Package: sights (via r-universe)

November 25, 2026

Type Package

Title Statistics and dIagnostic Graphs for HTS

Version 1.40.0

Author Elika Garg [aut, cre], Carl Murie [aut], Heydar Ensha [ctb], Robert Nadon [aut]

Maintainer Elika Garg <elika.garg@mail.mcgill.ca>

Description SIGHTS is a suite of normalization methods, statistical tests, and diagnostic graphical tools for high throughput screening (HTS) assays. HTS assays use microtitre plates to screen large libraries of compounds for their biological, chemical, or biochemical activity.

License GPL-3 | file LICENSE

URL https://eg-r.github.io/sights/

BugReports https://github.com/eg-r/sights/issues

Depends R(>= 3.3)

Imports MASS(>= 7.3), qvalue(>= 2.2), ggplot2(>= 2.0), reshape2(>= 1.4), lattice(>= 0.2), stats(>= 3.3)

LazyData TRUE

Suggests testthat, knitr, rmarkdown, ggthemes, gridExtra, xlsx

Collate 'ex_dataMatrix.R' 'inglese.R' 'internal.R' 'normZ.R' 'normRobZ.R' 'normR.R' 'normSPAWN.R' 'normLoess.R' 'normMedFil.R' 'normSights.R' 'plot3d.R' 'plotAutoco.R' 'plotBox.R' 'plotHeatmap.R' 'plotHist.R' 'plotScatter.R' 'statRVM.R' 'plotIGFit.R' 'plotSights.R' 'statFDR.R' 'statT.R' 'statSights.R'

VignetteBuilder knitr

RoxygenNote 5.0.1

biocViews ImmunoOncology, CellBasedAssays, MicrotitrePlateAssay, Normalization, MultipleComparison, Preprocessing, QualityControl, BatchEffect, Visualization

NeedsCompilation no

Config/pak/sysreqs libicu-dev

Repository Bioconductor 3.24

Date/Publication 2026-11-23 22:37:32 UTC

RemoteUrl https://github.com/bioc/sights

RemoteRef RELEASE_3_24

RemoteSha e359edf6182acda2a3574c6e40e937490c5210c2

Contents

ex_dataMatrix . . . 2
inglese . . . 4
normLoess . . . 5
normMedFil . . . 6
normR . . . 8
normRobZ . . . 9
normSights . . . 10
normSPAWN . . . 11
normZ . . . 13
plot3d . . . 14
plotAutoco . . . 15
plotBox . . . 16
plotHeatmap . . . 17
plotHist . . . 19
plotIGFit . . . 20
plotScatter . . . 21
plotSights . . . 22
statFDR . . . 24
statRVM . . . 25
statSights . . . 27
statT . . . 28

Index **30**

`ex_dataMatrix` *High-Throughput Screening example data - CMBA*

Description

An example dataset containing High-Throughput Screening (HTS) output and experimental design information. See References for details.

Usage

```
data(ex_dataMatrix)
```

Format

A data frame with 80 rows and 9 columns:

- Wells. Plate well numbers for each sample
- Rows. Plate row identifiers for each sample
- Columns. Plate column identifiers for each sample
- S1_R1. Screen 1 Replicate 1
- S1_R2. Screen 1 Replicate 2
- S1_R3. Screen 1 Replicate 3
- S2_R1. Screen 2 Replicate 1
- S2_R2. Screen 2 Replicate 2
- S2_R3. Screen 2 Replicate 3

This example data matrix consists of 6 plates with 80 wells each. Although these are 96-well plates, only 80 wells in each plate contained the active compounds. Therefore, the subsequent data matrix for this package excludes the inactive wells.

Details

The sights data format requires each plate matrix to be converted into a 1-dimensional vector. The plate wells in this vector should be arranged by row first. For example, this 3x3 plate matrix:

	Col 1	Col 2	Col 3
Row A	A1	A2	A3
Row B	B1	B2	B3
Row C	C1	C2	C3

can be converted into its vector form as:

Row	Col	Data
A	1	A1
A	2	A2
A	3	A3
B	1	B1
B	2	B2
B	3	B3
C	1	C1
C	2	C2
C	3	C3

Here, number of columns in a plate is 3, and number of rows is 3 as well. Each such plate vector should form a column in the data matrix before application of sights functions. Only the active wells should be included in the data matrix; inactive wells containing mock/control compounds should be marked as NAs, or if they are in entire rows/columns they can be removed completely as in this example dataset and the arguments plateRows and plateCols modified accordingly.

Value

Dataframe of 80 rows and 9 columns as explained in Format

References

CMBA Titration series 10uM Tyr samples. Murie et al. (2015). Improving detection of rare biological events in high-throughput screens. Journal of Biomolecular Screening, 20(2), 230-241.

Examples

```
## load dataset
data(ex_dataMatrix)

## structure of dataset
str(ex_dataMatrix)
## summary of dataset
summary(ex_dataMatrix)

## See help pages of SIGHTS functions for examples of using this dataset
```

`inglese` *High-Throughput Screening example data - Inglese*

Description

A published dataset containing High-Throughput Screening (HTS) output and experimental design information. See References for details.

Usage

```
data(inglese)
```

Format

A data frame with 1280 rows and 45 columns:

- Row. Plate row identifiers for each sample
- Col. Plate column identifiers for each sample
- Exp1R1. Screen 1 Replicate 1
- Exp1R2. Screen 1 Replicate 2
- Exp1R3. Screen 1 Replicate 3
- Exp2R1. Screen 2 Replicate 1
- Exp2R2. Screen 2 Replicate 2
- Exp2R3. Screen 2 Replicate 3

... and so on until Exp14 totaling to 14 screens in triplicate.

- Hits. Presence or absence of hits identified for each sample

Value

Dataframe of 1280 rows and 45 columns as explained in Format

Note

For information on how to arrange your dataset, please see (ex_dataMatrix)

References

Titration series samples. Inglese et al. (2006). Quantitative High-Throughput Screening: A Titration-Based Approach That Efficiently Identifies Biological Activities in Large Chemical Libraries. Proc. Natl. Acad. Sci. U. S. A., 103, 11473-11478.

Examples

```
## load dataset
data(inglese)

## structure of dataset
str(inglese)
## summary of dataset
summary(inglese)

## See SIGHTS vignette for examples of using this dataset and its anlaysis
```

---

normLoess *Normalization by loess method*

Description

Apply loess normalization to data

Usage

```
normLoess(dataMatrix, plateRows, plateCols, dataRows = NULL,
  dataCols = NULL)
```

Arguments

dataMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate.
dataRows, dataCols	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.

Details

Loess normalization adjusts each well by the fitted row and column values generated by calculating the loess curve for each row and column.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (ex_dataMatrix)

References

Baryshnikova et al. (2010). Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. Nature Methods, 7(12), 1017-1024.

See Also

Other normalization methods: normMedFil, normRobZ, normR, normSPAWN, normZ

Examples

```
## load dataset
data(ex_dataMatrix)

## apply Loess method
ex_normMatrix <- normLoess(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10)
```

---

normMedFil *Normalization by median filter method*

Description

Apply median filter normalization to data

Usage

```
normMedFil(dataMatrix, plateRows, plateCols, dataRows = NULL,
  dataCols = NULL, seqFilter = TRUE)
```

Arguments

`dataMatrix`	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
`plateRows, plateCols`	Number of rows/columns in plate.
`dataRows, dataCols`	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.
`seqFilter`	Optional logical. If TRUE apply initial row median filter then standard filter, else just apply standard filter.

Details

Median Filter normalization uses a two-step median filter process where each well is adjusted by the median score of a neighbouring group of wells [Bushway et al (2011)]. The first median filter uses a neighbour set based on the Manhattan distance to each well. The second median filter uses a neighbour set based on the proximity along each row or column.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (`ex_dataMatrix`)

References

Bushway et al.(2011). Optimization and application of median filter corrections to relieve diverse spatial patterns in microtiter plate data. Journal of Biomolecular Screening, 16(9), 1068-1080.

See Also

Other normalization methods: `normLoess`, `normRobZ`, `normR`, `normSPAWN`, `normZ`

Examples

```
## load dataset
data(ex_dataMatrix)

## apply standard median filter method
ex_normMatrix <- normMedFil(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10, seqFilter = FALSE)
## apply initial row median filter then standard filter
ex_normMatrix <- normMedFil(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10, seqFilter = TRUE)
```

normR *Normalization by R score method*

Description

Apply Robust Regression model separately to each plate

Usage

```
normR(dataMatrix, plateRows, plateCols, dataRows = NULL, dataCols = NULL)
```

Arguments

dataMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate.
dataRows, dataCols	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.

Details

R score normalization uses the robust regression method described by Wu et al (2008). Parameters are estimated through the rlm function. Data is pre-normalized by median normalization prior to applying the regression algorithm. R scores are the residuals produced by the model and rescaled by dividing with the standard deviation estimate from the regression function.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (ex_dataMatrix)

References

Wu et al. (2008). Quantitative Assessment of Hit Detection and Confirmation in Single and Duplicate High-Throughput Screenings. Journal of Biomolecular Screening, 13(2), 159-167.

See Also

Other normalization methods: normLoess, normMedFil, normRobZ, normSPAWN, normZ

Examples

```
## load dataset
data(ex_dataMatrix)

## apply R score
ex_normMatrix <- normR(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10)
```

---

normRobZ *Normalization by robust Z score method*

Description

Apply robust Z score to data

Usage

```
normRobZ(dataMatrix, dataRows = NULL, dataCols = NULL)
```

Arguments

`dataMatrix`	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
`dataRows, dataCols`	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.

Details

Robust Z score normalization subtracts the median of the raw well intensities of a given plate from the signal intensity of a given compound and divides it by the median absolute deviation of the raw well intensities of that plate.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (`ex_dataMatrix`)

References

Malo et al. (2006). Statistical practice in high-throughput screening data analysis. Nature Biotechnology, 24(2), 167-175.

See Also

Other normalization methods: normLoess, normMedFil, normR, normSPAWN, normZ

Examples

```
## load dataset
data(ex_dataMatrix)

## apply robust Z score
ex_normMatrix <- normRobZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
```

---

normSights *Normalization methods*

Description

Apply any of the available SIGHTS normalization methods

Usage

```
normSights(normMethod, dataMatrix, plateRows, plateCols, dataRows = NULL,
  dataCols = NULL, trimFactor = 0.2, wellCorrection = FALSE,
  biasMatrix = NULL, biasCols = NULL, seqFilter = TRUE)
```

Arguments

normMethod	Normalization method name from SIGHTS ('Z', 'RobZ', 'R', 'SPAWN', 'Loess', or 'MedFil')
dataMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate. Applies to normMethods 'R', 'SPAWN', 'Loess', and 'MedFil'.
dataRows, dataCols	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.
trimFactor	Optional trim value to be used in trimmed mean polish. It should be between 0 and 0.5. Default is 0.2. Applies to normMethod 'SPAWN'.
wellCorrection	Optional logical. If TRUE then individual wells are corrected based on spatial bias. Applies to normMethod 'SPAWN'.
biasMatrix	Optional data frame or numeric matrix, in the same format as dataMatrix and with the same plateRows and plateCols specifications. If NULL then normalized data is used as bias template. Applies to normMethod 'SPAWN'.

biasCols	Optional integer vector. Indicate which column numbers from biasMatrix or normalized dataMatrix (subset of dataCols) should be used to calculate bias template. Control plates or selection of dataMatrix plates to be used for well correction. If NULL then all plates of biasMatrix or normalized dataMatrix are used. Applies to normMethod 'SPAWN'.
seqFilter	Optional logical. If TRUE apply initial row median filter then standard filter, else just apply standard filter. Applies to normMethod 'MedFil'.

Details

One of the following SIGHTS normalization methods may be chosen: normZ, normRobZ, normR, normSPAWN, normLoess, or normMedFil. See their individual help pages for more details.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (ex_dataMatrix)

References

Murie et al. (2015). Improving detection of rare biological events in high-throughput screens. Journal of Biomolecular Screening, 20(2), 230-241.

See Also

Other SIGHTS functions: plotSights, statSights

Examples

```
## load dataset
data(ex_dataMatrix)

## choose a normalization method and provide relevant information
ex_normMatrix <- normSights(dataMatrix = ex_dataMatrix, dataCols = 5:10,
normMethod = 'RobZ')
```

---

normSPAWN *Normalization by SPAWN method*

Description

Apply trimmed mean polish to data

Usage

```
normSPAWN(dataMatrix, plateRows, plateCols, dataRows = NULL,
  dataCols = NULL, trimFactor = 0.2, wellCorrection = FALSE,
  biasMatrix = NULL, biasCols = NULL)
```

Arguments

`dataMatrix`	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
`plateRows, plateCols`	Number of rows/columns in plate.
`dataRows, dataCols`	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.
`trimFactor`	Optional trim value to be used in trimmed polish. It should be between 0 and 0.5. Default is 0.2.
`wellCorrection`	Optional logical. If TRUE then individual wells are corrected based on spatial bias.
`biasMatrix`	Optional data frame or numeric matrix, in the same format as dataMatrix and with the same plateRows and plateCols specifications. If NULL then normalized data is used as bias template.
`biasCols`	Optional integer vector. Indicate which column numbers from biasMatrix or normalized dataMatrix (subset of dataCols) should be used to calculate bias template. Control plates or selection of dataMatrix plates to be used for well correction. If NULL then all plates of biasMatrix or normalized dataMatrix are used.

Details

Spatial Polish And Well Normalization (SPAWN) uses a trimmed mean polish on individual plates to remove row and column effects. Data from each well location on each plate are initially fitted to the same model as the R score. Model parameters are estimated with an iterative polish technique but with a trimmed mean, rather than a median, as a measure of central tendency for row and column effects. The residuals are rescaled by dividing by the median average deviation of their respective plates. Well correction uses a bias template, which can either be the normalized plates themselves or be supplied externally (and SPAWN normalized before application). At each well location of this bias template, a median of all plates is calculated and subtracted from the normalized plates, thus correcting for well location bias.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (`ex_dataMatrix`)

References

SPAWN: Murie et al. (2015). Improving detection of rare biological events in high-throughput screens. Journal of Biomolecular Screening, 20(2), 230-241.

R score: Wu et al. (2008). Quantitative Assessment of Hit Detection and Confirmation in Single and Duplicate High-Throughput Screenings. Journal of Biomolecular Screening, 13(2), 159-167.

Trimmed Mean: Malo et al. (2010). Experimental design and statistical methods for improved hit detection in high-throughput screening. Journal of Biomolecular Screening, 15(8), 990-1000.

See Also

Other normalization methods: `normLoess`, `normMedFil`, `normRobZ`, `normR`, `normZ`

Examples

```
## load dataset
data(ex_dataMatrix)

## apply SPAWN method with default trim factor and without well correction
ex_normMatrix <- normSPAWN(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10, trimFactor = 0.2)
## apply SPAWN method with default trim factor and with well correction
ex_normMatrix <- normSPAWN(dataMatrix = ex_dataMatrix, dataCols = 5:10,
plateRows = 8, plateCols = 10, trimFactor = 0.2, wellCorrection = TRUE)
```

`normZ` *Normalization by Z score method*

Description

Apply Z score to data

Usage

```
normZ(dataMatrix, dataRows = NULL, dataCols = NULL)
```

Arguments

`dataMatrix`	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
`dataRows, dataCols`	Optional integer vector. Indicate which row/column numbers from the dataMatrix should be normalized. If NULL then all rows/columns from the dataMatrix are used.

Details

Z score normalization subtracts the mean of the raw well intensities of a given plate from the signal intensity of a given compound and divides it by the standard deviation of the raw well intensities of that plate.

Value

Numeric matrix of normalized data in the same format as dataMatrix

Note

For information on how to arrange your dataset for dataMatrix, please see (ex_dataMatrix)

See Also

Other normalization methods: normLoess, normMedFil, normRobZ, normR, normSPAWN

Examples

```
## load dataset
data(ex_dataMatrix)

## apply Z score
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
```

---

plot3d *3D plot*

Description

Plot a three-dimensional plot for each plate

Usage

```
plot3d(plotMatrix, plateRows, plateCols, plotRows = NULL, plotCols = NULL,
  plotName = NULL)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.

Details

3d plots can be used to assess the existence of spatial bias on a plate by plate basis. Spatial bias can be visually subtle, however, and sometimes difficult to detect with 3d plots. Auto-correlation plots (plotAutoco) can circumvent this problem.

Value

List of lattice objects

See Also

Other graphical devices: plotAutoco, plotBox, plotHeatmap, plotHist, plotIGFit, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## plot raw data
plot3d(plotMatrix = ex_dataMatrix, plotCols = 5:10,
plotName = 'Example', plateRows = 8, plateCols = 10)
## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## plot normalized data
plot3d(plotMatrix = ex_normMatrix, plotName = 'Example',
plateRows = 8, plateCols = 10)
```

---

plotAutoco *Auto-correlation*

Description

Plot auto-correlation for each plate

Usage

```
plotAutoco(plotMatrix, plateRows, plateCols, plotRows = NULL,
  plotCols = NULL, plotName = NULL, plotSep = TRUE, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.
plotSep	Optional logical. Should plots be presented in separate windows? Default is TRUE.
...	Optional. Additional parameters passed to geom_path.

Details

Auto-correlation plots can be used to identify spatial bias. Non-zero auto-correlations indicate within-plate bias, namely that proximal wells within-plates are correlated and that the measured intensity of a feature depends partially on its well location in the plate. Cyclical patterns of auto-correlation, in particular indicate within-plate spatial bias. Normalization methods that produce auto-correlations close to zero indicate the removal of spatial bias.

Value

Modifiable ggplot2 object or list of objects

See Also

Other graphical devices: plot3d, plotBox, plotHeatmap, plotHist, plotIGFit, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## plot raw data
plotAutoco(plotMatrix = ex_dataMatrix, plateRows = 8, plateCols = 10,
plotCols = 5:10, plotName = 'Example')
## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## plot normalized data
plotAutoco(plotMatrix = ex_normMatrix, plotName = 'Example',
plateRows = 8, plateCols = 10, plotSep = FALSE)
```

---

plotBox *Boxplot*

Description

Construct an ordered boxplot for each plate

Usage

```
plotBox(plotMatrix, plotRows = NULL, plotCols = NULL, plotName = NULL,
  repIndex = NULL, plotSep = TRUE, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.

plotName	Optional. Name of plotMatrix for plot title.
repIndex	Optional. Vector of labels indicating replicate group. Each index in the vector matches the corresponding column of plotMatrix. If NULL then all plates are plotted together without grouping.
plotSep	Optional logical. Should plots of different replicate groups be presented in separate windows? Default is TRUE. Does not apply if repIndex is NULL.
...	Optional. Additional parameters passed to geom_boxplot.

Details

Box plots can be used to identify scaling shifts among replicates and view the general distribution of data among all plates.

Value

Modifiable ggplot2 object or list of objects

See Also

Other graphical devices: plot3d, plotAutoco, plotHeatmap, plotHist, plotIGFit, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## plot raw data
plotBox(plotMatrix = ex_dataMatrix, repIndex = c(1,1,1,2,2,2), plotCols = 5:10,
plotName = 'Example')
## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## plot normalized data
plotBox(plotMatrix = ex_normMatrix, repIndex = c(1,1,1,2,2,2), plotName = 'Example')
```

---

plotHeatmap *Heat map*

Description

Plot heat map for each plate

Usage

```
plotHeatmap(plotMatrix, plateRows, plateCols, plotRows = NULL,
  plotCols = NULL, plotName = NULL, plotSep = TRUE, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
plateRows, plateCols	Number of rows/columns in plate.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.
plotSep	Optional logical. Should plots be presented in separate windows? Default is TRUE.
...	Optional. Additional parameters passed to geom_tile.

Details

Heat maps can be used to assess the existence of spatial bias on a plate by plate basis. Spatial bias can be visually subtle, however, and sometimes difficult to detect with heat maps. Auto-correlation plots (plotAutoco) can circumvent this problem.

Value

Modifiable ggplot2 object or list of objects

See Also

Other graphical devices: plot3d, plotAutoco, plotBox, plotHist, plotIGFit, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## plot raw data with graphs separated
plotHeatmap(plotMatrix = ex_dataMatrix, plotCols = 5:10,
plotName = 'Example', plateRows = 8, plateCols = 10)
## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## plot normalized data with graphs together
plotHeatmap(plotMatrix = ex_normMatrix, plotName = 'Example',
plateRows = 8, plateCols = 10, plotSep = FALSE)
```

plotHist *Histogram*

Description

Plot histogram of p-values or q-values for each plate or all plates together

Usage

```
plotHist(plotMatrix, plotRows = NULL, plotCols = NULL, plotAll = FALSE,
  plotSep = TRUE, plotName = NULL, colNames = NULL, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix consisting only of p-values or q-values. Columns are samples, and rows are plate wells.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotAll	Optional logical. Should all p-values or q-values be plotted together? Default is FALSE.
plotSep	Optional logical. If plotAll is FALSE, should plots be presented in separate windows? Default is TRUE.
plotName	Optional. Name of plotMatrix for plot title.
colNames	Optional. If plotAll is FALSE, names of plotCols for plot titles.
...	Optional. Additional parameters passed to geom_histogram.

Details

Histograms can be used to compare actual to expected p-value distributions obtained from statistical tests of replicated features. In the presence of rare biological events, the p-value distribution should be approximately uniformly distributed with somewhat more small p-values. Deviations from these patterns indicate that the activity measurements are incorrect and/or that the statistical model is incorrectly specified.

Value

Modifiable ggplot2 object or list of objects

Note

If using output from statT, statRVM, statFDR or statSights, please only select the plotCols corresponding to p-value and/or q-value columns, i.e., every 5th and/or 6th column in that output. Also, the x-axis label is derived from these column names indicating either 'p-values' or 'q-values'.

See Also

Other graphical devices: plot3d, plotAutoco, plotBox, plotHeatmap, plotIGFit, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## apply any test to normalized data and store in new variable
ex_testMatrix <- statRVM(normMatrix = ex_normMatrix,
repIndex = c(1,1,1,2,2,2))
## plot p-value data by selecting the p-value columns from test result matrix
plotHist(plotMatrix = ex_testMatrix, plotCols = c(5,10), plotName = 'Example',
colNames = c('Set_A', 'Set_B'))
```

---

plotIGFit *Inverse gamma*

Description

Plot an inverse gamma fit plot for all plates together

Usage

```
plotIGFit(plotMatrix, repIndex, plotRows = NULL, plotCols = NULL,
  plotName = NULL, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
repIndex	Optional. Vector of labels indicating replicate group. Each index in the vector matches the corresponding column of plotMatrix. If NULL then all plates are plotted together without grouping.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.
...	Optional. Additional parameters passed to geom_step.

Details

Inverse gamma fit plot can be used to check if RVM test (statRVM) assumptions are valid and it can be applied to the data.

Value

Modifiable ggplot2 object

See Also

Other graphical devices: plot3d, plotAutoco, plotBox, plotHeatmap, plotHist, plotScatter

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(dataMatrix = ex_dataMatrix, dataCols = 5:10,
normMethod = 'normZ')
## plot normalized data
plotIGFit(plotMatrix = ex_normMatrix, repIndex = c(1,1,1,2,2,2),
plotName = 'Example')
```

---

plotScatter *Scatter plot*

Description

Construct a scatter plot of all pairwise combinations of replicates

Usage

```
plotScatter(plotMatrix, repIndex, plotRows = NULL, plotCols = NULL,
  plotName = NULL, ...)
```

Arguments

plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells.
repIndex	Optional. Vector of labels indicating replicate group. Each index in the vector matches the corresponding column of plotMatrix. If NULL then all plates are plotted together without grouping.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.
...	Optional. Additional parameters passed to geom_point.

Details

Scatter plots with robust regression lines of replicate plates can reveal a kind of bias which acts independently of within-plate biases and which cannot be detected by heat maps (plotHeatmap) or auto-correlation plots (plotAutoco). A mixture of active and inactive features should produce a zero-correlation flat regression line within most of the range and a positively sloped line within the active range(s) at the extreme(s) of the distribution.

Value

List of modifiable ggplot2 objects

See Also

Other graphical devices: plot3d, plotAutoco, plotBox, plotHeatmap, plotHist, plotIGFit

Examples

```
## load dataset
data(ex_dataMatrix)

## plot raw data
plotScatter(plotMatrix = ex_dataMatrix, repIndex = c(1,1,1), plotCols = 5:7,
plotName = 'Example')
## normalize data matrix using any method and store in new variable
ex_normMatrix <- normZ(dataMatrix = ex_dataMatrix, dataCols = 5:10)
## plot normalized data
plotScatter(plotMatrix = ex_normMatrix, repIndex = c(1,1,1), plotCols = 1:3,
plotName = 'Example')
```

---

plotSights *Graphical devices*

Description

Apply any of the available SIGHTS graphical devices

Usage

```
plotSights(plotMethod, plotMatrix, plateRows, plateCols, repIndex = NULL,
  plotRows = NULL, plotCols = NULL, plotName = NULL, plotSep = TRUE,
  plotAll = FALSE, colNames = NULL, ...)
```

Arguments

plotMethod	Plotting method name from SIGHTS ('3d', 'Autoco', 'Box', 'Heatmap', 'Hist', 'IGFit', or 'Scatter').
plotMatrix	Data frame or numeric matrix. Columns are plates, and rows are plate wells. For plotMethod 'Hist', this is a p-value matrix with each column a single sample.
plateRows, plateCols	Number of rows/columns in plate. Applies to plotMethods '3d', 'Autoco' and 'Heatmap'.
repIndex	Vector of labels indicating replicate group. Each index in the vector matches the corresponding column of plotMatrix. Applies to plotMethods 'Box', 'Scatter' and 'IGFit'.
plotRows, plotCols	Optional integer vector. Indicate which row/column numbers from the plotMatrix should be plotted. If NULL then all rows/columns from the plotMatrix are used.
plotName	Optional. Name of plotMatrix for plot title.
plotSep	Optional logical. Should plots be presented in separate windows? Default is TRUE. Applies to plotMethods 'Autoco', 'Box', 'Hist' and 'Heatmap'. For 'Box', each replicate group is presented in a separate window and it only applies if repIndex is not NULL.
plotAll	Optional logical. Should all p-values be plotted together? Default is FALSE. Applies to plotMethod 'Hist'.
colNames	Optional. Names of plotCols for plot title. Applies to plotMethod 'Hist'.
...	Optional. Additional parameters passed to ggplot functions.

Details

One of the following SIGHTS graphical devices may be chosen: plot3d, plotAutoco, plotBox, plotHeatmap, plotHist, plotIGFit, or plotScatter. See their individual help pages for more details.

Value

List of lattice objects for 'plot3d'. Modifiable ggplot2 object or list of objects for all others.

References

Murie et al. (2015). Improving detection of rare biological events in high-throughput screens. Journal of Biomolecular Screening, 20(2), 230-241.

See Also

Other SIGHTS functions: normSights, statSights

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(normMethod = 'RobZ', dataMatrix = ex_dataMatrix,
dataCols = 5:10, wellCorrection = TRUE)
## choose a graphical device and provide relevant information
plotSights(plotMethod = 'Autoco', plotMatrix = ex_normMatrix,
plotName = 'Example', plateRows = 8, plateCols = 10)
```

`statFDR` *FDR control*

Description

Apply Storey's FDR control to p-values

Usage

```
statFDR(testMatrix, ctrlMethod = "smoother", ...)
```

Arguments

`testMatrix`	Data frame or numeric matrix consisting of output from `statT` or `statRVM` functions. P-value columns from this matrix are automatically selected for FDR calculation. Columns are samples, and rows are plate wells.
`ctrlMethod`	Optional. Method to use either 'smoother' or 'bootstrap' to estimate null. Default is 'smoother'.
`...`	Optional. Additional parameters passed to `qvalue` function.

Details

False Discovery Rate procedure is used to control the proportion of false positives in the results. This is an implementation of the positive false discovery (pFDR) procedure of the `qvalue` function.

Value

A matrix of parameters for each replicate group is returned:

`T-statistic` or `RVM T-statistic`
: Value of the t-statistic.

`Mean_Difference`
: Difference between the calculated and the true mean.

`Standard_Error`
: Standard error of the difference between means.

`Degrees_Of_Freedom`
: Degrees of freedom for the t-statistic.

P-value	P-value for the t-test.
q-value	FDR q-value for the P-value.

Note

Please install the package 'qvalue' from Bioconductor, if not already installed.

References

Storey (2002). A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B, 64, 479-498.

See Also

Other statistical methods: statRVM, statT

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(dataMatrix = ex_dataMatrix, dataCols = 5:10,
normMethod = 'normZ')
## test normalized data matrix using either the RVM or T test and store in new variable
ex_testMatrix <- statT(normMatrix = ex_normMatrix, trueMean = 0,
repIndex = c(1,1,1,2,2,2))
## apply FDR control to test matrix with bootstrap control method
ex_ctrlMatrix <- statFDR(testMatrix = ex_testMatrix,
ctrlMethod = 'bootstrap')
```

---

statRVM *RVM Test*

Description

Apply one-sample RVM t-test separately to each plate

Usage

```
statRVM(normMatrix, repIndex, normRows = NULL, normCols = NULL,
  testSide = "two.sided")
```

Arguments

normMatrix	Data frame or numeric matrix of normalized data. Columns are plates, and rows are plate wells.
repIndex	Integer vector indicating replicates in normMatrix. Which plates are replicates of each other? Provide the same number for plates belonging to a replicate group. Each index in the vector matches the corresponding column of normMatrix.
normRows, normCols	Optional integer vector. Indicate which row/column numbers from the normMatrix should be tested. If NULL then all rows/columns from the normMatrix are used.
testSide	Optional. Type of t-test: 'two.sided', 'less', or 'greater'. Default is 'two.sided'.

Details

Random Variance Model one-sample t-test is applied to the normalized data. RVM assumes that the across replicate variances are distributed according to an inverse gamma distribution. This can be checked by using the plotIGFit function.

Value

A matrix of parameters for each replicate group is returned:

RVM T-statistic	Value of the RVM t-statistic.
Mean_Difference	Difference between the calculated and the true mean.
Standard_Error	Standard error of the difference between means.
Degrees_Of_Freedom	Degrees of freedom for the t-statistic.
P-value	P-value for the RVM test.

References

Malo et al. (2006). Statistical practice in high-throughput screening data analysis. Nature Biotechnology, 24(2), 167-175.

Wright & Simon (2003). A random variance model for detection of differential gene expression in small microarray experiments. Bioinformatics, 19(18), 2448-2455.

See Also

Other statistical methods: statFDR, statT

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(dataMatrix = ex_dataMatrix, dataCols = 5:10,
normMethod = 'normZ')
## apply RVM test to normalized data matrix and get the p-values
ex_testMatrix <- statRVM(normMatrix = ex_normMatrix, repIndex = c(1,1,1,2,2,2))
```

`statSights` *Statistical methods*

Description

Apply any of the available SIGHTS statistical methods

Usage

```
statSights(statMethod, normMatrix, repIndex, normRows = NULL,
  normCols = NULL, ctrlMethod = NULL, testSide = "two.sided",
  trueMean = 0, ...)
```

Arguments

`statMethod`	Statistical testing method to use either 'T' or 'RVM'.
`normMatrix`	Data frame or numeric matrix of normalized data. Columns are plates, and rows are plate wells.
`repIndex`	Integer vector indicating replicates in normMatrix. Which plates are replicates of each other? Provide the same number for plates belonging to a replicate group. Each index in the vector matches the corresponding column of normMatrix.
`normRows, normCols`	Optional integer vector. Indicate which row/column numbers from the normMatrix should be tested. If NULL then all rows/columns from the normMatrix are used.
`ctrlMethod`	Optional. FDR method to use either 'smoother' or 'bootstrap' to estimate null. Default is NULL, which does not apply FDR control to the statistical testing output.
`testSide`	Optional. Type of t-test: 'two.sided', 'less', or 'greater'. Default is 'two.sided'.
`trueMean`	Optional. Number indicating true value of mean. Applies to statMethod 'T'. Default is 0.
`...`	Optional. Additional parameters passed to `qvalue` function.

Details

One of the two SIGHTS statistical testing methods may be chosen: statT or statRVM, and FDR control may be applied by statFDR. See their individual help pages for more details.

Value

A matrix of parameters for each replicate group including p-values and q-values, if FDR control is applied.

References

Murie et al. (2015). Improving detection of rare biological events in high-throughput screens. Journal of Biomolecular Screening, 20(2), 230-241.

See Also

Other SIGHTS functions: normSights, plotSights

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(normMethod = 'RobZ', dataMatrix = ex_dataMatrix,
dataCols = 5:10, wellCorrection = TRUE)
## choose a statistical testing method, indicate FDR control
## and provide relevant information
ex_statMatrix <- statSights(normMatrix = ex_normMatrix, statMethod = 'RVM',
ctrlMethod = 'smoother', repIndex = c(1,1,1,2,2,2))
```

---

statT *t-test*

Description

Apply one-sample t-test separately to each plate

Usage

```
statT(normMatrix, repIndex, normRows = NULL, normCols = NULL,
  testSide = "two.sided", trueMean = 0)
```

Arguments

normMatrix	Data frame or numeric matrix of normalized data. Columns are plates, and rows are plate wells.
repIndex	Integer vector indicating replicates in normMatrix. Which plates are replicates of each other? Provide the same number for plates belonging to a replicate group. Each index in the vector matches the corresponding column of normMatrix.
normRows, normCols	Optional integer vector. Indicate which row/column numbers from the normMatrix should be tested. If NULL then all rows/columns from the normMatrix are used.
testSide	Optional. Type of t-test: 'two.sided', 'less', or 'greater'. Default is 'two.sided'.
trueMean	Optional. Number indicating true value of mean. Default is 0.

Details

Standard one-sample t-test is applied to the normalized data.

Value

A matrix of parameters for each replicate group is returned:

T-statistic	Value of the t-statistic.
Mean_Difference	Difference between the calculated and the true mean.
Standard_Error	Standard error of the difference between means.
Degrees_Of_Freedom	Degrees of freedom for the t-statistic.
P-value	P-value for the t-test.

See Also

Other statistical methods: statFDR, statRVM

Examples

```
## load dataset
data(ex_dataMatrix)

## normalize data matrix using any method and store in new variable
ex_normMatrix <- normSights(dataMatrix = ex_dataMatrix, dataCols = 5:10,
normMethod = 'normZ')
## apply T test to normalized data matrix and get the p-values
ex_testMatrix <- statT(normMatrix = ex_normMatrix, trueMean = 0,
repIndex = c(1,1,1,2,2,2))
```

Index

∗ **datasets**
ex_dataMatrix, 2
inglese, 4

ex_dataMatrix, 2, *5–9*, *11, 12*, *14*

geom_boxplot, *17*
geom_histogram, *19*
geom_path, *15*
geom_point, *21*
geom_step, *20*
geom_tile, *18*
ggplot, *23*

inglese, 4

normLoess, 5, *7, 8*, *10, 11*, *13, 14*
normMedFil, 6, *6*, *8*, *10, 11*, *13, 14*
normR, *6, 7*, 8, *10, 11*, *13, 14*
normRobZ, *6–8*, 9, *11*, *13, 14*
normSights, 10, *23*, *28*
normSPAWN, *6–8*, *10*, 11, *11*, *14*
normZ, *6–8*, *10, 11*, 13, *13*

plot3d, 14, *16–18*, *20–23*
plotAutoco, *14*, 15, *15*, *17, 18*, *20–23*
plotBox, *15*, 16, *16*, *18*, *20–23*
plotHeatmap, *15, 16*, 17, *17*, *20–23*
plotHist, *15–18*, 19, *21–23*
plotIGFit, *15–18*, 20, *20*, *22, 23*, *26*
plotScatter, *15–18*, *20*, 21, *21*, *23*
plotSights, *11*, 22, *28*

qvalue, *24*, *27*

rlm, *8*

statFDR, *19*, 24, *26*, *28, 29*
statRVM, *19, 20*, *24*, 25, *25*, *28, 29*
statSights, *11*, *19*, *23*, 27
statT, *19*, *24–26*, 28, *28*