|-------------------------|------------------------------------------------------------------------------------------------------|
| <code>stats4trim</code> | <i>Minimum read mean quality and maximum percentage loss of reads estimations for trimming step.</i> |
|-------------------------|------------------------------------------------------------------------------------------------------|

Description

This function estimates how many reads would be lost if the sequences are filtered by a minimum read mean quality value. Also this function estimates what is the minimum read mean quality value for filtering and lose max percentage defined.

Usage

```
stats4trim(rqcResultSet, qmin, pmax)
```

Arguments

| | |
|---------------------------|--------------------------------------------------------------------------------------------|
| <code>rqcResultSet</code> | list of RqcResultSet objects created by <code>rqc</code> and <code>rqcQA</code> functions. |
| <code>qmin</code> | Minimum read mean quality value (between 0 and 41). |
| <code>pmax</code> | Maximum percentage of reads permitted been lost during trimming step. |

Value

A data frame containg estimated minimum quality and maximum percentage for each input file.

Author(s)

Welliton Souza

Examples

```
folder <- system.file(package="ShortRead", "extdata/E-MTAB-1147")
files <- list.files(folder, "fastq.gz", full.names=TRUE)
rqcResultSet <- rqcQA(files)
stats4trim(rqcResultSet, qmin=20)
stats4trim(rqcResultSet, pmax=10)
```

Index

*Topic **graphics**

- [rqc](#), [6](#)
- [rqcCycleAverageQualityPcaPlot](#), [9](#)
- [rqcCycleAverageQualityPlot](#), [10](#)
- [rqcCycleBaseCallsLinePlot](#), [12](#)
- [rqcCycleBaseCallsPlot](#), [12](#)
- [rqcCycleGCPlot](#), [14](#)
- [rqcCycleQualityBoxPlot](#), [16](#)
- [rqcCycleQualityPlot](#), [17](#)
- [rqcReadQualityBoxPlot](#), [20](#)
- [rqcReadQualityPlot](#), [21](#)
- [rqcReadWidthPlot](#), [23](#)

*Topic **package**

- [Rqc-package](#), [2](#)

*Topic **qc**

- [perCycleBasecall](#), [3](#)
- [perCycleQuality](#), [3](#)
- [perFileInformation](#), [4](#)
- [perReadQuality](#), [5](#)
- [perReadWidth](#), [6](#)
- [rqc](#), [6](#)
- [rqcCycleAverageQualityCalc](#), [8](#)
- [rqcCycleAverageQualityPcaCalc](#), [9](#)
- [rqcCycleAverageQualityPcaPlot](#), [9](#)
- [rqcCycleAverageQualityPlot](#), [10](#)
- [rqcCycleBaseCallsCalc](#), [11](#)
- [rqcCycleBaseCallsLinePlot](#), [12](#)
- [rqcCycleBaseCallsPlot](#), [12](#)
- [rqcCycleGCCalc](#), [13](#)
- [rqcCycleGCPlot](#), [14](#)
- [rqcCycleQualityBoxCalc](#), [15](#)
- [rqcCycleQualityBoxPlot](#), [16](#)
- [rqcCycleQualityCalc](#), [16](#)
- [rqcCycleQualityPlot](#), [17](#)
- [rqcReadQualityBoxCalc](#), [19](#)
- [rqcReadQualityBoxPlot](#), [20](#)
- [rqcReadQualityCalc](#), [21](#)
- [rqcReadQualityPlot](#), [21](#)
- [rqcReadWidthCalc](#), [22](#)

- [rqcReadWidthPlot](#), [23](#)

- [stats4trim](#), [25](#)

*Topic **trimming**

- [stats4trim](#), [25](#)

- [BiocParallel](#), [18](#)

- [bplapply](#), [18](#)

- [bpparam](#), [7](#), [18](#)

- [ggplot](#), [10](#), [12–14](#), [16](#), [17](#), [20](#), [22](#), [23](#)

- [knit](#), [24](#)

- [markdownToHTML](#), [24](#)

- [perCycleBasecall](#), [3](#), [25](#)

- [perCycleQuality](#), [3](#), [25](#)

- [perFileInformation](#), [4](#), [25](#)

- [perReadQuality](#), [5](#), [25](#)

- [perReadWidth](#), [6](#), [25](#)

- [Rqc \(Rqc-package\)](#), [2](#)

- [rqc](#), [3–6](#), [6](#), [8–17](#), [19–25](#)

- [Rqc-package](#), [2](#)

- [rqcCycleAverageQualityCalc](#), [8](#), [10](#)

- [rqcCycleAverageQualityPcaCalc](#), [9](#), [10](#)

- [rqcCycleAverageQualityPcaPlot](#), [9](#)

- [rqcCycleAverageQualityPlot](#), [8](#), [10](#)

- [rqcCycleBaseCallsCalc](#), [11](#), [12](#), [13](#)

- [rqcCycleBaseCallsLinePlot](#), [12](#)

- [rqcCycleBaseCallsPlot](#), [11](#), [12](#)

- [rqcCycleGCCalc](#), [13](#), [14](#)

- [rqcCycleGCPlot](#), [13](#), [14](#)

- [rqcCycleQualityBoxCalc](#), [15](#), [16](#)

- [rqcCycleQualityBoxPlot](#), [15](#), [16](#)

- [rqcCycleQualityCalc](#), [16](#), [18](#)

- [rqcCycleQualityPlot](#), [17](#), [17](#)

- [rqcQA](#), [3–17](#), [18](#), [19–25](#)

- [rqcReadQualityBoxCalc](#), [19](#), [20](#)

- [rqcReadQualityBoxPlot](#), [19](#), [20](#)

- [rqcReadQualityCalc](#), [21](#), [22](#)

rqcReadQualityPlot, [21](#), [21](#)
rqcReadWidthCalc, [22](#), [23](#)
rqcReadWidthPlot, [22](#), [23](#)
rqcReport, [24](#)
RqcResultSet (RqcResultSet-class), [25](#)
RqcResultSet-class, [25](#)

ShortRead, [18](#)
stats4trim, [25](#)