Package ‘arrayQualityMetrics’

April 9, 2015

Title Quality metrics report for microarray data sets

Version 3.22.1

Author Audrey Kauffmann, Wolfgang Huber

Maintainer Audrey Kauffmann <`akauffmann.bioinfo@gmail.com`>

VignetteBuilder knitr

Depends

Suggests ALLMLL, CCl4, BiocStyle, knitr

Imports affy, affyPLM (>= 1.27.3), beadarray, Biobase, Cairo (>= 1.4-6), genefilter, graphics, grDevices, grid, gridSVG (>= 1.4-3), Hmisc, hwriter, lattice, latticeExtra, limma, methods, RColorBrewer, setRNG, stats, SVGAnnotation (>= 0.9-0), utils, vsn (>= 3.23.3), XML

Description This package generates microarray quality metrics reports for data in Bioconductor microarray data containers (ExpressionSet, NChannelSet, AffyBatch). One and two color array platforms are supported.

biocViews Microarray, QualityControl, OneChannel, TwoChannel, ReportWriting

Collate classes.R affyspecific.r annotateSvg.R arrayQualityMetrics.r boxplot.r density.r globalParameters.R heatmap.r makeColors.R maplot.r meansd.r outlier.R pca.r prepdata.r probesmap.r spatial.r writereport.r XYfromGAL.r

R topics documented:

`addXYfromGAL` *Computing the coordinates of the spots on a slide*

Description

From the coordinates of the blocks of a microarray slide and the Row and Column locations of the spots within the blocks, `addXYfromGAL` computes the X and Y coordinates of the spots of a slide.

Usage

```
addXYfromGAL(x, gal.file, nBlocks, skip, ...)
```

Arguments

`x`	is an `AnnotatedDataFrame` representing the featureData of an object.
`gal.file`	name of the file .gal that contains the coordinates of the blocks.
`nBlocks`	number of blocks on the slide.
`skip`	number of header lines to skip when reading the gal.file.
`...`	Arguments that get passed on to `read.table`.

Value

The object `x` of class `AnnotatedDataFrame` will be returned with two added columns: X and Y corresponding to the absolute position of the probes on the array.

Author(s)

Audrey Kauffmann, Wolfgang Huber. Maintainer: <kauffmann@bergonie.org>

`aqm.writereport` *Write a quality report*

Description

`aqm.writereport` produces a quality report (HTML document with figures) from a list of `aqmReportModule` objects.

Usage

```
aqm.writereport(modules, arrayTable, reporttitle, outdir)
```

Arguments

`modules`	A list of `aqmReportModule` objects.
`arrayTable`	A data.frame with array (meta)data to be displayed in the report.
`reporttitle, outdir`	Report title and output directory - as in `arrayQualityMetrics`.

Value

A side effect of this function is the creation of the HTML report.

Author(s)

Audrey Kauffmann, Wolfgang Huber

`aqmReportModule` *Class to contain all the information to render a quality report module.*

Description

Please see the vignette *Advanced topics: Customizing arrayQualityMetrics reports and programmatic processing of the output.*

Creating Objects

Please see the manual page of the module generations functions, e.g. `aqm.boxplot`.

Author(s)

Audrey Kauffmann, Wolfgang Huber

`arrayQualityMetrics` *Quality metrics for microarray experiments*

Description

Produce an array quality metrics report. This is the main function of the package.

Usage

```
arrayQualityMetrics(expressionset,
  outdir = reporttitle,
  force = FALSE,
  do.logtransform = FALSE,
  intgroup = character(0),
  grouprep,
  spatial = TRUE,
 reporttitle = paste("arrayQualityMetrics report for", deparse(substitute(expressionset))),
  ...)
```

Arguments

`expressionset`	a Bioconductor microarray dataset container. This can be an object of class `ExpressionSet`, `AffyBatch`, `NChannelSet`, `ExpressionSetIllumina`, `RGList`, `MAList`.
`outdir`	the name of the directory in which the report is created; a character of length 1.
`force`	if the directory named by `outdir` already exists, then, if `force` is `TRUE`, the directory is overwritten, otherwise an error is thrown; if the directory does not exist, the value of `force` is irrelevant; a logical of length 1.
`do.logtransform`	indicates whether the data should be logarithm transformed before the analysis; a logical of length 1.
`intgroup`	the name of the sample covariate(s) used to draw a colour side bar next to the heatmap. The first element of `intgroup` is also used define sample groups in other plots (boxplots, densities). `intgroup` should be a character vector, and its elements need to match the columns names of `phenoData(expressionset)`. If its length is 0, then the plots are not decorated with sample covariate information.
`grouprep`	deprecated. Use argument `intgroup` instead.
`spatial`	indicates whether spatial plots should be made; a logical of length 1. This can be useful for large arrays (like Affymetrix hgu133Plus2) when CPU time and RAM resources of the machine would be limiting.
`reporttitle`	title for the report (character of length 1).
`...`	further arguments that will be passed on to the different `module functions`.

Details

See the arrayQualityMetrics vignette for examples of this function.

Value

A side effect of the function is the creation of directory named by `outdir` containing a HTML report `QMreport.html` and figures. The function also returns a list with R objects containing the report elements for subsequent programmatic processing.

Author(s)

Audrey Kauffmann and Wolfgang Huber.

`modulefunctions` *Functions for computing quality report modules.*

Description

These functions produce objects of class `aqmReportModule` representing the various modules of the quality report. Given a list of modules, the report is then rendered by the `aqm.writereport` function.

Most users will not call these functions directly, but will use the function `arrayQualityMetrics`, which in turns calls these functions. The function arguments can be provided through the `...` argument of `arrayQualityMetrics`.

Usage

```
aqm.boxplot(x, subsample=20000, outlierMethod = "KS", ...)
aqm.density(x, ...)
aqm.heatmap(x, ...)
aqm.pca(x, ...)
aqm.maplot(x, subsample=20000, Dthresh=0.15, maxNumArrays=8, nrColumns=4, ...)
aqm.spatial(x, scale="rank", channels = c("M", "R", "G"), maxNumArrays=8, nrColumns=4, ...)
aqm.meansd(x, ...)
aqm.probesmap(x, ...)

# Affymetrix specific sections

aqm.pmmm(x, ...)
aqm.rnadeg(expressionset, x, ...)
aqm.rle(x, outlierMethod = "KS", ...)
aqm.nuse(x, outlierMethod = "upperquartile", ...)
```

Arguments

`x`	An object resulting from a call to `prepdata(expressionset)`.
`expressionset`	An object of class `AffyBatch`.
`subsample`	For efficiency, some computations are performed not on the full set of features (which can be hundreds of thousands on some arrays), but on a randomly subset whose size is indicated by this number.
`outlierMethod`	As in `outliers`.
`Dthresh`	In `maplot`, the arrays with a Hoeffding D statistic larger than this value are called *outliers*. See also `hoeffd`.
`scale, channels`	In `aqm.spatial`, `scale` determines the choice of the false colour scale in the spatial plots. If the value is `"rank"`, then the colour is proportional to the ranks of the values; if it is `"direct"`, then it is proportional to the values themselves. `channels` determines for which elements of `x` spatial plots are made.

`maxNumArrays, nrColumns`	The parameter `maxNumArrays` determines the number of arrays for which a plot is produced. `nrColumns` determines the number of columns in the multi-panel plot. In `aqm.maplot`, first `maxNumArrays` is incremented to the next multiple of `maxNumArrays`. A value of `+Inf` is allowed. If this value is larger than or equal to the actual number of arrays in `x`, then plots are produced for all arrays. If it is smaller, then plots are shown for the `maxNumArrays/2` with the worst values of Hoeffding's D and for the `maxNumArrays/2` best.
`...`	Will be ignored - the dots are formal arguments which permit that all of these functions can be callled from `arrayQualityMetrics` with the same, overall set of arguments.

Details

For a simple example of the `aqm.*` functions, have a look at the source code of the `aqm.pca` function. Please see also the vignette *Advanced topics: Customizing arrayQualityMetrics reports and programmatic processing of the output*.

Value

An object of class `aqmReportModule`.

Author(s)

Audrey Kauffmann, Wolfgang Huber

`outlierDetection` *Represents the results from applying an outlier detection criterion to the arrays.*

Description

The class is described in the vignette *Advanced topics: Customizing arrayQualityMetrics reports and programmatic processing of the output.*

Author(s)

Audrey Kauffmann, Wolfgang Huber

See Also

`outliers`

`outliers` *Helper functions for outlier detection and reporting in arrayQualityMetrics*

Description

For an overview of outlier detection, please see the corresponding section in the vignette *Advanced topics: Customizing arrayQualityMetrics reports and programmatic processing of the output*. These two functions are helper functions used by the different report generating functions, such as `aqm.boxplot`.

Usage

```
outliers(exprs, method = c("KS", "sum", "upperquartile"))
boxplotOutliers(x, coef = 1.5)
```

Arguments

`exprs`	A matrix whose columns correspond to arrays, rows to the array features.
`method`	A character string specifying the summary statistic to be used for each column of `exprs`. See Details.
`x`	A vector of real numbers.
`coef`	A number is called an outlier if it is larger than the upper hinge plus `coef` times the interquartile range. Upper hinge and interquartile range are computed by `fivenum`.

Details

`outliers`: with argument `method="KS"`, the function first computes for each column of `exprs` (i.e. for each array) the value of the `ks.test` test statistic between its distribution of intensities and the pooled distribution of intensities from all arrays. With `"sum"` and `"upperquartile"`, it computes the sum or the 75 percent quantile. Subsequently, it calls `boxplotOutliers` on these values to identify the outlying arrays.

`boxplotOutliers` uses a criterion similar to that used in `boxplot.stats` to detect outliers in a set of real numbers. The main difference is that in `boxplotOutliers`, only the outliers to the right (i.e. extraordinarily large values) are detected.

Value

For `outliers`, an object of class `outlierDetection`. For `boxplotOutliers`, a list with two elements: `thresh`, the threshold against which `x` was compared, and `outliers`, an integer vector of indices.

Author(s)

Wolfgang Huber

`prepdata` *Compute useful summary statistics from a data object.*

Description

`prepdata` computes summary statistics that are useful for all platforms; `prepaffy` computes Affymetrix-specific ones. These are helper functions used by `arrayQualityMetrics`.

Usage

```
prepdata(expressionset, intgroup, do.logtransform)
prepaffy(expressionset, x)
```

Arguments

`expressionset`	An object of class `ExpressionSet` for one colour non Affymetrix data, `AffyBatch` for Affymetrix data, `NChannelSet` for two colour arrays, or `BeadLevelList` for Illumina bead arrays.
`intgroup, do.logtransform`	as in `arrayQualityMetrics`.
`x`	A list, typically the result from a prior call to `prepdata`.

Details

See the vignette *Working with arrayQualityMetrics report sections*.

Value

A list with various derived quantities. In the case of `prepaffy`, the returned list is `x` with the additional elements appended.

Author(s)

Audrey Kauffmann, Wolfgang Huber

Index