Package: scMetaTraj (via r-universe)

May 25, 2026

Type Package

Title Metabolic State Space and Trajectory Analysis for Single-Cell Data

Version 0.1.2

Date 2026-05-25

Description Provides a framework for modeling cellular metabolic states and continuous metabolic trajectories from single-cell RNA-seq data using pathway-level scoring. Enables lineage-restricted metabolic analysis, metabolic 'pseudotime' inference, module-level trend analysis, and visualization of metabolic state transitions.

License MIT + file LICENSE

URL https://github.com/Dr-xyGreg/scMetaTraj

BugReports https://github.com/Dr-xyGreg/scMetaTraj/issues

Encoding UTF-8

Depends R (>= 4.1.0)

Imports stats, igraph, ggplot2, uwot, RANN, Seurat, dplyr, tidyr

Suggests knitr, rmarkdown

VignetteBuilder knitr

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

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

Repository https://dr-xygreg.r-universe.dev

Date/Publication 2026-05-24 17:07:03 UTC

RemoteUrl https://github.com/dr-xygreg/scmetatraj

RemoteRef HEAD

RemoteSha 604357779a7cf7599fcb5fdad5c460ffc12d459c

Contents

`scMetaTraj_axis_score` *Calculate metabolic axis scores*

Description

Calculate metabolic axis scores

Usage

```
scMetaTraj_axis_score(scores, axis_list, scale = TRUE)
```

Arguments

`scores`	Cell-by-module metabolic score matrix
`axis_list`	Named list defining metabolic axes
`scale`	Logical, whether to scale axis scores (Z-score)

Value

A matrix of cell-by-axis scores

`scMetaTraj_cluster` *Cluster cells in metabolic PCA space*

Description

scMetaTraj_cluster() identifies metabolic subclusters by constructing a kNN graph in metabolic PCA space and applying community detection.

IMPORTANT DESIGN PRINCIPLES:

- Clustering is performed ONLY in metabolic PCA space.
- UMAP coordinates must NEVER be used for clustering.
- Results are independent of transcriptomic clustering.

Usage

```
scMetaTraj_cluster(
  embedding,
  k = 20,
  resolution = 0.5,
  method = c("leiden", "louvain"),
  seed = 123
)
```

Arguments

`embedding`	Numeric matrix (cells x PCs). Output of scMetaTraj_embed(method = "PCA").
`k`	Integer. Number of nearest neighbors for kNN graph.
`resolution`	Numeric. Resolution parameter for clustering (used for Leiden only).
`method`	Character. "leiden" (default) or "louvain".
`seed`	Integer. Random seed for reproducibility.

Value

A factor of length equal to number of cells, giving metabolic cluster labels per cell.

Examples

```
# Create example PCA embedding
set.seed(123)
n_cells <- 100
n_pcs <- 5

embedding <- matrix(rnorm(n_cells * n_pcs), nrow = n_cells, ncol = n_pcs)
rownames(embedding) <- paste0("Cell", 1:n_cells)
colnames(embedding) <- paste0("PC", 1:n_pcs)
```

```
# Perform clustering
clusters <- scMetaTraj_cluster(
  embedding = embedding,
  k = 20,
  method = "louvain"
)

# View results
table(clusters)
```

`scMetaTraj_cluster_profile` *Summarize metabolic profiles of clusters*

Description

scMetaTraj_cluster_profile() computes representative metabolic pathway activities for each metabolic cluster.

Usage

```
scMetaTraj_cluster_profile(
  scores,
  metabolic_cluster,
  stat = c("median", "mean"),
  scale = TRUE
)
```

Arguments

`scores`	Numeric matrix, cells x pathways.
`metabolic_cluster`	Factor or character vector, length = nrow(scores).
`stat`	Character. "median" (default) or "mean".
`scale`	Logical. Whether to z-score pathways across clusters.

Value

A data.frame: clusters x pathways.

`scMetaTraj_embed` *Embed cells in metabolic feature space*

Description

scMetaTraj_embed() constructs a low-dimensional representation of cells based on pathway-level metabolic scores.

DESIGN PRINCIPLES:

- PCA is the true analysis space (for graph construction).
- UMAP is ONLY for visualization.

Usage

```
scMetaTraj_embed(
  scores,
  method = c("PCA", "UMAP"),
  n_pcs = 10,
  umap_n_neighbors = 30,
  umap_min_dist = 0.3,
  seed = 123
)
```

Arguments

`scores`	Numeric matrix, cells x pathways.
`method`	Character. "PCA" (default) or "UMAP".
`n_pcs`	Integer. Number of PCs to return / use.
`umap_n_neighbors`	Integer. UMAP n_neighbors.
`umap_min_dist`	Numeric. UMAP min_dist.
`seed`	Integer. Random seed.

Value

A numeric matrix:

- PCA: cells x PCs
- UMAP: cells x 2

scMetaTraj_flow *Compute local directional consistency along mPT*

Description

Estimate local direction vectors pointing toward neighbors with higher metabolic pseudotime. Intended for visualization only.

Usage

```
scMetaTraj_flow(emb_pca, emb_umap, pseudotime, k = 15, min_delta = 0.02)
```

Arguments

emb_pca	Matrix (cells x PCs) used for neighborhood definition.
emb_umap	Matrix (cells x 2) used only for visualization.
pseudotime	Numeric vector of mPT.
k	Integer. Number of nearest neighbors.
min_delta	Minimum mPT difference to consider a neighbor "forward".

Value

Data.frame with UMAP coordinates and dx, dy vectors.

scMetaTraj_infer *Infer metabolic pseudotime from a k-nearest-neighbor graph*

Description

scMetaTraj_infer() builds a weighted k-nearest-neighbor graph in metabolic PCA space and computes graph distances from a selected root cell. The rescaled distances define metabolic pseudotime (mPT).

Usage

```
scMetaTraj_infer(
  embedding,
  k = 20,
  root_mode = c("pc1_min", "pc1_max", "axis_min", "axis_max", "manual"),
  axis_score = NULL,
  root_cell = NULL,
  scale = TRUE
)
```

Arguments

embedding	Numeric matrix of cells x PCs, usually returned by scMetaTraj_embed(method = "PCA").
k	Integer. Number of nearest neighbors used to define the graph.
root_mode	Character. Strategy used to choose the root cell. One of "pc1_min", "pc1_max", "axis_min", "axis_max", or "manual".
axis_score	Optional numeric vector used when root_mode is "axis_min" or "axis_max".
root_cell	Optional character scalar giving the row name of the root cell when root_mode = "manual".
scale	Logical. Whether to rescale graph distances to the interval [0, 1].

Value

A named list with elements:

- mPT: numeric vector of metabolic pseudotime values.
- root: selected root cell name.
- dist: raw graph distances from the root cell.

Examples

```
set.seed(123)
embedding <- matrix(rnorm(120 * 5), nrow = 120, ncol = 5)
rownames(embedding) <- paste0("Cell", seq_len(nrow(embedding)))
mpt <- scMetaTraj_infer(embedding, k = 15, root_mode = "pc1_min")
head(mpt$mPT)
```

---

scMetaTraj_mPT_distribution

Prepare mPT distribution data for metabolic subclusters

Description

Organize metabolic pseudotime distributions by cluster and automatically order clusters along median mPT.

Usage

```
scMetaTraj_mPT_distribution(mPT, cluster)
```

Arguments

mPT	Numeric vector of metabolic pseudotime.
cluster	Factor or character vector of metabolic cluster labels.

Value

Data.frame with columns mPT and cluster (ordered factor).

`scMetaTraj_plot_gradient` *Plot metabolic gradient on UMAP*

Description

Plot metabolic gradient on UMAP

Usage

```
scMetaTraj_plot_gradient(
  embedding,
  score,
  title = "Metabolic gradient",
  palette
)
```

Arguments

`embedding`	UMAP coordinates (data.frame or matrix)
`score`	Vector of metabolic axis score
`title`	Plot title
`palette`	Continuous color palette

Value

ggplot object

`scMetaTraj_plot_module_cards` *Plot metabolic module cards*

Description

Plot metabolic module cards

Usage

```
scMetaTraj_plot_module_cards(cluster_profile, module_axis_map, axis_palette)
```

Arguments

`cluster_profile`	Cluster-by-module matrix
`module_axis_map`	Data.frame with columns: module, axis
`axis_palette`	Named vector of colors for axes

Value

ggplot object

`scMetaTraj_plot_trend_by_cluster` *Plot module trends by metabolic cluster*

Description

Plot module trends by metabolic cluster

Plot module trends along mPT stratified by cluster

Usage

```
scMetaTraj_plot_trend_by_cluster(trend_by_cluster, palette)

scMetaTraj_plot_trend_by_cluster(trend_by_cluster, palette)
```

Arguments

`trend_by_cluster`	Output from scMetaTraj_trend_by_cluster().
`palette`	Named vector of cluster colors (e.g., scMetaTraj_palette_discrete).

Value

ggplot object.

ggplot object.

scMetaTraj_plot_trend_multi *Plot multiple module trends with switchpoints*

Description

Visualize metabolic module trends along pseudotime with identified switchpoints marked by vertical dashed lines.

Usage

```
scMetaTraj_plot_trend_multi(trend_long, switchpoints)
```

Arguments

trend_long	Data frame with columns: module, mPT_bin, score_smooth. Output from scMetaTraj_trend_multi()$trend_long.
switchpoints	Data frame with columns: module, mPT_switch. Output from scMetaTraj_trend_multi()$switchpoints.

Value

A ggplot2 object showing trends faceted by module.

Examples

```
# Create example trend data
trend_long <- data.frame(
  module = rep(c("Glycolysis", "OXPHOS"), each = 30),
  mPT_bin = rep(seq(0, 1, length.out = 30), 2),
  score_smooth = c(sin(seq(0, pi, length.out = 30)),
                   cos(seq(0, pi, length.out = 30)))
)

# Create example switchpoint data
switchpoints <- data.frame(
  module = c("Glycolysis", "OXPHOS"),
  mPT_switch = c(0.5, 0.3)
)

# Plot
p <- scMetaTraj_plot_trend_multi(trend_long, switchpoints)
print(p)
```

`scMetaTraj_score` *Calculate metabolic pathway scores for single cells*

Description

scMetaTraj_score() maps gene-level expression to pathway-level metabolic activity scores. The resulting matrix defines the metabolic feature space used by all downstream scMetaTraj modules.

IMPORTANT:

- Scores represent relative metabolic states, NOT metabolic flux.
- Designed to be robust to dropout in scRNA-seq data.

Usage

```
scMetaTraj_score(
  x,
  gene_sets,
  assay = "RNA",
  slot = "data",
  method = c("mean", "zscore"),
  min_genes = 3,
  scale = TRUE
)
```

Arguments

`x`	A Seurat object or a gene x cell expression matrix.
`gene_sets`	A named list: pathway -> character vector of genes.
`assay`	Character. Seurat assay to use. Default "RNA".
`slot`	Character. Expression slot. Default "data".
`method`	Character. Scoring method: "mean" or "zscore".
`min_genes`	Integer. Minimal number of detected genes per pathway.
`scale`	Logical. Whether to z-score pathway scores across cells.

Value

A numeric matrix: cells x pathways.

scMetaTraj_switchpoint *Identify metabolic trajectory switchpoint*

Description

Identifies the point along metabolic pseudotime where a module shows maximum change in trend (inflection point).

Usage

```
scMetaTraj_switchpoint(trend_df)
```

Arguments

trend_df	Data frame with columns: mPT_bin and score_smooth. Typically output from scMetaTraj_trend.

Value

A list with:

mPT_switch	Numeric. The mPT value at the switchpoint
index	Integer. The index (row number) of the switchpoint in trend_df

Examples

```
# Create example trend data
set.seed(456)
n_cells <- 200
mPT <- runif(n_cells, 0, 1)

# Simulate trend with switchpoint at mPT = 0.5
scores <- ifelse(mPT < 0.5,
                 0.3 + rnorm(n_cells, 0, 0.05),
                 0.7 + rnorm(n_cells, 0, 0.05))

# Compute trend
trend <- scMetaTraj_trend(scores, mPT, n_bins = 30, smooth = TRUE)

# Find switchpoint
switchpoint <- scMetaTraj_switchpoint(trend)
print(switchpoint$mPT_switch)

# Visualize
plot(trend$mPT_bin, trend$score_smooth, type = "l",
     xlab = "Metabolic pseudotime", ylab = "Module score")
abline(v = switchpoint$mPT_switch, col = "red", lty = 2)
```

scMetaTraj_transition_zone

Identify candidate transition zone along mPT

Description

Identify a candidate transition zone along mPT where the composition of metabolic subclusters changes most rapidly. Intended as a hypothesis-generating indicator.

Usage

```
scMetaTraj_transition_zone(mPT, cluster, n_bins = 30, top_frac = 0.2)
```

Arguments

mPT	Numeric vector of metabolic pseudotime.
cluster	Metabolic cluster labels.
n_bins	Integer. Number of bins along mPT.
top_frac	Fraction of bins with highest composition change to define zone.

Value

Named numeric vector with xmin and xmax.

scMetaTraj_trend

Compute metabolic module trend along pseudotime

Description

Bins cells along metabolic pseudotime (mPT) and computes mean module scores per bin, with optional loess smoothing.

Usage

```
scMetaTraj_trend(scores, mPT, n_bins = 30, smooth = TRUE, span = 0.3)
```

Arguments

scores	Numeric vector of module scores (length = n_cells).
mPT	Numeric vector of metabolic pseudotime values (length = n_cells).
n_bins	Integer. Number of bins along mPT for trend computation.
smooth	Logical. Whether to apply loess smoothing to the binned trend.
span	Numeric. Loess span parameter (only used if smooth = TRUE).

Value

A data frame with columns:

mPT_bin	Mid-point of each mPT bin
score	Mean score per bin
score_smooth	Smoothed score (if smooth = TRUE, otherwise same as score)

Examples

```
# Create example data
set.seed(123)
n_cells <- 200
mPT <- runif(n_cells, 0, 1)
scores <- sin(mPT * 2 * pi) + rnorm(n_cells, 0, 0.1)

# Compute trend
trend <- scMetaTraj_trend(
  scores = scores,
  mPT = mPT,
  n_bins = 30,
  smooth = TRUE,
  span = 0.3
)

# Plot trend
plot(trend$mPT_bin, trend$score_smooth, type = "l",
     xlab = "Metabolic pseudotime", ylab = "Module score")
```

---

scMetaTraj_trend_by_cluster

Compute module trends along mPT stratified by metabolic cluster

---

Description

Compute module trends along mPT stratified by metabolic cluster

Usage

```
scMetaTraj_trend_by_cluster(
  score_mat,
  mPT,
  cluster,
  modules,
  n_bins = 30,
  smooth = TRUE,
  span = 0.3,
  min_cells = 50
)
```

Arguments

score_mat	Matrix/data.frame (cells x modules).
mPT	Numeric vector.
cluster	Factor/character vector of cluster labels.
modules	Character vector of module names.
n_bins	Integer. Number of bins.
smooth	Logical. Whether to loess smooth.
span	Numeric. Loess span.
min_cells	Integer. Minimum cells per cluster to compute trends.

Value

Long-format data.frame with columns: cluster, module, mPT_bin, score, score_smooth, n_cells

---

scMetaTraj_trend_multi

Compute trends and switchpoints for multiple modules along mPT

Description

Compute trends and switchpoints for multiple modules along mPT

Usage

```
scMetaTraj_trend_multi(
  score_mat,
  mPT,
  modules,
  n_bins = 30,
  smooth = TRUE,
  span = 0.3
)
```

Arguments

score_mat	Matrix/data.frame (cells x modules). Row order must match mPT.
mPT	Numeric vector (length = n_cells).
modules	Character vector of module names (columns of score_mat).
n_bins	Integer. Number of mPT bins.
smooth	Logical. Whether to loess smooth.
span	Numeric. Loess span.

Value

A list with:

- trend_long: long-format data.frame for plotting
- switchpoints: data.frame of module-wise switchpoints

Index