Package ‘arrayMvout’

March 28, 2017

Title multivariate outlier detection for expression array QA
Version 1.32.0
Author Z. Gao, A. Asare, R. Wang, V. Carey

Description
This package supports the application of diverse quality metrics to AffyBatch instances, summarizing these metrics via PCA, and then performing parametric outlier detection on the PCs to identify aberrant arrays with a fixed Type I error rate

Depends R (>= 2.6.0), tools, methods, utils, parody, Biobase, affy, lumi

Suggests MAQCsubset, mvoutData, lumiBarnes, affyPLM, affydata, hgu133atagcdf

Imports simpleaffy, mdqc, affyContam,

Maintainer V. Carey <stvjc@channing.harvard.edu>
License Artistic-2.0
LazyLoad yes
biocViews Infrastructure, Microarray, QualityControl

needsCompilation no

R topics documented:

ArrayOutliers ................................................................. 1
ArrayOutliers-methods ................................................... 3
arrOutStruct-class ......................................................... 4

Index

ArrayOutliers Multivariate outlier detection based on PCA of QA statistics

Description
Multivariate outlier detection based on PCA of QA statistics
Usage

```
ArrayOutliers (data, alpha, alphaSeq = c(0.01, 0.05, 0.1), ... )
# # qcOutput = NULL, plmOutput = NULL, degOutput = NULL, prscale = TRUE,
# # pc2use = 1:3)
```

Arguments

data an (affy) AffyBatch instance with at least 11 samples
alpha false positive rate for outlier detection, adjusting for multiple comparisons according to Caroni and Prescott’s adaptation of Rosner (1983); full report based on this choice of alpha
alphaSeq vector of alpha candidates to be quickly tried for short report
... additional parameters, see below

Details

Additional parameters may be supplied

```
qcOutput optional result of simpleaffy qc() to speed computations
plmOutput optional result of affyPLM fitPLM() to speed computations
degOutput optional result of affy AffyRNAdeg() to speed computations
prscale scaling option for prcomp
pc2use selection of principal components to use for outlier detection
```

Data elements afxsubDEG, afxsubQC, s12cDEG, s12cQC are precomputed RNA degradation and simpleaffy qc() results; s12c is an AffyBatch with digital contamination of some samples. 

Data elements maqcQA and itnQA are affymetrix QC statistics on large collections of arrays. Data element ilmQA is a derived from a LumiBatch of the Illumina-submitted MAQC raw data, 19 arrays. (Conveyed by Leming Shi, personal communication). Data element spikQA is a 12x9 matrix of QA parameters obtained for 12 arrays from U133A spikein dataset, with first 2 arrays digitally contaminated as described in Asare et al. 

Data element fig3map gives the indices of the points labeled A-H in Figure 3 of the manuscript by Asare et al. associated with this package.

Value

an instance of arrOutStruct class, a list with a partition of samples into two data frames (inl and outl) with QA summary statistics

Author(s)

Z. Gao et al.

Examples

```
library(simpleaffy)
setQCEnvironment("hgu133acdf") # no CDF corresponding to tag array
if ( require("mvoutData") ) {
  data(s12c)
  data(s12cQC)
  data(s12cDEG)
```
library(affyPLM)
s12cPset = fitPLM(s12c)

ao = ArrayOutliers(s12c, alpha=0.05, qcOut=s12cQC, plmOut=s12cPset, degOut=s12cDEG)

if (require("lumiBarnes")) {
  library(lumiBarnes)
  data(lumiBarnes)
  ArrayOutliers(lumiBarnes, alpha=0.05)
  lb2 = lumiBarnes
  exprs(lb2)[1:20000,1:2] = 10000*exprs(lb2)[1:20000,1:2]
  ArrayOutliers(lb2, alpha=0.05)
}

data(maqcQA) # affy
ArrayOutliers(maqcQA[,,-c(1:2)], alpha=.05)
ArrayOutliers(maqcQA[,,-c(1:2)], alpha=.01)

data(ilmQA) # illumina
ArrayOutliers(data.frame(ilmQA), alpha=.01)
data(itnQA) # 507 arrays from ITN
ArrayOutliers(itnQA, alpha=.01)

---

ArrayOutliers – wrapper for platform-specific multivariate outlier detection for expression arrays

Description
wraps functions that perform multivariate outlier detection on dimension-reduced QA statistics of expression arrays

Methods

- `data = "ANY", alpha = "missing", alphaSeq = "missing"` fails; tells user that alpha is obligatory parameter
- `data = "AffyBatch", alpha = "numeric", alphaSeq = "ANY"` performs calibrated multivariate outlier detection on an AffyBatch instance using various affy-specific QA parameters
- `data = "LumiBatch", alpha = "numeric", alphaSeq = "ANY"` performs calibrated multivariate outlier detection on an LumiBatch instance using various illumina-specific QA parameters
- `data = "data.frame", alpha = "numeric", alphaSeq = "ANY"` performs calibrated outlier detection on QA statistics housed in data.frame – all columns of the data entity must be numeric QA statistics for the arrays.

Examples

example(ArrayOutliers)
**arrOutStruct-class**  

---

**Description**  

Class "arrOutStruct" container for ArrayOutliers output

**Objects from the Class**  

Objects can be created by calls of the form `new("arrOutStruct", ...)`. This class just extends `list` but has specialized `show` and `plot` methods.

**Extends**  

Class "list", from data part. Class "vector", by class "list", distance 2. Class `AssayData`, by class "list", distance 2.

**Methods**  

- `plot` signature(x = "arrOutStruct", y = "ANY"): a biplot of QA statistics
- `show` signature(object = "arrOutStruct"): summary report

**Author(s)**  

Vince Carey <stvjc@channing.harvard.edu>

**Examples**

data(maqcQA)
f1 = ArrayOutliers(maqcQA[,-c(1:2)], alpha=0.01)
names(f1)
f1
Index

*Topic classes
  arrOutStruct-class, 4

*Topic methods
  ArrayOutliers-methods, 3

*Topic models
  ArrayOutliers, 1

afxsubDEG(ArrayOutliers), 1
afxsubQC(ArrayOutliers), 1
ArrayOutliers, 1
ArrayOutliers, AffyBatch, numeric, ANY-method
  (ArrayOutliers-methods), 3
ArrayOutliers, ANY, missing, missing-method
  (ArrayOutliers-methods), 3
ArrayOutliers, data.frame, numeric, ANY-method
  (ArrayOutliers-methods), 3
ArrayOutliers, LumiBatch, numeric, ANY-method
  (ArrayOutliers-methods), 3
ArrayOutliers-methods, 3
arrOutStruct-class, 4
AssayData, 4

fig3map(ArrayOutliers), 1
ILM1 (ArrayOutliers), 1
ilmQA (ArrayOutliers), 1
itnQA (ArrayOutliers), 1

list, 4

maqcQA (ArrayOutliers), 1

plot, arrOutStruct, ANY-method
  (arrOutStruct-class), 4

s12c (ArrayOutliers), 1
s12cDEG (ArrayOutliers), 1
s12cQC (ArrayOutliers), 1

show, arrOutStruct-method
  (arrOutStruct-class), 4
spikQA (ArrayOutliers), 1

vector, 4