

# Package ‘spARI’

January 23, 2026

**Type** Package

**Title** Spatially Aware Adjusted Rand Index for Evaluating Spatial  
Transcriptomics Clustering

**Version** 1.0.0

**Description** The R package used in the manuscript ``Spatially Aware Adjusted Rand  
Index for Evaluating Spatial Transcriptomics Clustering".

**License** GPL (>= 2)

**Encoding** UTF-8

**Depends** R (>= 4.1.0)

**Imports** Rcpp, stats, Matrix, SpatialExperiment, SummarizedExperiment,  
BiocParallel (>= 1.0)

**LinkingTo** Rcpp

**LazyData** false

**RoxygenNote** 7.3.2

**VignetteBuilder** knitr

**Suggests** FNN, knitr, rmarkdown, BiocStyle, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**biocViews** Clustering, DataImport, GeneExpression, Transcriptomics,  
Spatial, Software

**git\_url** <https://git.bioconductor.org/packages/spARI>

**git\_branch** RELEASE\_3\_22

**git\_last\_commit** 4a3384b

**git\_last\_commit\_date** 2025-10-29

**Repository** Bioconductor 3.22

**Date/Publication** 2026-01-22

**Author** Yinqiao Yan [aut, cre],  
Xiangnan Feng [aut, fnd],  
Xiangyu Luo [aut, fnd]

**Maintainer** Yinqiao Yan <yinqiaoyan@bjut.edu.cn>

## Contents

|                                 |   |
|---------------------------------|---|
| generate_gs_pairs_int . . . . . | 2 |
| generate_sg_pairs_int . . . . . | 2 |
| perm_test . . . . .             | 3 |
| spARI . . . . .                 | 4 |
| spARI_example_data . . . . .    | 6 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>7</b> |
|--------------|----------|

---

generate\_gs\_pairs\_int    *Generate object pairs (gs pairs)*

---

### Description

This function identifies all object pairs that belong to the same group in the reference partition but different clusters in the clustering partition

### Arguments

c\_labels            An integer vector of clustering labels.  
r\_labels            An integer vector of reference labels.

### Value

An integer matrix with two columns, where each row is a pair of object indices.

### Examples

```
c_labels <- c(1,1,2,2,2,3,3)
r_labels <- c(1,1,1,2,2,3,1)
generate_gs_pairs_int(c_labels, r_labels)
```

---

generate\_sg\_pairs\_int    *Generate object pairs (sg pairs)*

---

### Description

This function identifies all object pairs that belong to the same cluster in the clustering partition but different groups in the reference partition

### Arguments

c\_labels            An integer vector of clustering labels.  
r\_labels            An integer vector of reference labels.

### Value

An integer matrix with two columns, where each row is a pair of object indices.

## Examples

```
c_labels <- c(1,1,2,2,2,3,3)
r_labels <- c(1,1,1,2,2,3,1)
generate_sg_pairs_int(c_labels, r_labels)
```

---

perm\_test

*The function perm\_test is the main function to carry out the hypothesis test procedure proposed in the paper.*

---

## Description

The function perm\_test is the main function to carry out the hypothesis test procedure proposed in the paper.

## Usage

```
perm_test(
  r_labels,
  c_labels,
  coords = NULL,
  dist_mat = NULL,
  f_func_input = NULL,
  h_func_input = NULL,
  alpha_val = 0.8,
  use_parallel = TRUE,
  replicate_times = 100,
  random_seed = 42,
  spe = NULL
)
```

## Arguments

|              |                                                                                                                                                                                                                                                                                                                |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| r_labels     | Annotated labels of all spots/cells. Can be numeric vector or character vector.                                                                                                                                                                                                                                |
| c_labels     | Estimated labels obtained by a certain spatial clustering method.                                                                                                                                                                                                                                              |
| coords       | Spatial coordinates (2 columns). 1st column: first dimension coordinate. 2nd column: second dimension coordinate. Default is NULL.                                                                                                                                                                             |
| dist_mat     | Distance matrix provided by users. If both coords and dist_mat are provided, we will directly use the distance matrix. Default is NULL. Please notice that if dist_mat is sparse, the weight function for object pairs without recorded distances degenerates to the setting used in the classical Rand index. |
| f_func_input | R function; function f provided by users.                                                                                                                                                                                                                                                                      |
| h_func_input | R function; function h provided by users.                                                                                                                                                                                                                                                                      |
| alpha_val    | Parameter in the default functions f and h, which belongs to the open interval (0, 1) to keep a positive gap between the maximal weight of the disagreement pair and the weight one of the agreement pair. Default is 0.8.                                                                                     |
| use_parallel | Logical; if TRUE, use parallel code to permute the two partitions. Default is TRUE.                                                                                                                                                                                                                            |

|                 |                                                                                                                                                                                                                                                                                                              |
|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| replicate_times | Number of permutations for both the reference and clustering partitions. Default is 100.                                                                                                                                                                                                                     |
| random_seed     | Random seed for reproducibility. Default is 42.                                                                                                                                                                                                                                                              |
| spe             | SpatialExperiment object; stores various components of spatial transcriptomics data, including spatialCoords: A matrix containing the spatial coordinates; colData\$cell_type: Annotated cell type labels for each spot or cell; colData\$cluster: Clustering labels for each spot or cell. Default is NULL. |

### Value

spARI returns an R numeric including the following information.

|           |                                                                       |
|-----------|-----------------------------------------------------------------------|
| spARI_obs | numeric, the observed spARI value calculated by r_labels and c_labels |
| p_value   | numeric, the p-value of the hypothesis testing                        |

### Examples

```
library(spARI)
data(spARI_example_data)
true_labels = spARI_example_data$true_labels
c1_labels = spARI_example_data$c1_labels
c2_labels = spARI_example_data$c2_labels
coords = spARI_example_data$coords
test_res1 = perm_test(r_labels=true_labels, c_labels=c1_labels, coords=coords,
                      use_parallel=FALSE)
test_res2 = perm_test(r_labels=true_labels, c_labels=c2_labels, coords=coords,
                      use_parallel=FALSE)
```

---

|       |                                                                                                          |
|-------|----------------------------------------------------------------------------------------------------------|
| spARI | <i>The function spARI is the main function to calculate spRI and spARI values proposed in the paper.</i> |
|-------|----------------------------------------------------------------------------------------------------------|

---

### Description

The function spARI is the main function to calculate spRI and spARI values proposed in the paper.

### Usage

```
spARI(
  r_labels,
  c_labels,
  coords = NULL,
  dist_mat = NULL,
  f_func_input = NULL,
  h_func_input = NULL,
  alpha_val = 0.8,
  spe = NULL
)
```

## Arguments

|                           |                                                                                                                                                                                                                                                                                                                                                        |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>r_labels</code>     | Annotated labels of all spots/cells. Can be numeric vector or character vector.                                                                                                                                                                                                                                                                        |
| <code>c_labels</code>     | Estimated labels obtained by a certain spatial clustering method.                                                                                                                                                                                                                                                                                      |
| <code>coords</code>       | Spatial coordinates (2 columns). 1st column: first dimension coordinate. 2nd column: second dimension coordinate. Default is NULL.                                                                                                                                                                                                                     |
| <code>dist_mat</code>     | Distance matrix provided by users. If both <code>coords</code> and <code>dist_mat</code> are provided, we will directly use the distance matrix. Default is NULL. Please notice that if <code>dist_mat</code> is sparse, the weight function for object pairs without recorded distances degenerates to the setting used in the classical Rand index.  |
| <code>f_func_input</code> | R function; function <code>f</code> provided by users.                                                                                                                                                                                                                                                                                                 |
| <code>h_func_input</code> | R function; function <code>h</code> provided by users.                                                                                                                                                                                                                                                                                                 |
| <code>alpha_val</code>    | Parameter in the default functions <code>f</code> and <code>h</code> , which belongs to the open interval (0, 1) to keep a positive gap between the maximal weight of the disagreement pair and the weight one of the agreement pair. Default is 0.8.                                                                                                  |
| <code>spe</code>          | SpatialExperiment object; stores various components of spatial transcriptomics data, including <code>spatialCoords</code> : A matrix containing the spatial coordinates; <code>colData\$cell_type</code> : Annotated cell type labels for each spot or cell; <code>colData\$cluster</code> : Clustering labels for each spot or cell. Default is NULL. |

## Value

spARI returns an R numeric including the following information.

|                          |                                                                                        |
|--------------------------|----------------------------------------------------------------------------------------|
| <code>spRI_value</code>  | numeric, the spRI value calculated by <code>r_labels</code> and <code>c_labels</code>  |
| <code>spARI_value</code> | numeric, the spARI value calculated by <code>r_labels</code> and <code>c_labels</code> |

## Examples

```
library(spARI)
### --- Import example data --- ###
# (1) true_labels: ground truth of 160 spots (input as r_labels)
# (2) c1_labels: one partition results of these spots (input as c_labels)
# (3) c2_labels: another partition results of these spots (input as c_labels)
# (4) coords: spatial coordinates of these spots (2 columns)
data(spARI_example_data)
true_labels = spARI_example_data$true_labels
c1_labels = spARI_example_data$c1_labels
c2_labels = spARI_example_data$c2_labels
coords = spARI_example_data$coords
### --- Compute spRI and spARI --- ###
res1 = spARI(r_labels=true_labels, c_labels=c1_labels, coords=coords)
res2 = spARI(r_labels=true_labels, c_labels=c2_labels, coords=coords)
cat(paste0("1st: spRI=", round(res1[1], 3), ", spARI=", round(res1[2], 3), "\n"))
cat(paste0("2nd: spRI=", round(res2[1], 3), ", spARI=", round(res2[2], 3), "\n"))
```

---

|                    |                               |
|--------------------|-------------------------------|
| spARI_example_data | <i>Example data for spARI</i> |
|--------------------|-------------------------------|

---

### Description

This dataset includes simulated spatial clustering results for demonstration purposes. It contains the following objects:

- **true\_labels**: Ground-truth labels of 160 spots.
- **c1\_labels**: First clustering result to compare.
- **c2\_labels**: Second clustering result to compare.
- **coords**: Spatial coordinates (matrix with 2 columns).

### Usage

```
data(spARI_example_data)
```

### Format

An environment containing 4 objects:

**true\_labels** Numeric or factor vector of length 160.

**c1\_labels** Numeric or factor vector of length 160.

**c2\_labels** Numeric or factor vector of length 160.

**coords** Numeric matrix of size 160x2.

# Index

## \* **datasets**

spARI\_example\_data, [6](#)

generate\_gs\_pairs\_int, [2](#)

generate\_sg\_pairs\_int, [2](#)

perm\_test, [3](#)

spARI, [4](#)

spARI\_example\_data, [6](#)