

# Package ‘varmixt’

April 19, 2009

**Version** 0.2-4

**Date** 2005-06-17

**Title** Mixture model on the Variance for the analysis of gene expression data

**Author** Paul Delmar and Julie Aubert

**Maintainer** Julie Aubert <Julie.Aubert@inapg.inra.fr>

**Depends** R (>= 2.0),MASS,stats, qvalue

**Description** This package performs mixture model on the variance for the analysis of gene expression data.

**License** GPL (>= 2)

**URL** <http://www.r-project.org>

**Repository** CRAN

**Date/Publication** 2006-02-08 12:07:35

## R topics documented:

apo.data.vm . . . . .	2
boxplotvm . . . . .	3
boxplotvm.g . . . . .	3
compare.method . . . . .	4
compare.res . . . . .	5
compare.var . . . . .	5
export.res . . . . .	6
fdr.an . . . . .	8
fdr.compare.method . . . . .	9
fdr.compare.res . . . . .	9
find.qval.index . . . . .	10
n.genes . . . . .	11
plotrm . . . . .	11
plotsdt . . . . .	12

plotsdt.VM . . . . .	12
plotvm . . . . .	13
pval.an . . . . .	13
qplot.vm . . . . .	14
qqplot.var.vm . . . . .	15
qqplot.var.vm.2 . . . . .	15
qqplot.vm . . . . .	16
qval.anova.vect . . . . .	16
qval.gene.vect . . . . .	17
qval.VM.vect . . . . .	18
qval.VM2.vect . . . . .	18
sbset.gene . . . . .	19
sbset.pval . . . . .	19
sbset.qval . . . . .	20
sd.param . . . . .	21
spleen.data.vm . . . . .	21
vm.analysis . . . . .	22
vm.analysis.paired . . . . .	24

## Index 27

---

apo.data.vm	<i>Apo Data set</i>
-------------	---------------------

---

### Description

An exemple data set for unpaired data analysis. A list with the normalized Apo data set as used in the original article.

### Usage

```
data(apo.data.vm)
```

### Format

apo.data.vm is a list with 3 elements :

apo.geneid A vector with gene names

apo.cond1 A Matrix with 6226 rows and 8 columns with normalized normal mice measurements

apo.cond2 A Matrix with 6226 rows and 8 columns with normalized KO mice measurements

### References

M.J. Callow, S. Dudoit, E.L. Gong, T.P. Speed, and E.M. Rubin. Microarray expression profiling identifies genes with altered expression in hdl-deficien mice. *Genome Res.*, 10(12) : 2022-9, 2000

### Examples

```
data(apo.data.vm)
```

---

`boxplotvm`*Boxplot of mean log-intensity in each variance group*

---

**Description**

Boxplot of mean log-intensity in each variance group

**Usage**

```
boxplotvm(x)
```

**Arguments**

`x`                      Gene expression data object

**Details**

The y axis is the standard deviation of mean log-ratio or mean difference. The y axis is the denominator of the test statistic

**Value**

A boxplot of the distribution of mean log-intensity in each variance group

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [plotrm](#), [qqplot.vm](#), [plotvm](#)

---

`boxplotvm.g`*Boxplot of mean log-intensity in each variance group*

---

**Description**

Boxplot of mean log-intensity in each variance group

**Usage**

```
boxplotvm.g(x)
```

**Arguments**

`x`                      Gene expression data object

**Details**

The y axis is the standard deviation of the gene.

**Value**

A boxplot of the distribution of mean log-intensity in each variance group

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [plotrm](#), [qqplot.vm](#), [plotvm](#)

---

`compare.method`

*Compare the different analysis methods*

---

**Description**

Compare the different analysis methods for a given p-value threshold

**Usage**

```
compare.method(data, pval = 0.05)
```

**Arguments**

<code>data</code>	gene expression data object
<code>pval</code>	p-value cut-off

**Value**

A data frame with each gene classified as regulated/not-regulated in for each method (gene-specific, homoscedastic, Varmixt). Also prints the tables of comparison.

**Author(s)**

Paul Delmar

**See Also**

[fdr.compare.method](#), [compare.res](#), [vm.analysis](#), [vm.analysis.paired](#)

---

compare.res	<i>Compare the results from 2 analysis</i>
-------------	--------------------------------------------

---

**Description**

Compare the results from 2 analysis with p-value criterion

**Usage**

```
compare.res(data1, data2, pval = 0.05)
```

**Arguments**

data1	gene expression data object (results)
data2	gene expression data object (results)
pval	p-value cut-off

**Value**

A data frame with each gene classified as regulated/not-regulated in the two data-set for each method (gene-specific, homoscedastic, Varmixt). Also prints the tables of comparison.

**Author(s)**

Paul Delmar

**See Also**

[fdr.compare.res](#), [compare.method](#), [vm.analysis](#), [vm.analysis.paired](#)

---

compare.var	<i>How to compare variance in the 2 conditions</i>
-------------	----------------------------------------------------

---

**Description**

This function compares variance in the 2 conditions

**Usage**

```
compare.var(data)
```

**Arguments**

data	gene expression data object
------	-----------------------------

**Value**

A data.frame with the columns :

var.treat	gene variance estimate in the "treated" condition
var.cont	gene variance estimate in the "control" condition
var.ratio	gene variances ratio
var.ratio.pval	p-value (H0 is {ratio==1})

**Author(s)**

Paul Delmar

---

export.res

*Format and export analysis results*

---

**Description**

This function builds a single data.frame with all the analysis results, computes the rank of genes based on their p-value and computes a ratio from the log-ratio.

**Usage**

```
export.res(data, filename=NULL, header=TRUE, comment=NULL, lambda=seq(0, 0.95, 0.05))
```

**Arguments**

data	A gene expression data object. The results of vm.analysis of vm.analysis.paired function
filename	the name of the output file. If NULL no file is written
header	a logical value indicating whether the file contains the names of the variables as its first line.
comment	
lambda	

**Value**

a data.frame with all the analysis results, computes the rank of genes based on their p-value and computes a ratio from the log-ratio.

geneid	gene name
ratio	ratio
qual	Number of low quality observations
pval.VM2	p-value with the variance mixture model

corrected.pval.VM2  
                   corrected p-value with the variance mixture model(Bonferroni)  
 test.stat.VM2  
                   test statistic with the variance mixture model  
 sigmadeltagVM2  
                   standard deviation of deltag with the variance mixture model  
 group  
                   variance component (variance mixture model)  
 rank.p.VM2  
                   rank according to the variance mixture model p-value  
 rank.p.VM  
                   rank according to the variance mixture model p-value without assigning gene to  
                   variance groups  
 rank.p.gene  
                   rank according to the gene-specific p-value  
 rank.p.anova  
                   rank according to the homoscedastic p-value  
 pval.gene  
                   p-value with the gene specific model  
 corrected.pval.gene  
                   corrected p-value with gene specific model (Bonferroni)  
 test.stat.gene  
                   test statistic with gene-specific model  
 sigmadeltag  
                   standard deviation of deltag with the gene-specific model  
 varg  
                   gene-specific variance of the gene  
 pval.anova  
                   p-value with the homoscedastic model  
 corrected.pval.anova  
                   corrected p-value with homoscedastic model (Bonferroni)  
 test.stat.anova  
                   test statistic with the homoscesatic model  
 sigmadeltaanova  
                   standard deviation of deltag with the homoscesatic model  
 pval.VM2  
                   p-value with the variance mixture model without assigning genes to variance  
                   groups  
 corrected.pval.VM2  
                   corrected p-value with the variance mixture model(Bonferroni) without assign-  
                   ing genes to variance groups  
 test.stat.VM  
                   test statistic with the continuous variance mixture model  
 sigmadeltagVM  
                   standard deviation of deltag with the continuous variance mixture model  
 rank.VM2  
                   rank according to the variance mixture model test statistic  
 rank.gene  
                   rank according to homoscedastic test statistic  
 rank.anova  
                   rank according to homoscedastic test statistic  
 rank.VM  
                   rank according to the continuous variance mixture model test statistic  
 deltag  
                   mean log2-ratio  
 meanint  
                   mean log2-intensity

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#)

---

fdr.an

*FDR analysis of variance mixture analysis results*

---

**Description**

This functions computes the number of regulated genes at given level of FDR

**Usage**

```
fdr.an(data, Q=0.05, display=TRUE, lambda=seq(0, 0.95, 0.05))
```

**Arguments**

data	A gene gene expression data object. The results of <code>vm.analysis</code> or <code>vm.analysis.paired</code> function
Q	the fdr level
display	if TRUE prints the number of regulated genes and corresponding p-values for the different methods on the screen
lambda	Argument to the q-value function.

**Details**

If `lambda=0`, then the FDR is computed (Benjamini Hochberg). If `lambda` is set to its default, the the pFDR according to Storey is computed.

**Value**

a data frame with the number of genes regulated with the different analysis methods and the corresponding p-value

**Author(s)**

Paul Delmar

**References**

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57(1), 289-300 Storey JD. (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, 64: 479-498

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [pval.an](#)

---

fdr.compare.method *Comparison of the different analysis methods*

---

### Description

This function compares the different analysis methods for a given FDR threshold

### Usage

```
fdr.compare.method(data, qval=0.05, display.all=FALSE, lambda=seq(0, 0.95, 0.05))
```

### Arguments

data	gene expression data object
qval	FDR cut-off
display.all	
lambda	Argument of the q-value function

### Value

A data frame with each gene classified as regulated/not-regulated in for each method (gene-specific, homoscedastic, Varmixt). Also prints the tables of comparison.

### Author(s)

Paul Delmar

### See Also

[fdr.compare.res](#), [compare.res](#), [vm.analysis](#), [vm.analysis.paired](#)

---

fdr.compare.res *Comparison of the results from 2 analysis*

---

### Description

This function compares the results from 2 analysis with FDR criterium

### Usage

```
fdr.compare.res(data1, data2, qval.1=0.05, qval.2=qval.1, display.all=FALSE, lambda=seq
```

**Arguments**

data1            gene expression data object (results)  
data2            gene expression data object (results)  
qval.1           FDR cut-off for data1  
qval.2           FDR cut-off for data2  
display.all  
lambda

**Value**

A data frame with each gene classified as regulated/not-regulated in the two data-set for each method (gene-specific, homoscedastic, Varmixt). Also prints the tables of comparison.

**Author(s)**

Paul Delmar

**See Also**

[fdr.compare.method](#), [compare.res](#), [vm.analysis](#), [vm.analysis.paired](#)

---

find.qval.index      *Subset of genes based on Q value*

---

**Description**

This function extracts a subset of genes based on Q value

**Usage**

```
find.qval.index(data, qval, method = c("gene", "VM", "VM2", "anova")[2], lambda = s
```

**Arguments**

data            gene expression data object  
qval            qvalue cut-off  
method          analysis method  
lambda          qvalue function parameter

**Author(s)**

Paul Delmar

---

n.genes	<i>Number of genes in a gene expression data object</i>
---------	---------------------------------------------------------

---

**Description**

Number of genes in a gene expression data object

**Usage**

```
n.genes(data)
```

**Arguments**

data                    gene expression data object

**Value**

This function returns the number of analyzed gene probes

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#)

---

plotrm	<i>Plot of mean log-ratio versus mean log-intensity</i>
--------	---------------------------------------------------------

---

**Description**

Plot of mean log-ratio versus mean log-intensity

**Usage**

```
plotrm(x, P = 0.05)
```

**Arguments**

x                    Gene expression data object. The results of vm.analysis of vm.analysis.paired function

P                    genes with a p-value less than P will be red in the plot

**Value**

Plot of mean log-ratio versus mean log-intensity with points with a p-value less than P in red

**Author(s)**

Paul Delmar

**See Also**[vm.analysis](#), [vm.analysis.paired](#), [plotvm](#), [qqplot.vm](#), [boxplotvm](#)

---

`plotsdt`*Plot Test statistic versus denominator of test statistic*

---

**Usage**

```
plotsdt(x, pval.f, test.stat.f, pval=0.05)
```

**Arguments**

<code>x</code>	gene expression data object
<code>pval.f</code>	
<code>test.stat.f</code>	
<code>pval</code>	pvalue cut-off

**Author(s)**

Paul Delmar

---

`plotsdt.VM`*Plot of VM test statistic versus denominator*

---

**Description**

Plot VM test statistic versus denominator

**Usage**

```
plotsdt.VM(x, pval=0.05)
```

**Arguments**

<code>x</code>	gene expression data object
<code>pval</code>	p value cut-off for coloring points

**Value**

A plot of the denominator of the VM test statistic versus the test static

**Author(s)**

Paul Delmar

---

plotvm *Plot of log-variance versus mean intensity*

---

**Description**

Plot of log-variance versus mean intensity

**Usage**

```
plotvm(x, colors=TRUE)
```

**Arguments**

x                    the results of vm.analysis of vm.analysis.paired function  
 colors              if TRUE a different color is used for each variance group

**Value**

Plot of log-variance versus mean intensity

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [plotrm](#), [qqplot.vm](#), [boxplotvm](#)

---

pval.an *Number of regulated genes*

---

**Description**

This function returns the number of regulated genes for a given significance level.

**Usage**

```
pval.an(data, P=.05, corrected=FALSE, display=TRUE)
```

**Arguments**

data                Gene expression data object  
 P                    the p-value criteria  
 corrected          If TRUE uses bonferonni corrected p-values.  
 display

**Value**

a data frame with the number of genes regulated with the different method at the given significance (p-value) level.

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [fdr.an](#)

---

qplot.vm

*Plot Diagnostic of FDR computation*

---

**Description**

Plot Diagnostic of FDR computation

**Usage**

```
qplot.vm(data, lambda = seq(0, 0.95, 0.05))
```

**Arguments**

data	gene expression data object
lambda	lambda parameter of qvalue function

**Value**

Plot Diagnostic of FDR computation

**Author(s)**

Paul Delmar

---

qqplot.var.vm      *QQ plot of the distribution of estimated variance*

---

**Description**

QQ plot of the distribution of estimated variance

**Usage**

```
qqplot.var.vm(data)
```

**Arguments**

data              Gene expression data object

**Value**

Group by group qq-plot of the distribution of estimated variance versus a theoretical Gamma distribution

**Author(s)**

Paul Delmar

---

qqplot.var.vm.2      *QQ plot of the variance versus gamma distribution in one variance group*

---

**Description**

QQ plot of the variance versus gamma distribution in one variance group

**Usage**

```
qqplot.var.vm.2(data, i)
```

**Arguments**

data              gene expression data object  
i                  component number

**Value**

QQ plot of the distribution of estimated variance of genes assigned to group i versus the theoretical quantiles of a gamma distribution.

**Author(s)**

Paul Delmar

---

`qqplot.vm`*QQ plot of the residuals*

---

**Description**

The function plots a global qqplots of all residuals and qqplots of residuals in each variance group

**Usage**

```
qqplot.vm(data, ...)
```

**Arguments**

<code>data</code>	Gene expression data object. The result of <code>vm.analysis</code> , <code>vm.analysis.2.Var</code> or <code>vm.analysis.paired</code> function
<code>...</code>	other arguments to the <code>qqnorm</code> function

**Value**

a global qqplots of all residuals and qqplots of residuals in each variance group

**Author(s)**

Paul Delmar

**See Also**

[vm.analysis](#), [vm.analysis.paired](#), [plotrm](#), [plotvm](#), [boxplotvm](#)

---

`qval.anova.vect`*Vector of homoscedastic q-values from vm.result data object*

---

**Description**

This function extracts the vector of homoscedastic q-values from `vm.result` data object

**Usage**

```
qval.anova.vect(data, lambda = seq(0, 0.95, 0.05), pi0.meth = "smoother", fdr.level
```

**Arguments**

data	Gene expression data object
lambda	parameter for qvalue function
pi0.meth	parameter for qvalue function
fdr.level	parameter for qvalue function
robust	parameter for qvalue function

**Value**

The vector of q value computed from the p values of the homoscedastic method

**Author(s)**

Paul Delmar

---

qval.gene.vect      *Vector of gene specific q-values from vm.result data object*

---

**Description**

This function extracts the vector of gene specific q-values from vm.result data object

**Usage**

```
qval.gene.vect(data, lambda = seq(0, 0.95, 0.05), pi0.meth = "smoother", fdr.level
```

**Arguments**

data	Gene expression data object
lambda	parameter for qvalue function
pi0.meth	parameter for qvalue function
fdr.level	parameter for qvalue function
robust	logical, parameter for qvalue function

**Value**

The vector of q value computed from the p values of the gene specific method

**Author(s)**

Paul Delmar

---

qval.VM.vect	<i>Vector of VM q-values from vm.result data object</i>
--------------	---------------------------------------------------------

---

**Description**

This function extracts the vector of VM q-values from vm.result data object

**Usage**

```
qval.VM.vect(data, lambda = seq(0, 0.95, 0.05), pi0.meth = "smoother", fdr.level =
```

**Arguments**

data	Gene expression data object
lambda	parameter for qvalue function
pi0.meth	parameter for qvalue function
fdr.level	parameter for qvalue function
robust	parameter for qvalue function

**Value**

The vector of q value computed from the p values of the VM method

**Author(s)**

Paul Delmar

---

qval.VM2.vect	<i>Vector of VM2 q-values from vm.result data object</i>
---------------	----------------------------------------------------------

---

**Description**

This function extracts the vector of VM2 q-values from vm.result data object

**Usage**

```
qval.VM2.vect(data, lambda = seq(0, 0.95, 0.05), pi0.meth = "smoother", fdr.level =
```

**Arguments**

data	Gene expression data object
lambda	parameter for qvalue function
pi0.meth	parameter for qvalue function
fdr.level	parameter for qvalue function
robust	parameter for qvalue function

**Value**

The vector of q value computed from the p values of the VM2 method

**Author(s)**

Paul Delmar

---

sbsset.gene

*Subset of vm.result object using gene IDs*

---

**Description**

This function extracts the subset of vm.result object using gene IDs

**Usage**

```
sbsset.gene(data, gene)
```

**Arguments**

data	gene expression data object
gene	a vector of gene names matchin the gene id of the data object

**Value**

A subset of the data with only those genes also in the gene argument.

**Author(s)**

Paul Delmar

---

sbsset.pval

*Subset of vm.result object using gene p-values*

---

**Description**

This function finds and returns the subset of vm.result object using gene p-values

**Usage**

```
sbsset.pval(data, pval, method = c("gene", "VM", "VM2", "anova")[2])
```

**Arguments**

data	gene expression data result object
pval	P value cut off
method	Method for computing the q-value. Must be "gene" or "VM" or "VM2" or "anova".

**Value**

A subset of the gene expression data object with the genes below a given p value criterion.

**Author(s)**

Paul Delmar

---

sbsset.qval                      *Subset of vm.result object using gene q-values*

---

**Description**

This function gets the subset of vm.result object using gene q-values

**Usage**

```
sbsset.qval(data, qval, method = c("gene", "VM", "VM2", "anova")[2], lambda = seq(0,
```

**Arguments**

data	gene expression data result object
qval	FDR cut off
method	Method for computing the q-value. Must be "gene" or "VM" or "VM2" or "anova".
lambda	lambda parameter for computing the q value

**Value**

A subset of the gene expression data object with the genes below a given q value criterion.

**Author(s)**

Paul Delmar

---

`sd.param`*Standard deviation of the variance mixture model parameters*

---

**Description**

This function computes the standard deviation of the variance mixture model parameters

**Usage**

```
sd.param(data)
```

**Arguments**

`data`            Gene expression data object

**Value**

A data.frame with the value of the standard deviation of the mixture model parameters

**Author(s)**

Paul Delmar

---

`spleen.data.vm`*Spleen Data set*

---

**Description**

An example data set for paired data analysis. A list with the normalized Spleen data set as used in the original article.

**Usage**

```
data(spleen.data.vm)
```

**Format**

`spleen.data.vm` is a list with 3 elements :

`spleen.geneid` A vector with gene names

`spleen.logratio` A Matrix with 4360 rows and 6 columns with normalized log-ratio

`spleen.meanint` A Matrix with 4360 rows and 6 columns with normalized mean log-intensity

## References

P. Delmar, Robin, S., Tronik-Le Roux S. and Daudin J.-J. (2005) Mixture model on the variance for the differential analysis of gene expression data, JRSS series C, 54(1), 31:50

## Examples

```
data(spleen.data.vm)
```

---

```
vm.analysis
```

```
Variance mixture analysis on unpaired data
```

---

## Description

Performs variance mixture analysis on unpaired data

## Usage

```
vm.analysis(geneId, cont, treat, filename=NULL, gene.annot=NULL, badqual=NULL, qualtol=NULL,
            center=TRUE, loess.cor=FALSE, min.rep=2, penalty=c("AIC", "BIC") [1],
            criterion.1=c("likelihood", "parameter") [1], criterion.2=c("likeli", "parameter") [1],
            stop.crit.2=1.e-8)
```

## Arguments

geneId	vector with the identifier of each gene (one gene per row)
cont	matrix of gene log-intensity in condition 1. The matrix has one gene per row and one replicate per column
treat	matrix of gene log-intensity in condition 1
filename	filename for export. If NULL no files are exported
gene.annot	data.frame with further gene annotations
badqual	integer matrix of number of bad quality observation per gene and replicate. Do not use badqual with different number of replicates in each condition
qualtol	integer. Genes with more than qualtol bad quality observations are removed from the analysis. This argument is ignored is Badqual is NULL
n.mixt	integer. Number of component in the mixture model. If n.mixt=NULL the optimal number of components is computed by the function
center	Logical. If True each array*condition is centered to have a 0 mean log-intensity.
loess.cor	Logical. if True then loess transformation of the data is performed. If False no loess transformation is performed
min.rep	Minimum number of non-missing value per gene and per condition. Must be at least 2.
penalty	Character. Either "BIC" or "AIC". The criterion for choosing the number of variance groups.

display	Boolean. Should the result of model fitting be displayed on the screen, on the fly.
stop.crit.1	The stopping relative precision limit for stopping EM algorithm
criterion.1	either "likelihood" or "parameter". The criterion for stopping EM algorithm while trying to determine the number of variance groups
criterion.2	Either "likelihood" or "parameter". The criterion for stopping EM algorithm while trying to estimate the parameters.
stop.crit.2	The stopping relative precision limit for stopping EM algorithm

### Details

We highly recommend the use of pre-normalized data. The function only centers the data set. It can only perform very simple global array by array lowess transform. Typically the badqual matrix is generated by counting the number of times of gene is under the detection threshold (background) or at the saturation level on each physical array. A gene with more than "qualtol" bad quality observation is removed from the analysis. The corrected p-value (controlling FWER) are computed using the Bonferroni correction. The EM algorithm stops when the relative increase in likelihood or relative maximum absolute difference in parameter value becomes less than the stop.crit value. If  $t$  is the iteration number, and the criterion is likelihood then the stop criterion is  $(\log\text{-like}[t]-\log\text{-like}[t-1])/\log\text{-lik}[t-1]<\text{stop.crit}$

### Value

A gene expression data object with the results of the variance mixture analysis

geneid	the vector of gene names
raw.cond1	a matrix of the raw log-intensity in condition 1
raw.cond2	a matrix of the raw log-intensity in condition 2
cond1	a matrix of the normalized log-intensity in condition 1
cond2	a matrix of the normalized log-intensity in condition 2
stat1	a data.frame of results of homoscedastic and gene-specific model
stat1.call	call that generated the stat1 data.frame
df	number of degrees of freedom
var.to.vardelta	factor for computing the test statistic variance given the gene variance
residual	matrix of the residual
stat2	a data.frame of results of the variance mixture model analysis
stat2.call	call that generated the stat2 data.frame
param	a data.frame of value of the variance mixture paramters
call	The call to the function
ppost	The matrix of posterior probability that a gene belongs to each variance component. One row per gene, one column per variance component.
choose.nmixt	A data frame with some information of models with growing number of components

**Author(s)**

Paul Delmar and Julie Aubert

**References**

P. Delmar, Robin, S., Tronik-Le Roux S. and Daudin J.-J. (2005) Mixture model on the variance for the differential analysis of gene expression data, JRSS series C, 54(1), 31:50

**See Also**

[fdr.an](#), [export.res](#), [plotvm](#), [qqplot.vm](#), [plotrm](#), [boxplotvm](#), [compute.dif](#)

**Examples**

```
data(apo.data.vm)
res.apo<-vm.analysis(geneId=apo.data.vm$apo.geneid, cont=apo.data.vm$apo.cond1,
                    treat=apo.data.vm$apo.cond2)

fdr.an(res.apo, 0.01)
res.apo.data.frame<-export.res(res.apo)
par(mfrow=c(2,2))
plotvm(res.apo)
plotrm(res.apo)
boxplotvm(res.apo)
qqplot.vm(res.apo)
```

---

vm.analysis.paired *Variance mixture analysis on paired data*

---

**Description**

Performs variance mixture analysis on paired data

**Usage**

```
vm.analysis.paired(geneId, ratio, meanint=NULL, filename=NULL, gene.anot=NULL, badqual=N,
                  center=TRUE, loess.cor=FALSE, min.rep=2, penalty=c("AIC",
                  criterion.1=c("likelihood", "parameter")[1], stop.crit.
                  criterion.2=c("likelihood", "parameter")[1], stop.crit.
```

**Arguments**

geneId	vector with the identifier of each gene (one gene per row)
ratio	matrix of gene log-ratio. The matrix has one gene per row and one replicate per column
meanint	matrix of gene log-intensity
filename	filename for export. If NULL no files are exported
gene.anot	data.frame with further gene anotations

badqual	integer matrix of number of bad quality observation per gene and replicate
qualtol	integer. Genes with more than qualtol bad quality observations are removed from the analysis. This argument is ignored if Badqual is NULL
n.mixt	number of component in the mixture model. If n.mixt=NULL the optimal number of components is computed by the function
center	Logical. If True each array is centered to have a 0 mean log-ratio
loess.cor	Logical. If True then loess transformation of the data is performed. If False no loess transformation is performed
min.rep	Minimum number of non-missing value per gene and per condition. Must be at least 2
penalty	Either "BIC" or "AIC". The criterion for choosing the number of variance groups.
display	Boolean. Should the result of model fitting be displayed on the screen, on the fly.
criterion.1	either "likelihood" or "parameter". The criterion for stopping EM algorithm while trying to determine the number of variance groups
stop.crit.1	The stopping relative precision limit for stopping EM algorithm
criterion.2	Either "likelihood" or "parameter". The criterion for stopping EM algorithm while trying to estimate the parameters.
stop.crit.2	The stopping relative precision limit for stopping EM algorithm

### Details

We highly recommend the use of pre-normalized data. The function only centers the data set. It can only perform very simple global array by array lowess transform. Typically the badqual matrix is generated by counting the number of times of gene is under the detection threshold (background) or at the saturation level on each physical array. A gene with more than "qualtol" bad quality observation is removed from the analysis. The corrected p-value (controlling FWER) are computed using the Bonferroni correction. The EM algorithm stops when the relative increase in likelihood or relative maximum absolute difference in parameter value becomes less than the stop.crit value. If  $t$  is the iteration number, and the criterion is likelihood then the stop criterion is  $(\log\text{-like}[t]-\log\text{-like}[t-1])/\log\text{-lik}[t-1]<\text{stop.crit}$

### Value

a gene expression data object with the following elements :

geneid	a vector of the vector of gene names
raw.ratio	a matrix of the raw log-ratio
raw.meanint	a matrix of the raw mean log-intensity in condition 2
log.ratio	a matrix of the normalized log-ratio
meanint	a matrix of the normalized mean log-intensity
stat1	a data.frame of results of homoscedastic and gene-specific model
stat1.call	call that generated the stat1 data.frame

df	number of degrees of freedom
var.to.vardelta	factor for computing the test statistic variance given the gene variance
residual	matrix of the residual
stat2	a data frame of results of the variance mixture model analysis
stat2.call	call that generated the stat2 data.frame
param	a data.frame of value of the variance mixture paramters
call	The call to the function
ppost	The matrix of posterior probability that a gene belongs to each variance component. One row per gene, one column per variance component.
choose.nmixt	A data frame with some information of models with growing number of components

**Author(s)**

Paul Delmar and Julie Aubert

**References**

P. Delmar, Robin, S., Tronik-Le Roux S. and Daudin J.-J. (2005) Mixture model on the variance for the differential analysis of gene expression data, JRSS series C, 54(1), 31:50

**See Also**

[fdr.an](#), [export.res](#), [plotvm](#), [qqplot.vm](#), [plotrm](#), [boxplotvm](#), [compute.dif.paired](#)

**Examples**

```
data(spleen.data.vm)
res.spleen<-vm.analysis.paired(geneId=spleen.data.vm$spleen.geneid, ratio=spleen.data.vm$spleen.meanint)
fdr.an(res.spleen, 0.01)
res.spleen.data.frame<-export.res(res.spleen)
par(mfrow=c(2,2))
plotvm(res.spleen)
plotrm(res.spleen)
boxplotvm(res.spleen)
qqplot.vm(res.spleen)
```

# Index

## \*Topic **datasets**

apo.data.vm, 1  
spleen.data.vm, 21

## \*Topic **htest**

boxplotvm, 2  
boxplotvm.g, 3  
compare.method, 3  
compare.res, 4  
compare.var, 5  
export.res, 5  
fdr.an, 7  
fdr.compare.method, 8  
fdr.compare.res, 9  
find.qval.index, 10  
n.genes, 10  
plotrm, 11  
plotsdt, 11  
plotsdt.VM, 12  
plotvm, 12  
pval.an, 13  
qqplot.vm, 14  
qqplot.var.vm, 14  
qqplot.var.vm.2, 15  
qqplot.vm, 15  
qval.anova.vect, 16  
qval.gene.vect, 17  
qval.VM.vect, 17  
qval.VM2.vect, 18  
sbsset.gene, 19  
sbsset.pval, 19  
sbsset.qval, 20  
sd.param, 20  
vm.analysis, 22  
vm.analysis.paired, 24

apo.data.vm, 1

boxplotvm, 2, 11, 13, 16, 24, 26  
boxplotvm.g, 3

compare.method, 3, 4  
compare.res, 4, 4, 9  
compare.var, 5  
compute.dif, 24  
compute.dif.paired, 26

export.res, 5, 24, 26

fdr.an, 7, 13, 24, 26  
fdr.compare.method, 4, 8, 9  
fdr.compare.res, 4, 9, 9  
find.qval.index, 10

n.genes, 10

plotrm, 2, 3, 11, 13, 16, 24, 26  
plotsdt, 11  
plotsdt.VM, 12  
plotvm, 2, 3, 11, 12, 16, 24, 26  
pval.an, 8, 13

qqplot.vm, 14  
qqplot.var.vm, 14  
qqplot.var.vm.2, 15  
qqplot.vm, 2, 3, 11, 13, 15, 24, 26  
qval.anova.vect, 16  
qval.gene.vect, 17  
qval.VM.vect, 17  
qval.VM2.vect, 18

sbsset.gene, 19  
sbsset.pval, 19  
sbsset.qval, 20  
sd.param, 20  
spleen.data.vm, 21

vm.analysis, 2-4, 7-11, 13, 16, 22  
vm.analysis.paired, 2-4, 7-11, 13, 16,  
24