

Package: scMAGeCK (via r-universe)

July 10, 2026

Type Package

Title Identify genes associated with multiple expression phenotypes in single-cell CRISPR screening data

Version 0.99.1

Date 2026-07-09

Description scMAGeCK is a computational model to identify genes associated with multiple expression phenotypes from CRISPR screening coupled with single-cell RNA sequencing data (e.g. CROP-seq, Perturb-seq). It provides two complementary modules: a rank-based test (scMAGeCK-RRA) that links a perturbation to a single marker or gene signature, and a linear-regression test (scMAGeCK-LR) that estimates the effect of each perturbation on the expression of all genes. The package also estimates single-cell perturbation-response scores to capture heterogeneous perturbation effects across individual cells.

License BSD_2_clause + file LICENSE

URL <https://github.com/weili-lab/scMAGeCK>

BugReports <https://github.com/weili-lab/scMAGeCK/issues>

biocViews CRISPR, SingleCell, RNASeq, Sequencing, PooledScreens, Transcriptomics, GeneExpression, Regression

NeedsCompilation yes

Imports Seurat (>= 5.0.0), ggplot2, stats, utils, methods, graphics, grDevices, Rcpp

LinkingTo Rcpp

Suggests knitr, rmarkdown

VignetteBuilder knitr

Config/pak/sysreqs cmake libglpk-dev make libicu-dev libpng-dev libuv1-dev libxml2-dev libssl-dev python3 zlib1g-dev

Repository <https://biocstaging.r-universe.dev>

Date/Publication 2026-07-10 03:47:54 UTC

RemoteUrl <https://github.com/BiocStaging/scMAGeCK>

RemoteRef HEAD

RemoteSha e24aa0f3c9734c26ffbc200107adf0e199b26208

Contents

scMAGeCK-package	2
assign_cell_identity	3
featurePlot	4
guidematrix_to_triplet	6
pre_processRDS	6
read_gmt_file	7
scmageck_best_lambda	8
scmageck_eff_estimate	9
scmageck_lr	11
scmageck_rra	13
selectPlot	14

Index **16**

scMAGeCK-package	<i>Identify genes associated with multiple expression phenotypes in single-cell CRISPR screening data</i>
------------------	---

Description

scMAGeCK is a computational model to identify genes associated with multiple expression phenotypes from CRISPR screening coupled with single-cell RNA sequencing data (e.g. CROP-seq, Perturb-seq). It provides two complementary modules: a rank-based test (scMAGeCK-RRA) that links a perturbation to a single marker or gene signature, and a linear-regression test (scMAGeCK-LR) that estimates the effect of each perturbation on the expression of all genes. The package also estimates single-cell perturbation-response scores to capture heterogeneous perturbation effects across individual cells.

Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.

scMAGeCK is a computational model to identify genes associated with multiple expression phenotypes from CRISPR screening coupled with single-cell RNA sequencing data (CROP-seq).scMAGeCK is based on our previous MAGeCK and MAGeCK-VISPR models for pooled CRISPR screens.

The scMAGeCK manuscript can be found at bioRxiv(<https://www.biorxiv.org/content/10.1101/658146v1/>).

Author(s)

Wei Li [aut, cre], Xiaolong Cheng [aut], Lin Yang [aut]

Maintainer: Wei Li <li.david.wei@gmail.com>

Examples

```
### BARCODE file contains cell identity information, generated from
### the cell identity collection step
BARCODE <- system.file("extdata","barcode_rec.txt",package = "scMAGeCK")
### RDS can be a Seurat object or local RDS file path that contains
### the scRNA-seq dataset
RDS <- system.file("extdata","singles_dox_mki67_v3.RDS",package = "scMAGeCK")

target_gene <- "MKI67"

### Set RRA executable file path or activate scmageck env if needed (see https://bitbucket.org/weililab/scmageck)
RRAPATH <- NULL

rra_result <- scmageck_rra(BARCODE=BARCODE, RDS=RDS, GENE=target_gene,
                          RRAPATH=RRAPATH, LABEL='dox_mki67',
                          NEGCTRL=NULL, KEPTMP=FALSE,
                          PATHWAY=FALSE, SAVEPATH=NULL)

head(rra_result)

lr_result <- scmageck_lr(BARCODE=BARCODE, RDS=RDS, LABEL='dox_scmageck_lr',
                        NEGCTRL = 'NonTargetingControlGuideForHuman', PERMUTATION = 1000,
                        SAVEPATH=NULL, LAMBDA=0.01)
lr_score <- lr_result[1]
lr_score_pval <- lr_result[2]
head(lr_score_pval)
```

assign_cell_identity *Assign single cell identity to genes to be perturbed.*

Description

assign the identity of single cells, by adding a gene column to the metadata.

Usage

```
assign_cell_identity(BARCODE, RDS, ASSIGNMETHOD='unique')
```

Arguments

BARCODE	A txt file to include cell identity information, generated from the cell identity collection step; or a corresponding data.frame.
RDS	RDS object from the pre-processRDS step; or a path to RDS file.
ASSIGNMETHOD	Determine the way to assign cell identity. "unique" only assigns single cells that only have 1 guide detected (others are labeled as NA or doublet). "largest" will assign single cells with the largest umi count.

Value

The input Seurat object with per-cell identity added to its metadata: gene, sgrna and umi_count columns (cells with multiple guides are labeled "doublet" under the "unique" method).

Examples

```
### Loading required package
require(Seurat)

### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")

### For using the featurePlot function, it needs to do the preprocessRDS first
RDS <- assign_cell_identity(BARCODE, RDS)
```

featurePlot	<i>Detect the sgRNA distribution and generate Vlnplot to identify gene regulation between different cells.</i>
-------------	--

Description

identify how many sgRNAs entered the cells, visualize the counts of sgRNAs distribution and gene regulation.

Usage

```
featurePlot(RDS, TYPE = plot.type, BARCODE = NULL, sgRNA = NULL, GENE = NULL, CONTROL = NULL, GROUP2=NULL)

plot.type
# c("Dis", "Vln", "Den")
```

Arguments

RDS	RDS object from the pre-processRDS step
TYPE	Type of the plot.
BARCODE	A txt file to include cell identity information, generated from the cell identity collection step.
sgRNA	Generate whole sgRNAs distribution when sgRNA = NULL, add sgRNAs to see the specific sgRNA distribution. Mutiple sgRNAs can be provided, separated by ",". For example, "APC,TP53".

GENE	Genes whose expressions are to be compared under different cell conditions. Multiple genes can be provided, as a vector or as a string separated by ",". For example, "APC,TP53". when provide mutiple genes, it would show the average gene expression.
CONTROL	Set up the sepecific clusters to compare the gene expression. it would compare gene expression across the dataset when CONTROL = NULL.
GROUP2	Set up the gene name as the control group. it would compare gene expression across the dataset when GROUP2= NULL.
SLOT	Use the slot in Seurat object for plot. May choose 'data', 'scale.data' or 'count'. See GetAssayData funciton in Seurat.
palette	The color palette to change the color of VlnPlot.
label.size	Text size of label.
axis.size	Text size of axis.text.
title.size	Text size of axis/pics' title.
legend.text	Text size of figure legend.
fill	Fill colour.

Value

A ggplot object with the requested plot (sgRNA distribution, violin/ECDF of gene expression by perturbation, or a density plot). Returns NULL (with a message) if TYPE is not recognized.

Examples

```
### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")

library(Seurat)

### For using the featurePlot function, it needs to do the preprocessRDS first
Demo <- pre_processRDS(BARCODE = BARCODE, RDS = RDS)

### For the sgRNA distribution
featurePlot(BARCODE = BARCODE, RDS = Demo, TYPE = "Dis")

### For the density of sgRNA, clustering needed to be done first.
Demo <- RunUMAP(Demo, dims = 1:10)
featurePlot(RDS = Demo, sgRNA = NULL, TYPE = "Den")

### For the Vlnplot to display the gene regulation, take MKI67 for examples
featurePlot(RDS = Demo, GENE = "MKI67", sgRNA = "TP53", TYPE = "Vln")
```

guidematrix_to_triplet

Convert guide matrix into a dataframe containing cell, barcode, read and UMI count.

Description

Convert guide matrix into a dataframe, where columns include cell, barcode, read count and UMI count.

Usage

```
guidematrix_to_triplet(count_mat, RDS)
```

Arguments

`count_mat` A guide count matrix where columns are cell names and rows are guide names

`RDS` A Seurat object or local RDS file path that contains the scRNA-seq dataset. Note that the dataset has to be normalized and scaled.

Value

A data frame that contains cell names, guide names, read count and UMI count as column.

Examples

```
### a guide-by-cell count matrix; column names must be cell names in the Seurat object
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")
obj <- readRDS(RDS)
cells <- head(colnames(obj), 4)
count_mat <- matrix(c(3, 0, 2, 0,
                     0, 5, 0, 1),
                   nrow = 2, byrow = TRUE,
                   dimnames = list(c("sgRNA_A", "sgRNA_B"), cells))
triplets <- guidematrix_to_triplet(count_mat, obj)
head(triplets)
```

pre_processRDS

Integrate the information of sgRNA into RDS file for the further analysis.

Description

Pre-process the sgRNA count from previous step, and generate the sgRNA expression matrix.

Usage

```
pre_processRDS(BARCODE, RDS, normalize = TRUE, scale = TRUE)
```

Arguments

BARCODE	A txt file to include cell identity information, generated from the cell identity collection step.
RDS	A Seurat object or local RDS file path that contains the scRNA-seq dataset. Note that the dataset has to be normalized and scaled.
normalize	Whether to perform normalization on sgRNA count matrix
scale	Whether to scale the normalized sgRNA count matrix

Value

Returns an updated Seurat object, with the following modifications:

- An added "sgrna_guide" assay that contains the normalized and scaled guide counts;
- An added "sgrna" assay that contains the normalized and scaled target gene counts. Counts from different sgRNAs targeting the same gene are merged;
- Several columns added to metadata to describe guide assignments (which can be overwritten by calling `assign_cell_identity` function later).

Examples

```
### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")

Demo <- pre_processRDS(BARCODE = BARCODE, RDS = RDS)
```

read_gmt_file	<i>Read a gene set file in GMT format.</i>
---------------	--

Description

Read a gene set file in the Broad Institute GMT format, where each line describes one gene set (name, description, followed by its member genes, tab-separated), and return it as a data frame.

Usage

```
read_gmt_file(gmt_path)
```

Arguments

gmt_path	Path to a gene set file in GMT format.
----------	--

Value

A data frame with one row per gene set; the first column is the gene set name and the remaining columns are its member genes (padded with empty strings for sets with fewer genes).

Examples

```
gmt <- system.file("extdata", "test_symbols.gmt.txt", package = "scMAGeCK")
gene_sets <- read_gmt_file(gmt)
```

scmageck_best_lambda *Find the best lambda using negative controls*

Description

This function identifies the best value for lambda, a tuning parameter used in a modeling approach that uses negative controls. Negative controls are samples or features that should not have any meaningful association with the outcome of interest. This approach can be useful in situations where there may be confounding factors or batch effects that can affect the results of a model.

Usage

```
scmageck_best_lambda(
  rds_object,
  bc_frame,
  non_target_ctrl = "NT",
  lambda_seq = 10^seq(-3, 3, length = 10),
  pseudogene_label = "PSEUDO_GENE",
  pseudogene_num = 250
)
```

Arguments

rds_object	Seurat object or local RDS file path that contains the scRNA-seq dataset. Alternatively, you can provide the path to an RDS file. The dataset should only contain cells expressing non-targeting gRNA, which will be used as negative controls for lambda tuning.
bc_frame	A txt file or corresponding data.frame that includes cell identity information generated from the cell identity collection step. The bc_frame should have the same number of rows as the rds_object input, and each row should contain a unique cell barcode or identifier along with its corresponding cell identity.
non_target_ctrl	A character vector or string containing a comma-separated list of genes that serve as negative controls.
lambda_seq	A sequence of lambda values to test. Defaults to 10 ^{seq(-3, 3, length = 10)} .
pseudogene_label	A character string to be used as a pseudogene to evaluate the false positive rate. The default value for pseudogene_label is PSEUDO_GENE.

`pseudogene_num` An integer value indicating the number of cells that should be selected as pseudogenes for false positive rate evaluation. The default value for `pseudogene_num` is 250.

Value

The output of the function is a data frame containing the lambda value and the corresponding false positive rate on the negative control samples. Additionally, the function generates a plot showing the relationship between lambda and the false positive rate.

See Also

[scmageck_eff_estimate](#) [scMAGeCK](#) [BitBucket](#)

Examples

```
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")
lambda_seq <- 10^seq(-2, 2, length = 5)

# the cross-validation over lambda is time consuming

obj <- assign_cell_identity(BARCODE, readRDS(RDS))
bc_frame <- read.table(BARCODE, header = TRUE, as.is = TRUE)
lambda_df <- scmageck_best_lambda(obj, bc_frame,
                                non_target_ctrl = "NonTargetingControlGuideForHuman",
                                lambda_seq = lambda_seq, pseudogene_num = 100)
```

`scmageck_eff_estimate` *Detect heterogenous perturbation responses from Perturb-seq using perturbation-response score (PS)*

Description

"This function uses constrained linear least squares to calculate perturbation-response score (PS), which measures heterogenous perturbation effects from single-cell CRISPR screens (e.g., Perturb-seq, CROP-seq)."

Usage

```
scmageck_eff_estimate(
  rds_object,
  bc_frame,
  perturb_gene,
  non_target_ctrl,
  perturb_target_gene = NULL,
  scale_factor = 3,
  target_gene_min = 10,
```

```

target_gene_max = 500,
assay_for_cor = 'RNA',
subset_rds = TRUE,
scale_score = TRUE,
perturb_gene_exp_id_list = NULL,
lambda = 0,
background_correction=FALSE,
use_perturb_exp=TRUE,
logfc.threshold=0.1
)

```

Arguments

rds_object	A Seurat object or local RDS file path that contains the scRNA-seq dataset; or a path to RDS file. Note that the dataset has to be normalized and scaled.
bc_frame	A txt file to include cell identity information, generated from the cell identity collection step; or a corresponding data.frame.
perturb_gene	The list of perturbed genes. By default, all genes in the table are subject to regression.
non_target_ctrl	The list of genes (separated by ",") served as negative controls.
perturb_target_gene	The list of target genes for modeling. If null, will automatic search and identify the target genes.
scale_factor	The upper bound of the constraints. Must be a positive value. Default 3. Assign a higher value for a more continuous distribution of the signa scores.
target_gene_min	The minimum number of genes selected for target genes. If DEG analysis does not provide enough number of genes that reach this number, the algorithm will iteratively (at most 3 times) decrease LFC threshold to select more genes.
target_gene_max	The maximum number of genes selected for target genes.
assay_for_cor	The assays used for estimating correlation. Default: RNA.
subset_rds	Whether to return an R object that only contains cells that express guides targeting perturbed genes (or negative control genes). If TRUE, a gene column in metadata will be added (or updated) that assigns perturbed genes to single cells. Default: TRUE
scale_score	Whether to scale the scores for each gene to 1. Default: TRUE
perturb_gene_exp_id_list	If the perturbed_gene id is different from expression feature id, use this parameter to provide the corresponding expression features. Must be the same length as perturb_gene. Default: NULL
lambda	Sparse penalty (similar with the lambda value in LASSO regression). Must be non-negative. Default: 0

background_correction
Whether to extract background gene expression, which is estimated from negative control cells. Turn this option on will reduce false positives in datasets containing multiple cell types, where gene expressions may be largely different from different cell types. Default: False

use_perturb_exp
Whether to use the expression of perturbed gene in the estimation of efficiency score. If False, the perturbed gene (or the corresponding id in `perturb_gene_exp_id_list`) will be removed from estimating the score. Default: True

logfc.threshold
logfc threshold when determining the downstream targets of a perturbed gene. Default: 0.1

Value

Returns a list of several items: `eff_matrix`: the PS score matrix containing the PS scores of each cells for each perturbed gene `rds`: the R object if `subset_rds` is set as TRUE `optimization_matrix`: the matrix used for actually performing the constrained optimization `target_gene_search_result`: the results of target gene search for each perturbed gene

Examples

```
# set the BARCODE and RDS file path
# if you have a guide matrix, use guidematrix_to_triplet() to convert it to a BARCODE file
BARCODE = system.file("extdata", "barcode_rec.txt", package = "scMAGECK")
bc_frame = read.table(BARCODE, header = TRUE, as.is = TRUE)
# cell identity may need to be fixed to match the expression matrix
bc_frame$cell = sub('-1', '', bc_frame$cell)
RDS = system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGECK")

# the full estimation (DE search + quadratic optimization) is time consuming

rds_object = readRDS(RDS)
eff_object <- scmageck_eff_estimate(rds_object, bc_frame, perturb_gene = 'TP53',
                                   non_target_ctrl = 'NonTargetingControlGuideForHuman',
                                   assay_for_cor = 'RNA')
eff_estimat = eff_object$eff_matrix
rds_subset = eff_object$rds
```

scmageck_lr

Use linear regression to test the association of gene knockout with all possible genes

Description

echo "Use linear regression to test the association of gene knockout with all possible genes"

Usage

```
scmageck_lr(BARCODE, RDS, NEGCTRL, SELECT_GENE=NULL, LABEL = NULL,
PERMUTATION = NULL, SIGNATURE = NULL, SAVEPATH = "./", LAMBDA=0.01, GENE_FRAC=0.01, SLOT='scale.data')
```

Arguments

BARCODE	A txt file to include cell identity information, generated from the cell identity collection step; or a corresponding data.frame.
RDS	A Seurat object or local RDS file path that contains the scRNA-seq dataset; or a path to RDS file. Note that the dataset has to be normalized and scaled.
NEGCTRL	The name of the genes (separated by ",") served as negative controls.
SELECT_GENE	The list of genes for regression. By default, all genes in the table are subject to regression.
LABEL	The label of the output file.
PERMUTATION	The number of permutations for p value calculation.
SAVEPATH	The save path of result. Default save path is the current working directory. If you don't need save the result, set SAVEPATH as NULL.
LAMBDA	A paramter for the LR model for ridge regression. Default: 0.01.
GENE_FRAC	A paramter for filtering low expressed genes. By default, only genes that have expressions in at least that fractions of cells (in raw count table) are kept. If raw count table is not found, will look into scaled expression, and only keep genes whose expression is greater than zero in at least that fraction of cells. Default: 0.01.
SIGNATURE	A GMT text file, the format must be as follows:(1)Column 1: name of gene set; (2)Colum 2: Empty, or the information about gene set e.g. the source of the gene set; (3)Column 3 and onwards: ids of genes belonging to a particular gene set. Note that this parameter for applying LR model for the gene set analysis. Reference: http://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats
SLOT	Use the slot in Seurat object for plot. May choose 'data', 'scale.data' or 'count'. See GetAssayData funciton in Seurat.

Value

If SIGNATURE is set as NULL (default): returns a list of lr results, including: beta score (as a data.frame), the associated p value (as a data.frame), and the regression matrix that is used for linear regression.

If SIGNATURE is set: returns a list of signature results, including: signature result (as a data.frame), and the regression matrix that is used for linear regression.

Examples

```
### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")
```

```
lr_result <- scmageck_lr(BARCODE=BARCODE, RDS=RDS, LABEL='dox_scmageck_lr', SIGNATURE = NULL,
  NEGCTRL = 'NonTargetingControlGuideForHuman', PERMUTATION = 1000, SAVEPATH=NULL, LAMBDA=0.01)
lr_score <- lr_result[1]
lr_score_pval <- lr_result[2]
head(lr_score_pval)
```

scmageck_rra	<i>Use RRA to test the association of gene knockout with certain marker expression</i>
--------------	--

Description

echo "Use RRA to test the association of gene knockout with certain marker expression"

Usage

```
scmageck_rra(BARCODE, RDS, GENE, RRAPATH = NULL, LABEL = NULL, NEGCTRL = NULL,
SIGNATURE = NULL, KEEPTMP = FALSE, PATHWAY = FALSE, SAVEPATH = "./", ASSIGNMETHOD = "largest", SLOT='scale')
```

Arguments

BARCODE	A txt file to include cell identity information, generated from the cell identity collection step; or a corresponding data.frame.
RDS	A Seurat object or local RDS file path that contains the scRNA-seq dataset; or a path to RDS file. Note that the dataset has to be normalized and scaled.
GENE	Genes whose expressions are to be tested. Multiple genes can be provided, separated by ",". For example, "MKI67,TP53"
RRAPATH	The path to the RRA program, if RRA cannot be found in the PATH environment variable. Depreciated in 1.5.1.
LABEL	The label of the output file.
NEGCTRL	The name of the negative control gene. For example, "NonTargetingControl-GuideForHuman". Default is NULL (do not use any negative controls).
KEEPTMP	Keep temporary files.
PATHWAY	Treat genes in –GENE option as a pathway. In other words, the averaged expression of these genes will be used for testing.
SAVEPATH	The save path of result. Default save path is the current working directory. If you don't need save the result, set SAVEPATH as NULL.
ASSIGNMETHOD	Method used to assign the sgRNA identity to each cell. Can be either unique or largest. See assign_cell_identity for more details.
SIGNATURE	A GMT text file, the format must be as follows:(1)Column 1: name of gene set; (2)Column 2: Empty, or the information about gene set e.g. the source of the gene set; (3)Column 3 and onwards: ids of genes belonging to a particular gene set. Note that if you don't set the parameter "SAVEPATH", this parameter would create a folder called "GENE_SET" in the current working directory to store the results from applying RRA program to do gene set analysis. Reference: http://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats

SLOT Use the slot in Seurat object for plot. May choose 'data', 'scale.data' or 'count'. See GetAssayData function in Seurat.

Value

A data frame of RRA results.

Examples

```
### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")

target_gene <- "MKI67"
### Set RRA executable file path or activate scmageck env if needed (see https://bitbucket.org/weililab/scmageck)
RRAPATH <- NULL

rra_result <- scmageck_rra(BARCODE=BARCODE, RDS=RDS, GENE=target_gene,
                           RRAPATH=RRAPATH, LABEL='dox_mki67',
                           NEGCTRL=NULL, KEEPTMP=FALSE, SIGNATURE = NULL,
                           PATHWAY=FALSE, SAVEPATH=NULL)

head(rra_result)
```

selectPlot

generate the selection plot

Description

detect the gene regulation relationship between genes and perturbation by using RRA or LR test.

Usage

```
selectPlot(GENE = NULL, lr_result = NULL, CUTOFF = 0.05, ADJ = "fdr",
           RRA_re1 = NULL, RRA_re2 = NULL, TYPE = select.type, QUALITY = 10)
```

```
select.type
#c("lr", "rra")
```

Arguments

GENE Genes whose expressions are to be tested under the LR test. Multiple genes can be provided, separated by ",". For example, "MKI67,TP53"

lr_result The result from the scmageck-lr step.

CUTOFF Determine the significant pvalue.

ADJ	P.adjust.methods. Choose one of correction method. c("holm","hochberg", "holmel", "bonferroni", "BH", "BY", "fdr", "none")
RRA_re1	RRA result from the scmageck-rra step.
RRA_re2	Optional input. The second RRA result from the scmageck-rra step. Add this input to visualize gene selection under two different cell condition.
TYPE	The type of the scMAGeCK results. Can be either "rra" or "lr".
QUALITY	The number of single-cells that passes the threshold when use the RRA test, default is 10. Lower quality could improve the sensitivity but reduce accuracy.

Value

A ggplot object showing the selected genes from the scMAGeCK-RRA or scMAGeCK-LR results. Returns NULL (with a message) if no plot can be generated for the requested inputs.

Examples

```
### by using RRA test, take MKI67 for example
### only works if you have RRA installed
#selectPlot(RRA_re1 = rra_result, CUTOFF = 0.05, QUALITY = 10, ADJ = "fdr", TYPE = "rra")

### by using LR test, take MKI67 for example

### BARCODE file contains cell identity information, generated from the cell identity collection step
BARCODE <- system.file("extdata", "barcode_rec.txt", package = "scMAGeCK")

### RDS can be a Seurat object or local RDS file path that contains the scRNA-seq dataset
RDS <- system.file("extdata", "singles_dox_mki67_v3.RDS", package = "scMAGeCK")

lr_result <- scmageck_lr(BARCODE=BARCODE, RDS=RDS, LABEL='dox_scmageck_lr', SIGNATURE = NULL,
  NEGCTRL = 'NonTargetingControlGuideForHuman', PERMUTATION = 1000, SAVEPATH=NULL, LAMBDA=0.01)
selectPlot(GENE = "MKI67", lr_result = lr_result, CUTOFF = 0.05, ADJ = "fdr", TYPE = "lr")
```

Index

* package

scMAGeCK-package, 2

assign_cell_identity, 3

featurePlot, 4

guidematrix_to_triplet, 6

pre_processRDS, 6

read_gmt_file, 7

scmageck (scMAGeCK-package), 2

scMAGeCK-package, 2

scmageck_best_lambda, 8

scmageck_eff_estimate, 9, 9

scmageck_lr, 11

scmageck_rra, 13

selectPlot, 14