Package: Polytect (via r-universe)

January 11, 2025

Title An R package for digital data clustering

Version 0.99.5

Description Polytect is an advanced computational tool designed for the analysis of multi-color digital PCR data. It provides automatic clustering and labeling of partitions into distinct groups based on clusters first identified by the flowPeaks algorithm. Polytect is particularly useful for researchers in molecular biology and bioinformatics, enabling them to gain deeper insights into their experimental results through precise partition classification and data visualization.

biocViews ddPCR, Clustering, MultiChannel, Classification

License Artistic-2.0

URL https://github.com/emmachenlingo/Polytect

BugReports https://github.com/emmachenlingo/Polytect/issues

Encoding UTF-8

LazyData false

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Depends R (>= 4.4.0)

Imports stats, utils, grDevices, mvtnorm, sn, dplyr, flowPeaks, ggplot2, tidyverse, cowplot, mlrMBO, DiceKriging, smoof, ParamHelpers, lhs, rgenoud, BiocManager

Suggests testthat (>= 3.0.0), knitr, rmarkdown, ddPCRclust

VignetteBuilder knitr

Config/testthat/edition 3

Config/pak/sysreqs libfontconfig1-dev libfreetype6-dev libfribidi-dev libgdal-dev gdal-bin libgeos-dev libglu1-mesa-dev libgmp3-dev make libgsl0-dev libharfbuzz-dev jags libicu-dev libjpeg-dev libpng-dev libtiff-dev libxml2-dev libmpfr-dev libopenmpi-dev libssl-dev libproj-dev libx11-dev zlib1g-dev

Repository https://bioc.r-universe.dev

RemoteUrl https://github.com/bioc/Polytect

RemoteRef HEAD

RemoteSha ff9290d10fa6c4ec0b74fc98ec3d19bb08e99e6a

Contents

BPV *BPV data*

Description

A 3-color dPCR data of bovine papilloma virus assay

Usage

```
data(BPV)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

channel3 fluorescence intensities of color 3

Examples

```
data(BPV)
head(BPV)
```

CA *CA data*

Description

2-color competitive assay of competition BRAF V600E assay with 1% mutant

Usage

```
data(CA)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel. data is not orthogonal.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

Examples

```
data(CA)
head(CA)
```

CNV5plex *CNV 5-plex data*

Description

CNV 5-plex universal probes

Usage

```
data(CNV5plex)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

channel3 fluorescence intensities of color 3

channel4 fluorescence intensities of color 4

channel5 fluorescence intensities of color 5

Examples

```
data(CNV5plex)
head(CNV5plex)
```

`CNV6plex` *CNV 6-plex data*

Description

CNV 6-plex universal probes

Usage

```
data(CNV6plex)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

channel3 fluorescence intensities of color 3

channel4 fluorescence intensities of color 4

channel5 fluorescence intensities of color 5

channel6 fluorescence intensities of color 6

Examples

```
data(CNV6plex)
head(CNV6plex)
```

`conc_cal` *concentration calculation function*

Description

This function takes a data frame of fluorescence intensities and partition clusters as input. It can be results from polytect_clust or polytect_merge. It will give the target concentration as output.

Usage

```
conc_cal(df_data, cluster_num, sampvol = 0.91, volmix = 20, voltemp = 20)
```

Arguments

df_data	A data frame containing partition fluorescence intensities and corresponding cluster label. This can be the output of polytect_merge or any data frame containing the above information.
cluster_num	the expected number of clusters
sampvol	The sample volume in microliters (µL)
volmix	The volume of the mixture
voltemp	The volume of the template

Value

a data frame of target concentration.

Examples

```
data(HR)
df_data<-polytect_clust(HR,4)
conc_cal(df_data,4)
```

HIV *HIV data*

Description

A 4-color dPCR data of intact HIV-1 proviruses

Usage

```
data(HIV)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

channel3 fluorescence intensities of color 3

channel4 fluorescence intensities of color 4

Source

https://www.biorxiv.org/content/10.1101/2023.08.18.553846v1

Examples

```
data(HIV)
head(HIV)
```

HR *HR data*

Description

A high-resolution 2-color dPCR data of RPP30 genomic DNA assay

Usage

```
data(HR)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel. good separation but some crosstalk.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

Source

https://pubmed.ncbi.nlm.nih.gov/33992770/

Examples

```
data(HR)
head(HR)
```

LR *LR data*

Description

A low-resolution 2-color dPCR data of development of genotyping assays for plants various

Usage

```
data(LR)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel. barely separable on x-axis.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

Examples

```
data(LR)
head(LR)
```

`MM` *MM data*

Description

A multi-mode 2-color dPCR data of HIV gBlock sequences

Usage

```
data(MM)
```

Format

A data frame of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel. obvious multimodality.

channel1 fluorescence intensities of color 1

channel2 fluorescence intensities of color 2

Source

https://pubmed.ncbi.nlm.nih.gov/37827643/

Examples

```
data(MM)
head(MM)
```

`polytect_clust` *Main function for clustering*

Description

This is the main function for clustering. The function will start with flowPeaks, then merge the excess clusters. It will return a data frame of fluorescence intensities and partition labels.

Usage

```
polytect_clust(
  data,
  cluster_num,
  fp_par = "default",
  fp_optim = c(0.1, 1, 1.5),
  lambdas = rep(2, 64 - log2(64)),
  coefs = rep(1, 6)
)
```

Arguments

data	A matrix of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.
cluster_num	The expected maximum number of clusters.
fp_par	The parameters for flowPeaks. fp_par=c("default","manual","auto"). When "default" is chosen, the default parameters of flowPeaks will be used. With "manual", you have to fill in fp_optim.
fp_optim	The paramters for flowPeaks that users have to fill in manually when fp_par is set at "manual".
lambdas	The penalty terms for the deviation from the expected cluster centers. Higher lambdas penalizes the deviation more.
coefs	The coefficients to adjust for the expected cluster centers. The default is 1 which can be used for common assay designs and has to be modified for special assays such as competing assays.

Value

A data frame containing the original fluorescence intensity and the cluster labels.

Examples

```
data(HR)
head(polytect_clust(HR, 4))
```

---

polytect_merge *Function for merging*

Description

This function takes the clustering result as input. Users can first perform any clustering algorithm, then use this function. It will return a data frame of fluorescence intensities and partition labels.

Usage

```
polytect_merge(
  data,
  cluster_num,
  base_clust,
  lambdas = rep(2, 64 - log2(64)),
  coefs = rep(1, 6)
)
```

Arguments

data	A matrix of fluorescence intensities in each channel. Each row represents each partitions, and each column each channel.
cluster_num	The expected maximum number of clusters.
base_clust	A list that contains partition labels given by initial clustering.
lambdas	The penalty terms for the deviation from the expected cluster centers. Higher lambdas penalizes the deviation more.
coefs	The coefficients to adjust for the expected cluster centers. The default is 1 which can be used for common assay designs and has to be modified for special assays such as competing assays.

Value

A data frame containing the original fluorescence intensity and the cluster labels.

Examples

```
data(HR)
dist_matrix <- dist(HR)
hc <- hclust(dist_matrix, method = "ward.D2")
hc_clusters <- cutree(hc, k = 6)
base_clust<-list()
base_clust$cluster<-hc_clusters
head(polytect_merge(HR, 4, base_clust))
```

---

polytect_plot *Plotting function for clustering results*

Description

This function takes results from polytect_clust and polytect_merge, or a data frame containing flurescence intensities and partition labels. It will output all combination of 2-color plots.

Usage

```
polytect_plot(df_data, cluster_num, cluster_selected = TRUE)
```

Arguments

df_data	A data frame containing partition fluorescence intensities and corresponding cluster label. This can be the output of polytect_clust and polytect_merge or any data frame containing the above information.
cluster_num	the expected number of clusters
cluster_selected	Indicator of whether all the clusters are present in the plots. If TRUE, then only selected ones (the ones only positive in the selected 2 dimensions) are shown. The default value is "TRUE".

Value

2-color plots.

Examples

```
data(HR)
df_data<-polytect_clust(HR,4)
polytect_plot(df_data,4)
```

---

polytect_summary *Summary function*

Description

This function takes results from polytect_clust and polytect_merge, or a data frame containing flurescence intensities and partition labels. It will summarise cluster centers, cluster sizes and cluster silhouette coefficients.

Usage

```
polytect_summary(df_data)
```

Arguments

df_data	A data frame containing partition fluorescence intensities and corresponding cluster label. This can be the output of polytect_clust and polytect_merge or any data frame containing the above information.

Value

a data frame of the summary of cluster centers, cluster sizes and cluster silhouette coefficients.

Examples

```
data(HR)
df_data<-polytect_clust(HR,4)
polytect_summary(df_data)
```

`sil_plot` *Plotting function for silhouette coefficients*

Description

This function takes results from polytect_clust and polytect_merge, or a data frame containing flurescence intensities and partition labels. It will output the silhouette coefficients of each cluster.

Usage

```
sil_plot(df_data)
```

Arguments

`df_data`	A data frame containing partition fluorescence intensities and corresponding cluster label. This can be the output of `polytect_clust` and `polytect_merge` or any data frame containing the above information.

Value

plot of silhouette coefficients for each cluster.

Examples

```
data(HR)
df_data<-polytect_clust(HR,4)
sil_plot(df_data)
```

Index