Package 'pmp'

June 26, 2018

Type Package

Title Peak matrix processing

Version 0.1.0

Author Ralf Weber and Andris Jankevics

Maintainer Andris Jankevics <a.jankevics@bham.ac.uk>

Description Tools and filters for peak matrix scaling, normalisation and filtering.

License GPL

biocViews

Imports impute, pcaMethods, missForest, ggplot2, scales, ggthemes, gridExtra, codetools

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

R topics documented:

`createClassAndColors` *Function to create sorted class labels and colors for reproducable Ggplot objects*

Description

Function to create sorted class labels and colors for reproducable Ggplot objects

Usage

```
createClassAndColors(class, QC_label = "QC", Blank_label = "Blank",
  QC_color = "#000000", Blank_color = "#A65628",
  manual_color = c("#386cb0", "#ef3b2c", "#7fc97f", "#fdb462", "#984ea3",
  "#a6cee3", "#778899", "#fb9a99", "#ffff33"))
```

Arguments

`class`	Vector of class labels.
`QC_label`	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
`Blank_label`	Label used for blank samples, if set to NULL no samples will be removed
`QC_color`	Color to use for QC samples
`Blank_color`	Color to use for blank samples
`manual_color`	Colors to usef for samples classes

Value

List of processed data table and RSD

`doPCA` *Perform PCA analysis ot single data set of list of data sets*

Description

Perform PCA analysis ot single data set of list of data sets

Usage

```
doPCA(Data, classes, plotTitle = "PCA", PQN, mv_impute, glogScaling,
  scale = F, labels = "QC", qc_label, qc_shape = 17, base_size = 12,
  pccomp = c(1, 2), subtitle = NULL, loadings = F, loadingsCol = NULL)
```

Arguments

Data	Data frame.
classes	Vector of class labels.
plotTitle	Character value to display as the plot title
PQN	Can be set to T or F, to perform PQN normalisation
mv_impute	T or F, indicates if missing value imputation has to be carried
glogScaling	T or F, applie glog transformation to the given data
scale	Perform UV scaling on data
labels	Can be set to "QC" to label only QC samples. "none" to no include labels. If set to any other value will use column names of Data.
qc_label	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
qc_shape	Shape symbol to use for QC samples
base_size	Font size for plot fonts
pccomp	PCA components to plot
subtitle	Subtitle to include in PCA plot
loadings	T or F, to include PCA loadings plot or not
loadingsCol	Colors to use for loadings plot

Value

Ggplot object with plot(s)

---

doRSD *Calculate RSD% values per sample group*

Description

Calculate RSD% values per sample group

Usage

```
doRSD(Data, classes)
```

Arguments

Data	Data frame
classes	Vector of class labels

Value

List of RSD% values for each variable for each sample group

doRSDplot *Plot violin plots from the doRSD fucntion output*

Description

Plot violin plots from the doRSD fucntion output

Usage

```
doRSDplot(RSD, plotTitle = NULL, base_size = 12, subtitle = NULL)
```

Arguments

RSD	OUtput of doRSD function
plotTitle	Main tilte for the plot
base_size	Font size for plot fonts
subtitle	Subtitle to include in PCA plot

Value

Ggplot object with plot(s)

doRSDtable *Plot violin plots from the doRSD fucntion output*

Description

Plot violin plots from the doRSD fucntion output

Usage

```
doRSDtable(RSD, QC_label = "QC", Blank_label = "Blank")
```

Arguments

RSD	OUtput of doRSD function
QC_label	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
Blank_label	Label used for Blank samples

Value

Table of RSD

doSummaryPlot *Wrapper function to generate PCA plot and RSD statistics plot*

Description

Wrapper function to generate PCA plot and RSD statistics plot

Usage

```
doSummaryPlot(Data, classes, plotTitle = "PCA", blank = "BLANK", PQN = F,
  mv_impute = T, glogScaling = T, scale = T, qc_label = "QC",
  ignorelabel = "Removed", output = "PCA_plot.pdf", labels = "QC",
  qc_shape = 17, base_size = 12, pccomp = c(1, 2), plot = T)
```

Arguments

Data	Data frame.
classes	Vector of class labels.
plotTitle	Title to use for plot output
blank	Label used for blank samples, if set to NULL no samples will be removed
PQN	Can be set to T or F, to perform PQN normalisation
mv_impute	T or F, indicates if missing value imputation has to be carried
glogScaling	T or F, applie glog transformation to the given data
scale	Perform UV scaling on data
qc_label	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
ignorelabel	Label for samples which should be excluded from processed data
output	File name for pdf output
labels	Can be set to "QC" to label only QC samples. "none" to no include labels. If set to any other value will use column names of Data.
qc_shape	Shape symbol to use for QC samples
base_size	Ggplot font size
pccomp	PCA components to plot
plot	If set to T will output pdf file, otherwise ggplot object

Value

Summary plot as ggplot object or directly to the pdf file.

filter_peaks_by_blank *Blank filter*

Description

Blank filter

Usage

```
filter_peaks_by_blank(df, fold_change, classes, blank_label, qc_label = NULL,
  remove = NULL)
```

Arguments

df	Input data frame.
fold_change	Threshold of fold change for blank filter.
classes	Vector of class labels.
blank_label	Class label for blank samples
qc_label	Class label for QC sample
remove	If features abobe threshold shoule be removed or not

filter_peaks_by_fraction *Filter by fraction of detected features in QC sample*

Description

Filter by fraction of detected features in QC sample

Usage

```
filter_peaks_by_fraction(df, min_frac, classes = NULL, method = "QC",
  qc_label = "QC")
```

Arguments

df	Input data frame.
min_frac	Threshold of fraction of detection
classes	Vector of class labels.
method	Method to use, default is QC
qc_label	Class label for QC sample

filter_peaks_by_rsd *Filter features by QC samples RSD*

Description

Filter features by QC samples RSD

Usage

```
filter_peaks_by_rsd(df, max_rsd, classes, qc_label)
```

Arguments

df	Input data frame.
max_rsd	Threshold of QC RSD value.
classes	Vector of class labels.
qc_label	Class label for QC sample.

filter_samples_by_mv *Missing values filter*

Description

Missing values filter

Usage

```
filter_samples_by_mv(df, max_perc_mv)
```

Arguments

df	Input data frame.
max_perc_mv	Threshold of missing value percentage.

glog *Internal glog function*

Description

Internal glog function

Usage

```
glog(y, alpha, lambda)
```

Arguments

y	values.
alpha	alpha.
lambda	lambda.

`glog_transformation` *Performs glog transformation on the data set. Using QC samples to estimate lamda and alpha parameters.*

Description

Performs glog transformation on the data set. Using QC samples to estimate lamda and alpha parameters.

Usage

```
glog_transformation(df, classes, qc_label)
```

Arguments

`df`	Data frame.
`classes`	Vector of sample classes.
`qc_label`	Lable for QC sample.

`jglog` *Internal glog function*

Description

Internal glog function

Usage

```
jglog(y, y0, lambda)
```

Arguments

`y`	values.
`y0`	.
`lambda`	lambda.

multiplot *Plot multiple ggplot objects*

Description

Plot multiple ggplot objects

Usage

```
multiplot(..., plotlist = NULL, file, cols = 1, layout = NULL)
```

Arguments

...	ggplot objects
plotlist	List of ggplot objects
file	name of output pdf files
cols	number of columns
layout	layout matrix

mv_imputation *Missing value imputation using different algorithms*

Description

Missing value imputation using different algorithms

Usage

```
mv_imputation(df, method, k = 10, rowmax = 0.5, colmax = 0.5)
```

Arguments

df	Data frame.
method	Missing value imputation method.
k	Number of neighbour values to use.
rowmax	Fraction of missing values per row.
colmax	Fraction of missing values per column.

pqn_normalisation *PQN normalisation*

Description

PQN normalisation

Usage

```
pqn_normalisation(df, classes, qc_label)
```

Arguments

df	Data frame.
classes	Vector of class labels.
qc_label	Label used for QC samples.

Value

List of normalised data set and correction coefficients

prepareData *Wrapper function to transform data for statistical analysis*

Description

Wrapper function to transform data for statistical analysis

Usage

```
prepareData(Data, classes, blank = "BLANK", PQN = F, mv_impute = T,
  glogScaling = T, qc_label = "QC", ignorelabel = "Removed",
  checkNA = T)
```

Arguments

Data	Data frame.
classes	Vector of class labels.
blank	Label used for blank samples, if set to NULL no samples will be removed
PQN	Can be set to T or F, to perform PQN normalisation
mv_impute	T or F, indicates if missing value imputation has to be carried
glogScaling	T or F, applie glog transformation to the given data
qc_label	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
ignorelabel	Label for samples which should be excluded from processed data
checkNA	removes rows, columns containing all NA's

Value

List of processed data table and RSD

reorderFactorLevels *Function to create sorted class labels from claaa names*

Description

Function to create sorted class labels from claaa names

Usage

```
reorderFactorLevels(labels, QC_label = "QC", Blank_label = "Blank")
```

Arguments

labels	Vector of class labels.
QC_label	Label used for QC samples. If set to NULL, assumes that no QC samples are present in data set
Blank_label	Label used for blank samples, if set to NULL no samples will be removed

Value

Vector of reordered class labels

scale_colour_Publication *Ggplot colors*

Description

Ggplot colors

Usage

```
scale_colour_Publication(...)
```

Arguments

...	No parameters to specify

scale_fill_Publication *Ggplot fill colors*

Description

Ggplot fill colors

Usage

```
scale_fill_Publication(...)
```

Arguments

...	No parameters to specify

set_scale *Ggplot color scales function*

Description

Ggplot color scales function

Usage

```
set_scale(y = NULL, name = "PCA", ...)
```

Arguments

y	vector or classes
name	type of plotted object
...	Other parameters

smu_normalisation *SMU normalisation*

Description

SMU normalisation

Usage

```
smu_normalisation(df)
```

Arguments

df	Data frame.

Value

Normalised data frame.

SSE *Internal glog function*

Description

Internal glog function

Usage

```
SSE(lambda, alpha, y)
```

Arguments

lambda	lambda.
alpha	alpha.
y	values.

theme_Publication *Ggplog theme, source ...*

Description

Ggplog theme, source ...

Usage

```
theme_Publication(base_size = 14)
```

Arguments

base_size	Ggplot font size

Index