Package 'cytoKernel'

August 21, 2025

Type Package

Title Differential expression using kernel-based score test

Version 1.15.0

Date 2021-09-27

Imports Rcpp, SummarizedExperiment, utils, methods, ComplexHeatmap, circlize, ashr, data.table, BiocParallel, dplyr, stats, magrittr, rlang, S4Vectors

LinkingTo Rcpp

Depends R (>= 4.1)

Suggests knitr, rmarkdown, BiocStyle, testthat

VignetteBuilder knitr

Encoding UTF-8

License GPL-3

Description cytoKernel implements a kernel-based score test to identify differentially expressed features in high-dimensional biological experiments. This approach can be applied across many different high-dimensional biological data including gene expression data and dimensionally reduced cytometry-based marker expression data. In this R package, we implement functions that compute the feature-wise p values and their corresponding adjusted p values. Additionally, it also computes the feature-wise shrunk effect sizes and their corresponding shrunken effect size. Further, it calculates the percent of differentially expressed features and plots user-friendly heatmap of the top differentially expressed features on the rows and samples on the columns.

biocViews ImmunoOncology, Proteomics, SingleCell, Software, OneChannel, FlowCytometry, DifferentialExpression, GeneExpression, Clustering

BugReports https://github.com/Ghoshlab/cytoKernel/issues

RoxygenNote 7.1.2

git_url https://git.bioconductor.org/packages/cytoKernel

git_branch devel

git_last_commit 5dc28e4

git_last_commit_date 2025-04-15

Repository Bioconductor 3.22

Date/Publication 2025-08-21

Author Tusharkanti Ghosh [aut, cre],
Victor Lui [aut],
Pratyaydipta Rudra [aut],
Souvik Seal [aut],
Thao Vu [aut],
Elena Hsieh [aut],
Debashis Ghosh [aut, cph]

Maintainer Tusharkanti Ghosh <tusharkantighosh30@gmail.com>

Contents

`cytoHDBMW` *Example of processed dimensionally reduced flow cytometry (marker median intensities) Bodenmiller_BCR_XL_flowSet() expression dataset from HDCytoData Bioconductor data package.*

Description

The raw data (fcs files) were pre-processed using CATALYST, scuttle, scran Bioconductor packages and igraph CRAN package. The data processing package includes 4 steps and they are as follows: 1. Creating a SingleCellExperiment Object: the flowSet data object along with the metadata are converted into a SingleCellExperiment object using the CATALYST R/Bioconductor package. 2. Clustering: We apply Louvain algorithm using the R package igraph to cluster the expression values by the type markers (surface markers). 3. Median: Medians are calculated within a cluster for every signaling marker and subject. 4. Aggregating and converting the data: We convert the

aggregated data into a SummarizedExperiment. The row meta-data indicates "cluster" corresponding to the cluster id for each protein marker. The colData represents the "sample_id", "condition", "patient_id", "ids" The remaining columns indicate median expression intensities for each of the 126 (14 markers * 9 clusters) cluster combination for each sample.

Usage

```
data(cytoHDBMW)
```

Format

SummarizedExperiment assay object containing 126 cluster-marker median expression intenities (features) of 8 subjects (samples).

Details

The HDCytoData package is an extensible resource containing a set of publicly available high-dimensional flow cytometry and mass cytometry (CyTOF) benchmark datasets hosted on Bioconductor's ExperimentHub platform.

References

Weber, M L, Soneson, Charlotte (2019). "HDCytoData: Collection of high-dimensional cytometry benchmark datasets in Bioconductor object formats." F1000Research, 8(v2), 1459.

Examples

```
data(cytoHDBMW)
```

`CytoK` *CytoK*

Description

This function applies a kernel-based score test for identifying differentially expressed features in high-throughput experiments, called the the CytoK procedure. This function also defines the CytoK class and constructor.

Usage

```
CytoK(
  object,
  group_factor,
  lowerRho = 2,
  upperRho = 12,
  gridRho = 4,
  alpha = 0.05,
  featureVars = NULL
)
```

Arguments

object	an object which is a matrix or data.frame with features (e.g. cluster-marker combinations or genes) on the rows and samples as the columns. Alternatively, a user can provide a SummarizedExperiment object and the assay(object) will be used as input for the CytoK procedure.
group_factor	a group level binary categorical response associated with each sample or column in the object. The order of the group_factor must match the order of the columns in object.
lowerRho	(Optional) lower bound of the kernel parameter.
upperRho	(Optional) upper bound of the kernel parameter.
gridRho	(Optional) number of grid points in the interval [lowerRho, upperRho].
alpha	(Optional) level of significance to control the False Discovery Rate (FDR). Default is 0.05.
featureVars	(Optional) Vector of the columns which identify features. If a 'SummarizedExperiment' is used for 'data', row variables will be used.

Details

CytoK (Kernel-based score test in biological feature differential analysis) is a nonlinear approach, which identifies differentially expressed features in high-dimensional biological experiments. This approach can be applied across many different high-dimensional biological data including Flow/Mass Cytometry data and other variety of gene expression data. The CytoK procedure employs a kernel-based score test to identify differentially expressed features. This procedure can be easily applied to a variety of measurement types since it uses a Gaussian distance based kernel.

This function computes the feature-wise p values and their corresponding adjusted p values. Additionally, it also computes the feature-wise shrunk effect sizes and their corresponding shrunk effect size sd's. Further, it calculates the percent of differentially expressed features. See the vignette for more details.

Value

A object of the class CytoK that contains a data.frame of the CytoK features in the CytoKFeatures slot, a data.frame of the CytoK features in the CytoKFeaturesOrdered slot ordered by adjusted p values from low to high, a numeric value of the CytoK differentially expressed features CytoKDEfeatures slot, a data.frame or SummarizedExperiment original data objject in the CytoKData slot, a numeric value of the level of significance in the CytoKalpha slot and (optional) a vector of the columns which identify features in the CytoKfeatureVars slot.

References

Liu D, Ghosh D, Lin X. Estimation and testing for the effect of a genetic pathway on a disease outcome using logistic kernel machine regression via logistic mixed models. BMC Bioinf. 2008; 9(1):292.

Zhan X, Ghosh D. Incorporating auxiliary information for improved prediction using combination of kernel machines. Stat Methodol. 2015; 22:47–57.

Zhan, X., Patterson, A.D. & Ghosh, D. Kernel approaches for differential expression analysis of mass spectrometry-based metabolomics data. BMC Bioinformatics 16, 77 (2015). https://doi.org/10.1186/s12859-015-0506-3

Matthew Stephens, False discovery rates: a new deal, Biostatistics, Volume 18, Issue 2, April 2017, Pages 275–294, https://doi.org/10.1093/biostatistics/kxw041

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
data("cytoHDBMW")
data_CytoK_HD <- CytoK(object=cytoHDBMW,
group_factor = rep(c(0, 1), c(4, 4)), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
```

CytoK-class *the CytoK class*

Description

Objects of this class store needed information to work with a CytoK object

Value

`CytoKFeatures` returns the data.frame with shrunk effect size, shrunk effect size sd, unadjusted p value and adjusted p value for each feature, `CytoKFeaturesOrdered` returns the data.frame with shrunk effect size, shrunk effect size sd, unadjusted p value and adjusted p value for each feature ordered by unadjusted p value from low to high, `CytoKDEfeatures` returns the percent of differentially expressed features based on alpha (level of significance), `CytoKData` returns the original data object, `CytoKalpha` returns the specified level of significance. Default is alpha=0.05. `CytoKFeatureVars` returns the value of featureVars. Default is NULL.

Slots

`CytoKFeatures` CytoK features

`CytoKFeaturesOrdered` CytoK features ordered by adjusted p values

`CytoKDEfeatures` Percent of Differentially Expressed CytoK features

`CytoKData` Original data object passed to `CytoK`

`CytoKalpha` Value of `alpha` argument passed to `CytoK`

`CytoKFeatureVars` Value of `featureVars` passed to `CytoK`. NULL if `featureVars` is left blank

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
```

CytoKalpha *Generic function that returns the CytoK level of significance (alpha) Given a CytoK object, this function returns the CytoK alpha*

Description

Accessors for the 'CytoKalpha' slot of a CytoK object.

Usage

```
CytoKalpha(object)

## S4 method for signature 'CytoK'
CytoKalpha(object)
```

Arguments

object	an object of class CytoK.

Value

Value of CytoKalpha argument passed to CytoK

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKalpha(data_CytoK)
```

CytoKData *Generic function that returns the CytoK Data*

Description

Given a CytoK object, this function returns the CytoK Data

Accessors for the 'CytoKData' slot of a CytoK object.

Usage

```
CytoKData(object)

## S4 method for signature 'CytoK'
CytoKData(object)
```

Arguments

object	an object of class CytoK.

Value

Original data object passed to CytoK.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKData(data_CytoK)
```

CytoKDEData *Differentially expressed data by cytoKernel*

Description

Select CytoK object according to the differentially expressed features identified by cytoKernel. Features are filtered if their adjusted p values are greater than CytoKalpha.

Usage

```
CytoKDEData(object, by = c("features"))
```

Arguments

object	a CytoK object from CytoK
by	String specifying which adjusted p values of the features to filter by. Default is "features".

Value

A list of data.frame's or a SummarizedExperiment. If a data.frame was originally input into the CytoK function, a list with two elements, DEdata, nonDEfeatures, will be returned. If a SummarizedExperiment was originally input, output will be a SummarizedExperiment with the filtered assay with one metadata object nonDEfeatures and four row meta-data EffectSize, EffectSizeSD, pvalue and padj.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKDEData(data_CytoK, by = "features")
```

`CytoKDEfeatures` *Generic function that returns the CytoK Differentially Expressed (DE) features Given a CytoK object, this function returns the CytoK DE features*

Description

Accessors for the 'CytoKDEfeatures' slot of a CytoK object.

Usage

```
CytoKDEfeatures(object)

## S4 method for signature 'CytoK'
CytoKDEfeatures(object)
```

Arguments

`object`	an object of class `CytoK`.

Value

The percent of differentially expressed features based on alpha (level of significance).

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKDEfeatures(data_CytoK)
```

CytoKFeatures *Generic function that returns the CytoK features*

Description

Given a CytoK object, this function returns the CytoK features

Accessors for the 'CytoKFeatures' slot of a CytoK object.

Usage

```
CytoKFeatures(object)

## S4 method for signature 'CytoK'
CytoKFeatures(object)
```

Arguments

object	an object of class CytoK.

Value

The data.frame with shrunk effect size, shrunk effect size sd, unadjusted p value and adjusted p value for each feature.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKFeatures(data_CytoK)
```

CytoKFeaturesOrdered *Generic function that returns the ordered CytoK features*

Description

Given a CytoK object, this function returns the CytoK features ordered by adjusted p values

Accessors for the 'CytoKFeaturesOrdered' slot of a CytoK object.

Usage

```
CytoKFeaturesOrdered(object)

## S4 method for signature 'CytoK'
CytoKFeaturesOrdered(object)
```

Arguments

object	an object of class CytoK.

Value

the data.frame with shrunk effect size, shrunk effect size sd, unadjusted p value and adjusted p value for each feature ordered by unadjusted p value from low to high.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKFeaturesOrdered(data_CytoK)
```

CytoKFeatureVars *Generic function that returns the CytoK Feature Vars*

Description

Given a CytoK object, this function returns the CytoK Feature Vars

Accessors for the 'CytoKFeatureVars' slot of a CytoK object.

Usage

```
CytoKFeatureVars(object)

## S4 method for signature 'CytoK'
CytoKFeatureVars(object)
```

Arguments

object	an object of class CytoK.

Value

Value of featureVars passed to CytoK. NULL if featureVars was left blank

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
CytoKFeatureVars(data_CytoK)
```

CytoKProc *CytoKProc*

Description

This function is a helper function that computes the shrunk effective size mean, shrunk effective size standard deviation (sd), p value and adjusted p value of each feature for the function CytoK.

Usage

```
CytoKProc(object, group_factor, lowerRho = 2, upperRho = 12, gridRho = 4)
```

Arguments

object	an object which is a matrix or data.frame with features (e.g. cluster-marker combinations or genes) on the rows and samples (group factors) as the columns. Alternatively, a user can provide a SummarizedExperiment object and the assay(object) will be used as input for the CytoK procedure.
group_factor	a group level binary categorical response associated with each sample or column in the object. The order of the group_factor must match the order of the columns in object.
lowerRho	(Optional) lower bound of the kernel parameter.
upperRho	(Optional) upper bound of the kernel parameter.
gridRho	(Optional) number of grid points in the interval [lowerRho, upperRho].

Value

A list of CytoK statistics including

shrunkEffectSizeMean	the shrunk effective size posterior mean per feature
shrunkEffectSizeSD	the shrunk effective size posterior sd per feature
vec_pValue	the unadjusted p value per feature
AdjPvalue_features	the adjusted p value per feature using the Benjamini-Hochberg procedure.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoKProc <- CytoKProc(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4)
```

data.frameORSummarizedExperiment-class *S4 Class union*

Description

Class union allowing CytoKData slot to be a data.frame or Summarized Experiment

plotCytoK *Heatmap of the differentially expressed data with features on the rows and samples (group factors) as the columns from* CytoK *and* CytoKDEData *function.*

Description

This function plots a heatmap of the expression matrix with features (e.g., cluster-marker combinations) on the rows and samples (group factors) as the columns.

Usage

```
plotCytoK(object, group_factor, topK, featureVars = NULL)
```

Arguments

object	a CytoK object from CytoK
group_factor	a group level binary categorical response associated with each sample or column in the object. The order of the group_factor must match the order of the columns in object.
topK	top K differentially expressed features.
featureVars	(Optional) Vector of the columns which identify features. If a 'SummarizedExperiment' is used for 'data', row variables will be used.

Value

A heatmap will be created showing the samples on the columns and features on the rows.

Examples

```
data <- cbind(matrix(rnorm(1200,mean=2, sd=1.5),
nrow=200, ncol=6), matrix(rnorm(1200,mean=5, sd=1.9),
nrow=200, ncol=6))
data_CytoK <- CytoK(object=data,
group_factor = rep(c(0,1), each=6), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
data("cytoHDBMW")
data_CytoK_HD <- CytoK(object=cytoHDBMW,
group_factor = rep(c(0, 1), c(4, 4)), lowerRho=2,
upperRho=12,gridRho=4,alpha = 0.05,
featureVars = NULL)
```

```
plotCytoK(data_CytoK_HD,
group_factor = rep(c(0, 1), c(4, 4)),topK=10,
featureVars = NULL)
```

`vectorORNull-class` *S4 Class union*

Description

Class union allowing `CytoKFeatureVars` slot to be a vector or NULL

Index