Package 'flowAI'

April 15, 2019

Type Package

Title Automatic and interactive quality control for flow cytometry data

Version 1.12.7

Date 2019-03-04

Author Gianni Monaco, Hao Chen

Description The package is able to perform an automatic or interactive quality control on FCS data acquired using flow cytometry instruments. By evaluating three different properties: 1) flow rate, 2) signal acquisition, 3) dynamic range, the quality control enables the detection and removal of anomalies.

Depends R (>= 3.4)

License GPL (>= 2)

LazyData TRUE

biocViews FlowCytometry, QualityControl, BiomedicalInformatics, ImmunoOncology

Imports ggplot2, flowCore, plyr, changepoint, knitr, reshape2, RColorBrewer, scales, methods, graphics, stats, utils, rmarkdown

Suggests testthat, shiny

Maintainer Gianni Monaco <gianni.monaco@unibas.ch>

VignetteBuilder knitr

RoxygenNote 6.1.1

NeedsCompilation no

git_url https://git.bioconductor.org/packages/flowAI

git_branch RELEASE_3_8

git_last_commit 25587df

git_last_commit_date 2019-03-04

Date/Publication 2019-04-15

R topics documented:

`Bcells` *flowSet of B cells.*

Description

This data set contain three flowFrame objects created by subsetting three FCS files of an Aging study made in Singapore. The samples were stained with a panel aimed to identify B cell subpopulations. The data is stored as a flowSet object, a class implemented in the flowCore package to handle FCS files in R.

Usage

```
data(Bcells)
```

Format

A flowSet containing 3 flowFrames

`flow_auto_qc` *Automatic quality control of flow cytometry data.*

Description

For a set of FCS files, *flow_auto_qc* performs a complete and automatic quality control. It consists in the detection and removal of anomalies by checking three properties of flow cytometry: 1) flow rate, 2) signal acquisition, 3) dynamic range.

Usage

```
flow_auto_qc(fcsfiles, remove_from = "all", output = 1,
  timeCh = NULL, second_fractionFR = 0.1, alphaFR = 0.01,
  decompFR = TRUE, ChExcludeFS = c("FSC", "SSC"),
  outlier_binsFS = FALSE, pen_valueFS = 500, max_cptFS = 3,
  ChExcludeFM = c("FSC", "SSC"), sideFM = "both", neg_valuesFM = 1,
  html_report = "_QC", mini_report = "QCmini", fcs_QC = "_QC",
  fcs_highQ = FALSE, fcs_lowQ = FALSE, folder_results = "resultsQC")
```

Arguments

`fcsfiles`	It can be a character vector with the filenames of the FCS files, a flowSet or a flowFrame.
`remove_from`	Select from which of the three steps the anomalies have to be excluded in the high quality FCS file. The default option `"all"` removes the anomalies from all the three steps. Alternatively, you can use: `"FR_FS"`, `"FR_FM"`, `"FS_FM"`, `"FR"`, `"FS"`, `"FM"`, to remove the anomalies only on a subset of the steps where *FR* stands for the flow rate, *FS* stands for signal acquisition and *FM* stands for dynamic range.

output — Set it to 1 to return a flowFrame or a flowSet with high quality events only. Set it to 2 to return a flowFrame or a flowSet with an additional parameter where the low quality events have a value higher than 10,000. Set it to 3 to return a list with the IDs of low quality cells. Set it to any other value if no R object has to be returned. Default is 1.

timeCh — Character string corresponding to the name of the Time Channel in the set of FCS files. By default is NULL and the name is retrieved automatically.

second_fractionFR — The fraction of a second that is used to split the time channel in order to recreate the flow rate. Set it to "timestep" if you wish to recreate the flow rate at the maximum resolution allowed by the flow cytometry instrument. Usually, the timestep corresponds to 0.01, however, to shorten the running time of the analysis the fraction used by default is 0.1, corresponding to 1/10 of a second.

alphaFR — The level of statistical significance used to accept anomalies detected by the ESD method. The default value is 0.01.

decompFR — Logical indicating whether the flow rate should be decomposed in the trend and cyclical components. Default is TRUE and the ESD outlier detection will be executed on the trend component penalized by the magnitude of the cyclical component. If it is FALSE the ESD outlier detection will be executed on the original flow rate.

ChExcludeFS — Character vector with the names or name patterns of the channels that you want to exclude from the signal acquisition check. The default option, c("FSC", "SSC"), excludes the scatter parameters. If you want to include all the parameters in the analysis use NULL.

outlier_binsFS — logical indicating whether outlier bins (not events) have to be removed before the changepoint detection of the signal acquisition check. The default is FALSE.

pen_valueFS — The value of the penalty for the changepoint detection algorithm. This can be a numeric value or text giving the formula to use; for instance, you can use the character string "1.5*log(n)", where n indicates the number of cells in the FCS file. The higher the penalty value the less strict is the detection of the anomalies. The default is 500.

max_cptFS — The maximum number of changepoints that can be detected for each channel. The default is 3.

ChExcludeFM — Character vector with the names or name patterns of the channels that you want to exclude from the signal acquisition check. The default option, c("FSC", "SSC"), excludes the scatter parameters. If you want to include all the parameters in the analysis use NULL.

sideFM — Select whether the dynamic range check has to be executed on both limits, the upper limit or the lower limit. Use one of the options: "both", "upper", "lower". The default is "both".

neg_valuesFM — Scalar indicating the method to use for the removal of the anomalies from the lower limit of the dynamic range. Use 1 to remove negative outliers or use 2 to truncate the negative values to the cut-off indicated in the FCS file.

html_report — Suffix to be added to the FCS filename to name the HTML report of the quality control. The default is "_QC". If you do not want to generate a report use FALSE.

mini_report — Name for the TXT file containing the percentage of anomalies detected in the set of FCS files analyzed. The default is "_QCmini". If you do not want to generate the mini report use FALSE.

fcs_QC	Suffix to be added for the filename of the new FCS containing a new parameter where the low quality events only have a value higher than 10,000. The default is "_QC". If you do not want to generate the high quality FCS file use FALSE.
fcs_highQ	Suffix to be added for the filename of the new FCS containing only the events that passed the quality control. The default is FALSE and hence the high quality FCS file is not generated.
fcs_lowQ	Suffix to be added for the filename of the new FCS containing only the events that did not pass the quality control. The default is FALSE and hence the low quality FCS file is not generated.
folder_results	Character string used to name the directory that contains the results. The default is "resultsQC". If you intend to return the results in the working directory use FALSE.

Value

A complete quality control is performed on flow cytometry data in FCS format. By default the analysis returns:

1. a flowFrame or flowSet object containing new FCS files with only high quality events

and a directory named *resultsQC* containing:

1. a set of new FCS files with a new parameter to gate out the low quality events a value larger than 10,000 is assigned to them only,

2. a set of HTML reports, one for each FCS file, that include graphs and table indicating where the anomalies were detected,

3. a single TXT file reporting the percentage of events removed in each FCS file.

Author(s)

Gianni Monaco, Chen Hao

Examples

```
## a sample dataset as flowSet object
data(Bcells)

## quality control on a flowFrame object
resQC <- flow_auto_qc(Bcells[[1]], html_report = FALSE, mini_report = FALSE, fcs_QC = FALSE, folder_results =
```

---

flow_iQC *Interactive quality control of Flow Cytometry Data*

Description

The call of the flow_iQC function opens a Shiny application that allows to perfom a complete and interactive quality control of an FCS file. The framework of the interactive quality control is complementary to the automatic one of the flow_auto_qc function. Hence, the anomalies are manually selected from the evaluation of three main properties of flow cytometry: 1) flow rate, 2) signal acquisition, 3) dynamic range.

Usage

```
flow_iQC()
```

Author(s)

Chen Hao, Gianni Monaco

Examples

```
if (interactive()) flowAI::flow_iQC()
```

Index