Package ‘cytofkit’

April 22, 2016

Type Package

Title cytofkit: an integrated analysis pipeline for mass cytometry data

Version 1.2.4

Date 2015-11-12

Author Jinmiao Chen, Hao Chen

Description An integrated mass cytometry data analysis pipeline that enables simultaneous illustration of cellular diversity and progression.

License Artistic-2.0

Maintainer
Jinmiao Chen <Chen_Jinmiao@immunol.a-star.edu.sg>, Hao Chen <chen_hao@immunol.a-star.edu.sg>

biocViews FlowCytometry, GUI, CellBiology, Clustering, DimensionReduction, BiomedicalInformatics

Depends R (>= 2.10.0), ggplot2(>= 0.9.3.1), plyr

VignetteBuilder knitr

Suggests knitr, RUnit, testthat, BiocGenerics

Imports tcltk, stats, Rtsne, e1071, flowCore, gplots, VGAM, reshape, reshape2, shiny, vegan, Biobase, doParallel, pdist, igraph(>= 1.0.1), RANN(>= 2.5), Rcpp (>= 0.12.0)

LinkingTo Rcpp

RoxygenNote 5.0.1

NeedsCompilation yes

R topics documented:

cytofkit-package 2
applyComp 4
auto_lgcl 4
ClusterX 5
cytofkit 6
cytofkitNews 8

cytofkitShinyAPP . . . 8
cytofkit_GUI . . . 9
cytof_addToFCS . . . 9
cytof_cluster . . . 10
cytof_clusterPlot . . . 11
cytof_dimReduction . . . 12
cytof_exprsExtract . . . 13
cytof_exprsMerge . . . 14
cytof_heatmap . . . 15
cytof_progression . . . 16
cytof_progressionPlot . . . 17
cytof_writeResults . . . 18
densVM . . . 19
find_neighbors . . . 19
getParameters_GUI . . . 20
launchShinyAPP_GUI . . . 21
peaksGamma_plot . . . 21
Rphenograph . . . 22
scaleData . . . 23

Index **24**

`cytofkit-package` *cytofkit: an integrated analysis pipeline for mass cytometry data*

Description

This package is designed to facilitate the analysis workflow of mass cytometry data with automatic subset identification and mapping of cellular progression. Both command line and a GUI client are provided for executing the workflow easily.

Details

This package integrates merging methods of multiple FCS files, dimension reduction methods (PCA, t-SNE and ISOMAP) and clustering methods (DensVM, densityClustX, and Rphenograph) for rapid subset detection. Cell subsets can be visualized in scatter plot and heat map. The method isomap is also provided to map the cellular progression. This workflow can be easily executed with the main function `cytofkit` or through the GUI client `cytofkit_GUI`.

Pre-processing

Using function `cytof_exprsMerge`, one or multiple FCS files will be loaded via the *read.FCS* function in the *flowCore* package. Then transformation was applied to the expression value of selected markers of each FCS file. Transformation methods include `auto_lgcl`, `fixed_lgcl`, `arcsin` and `biexp`, where `auto_lgcl` is the default.Then mutilple FCS files are merged using method `all`, `min`, `fixed` or `ceil`.

Dimensionality reduction

Using function `cytof_dimReduction`, t-Distributed Stochastic Neighbor Embedding (`tsne`) is suggested for dimensionality reduction although we also provide methods like `isomap` and `pca`.

Cluster

Using function cytof_cluster, three cluster method are provided, densVM, ClusterX and Rphenograph. densVM, densityClustX are performend on the dimension reducted data, while Rphenograph is peroformed directed on the high dimensional expression data.

Post-processing

- Using function cytof_clusterPlot to visualize the cluster results in a catter plot, in which dots represent cells, colours indicate their assigned clusters and point shapes represent their belonging samples.

- Using function cytof_heatmap to generate heat map to visualize the mean expression of every marker in every cluster. This heat maps is useful to interrogate marker expression to identify each cluster's defining markers.

- Using function cytof_progressionPlot to visualize the expression patter of selected markers against the estimated cellular progression order.

- Using function cytof_addToFCS to add any dimension reduced data, cluster results, progression data into the original FCS files, new FCS files will be saved for easy checking with other softwares like FlowJo.

All the above post processing can be automatically implemented and saved using one function cytof_writeResults.

References

http://signbioinfo.github.io/cytofkit/

See Also

cytofkit, cytofkit_GUI

Examples

```
## Run on GUI
#cytofkit_GUI()  # remove the hash symbol to launch the GUI

## Run on command
dir <- system.file('extdata',package='cytofkit')
file <- list.files(dir, pattern='.fcs$', full=TRUE)
parameters <- list.files(dir, pattern='.txt$', full=TRUE)
## remove the hash symbol to run the following command
#cytofkit(fcsFile = file, markers = parameters, projectName = 'test')

## Checking the vignettes for more details
if(interactive()) browseVignettes(package = 'cytofkit')
```

applyComp — *apply compensation on the FCS expression data*

Description

apply compensation on the FCS expression data

Usage

```
applyComp(fcs, keyword)
```

Arguments

fcs	FCS file.
keyword	Keywords.

auto_lgcl — *a modified version of "estimateLogicle" from flowCore*

Description

a modified version of "estimateLogicle" from flowCore

Usage

```
auto_lgcl(x, channels, m = 4.5, q = 0.05)
```

Arguments

x	Data.
channels	Channel names.
m	Para m.
q	Para q.

ClusterX *Fast clustering by automaticly search and find of density peaks*

Description

This package implement the clustering algorithm described by Alex Rodriguez and Alessandro Laio (2014) with improvements of automatic peak detection and parallel implementation

Usage

```
ClusterX(data, dimReduction = NULL, outDim = 2, dc, gaussian = TRUE,
  alpha = 0.001, detectHalos = FALSE, parallel = FALSE, nCore = 4)
```

Arguments

data	A data matrix for clustering.
dimReduction	Dimenionality reduciton method.
outDim	Number of dimensions will be used.
dc	Distance cutoff value.
gaussian	If apply gaussian to esitmate the density.
alpha	Signance level for peak detection.
detectHalos	If detect the halos.
parallel	If run the algorithm in parallel.
nCore	Number of cores umployed for parallel compution.

Value

a object of `ClusterX` class

Author(s)

Chen Hao

Examples

```
iris_unique <- unique(iris) # Remove duplicates
data <- as.matrix(iris_unique[,1:4])
ClusterXRes <- ClusterX(data)
```

cytofkit — *cytofkit: an integrated analysis pipeline for mass cytometry data*

Description

A user friendly GUI is provided for easy usage of cytofkit, cytofkit_GUI.

Usage

```
cytofkit(fcsFiles = getwd(), markers = NULL, projectName = "cytofkit",
  mergeMethod = "ceil", fixedNum = 10000, ifCompensation = FALSE,
  transformMethod = "auto_lgcl", dimReductionMethod = "tsne",
  clusterMethods = "ClusterX", visualizationMethods = "tsne",
  progressionMethod = NULL, uniformClusterSize = 500, resultDir = getwd(),
  saveResults = TRUE, saveObject = TRUE, saveToFCS = TRUE,
  scaleTo = NULL, q = 0.05, ...)
```

Arguments

- `fcsFiles` — it can be either the name of the path where stores your FCS files or a vector of FCS file names.
- `markers` — it can be either a text file that specifies the makers to be used for analysis or a vector of the marker names.
- `projectName` — a prefix that will be added to the names of result files.
- `mergeMethod` — when multiple fcs files are selected, cells can be combined using one of the four different methods including `ceil`, `all`, `min`, `fixed`. The default option is `ceil`, up to a fixed number (specified by `fixedNum`) of cells are sampled without replacement from each fcs file and combined for analysis. `all`: all cells from each fcs file are combined for analysis. `min`: The minimum number of cells among all the selected fcs files are sampled from each fcs file and combined for analysis. `fixed`: a fixed num (specified by fixedNum) of cells are sampled (with replacement when the total number of cell is less than fixedNum) from each fcs file and combined for analysis.
- `fixedNum` — up to fixedNum of cells from each fcs file are used for analysis.
- `ifCompensation` — Boolean value to decide if do compensation. This will be applied to flow cytometry data.
- `transformMethod` — dat transformation method, either `auto_lgcl`, `fixed_lgcl`, `arcsin` or `biexp`.
- `dimReductionMethod` — the method used for dimensionality reduction, including `tsne`, `pca` and `isomap`.
- `clusterMethods` — the clustering method(s) used for subpopulation detection, including `densVM`, `ClusterX` and `Rphenograph`. Multiple selection are accepted.
- `visualizationMethods` — the method(s) used for visualize the cluster data, including `tsne`, `pca` and `isomap`. Multiple selection are accepted.

`progressionMethod`
: use the first ordination score of `isomap` to estimated the preogression order of cells, choose `NULL` to ignore.

`uniformClusterSize`
: the uniform size of each cluster.

`resultDir`
: the directory where result files will be generated.

`saveResults`
: if save the results, and the post-processing results including scatter plot, heatmap, and statistical results.

`saveObject`
: save the resutls into RData objects for loading back to R for further analysis

`saveToFCS`
: save the results back to the FCS files, new FCS files will be generated.

`scaleTo`
: scale the expression values to the same scale after transformation, default is NULL, should be a vector of two numbers if scale.

`q`
: quantile of negative values removed for auto w estimation in logicle transformation, default is 0.05.

`...`
: more arguments contral the logicle transformation

Details

`cytofkit` provides a workflow for one or multiple CyTOF data analysis, including data preprocess with merging methods of multiple fcs file, expression data transformation, dimension reduction with PCA, isomap or tsne(default), clustering methods(densVM, ClusterX, Rphenograph) for subpopulation detection, and estimation of cellular progression with isomap. The analysis results can be visualized with scatter plot, heatmap plot or progression plot. Moreover theses results can be saved back to FCS files. By default the results will be automatically saved for further annotation. An interactive web application is provided for interactive exploration of the analysis results, `cytofkitShinyAPP`.

Value

a list containing `expressionData`, `dimReductionMethod`, `visualizationMethods`, `dimReducedRes`, `clusterRes` and `progressionRes`. If choose 'saveResults = TRUE', results will be saved into files under `resultDir`

Author(s)

Chen Jinmiao, Chen Hao

References

http://signbioinfo.github.io/cytofkit/

See Also

`cytofkit`, `cytofkit_GUI`

Examples

```
dir <- system.file('extdata',package='cytofkit')
file <- list.files(dir, pattern='.fcs$', full=TRUE)
parameters <- list.files(dir, pattern='.txt$', full=TRUE)
## remove the hash symbol to run the following command
#cytofkit(fcsFile = file, markers = parameters, projectName = 'test')
```

`cytofkitNews` *check the package update news*

Description

check the package update news

Usage

```
cytofkitNews()
```

`cytofkitShinyAPP` *A Shiny app to interactively visualize the analysis results*

Description

Load the RData object saved by cytofkit, explore the analysis results with interactive control

Usage

```
cytofkitShinyAPP()
```

Examples

```
if (interactive()) cytofkit::cytofkitShinyAPP()
```

cytofkit_GUI *The user friendly GUI client for* `cytofkit-package`

Description

This GUI provides an easy way for CyToF data analysis using `cytofkit` package. Main parameters for running 'cytofkit' were integrated in this GUI, and each parameter has a help button to show the instruction. `cytofkit` analysis will be launched after submitting.

Usage

```
cytofkit_GUI()
```

Value

the GUI for `cytofkit-package`

Author(s)

Chen Jinmiao, Chen Hao

References

http://signbioinfo.github.io/cytofkit/

See Also

`cytofkit-package`, `cytofkit`

Examples

```
#cytofkit_GUI()  # remove the comment hash to run
```

cytof_addToFCS *Add data to the original FCS files*

Description

Store the new dimension transformed data and cluster data into the exprs matrix in new fcs files under `analyzedFCSdir`

Usage

```
cytof_addToFCS(data, rawFCSdir, analyzedFCSdir, transformed_cols = c("tsne_1",
  "tsne_2"), cluster_cols = c("cluster"), inLgclTrans = TRUE)
```

Arguments

`data`	The new data matrix to be added in.
`rawFCSdir`	The directory containing the original fcs files.
`analyzedFCSdir`	The directory to store the new fcs files.
`transformed_cols`	the column name of the dimension transformend data in `data`.
`cluster_cols`	the column name of the cluster data in `data`.
`inLgclTrans`	Boolean value decides if apply the inverse lgcl transformation to the data before saving

Value

new fcs files stored under `analyzedFCSdir`

`cytof_cluster` *Subset detection by clustering*

Description

Apply clustering algorithms to detect cell subsets. `densVM` and `densityClustX` clustering is based on the transformend ydata; `Rphenograph` is directly applied on the high dimemnional xdata. And `densVM` need the xdata to train the VM model.

Usage

```
cytof_cluster(ydata = NULL, xdata = NULL, method = "densVM")
```

Arguments

`ydata`	a matrix of the dimension reduced(transformed) data
`xdata`	a matrix of the expression data
`method`	cluster method including `densVM`, `densityClustX` and `Rphenograph`.

Value

a vector of the clusters assigned for each row of the ydata

Examples

```
d<-system.file('extdata', package='cytofkit')
fcsFile <- list.files(d, pattern='.fcs$', full=TRUE)
xdata <- cytof_exprsMerge(fcsFile, mergeMethod = 'fixed', fixedNum = 100)
ydata <- cytof_dimReduction(xdata, method = "tsne")
clusters <- cytof_cluster(ydata, xdata, method = "densVM")
```

cytof_clusterPlot *Scatter plot of the cluster results*

Description

Dot plot visualization of the cluster results, with color indicating different clusters, and shape of different samples.

Usage

```
cytof_clusterPlot(data, xlab, ylab, cluster, sample, title = "cluster",
  type = 1, point_size = NULL)
```

Arguments

data	The data frame of cluster results, which should contains at least xlab, ylab and cluster
xlab	The column name of the x axis in input data
ylab	The column name of the y axis in input data
cluster	The column name of cluster in input data
sample	the column name of the sample in input data
title	the title of the plot
type	plot type, 1 indicates combined plot, 2 indicated grid facet plot seperated by samples
point_size	the size of the dot.

Value

the ggplot object of the scatter cluster plot

Examples

```
x <- c(rnorm(100, mean = 1), rnorm(100, mean = 3), rnorm(100, mean = 9))
y <- c(rnorm(100, mean = 2), rnorm(100, mean = 8), rnorm(100, mean = 5))
c <- c(rep(1,100), rep(2,100), rep(3,100))
rnames <- paste(paste('sample_', c('A','B','C'), sep = ''), rep(1:100,each = 3), sep='_')
data <- data.frame(dim1 = x, dim2 = y, cluster = c)
rownames(data) <- rnames
data$sample <- "data"
cytof_clusterPlot(data, xlab="dim1", ylab="dim2", cluster="cluster", sample = "sample")
```

cytof_dimReduction *Dimension reduction of cytof expression data*

Description

Apply dimension reduction on the cytof expression data, with method `isomap`, `pca`, or `tsne`.

Usage

```
cytof_dimReduction(data, method = "tsne", distMethod = "euclidean",
  out_dim = 2, isomap_k = 5, isomap_ndim = NULL,
  isomapFragmentOK = TRUE)
```

Arguments

`data`	An expression data matrix.
`method`	Method chosed for dimensition reduction, must be one of `isomap`, `pca` or `tsne`.
`distMethod`	Method for distance calcualtion.
`out_dim`	The dimensionality of the output.
`isomap_k`	Number of shortest dissimilarities retained for a point, parameter for `isomap` method.
`isomap_ndim`	Number of axes in metric scaling, parameter for `isomap` method.
`isomapFragmentOK`	What to do if dissimilarity matrix is fragmented, parameter for `isomap` method.

Value

a matrix of the dimension reducted data, with colnames and rownames(if have, same as the input).

Author(s)

Chen Jinmiao

Examples

```
data(iris)
in_data <- iris[, 1:4]
out_data <- cytof_dimReduction(in_data)
```

cytof_exprsExtract — *Extract the expression matrix of the FCS data*

Description

Extract the FCS expresssion data and apply the transformation

Usage

```
cytof_exprsExtract(fcsFile, comp = FALSE, verbose = FALSE, markers = NULL,
  transformMethod = "auto_lgcl", scaleTo = NULL, w = 0.1, t = 4000,
  m = 4.5, a = 0, q = 0.05)
```

Arguments

fcsFile	The name of the FCS file
comp	Boolean value tells if do compensation
verbose	Boolean value detecides if print the massage details
markers	Selected markers for analysis, either from names or from description
transformMethod	transformation method, auto_lgcl, fixed_lgcl, arcsin or biexp
scaleTo	scale the expression to same scale, default is NULL, should be a vector of two numbers if scale
w	Linearization width in asymptotic decades
t	Top of the scale data value
m	Full width of the transformed display in asymptotic decades
a	Additional negative range to be included in the display in asymptotic decades
q	quantile of negative values removed for auto w estimation, default is 0.05

Value

The transformend expression data matrix with selected markers

Examples

```
d<-system.file('extdata',package='cytofkit')
fcsFile <- list.files(d,pattern='.fcs$',full=TRUE)
transformed <- cytof_exprsExtract(fcsFile)
```

`cytof_exprsMerge` *Merge the transformed expression data of FCS file(s) of selected markers*

Description

Apply transformation of selected markers of each FCS file, arcsin, biexponential, auto logicle transformation and fixed logicle transformation are provided, then mutilple FCS files are merged using method `all`, `min`, `fixed` or `ceil`

Usage

```
cytof_exprsMerge(fcsFiles, comp = FALSE, verbose = FALSE, markers = NULL,
  transformMethod = "auto_lgcl", scaleTo = NULL, mergeMethod = "ceil",
  fixedNum = 10000, w = 0.1, t = 4000, m = 4.5, a = 0, q = 0.05)
```

Arguments

`fcsFiles`	the input fcsFiles (usually more than 1 file)
`comp`	Boolean value tells if do compensation
`verbose`	Boolean value detecides if print the massage details
`markers`	Selected markers for analysis, either from names or from description
`transformMethod`	transformation method, `auto_lgcl`, `fixed_lgcl`, `arcsin` or `biexp`
`scaleTo`	scale the expression to same scale, default is NULL, should be a vector of two numbers if scale
`mergeMethod`	merge method for mutiple FCS expression data. cells can be combined using one of the four different methods including `ceil`, `all`, `min`, `fixed`. The default option is `ceil`, up to a fixed number (specified by `fixedNum`) of cells are sampled without replacement from each fcs file and combined for analysis. `all`: all cells from each fcs file are combined for analysis. `min`: The minimum number of cells among all the selected fcs files are sampled from each fcs file and combined for analysis. `fixed`: a fixed num (specified by fixedNum) of cells are sampled (with replacement when the total number of cell is less than fixedNum) from each fcs file and combined for analysis.
`fixedNum`	the fixed number of cells for merging multiple FCSs
`w`	Linearization width in asymptotic decades
`t`	Top of the scale data value
`m`	Full width of the transformed display in asymptotic decades
`a`	Additional negative range to be included in the display in asymptotic decades
`q`	quantile of negative values removed for auto w estimation, default is 0.05

Value

Merged FCS expression data matrix of selected markers after transformation

Examples

```
d<-system.file('extdata',package='cytofkit')
fcsFiles <- list.files(d,pattern='.fcs$',full=TRUE)
merged <- cytof_exprsMerge(fcsFiles)
```

`cytof_heatmap` *Heatmap plot of cluster mean value results*

Description

Heatmap plot of cluster mean value results

Usage

```
cytof_heatmap(data, baseName = "Cluster", scaleMethod = "none")
```

Arguments

`data`	a matrix with rownames and colnames
`baseName`	The name as a prefix in the title of the heatmap.
`scaleMethod`	character indicating if the values should be centered and scaled in either the row direction or the column direction, or none. The default is 'none'.

Value

a heatmap object from `gplots`

Examples

```
m1 <- c(rnorm(300, 10, 2), rnorm(400, 4, 2), rnorm(300, 7))
m2 <- c(rnorm(300, 4), rnorm(400, 16), rnorm(300, 10, 3))
m3 <- c(rnorm(300, 16), rnorm(400, 40, 3), rnorm(300, 10))
m4 <- c(rnorm(300, 7, 3), rnorm(400, 30, 2), rnorm(300, 10))
m5 <- c(rnorm(300, 27), rnorm(400, 40, 1),rnorm(300, 10))
c <- c(rep(1,300), rep(2,400), rep(3,300))
rnames <- paste(paste('sample_', c('A','B','C','D'), sep = ''),
rep(1:250,each = 4), sep='_')
exprs_cluster <- data.frame(cluster = c, m1 = m1, m2 = m2, m3 = m3, m4 = m4, m5 = m5)
row.names(exprs_cluster) <- sample(rnames, 1000)
cluster_mean <- aggregate(. ~ cluster, data = exprs_cluster, mean)
rownames(cluster_mean) <- paste("cluster_", cluster_mean$cluster, sep = "")
cytof_heatmap(cluster_mean[, -which(colnames(cluster_mean) == "cluster")])
```

cytof_progression — *Progression estimation of cytof expression data*

Description

Apply isomap to estimate the relationship of cell progression

Usage

```
cytof_progression(data, cluster, method = "isomap",
  uniformClusterSize = 500, seed = 500)
```

Arguments

`data`	Expression data matrix.
`cluster`	A vector of cluster results for the data.
`method`	Method for estimation of cell progression, isomap by default, tsne or pca.
`uniformClusterSize`	The down sampled size of each cluster.
`seed`	The seed for random down sample of the clusters.

Value

a list includes sampleData, sampleCluster and progressionData.

Author(s)

Chen Jinmiao

Examples

```
data(iris)
in_data <- iris[, 1:4]
out_data <- cytof_progression(in_data, cluster = iris[,5], uniformClusterSize = 50)
```

cytof_progressionPlot *Progression plot*

Description

Plot the expression trend along the estimated cell progressing order

Usage

```
cytof_progressionPlot(data, markers, orderCol = "isomap_1",
  clusterCol = "cluster", min_expr = NULL,
  trend_formula = "expression ~ sm.ns(Pseudotime, df=3)")
```

Arguments

data	The data frame for progression plot.
markers	The column names of the selected markers for visualization.
orderCol	The column name of the estimated cell progression order.
clusterCol	The column name of the cluster results
min_expr	the threshold of the minimal expression value for markers
trend_formula	a symbolic description of the model to be fit.

Value

a heatmap object from `gplots`

Examples

```
m1 <- c(rnorm(300, 10, 2), rnorm(400, 4, 2), rnorm(300, 7))
m2 <- c(rnorm(300, 4), rnorm(400, 16), rnorm(300, 10, 3))
m3 <- c(rnorm(300, 16), rnorm(400, 40, 3), rnorm(300, 10))
m4 <- c(rnorm(300, 7, 3), rnorm(400, 30, 2), rnorm(300, 10))
m5 <- c(rnorm(300, 27), rnorm(400, 40, 1),rnorm(300, 10))
c <- c(rep(1,300), rep(2,400), rep(3,300))
rnames <- paste(paste('sample_', c('A','B','C','D'), sep = ''),
rep(1:250,each = 4), sep='_')
exprs_cluster <- data.frame(cluster = c, m1 = m1, m2 = m2, m3 = m3, m4 = m4, isomap_1 = m5)
row.names(exprs_cluster) <- sample(rnames, 1000)
cytof_progressionPlot(exprs_cluster, markers = c("m1","m2","m3","m4"))
```

cytof_writeResults *Save the cytofkit analysis results*

Description

Scatter dot plot and heatmap of the cluster results, and all intermediate files will be generated and saved in the `resultDir`

Usage

```
cytof_writeResults(analysis_results, projectName = "cytofkit",
  resultDir = getwd(), saveToFCS = TRUE, rawFCSdir = getwd())
```

Arguments

`analysis_results`	result data from output of `cytofkit`
`projectName`	a prefix that will be added to the names of result files.
`resultDir`	the directory where result files will be generated.
`saveToFCS`	save the results back to the FCS files, new FCS files will be generated.
`rawFCSdir`	the directory that contains fcs files to be analysed.

Value

save all results in the `resultDir`

See Also

`cytofkit`

Examples

```
d <- system.file('extdata',package='cytofkit')
f <- list.files(d, pattern='.fcs$', full=TRUE)
p <- list.files(d, pattern='.txt$', full=TRUE)
#tr <- cytofkit(fcsFile=f,markers=p,projectName='t',saveResults=FALSE)
#cytof_write_results(tr,projectName = 'test',resultDir=d,rawFCSdir =d)
```

densVM *Density-based local maxima cluster with SVM*

Description

Density-based local maxima peak finding, subpopulation assigning with the power of SVM

Usage

```
densVM(ydata, xdata)
```

Arguments

ydata	a matrix of the dimension reduced(transformed) data
xdata	a matrix of the expression data

Value

a list contains a matrix `peakdata` of the peak numbers with different kernel bandwidth, and a matrix `clusters` of the cluster results

Author(s)

Chen Jinmiao

Examples

```
d<-system.file('extdata',package='cytofkit')
fcsFile <- list.files(d,pattern='.fcs$',full=TRUE)
xdata <- cytof_exprsMerge(fcsFile, mergeMethod = 'fixed', fixedNum = 100)
ydata <- cytof_dimReduction(xdata)
#clusters <- densVM(ydata, xdata)
```

find_neighbors *K Nearest Neighbour Search*

Description

Uses a kd-tree to find the p number of near neighbours for each point in an input/output dataset.

Usage

```
find_neighbors(data, k)
```

Arguments

`data`	Input data matrix.
`k`	Number of nearest neighbours.

Details

Use the nn2 function from the RANN package, utilizes the Approximate Near Neighbor (ANN) C++ library, which can give the exact near neighbours or (as the name suggests) approximate near neighbours to within a specified error bound. For more information on the ANN library please visit http://www.cs.umd.edu/~mount/ANN/.

Value

a n-by-k matrix of neighbor indices

Examples

```
iris_unique <- unique(iris) # Remove duplicates
data <- as.matrix(iris_unique[,1:4])
neighbors <- find_neighbors(data, k=10)
```

`getParameters_GUI` *GUI for marker selection*

Description

Extract the markers from the fcsfiles

Usage

```
getParameters_GUI(fcsFile, rawFCSdir)
```

Arguments

`fcsFile`	The name of the FCS file
`rawFCSdir`	The path of the FCS file

Examples

```
#getParameters_GUI()
```

`launchShinyAPP_GUI` *GUI for launching shiny APP*

Description

A shiny APP for interactive exploration of the analysis results

Usage

```
launchShinyAPP_GUI(message)
```

Arguments

`message` A message to determine if open the shiny APP

Examples

```
# launchShinyAPP_GUI()
```

`peaksGamma_plot` *Plot varaition of peak nums with increasing gamma*

Description

Plot varaition of peak nums with increasing gamma

Usage

```
peaksGamma_plot(peakdata)
```

Arguments

`peakdata` a matrix of `peakdata` returned from `densVM_cluster`

Value

a line graph of peak nums vs. increasing gamma

Examples

```
x <- seq(0, 1, length.out = 20)
y <- c(20:6, 6, 6, 5:3)
peakdata <- data.frame(sig_range = x, numpeaks = y)
peaksGamma_plot(peakdata)
```

`Rphenograph` *RphenoGraph clustering*

Description

R implementation of the phenograph algorithm

Usage

```
Rphenograph(data, k = 30)
```

Arguments

`data`	Input data matrix.
`k`	Number of nearest neighbours, default is 30.

Details

A simple R implementation of the phenograph [PhenoGraph](http://www.cell.com/cell/abstract/S0092-8674(15)00637-6) algorithm, which is a clustering method designed for high-dimensional single-cell data analysis. It works by creating a graph ("network") representing phenotypic similarities between cells by calclating the Jaccard coefficient between nearest-neighbor sets, and then identifying communities using the well known [Louvain method](https://sites.google.com/site/findcommunities/) in this graph.

Value

a communities object, the operations of this class contains:

`print`	returns the communities object itself, invisibly.
`length`	returns an integer scalar.
`sizes`	returns a numeric vector.
`membership`	returns a numeric vector, one number for each vertex in the graph that was the input of the community detection.
`modularity`	returns a numeric scalar.
`algorithm`	returns a character scalar.
`crossing`	returns a logical vector.
`is_hierarchical`	returns a logical scalar.
`merges`	returns a two-column numeric matrix.
`cut_at`	returns a numeric vector, the membership vector of the vertices.
`as.dendrogram`	returns a dendrogram object.
`show_trace`	returns a character vector.
`code_len`	returns a numeric scalar for communities found with the InfoMAP method and NULL for other methods.
`plot`	for communities objects returns NULL, invisibly.

Author(s)

Chen Hao

References

Jacob H. Levine and et.al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell, 2015.

Examples

```
iris_unique <- unique(iris) # Remove duplicates
data <- as.matrix(iris_unique[,1:4])
Rphenograph_out <- Rphenograph(data, k = 45)
```

`scaleData` *rescale the data*

Description

rescale the data

Usage

```
scaleData(x, range = c(0, 4.5))
```

Arguments

`x`	data.
`range`	The range of the data.

Index

applyComp, 4
auto_lgcl, 4

ClusterX, 5
cytof_addToFCS, *3*, 9
cytof_cluster, *3*, 10
cytof_clusterPlot, *3*, 11
cytof_dimReduction, *2*, 12
cytof_exprsExtract, 13
cytof_exprsMerge, *2*, 14
cytof_heatmap, *3*, 15
cytof_progression, 16
cytof_progressionPlot, *3*, 17
cytof_writeResults, *3*, 18
cytofkit, *2, 3*, 6, *7*, *9*, *18*
cytofkit-package, 2
cytofkit_GUI, *2, 3*, *6, 7*, 9
cytofkitNews, 8
cytofkitShinyAPP, 8

densVM, 19

find_neighbors, 19

getParameters_GUI, 20

launchShinyAPP_GUI, 21

peaksGamma_plot, 21

Rphenograph, 22

scaleData, 23