

Package: COSG (via r-universe)

October 14, 2026

Title Marker Gene Identification for Single Cell Sequencing Data

Description COSG is designed for accurate and ultra-fast marker gene identification in single-cell sequencing data. COSG can not only be applied to single-cell RNA-seq data, but also single-cell ATAC-seq data and spatial transcriptomics.

Version 1.0.0

Date 2021-06-17

URL <https://github.com/genecell/COSGR>

BugReports <https://github.com/genecell/COSGR/issues>

License BSD 3-Clause | LICENSE

Encoding UTF-8

LazyData true

Imports data.table, Matrix, proxyC, SeuratObject (>= 3.0.0), methods, utils

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Suggests Seurat, testthat

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

Date/Publication 2026-08-25 16:20:50 UTC

RemoteUrl <https://github.com/genecell/COSGR>

RemoteRef HEAD

RemoteSha a1d58ca262a1138cab758ee16b343964b4ec4aae

Contents

cosg	2
pbmc_small	3
Index	5

cosg

*Accurate and fast cell marker gene identification with COSG***Description**

Marker gene identification for cell groups in a given dataset.

Usage

```
cosg(
  object,
  groups = "all",
  assay = "RNA",
  slot = "data",
  layer = NULL,
  mu = 1,
  remove_lowly_expressed = TRUE,
  expressed_pct = 0.1,
  n_genes_user = 100
)
```

Arguments

object	A Seurat object
groups	Character vector of cluster names to compare, or 'all' for all clusters
assay	Assay to use in marker gene identification
slot	Character. The slot name for expression data extraction (Seurat v3/v4). For Seurat v5, use the 'layer' parameter instead. Default: 'data'.
layer	Character. The layer name for expression data extraction (Seurat v5). If NULL, falls back to the 'slot' parameter for backward compatibility. Default: NULL.
mu	The penalty factor to penalize gene expression in cells not belonging to the cluster of interest
remove_lowly_expressed	Logical. If TRUE, genes that express a percentage of target cells smaller than a specific value (expressed_pct) are not considered as marker genes for the target cells. Default: TRUE.
expressed_pct	Numeric. When remove_lowly_expressed is TRUE, genes that express a percentage of target cells smaller than this value are not considered as marker genes. Default: 0.1.
n_genes_user	Number of top ranked genes returned in the result

Value

A list containing two dataframes for ranked marker genes' names and scores, respectively

Note

This function is compatible with Seurat v3, v4, and v5. For Seurat v5 users, we recommend using the 'layer' parameter. The 'slot' parameter is kept for backward compatibility with Seurat v3/v4.

Examples

```
library(COSG)
library(Seurat)
data('pbmc_small')

# Check cell groups:
table(Idsents(pbmc_small))

# Run COSG (Seurat v5 - use layer parameter):
marker_cosg <- cosg(
  pbmc_small,
  groups='all',
  assay='RNA',
  layer='data',
  mu=1,
  n_genes_user=100)

# Run COSG (Seurat v3/v4 - use slot parameter):
# marker_cosg <- cosg(
#   pbmc_small,
#   groups='all',
#   assay='RNA',
#   slot='data',
#   mu=1,
#   n_genes_user=100)

# Check the marker genes:
head(marker_cosg$names)
head(marker_cosg$scores)
```

pbmc_small*A small example Seurat object for testing*

Description

A subsetted version of a 10X Genomics PBMC dataset, containing 80 cells and 230 genes. This is the same dataset as SeuratObject::pbmc_small.

Usage

```
pbmc_small
```

Format

A Seurat object with:

assays RNA assay with counts and data layers

meta.data Cell metadata including cell type annotations

Source

SeuratObject package

Examples

```
data(pbmc_small)
pbmc_small
```

Index

* **datasets**

pbmc_small, [3](#)

cosg, [2](#)

pbmc_small, [3](#)