# Package 'scReClassify'

November 12, 2026

**Type** Package

**Title** scReClassify: post hoc cell type classification of single-cell RNA-seq data

**Version** 1.18.0

**Description** A post hoc cell type classification tool to fine-tune cell type annotations generated by any cell type classification procedure with semi-supervised learning algorithm AdaSampling technique. The current version of scReClassify supports Support Vector Machine and Random Forest as a base classifier.

**License** GPL-3 + file LICENSE

**BugReports** https://github.com/SydneyBioX/scReClassify/issues

**URL** https://github.com/SydneyBioX/scReClassify, http://www.bioconductor.org/packages/release/bioc/html/scReClassify.html

**Depends** R (>= 4.1)

**Encoding** UTF-8

**LazyData** false

**RoxygenNote** 7.2.1

**Roxygen** list(markdown = TRUE)

**Imports** randomForest, e1071, stats, SummarizedExperiment, SingleCellExperiment, methods

**VignetteBuilder** knitr

**biocViews** Software, Transcriptomics, SingleCell, Classification, SupportVectorMachine

**Suggests** testthat, knitr, BiocStyle, rmarkdown, DT, mclust, dplyr

**git_url** https://git.bioconductor.org/packages/scReClassify

**git_branch** RELEASE_3_23

**git_last_commit** a6af3cf

**git_last_commit_date** 2026-04-28

**Repository** Bioconductor 3.23

**Date/Publication** 2026-11-11

**Author** Pengyi Yang [aut] (ORCID: <https://orcid.org/0000-0003-1098-3138>),
Taiyun Kim [aut, cre] (ORCID: <https://orcid.org/0000-0002-5028-836X>)

**Maintainer** Taiyun Kim <taiyun.kim91@gmail.com>

# Contents

bAccuracy . . . . . 2
gse87795_subset_sce . . . . . 3
matPCs . . . . . 3
multiAdaSampling . . . . . 4
scReClassify . . . . . 6

**Index** 7

---

`bAccuracy` *bAccuracy*

---

## Description

This function calculates the accuracy of the prediction to the true label.

## Usage

```
bAccuracy(cls.truth, final)
```

## Arguments

| | |
|---|---|
| `cls.truth` | A character vector of true class label. |
| `final` | A vector of final classified label prediction from `multiAdaSampling`. |

## Value

An accuracy value.

## Author(s)

Pengyi Yang, Taiyun Kim

## Examples

```
data("gse87795_subset_sce")

mat.expr <- gse87795_subset_sce
cellTypes <- gse87795_subset_sce$cellTypes

# Get dimension reduced matrix. We are using `logNorm` assay from `mat.expr`.
mat.pc <- matPCs(mat.expr, assay = "logNorm")
```

```
# Here we are using Support Vector Machine as a base classifier.
result <- multiAdaSampling(mat.pc, cellTypes, classifier = "svm",
percent = 1, L = 10)

final <- result$final

# Balanced accuracy
bacc <- bAccuracy(cellTypes, final)
```

---

`gse87795_subset_sce` *GSE827795 subset data*

---

## Description

A SingleCellExperiment object containing a subset expression matrix of GSE827795. The data contains log2 transformed FPKM expression.

GSE87795 is a mouse fetal liver development data containing 1000 genes, 367 cells and 6 cell types.

The original GSE87795 data and the study details can be found at this link

## Usage

```
gse87795_subset_sce
```

## Format

An object of class `SingleCellExperiment` with 1000 rows and 367 columns.

---

`matPCs` *matPCs function*

---

## Description

Performs PCA on a given matrix and returns a dimension reduced matrix which captures at least 80% (default) of overall variability.

## Usage

```
matPCs(data, assay = NULL, percentVar = 0.8)
```

### Arguments

| | |
|---|---|
| data | An expression matrix or a SingleCellExperiment object. |
| assay | An assay to select if data is a SingleCellExperiment object |
| percentVar | The percentage of variance threshold. This is used to select number of Principal Components. |

### Details

This function performs PCA to reduce the dimension of the gene expression matrix limited from 10 to 20 PCs.

### Value

Dimensionally reduced matrix.

### Author(s)

Pengyi Yang, Taiyun Kim

### Examples

```
data("gse87795_subset_sce")

mat.expr <- gse87795_subset_sce

mat.pc <- matPCs(mat.expr, assay = "logNorm")

# to capture at least 70% of overall variability in the dataset,
mat.dim.reduct.70 <- matPCs(mat.expr, assay = "logNorm", 0.7)
```

---

## multiAdaSampling *multi Adaptive Sampling function*

---

### Description

Performs multiple adaptive sampling to train a classifier model.

### Usage

```
multiAdaSampling(
  data,
  label,
  reducedDimName = NULL,
  classifier = "svm",
  percent = 1,
  L = 10,
```

```
  prob = FALSE,
  balance = TRUE,
  iter = 3
)
```

## Arguments

| | |
|---|---|
| data | A dimension reduced matrix from matPCs. |
| label | A named vector of label information for each sample. The names should match the sample names of data |
| reducedDimName | A name of the reducedDim to use. This must be specified if data is a SingleCellExperiment object. |
| classifier | Base classifier model, either "SVM" (svm) or "RF" 'rf' is supported. |
| percent | Percentage of samples to select at each iteration. |
| L | Number of ensembles. Default to 10. |
| prob | logical flag to return sample's probabilities to each class. |
| balance | logical flag to if the cell types are balanced. If FALSE, down sample large cell types classes to the median of all class sizes. |
| iter | A number of iterations to perform adaSampling. |

## Value

A final prediction, probabilities for each cell type and the model are returned as a list.

## Author(s)

Pengyi Yang, Taiyun Kim

## Examples

```
library(SingleCellExperiment)

# Loading the data
data("gse87795_subset_sce")

mat.expr <- gse87795_subset_sce
cellTypes <- gse87795_subset_sce$cellTypes

# Get dimension reduced matrix. We are using `logNorm` assay from `mat.expr`.
reducedDim(mat.expr, "matPCs") <- matPCs(mat.expr, assay = "logNorm")

# Here we are using Support Vector Machine as a base classifier.
result <- multiAdaSampling(mat.expr, cellTypes, reducedDimName = "matPCs",
classifier = "svm", percent = 1, L = 10)
```

`scReClassify` *scReClassify: a package for post hoc cell type classification of single-cell RNA-sequencing data.*

## Description

A post hoc cell type classification tool to fine-tune cell type annotations generated by any cell type classification procedure with semi-supervised learning algorithm AdaSampling technique.

The current version of scReClassify supports Support Vector Machine and Random Forest as a base classifier.

## Author(s)

**Maintainer**:

- Taiyun Kim (ORCID:0000-0002-5028-836X)
  - Email: taiyun.kim91@gmail.com

Authors:

- Pengyi Yang (ORCID: 0000-0003-1098-3138)

## See Also

Useful links:

- Vignette available at: https://sydneybiox.github.io/scdney/

# Index

∗ **datasets**
gse87795_subset_sce, 3

bAccuracy, 2

gse87795_subset_sce, 3

matPCs, 3
multiAdaSampling, 4

scReClassify, 6