Package 'PopComm'

May 7, 2026

Title Population-Level Cell-Cell Communication Analysis Tools

Version 1.0.0

Description Facilitates population-level analysis of ligand-receptor (LR) interactions using large-scale single-cell transcriptomic data. Identifies significant LR pairs and quantifies their interactions through correlation-based filtering and projection score computations. Designed for large-sample single-cell studies, the package employs statistical modeling, including linear regression, to investigate LR relationships between cell types. It provides a systematic framework for understanding cell-cell communication, uncovering regulatory interactions and signaling mechanisms. Offers tools for LR pair-level, sample-level, and differential interaction analyses, with comprehensive visualization support to aid biological interpretation. The methodology is described in a manuscript currently under review and will be referenced here once published or publicly available.

URL https://github.com/JusticeGO/PopComm

BugReports https://github.com/JusticeGO/PopComm/issues

Depends R (>= 4.1.0)

Imports Seurat (>= 4.1.0), broom (>= 1.0.0), dplyr (>= 1.0.0), purrr (>= 0.3.0), rlang, stringr (>= 1.4.0), tibble (>= 3.0.0), tidyr (>= 1.0.0), tidyselect (>= 1.1.0), Matrix (>= 1.2-0), ggplot2 (>= 3.3.0), ggpubr (>= 0.6.0), pheatmap (>= 1.0.12), RColorBrewer, reshape2 (>= 1.4.1), scales (>= 1.1.1), igraph (>= 2.0.0), parallel, pbmcapply (>= 1.5.0), grDevices, cowplot (>= 1.1.1)

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

LazyData true

NeedsCompilation no

Author Zheng Gong [aut, cre] (ORCID: <https://orcid.org/0000-0001-7285-2630>), Mengwei Li [aut, ctb] (ORCID: <https://orcid.org/0000-0001-6163-2827>)

Maintainer Zheng Gong <gongzheng131@hotmail.com>

Repository CRAN

Date/Publication 2025-11-01 22:00:02 UTC

Contents

`boxplot_lr_group_comparison` *Boxplot Comparison of Ligand-Receptor Interaction Scores Across Groups*

Description

Generates a boxplot comparing ligand-receptor (LR) interaction scores across sample groups with optional significance testing (t-test or Wilcoxon).

Usage

```
boxplot_lr_group_comparison(
  lr_scores,
  metadata,
  ligand,
  receptor,
  sender,
  receiver,
  group_by,
  score = c("normalized", "raw"),
```

```
  test = TRUE,
  paired = FALSE,
  test_method = c("wilcox.test", "t.test"),
  colors = c("#5fa9d1", "#154778"),
  title = NULL
)
```

Arguments

lr_scores	Data frame containing LR interaction scores per sample (data frame).
metadata	Data frame containing sample metadata (data frame).
ligand	Ligand gene name to filter (character).
receptor	Receptor gene name to filter (character).
sender	Sender cell type to filter (character).
receiver	Receiver cell type to filter (character).
group_by	Column name in metadata to group samples (character).
score	Use 'normalized' or 'raw' score (default: "normalized") (character).
test	Whether to add a statistical test annotation (logical, default: TRUE).
paired	Whether to treat the comparison as paired (logical, default: FALSE).
test_method	Statistical test to use: "t.test" or "wilcox.test" (default = "wilcox.test") (character).
colors	Vector of colors for groups (default: c("#5fa9d1", "#154778")).
title	Custom plot title (optional).

Value

A list containing:

- plot - ggplot object of the boxplot
- df - data frame used for plotting

Examples

```
# Boxplot of LR Score by group
data(lr_scores_eg)
data(metadata_eg)

res <- boxplot_lr_group_comparison(
  lr_scores = lr_scores_eg,
  metadata = metadata_eg,
  ligand = "PSAP",
  receptor = "LRP1",
  sender = "Perivascular",
  receiver = "Fibroblast",
  group_by = "IFN_type",
  score = "normalized"
  )
```

```
print(res$plot)
head(res$df)
```

`circle_plot` *Plot Circular Ligand-Receptor Interaction Network*

Description

Plots a circular ligand-receptor (LR) interaction network with curved directed edges. Nodes are arranged in a circle, and edge widths and colors represent interaction strengths.

Usage

```
circle_plot(
  filtered_lr,
  edge_width = c("count", "cor"),
  node_colors = NULL,
  show_self_interactions = TRUE,
  cutoff = 0
)
```

Arguments

`filtered_lr`	A data frame of ligand-receptor pairs from prior analysis (e.g., output of `filter_lr_all`), containing at least the columns "sender", "receiver", and "cor".
`edge_width`	Determines edge weights, either "cor" (correlation) or "count" (interaction count) (default: "count").
`node_colors`	Named vector mapping cell types to colors. Example: c("Cardiac" = "#E41A1C", "Fibroblast" = "#377EB8"). If NULL, uses default palette.
`show_self_interactions`	Logical indicating whether to display self-interactions (logical, default: TRUE).
`cutoff`	Minimum edge weight to display (numeric, default: 0).

Value

A ggplot object representing the network plot.

Examples

```
# Plot Circular Cell-Cell Interaction Network
data(filtered_lr_eg)

p <- circle_plot(
  filtered_lr = filtered_lr_eg,
  edge_width = "count",
  show_self_interactions = TRUE
```

```
)

print(p)
```

---

dotplot_lr_continuous_group *Dotplot of Ligand-Receptor Interaction Scores Against Continuous Group Variable*

Description

Creates a dotplot (scatter plot) of ligand-receptor (LR) interaction scores against a continuous variable with optional regression line.

Usage

```
dotplot_lr_continuous_group(
  lr_scores,
  metadata,
  ligand,
  receptor,
  sender,
  receiver,
  group_by,
  score = c("normalized", "raw"),
  point_size = 3,
  point_color = "dodgerblue4",
  add_regression = TRUE,
  title = NULL
)
```

Arguments

lr_scores	Data frame containing LR interaction scores per sample (data frame).
metadata	Data frame containing sample metadata (data frame).
ligand	Ligand gene name to filter (character).
receptor	Receptor gene name to filter (character).
sender	Sender cell type to filter (character).
receiver	Receiver cell type to filter (character).
group_by	Continuous variable column in metadata (e.g., age, severity score) (character).
score	Use 'normalized' or 'raw' score (default: "normalized") (character).
point_size	Size of the points in the plot (numeric, default: 3).
point_color	Color of the points in the plot (default: "dodgerblue4").
add_regression	Whether to add regression line (logical, default: TRUE).
title	Custom plot title (optional).

Value

A list containing:

- plot - ggplot object of the dotplot
- df - data frame used for plotting

Examples

```
# Dotplot of LR Score Against Continuous Group Variable
data(lr_scores_eg)
data(metadata_eg)

res <- dotplot_lr_continuous_group(
  lr_scores = lr_scores_eg,
  metadata = metadata_eg,
  ligand = "HLA-A",
  receptor = "LILRB2",
  sender = "Lymphoid",
  receiver = "Myeloid",
  group_by = "IFNscore"
)

print(res$plot)
head(res$df)
```

---

dot_plot *Create Ligand-Receptor Interaction Dot Plot*

Description

Generates a dot plot to visualize ligand-receptor (LR) interaction. Dot sizes are scaled by the correlation coefficient and dot colors represent -log10(adjust.p). The function supports plotting the top interactions per sender-receiver pair or user-specified ligand-receptor pairs.

Usage

```
dot_plot(
  filtered_lr,
  top_n = 5,
  axis = c("LR-SR", "SR-LR"),
  type_scale = c("size", "radius"),
  selected_LR = NULL
)
```

Arguments

filtered_lr	A data frame containing ligand-receptor interaction data.
top_n	Integer specifying the number of top interactions to select per sender-receiver pair (numeric, default: 5).
axis	Character indicating the configuration of rows and columns in the plot. Options: "LR-SR" (default, rows = ligand-receptor pairs, columns = sender-receiver pairs) or "SR-LR".
type_scale	Character indicating the scaling method for the plot. Options: "size" (default, uses scale_size() for point scaling) or "radius" (uses scale_radius() for point scaling).
selected_LR	Optional character vector of ligand-receptor pair identifiers (e.g., c("TIMP1_CD63", "DSCAM_DCC")). If NULL, the top_n interactions per sender-receiver pair are used.

Value

A ggplot object representing the dot plot.

Examples

```
# Plot LR Interaction Dot Plot
data(filtered_lr_eg)

p <- dot_plot(
  filtered_lr = filtered_lr_eg,
  top_n = 3,
  axis = "LR-SR",
  type_scale = "size",
  )

print(p)
```

---

filtered_lr_eg *Example for filtered_lr*

Description

Example for filtered_lr

Usage

```
filtered_lr_eg
```

Format

An object of class data.frame with 1513 rows and 11 columns.

`filter_lr_all` *Filter and Analyze Ligand-Receptor Pair Correlations (All Cell Types)*

Description

Filters ligand-receptor (LR) pairs and analyzes their correlations for all possible cell type pairs, and returns significant LR pairs based on user-defined thresholds. This function supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names).

Usage

```
filter_lr_all(
  rna,
  lr_database = PopComm::lr_db,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  min_samples = 10,
  min_cell_ratio = 0.1,
  min_sample_ratio = 0.1,
  cor_method = "spearman",
  adjust_method = "BH",
  min_adjust_p = 0.05,
  min_cor = 0,
  min_r2 = 0,
  min_fstat = 0,
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

`rna`	A Seurat object or a matrix containing single-cell RNA expression data.
`lr_database`	A data frame of ligand-receptor pairs with columns "ligand_gene_symbol" and "receptor_gene_symbol".
`sample_col`	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).
`cell_type_col`	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
`id_sep`	Separator used in matrix column names to split sample and cell type (e.g., `--` for "Cardiac–sample1"). Only used in Matrix mode.
`min_cells`	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.

`min_samples`	Minimum number of valid samples to proceed (numeric, default 10).
`min_cell_ratio`	Minimum ratio of cells expressing ligand and receptor genes in sender or receiver cells (numeric, default 0.1). Only used in Seurat mode.
`min_sample_ratio`	Minimum ratio of samples in which both the ligand and receptor genes must be expressed (numeric, default 0.1).
`cor_method`	Correlation method: "spearman" (default), "pearson", or "kendall".
`adjust_method`	P-value adjustment method (default "BH" for Benjamini-Hochberg). Options: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
`min_adjust_p`	Adjusted p-value threshold for significance (numeric, default 0.05).
`min_cor`	Minimum correlation coefficient threshold (numeric, default 0). Must be ≥ 0.
`min_r2`	Minimum R-squared threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
`min_fstat`	Minimum F-statistic threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
`num_cores`	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
`verbose`	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

A data frame includes LR pairs with sufficient correlation and expression support across samples.

`ligand`, `receptor`	Ligand and receptor gene symbols.
`cor`	Correlation coefficient.
`p_val`	Raw p-value.
`adjust.p`	Adjusted p-value.
`sender`, `receiver`	Sender and receiver cell types.
`slope`	Slope of the linear regression model.
`intercept`	Intercept of the linear regression model.
`r2`	R-squared of the linear regression model.
`fstat`	F-statistic of the linear regression model.

Rows are ordered by ascending `adjust.p` and descending `cor`.

Returns `NULL` if:

- No cell types are found in the metadata.
- Insufficient samples or cells remain after filtering.
- No ligand-receptor pairs pass the filtering thresholds.

Examples

```
data(matrix_object)
data(lr_db)

# Analyzing ligand-receptor interactions between all cell types
result01a <- filter_lr_all(
  rna = matrix_object,
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

if (!is.null(result01a)) {
print(head(result01a))
}
```

`filter_lr_single` *Filter and Analyze Ligand-Receptor Pair Correlations (Specified Sender and Receiver)*

Description

Filters ligand-receptor (LR) pairs and analyzes their correlations for specified sender and receiver cell types, and returns significant LR pairs based on user-defined thresholds. This function supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names).

Usage

```
filter_lr_single(
  rna,
  sender,
  receiver,
  lr_database = PopComm::lr_db,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  min_samples = 10,
  min_cell_ratio = 0.1,
  min_sample_ratio = 0.1,
```

```
  cor_method = "spearman",
  adjust_method = "BH",
  min_adjust_p = 0.05,
  min_cor = 0,
  min_r2 = 0,
  min_fstat = 0,
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

rna	A Seurat object or a matrix containing single-cell RNA expression data.
sender	Cell type designated as the ligand sender (character).
receiver	Cell type designated as the receptor receiver (character).
lr_database	A data frame of ligand-receptor pairs with columns "ligand_gene_symbol" and "receptor_gene_symbol".
sample_col	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).
cell_type_col	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
id_sep	Separator used in matrix column names to split sample and cell type (e.g., -- for "Cardiac–sample1"). Only used in Matrix mode.
min_cells	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.
min_samples	Minimum number of valid samples to proceed (numeric, default 10).
min_cell_ratio	Minimum ratio of cells expressing ligand and receptor genes in sender or receiver cells (numeric, default 0.1). Only used in Seurat mode.
min_sample_ratio	Minimum ratio of samples in which both the ligand and receptor genes must be expressed (numeric, default 0.1).
cor_method	Correlation method: "spearman" (default), "pearson", or "kendall".
adjust_method	P-value adjustment method (default "BH" for Benjamini-Hochberg). Options: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
min_adjust_p	Adjusted p-value threshold for significance (numeric, default 0.05).
min_cor	Minimum correlation coefficient threshold (numeric, default 0). Must be ≥ 0.
min_r2	Minimum R-squared threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
min_fstat	Minimum F-statistic threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
num_cores	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
verbose	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

A data frame includes LR pairs with sufficient correlation and expression support across samples.

ligand, receptor	Ligand and receptor gene symbols.
cor	Correlation coefficient.
p_val	Raw p-value.
adjust.p	Adjusted p-value.
sender, receiver	Sender and receiver cell types.
slope	Slope of the linear regression model.
intercept	Intercept of the linear regression model.
r2	R-squared of the linear regression model.
fstat	F-statistic of the linear regression model.

Rows are ordered by ascending adjust.p and descending cor.

Returns NULL if:

- No cell types are found in the metadata.
- Insufficient samples or cells remain after filtering.
- No ligand-receptor pairs pass the filtering thresholds.

Examples

```
data(matrix_object)
data(lr_db)

# Analyzing ligand-receptor interactions: Perivascular -> Endothelial
result01s <- filter_lr_single(
  rna = matrix_object,
  sender = "Perivascular",
  receiver = "Endothelial",
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

if (!is.null(result01s)) {
print(head(result01s))
}
```

heatmap_sample *Generate Heatmap of Ligand-Receptor Interaction Scores*

Description

This function generates a heatmap to visualize the ligand-receptor (LR) interaction scores across samples. Rows represent LR pairs and columns represent samples. Optionally, sample metadata can be used to annotate the columns.

Usage

```
heatmap_sample(
  lr_scores,
  metadata,
  score = c("normalized", "raw"),
  selected_sender = NULL,
  selected_receiver = NULL,
  selected_metadata = NULL,
  treeheight_row = 50,
  treeheight_col = 50,
  show_LR = FALSE,
  show_sample = FALSE,
  basic_title = NULL
)
```

Arguments

lr_scores	Data frame containing LR interaction scores per sample (data frame).
metadata	Data frame containing sample metadata (data frame).
score	Character string indicating which score to use: "normalized" (default) or "raw" .
selected_sender	Specific sender cell type to filter, default is None (use all) (character).
selected_receiver	Specific receiver cell type to filter, default is None (use all) (character).
selected_metadata	List of column names in metadata to annotate samples (default: None, use all)(character vector).
treeheight_row	The height of a tree for rows (numeric, default: 50).
treeheight_col	The height of a tree for columns (numeric, default: 50).
show_LR	Whether to display ligand-receptor names on rows (logical, default: FALSE).
show_sample	Whether to display sample names on columns (logical, default: FALSE).
basic_title	Custom heatmap title (optional).

Value

A pheatmap object.

Examples

```
# Heatmap of LR Interaction Scores
data(lr_scores_eg)
data(metadata_eg)

p <- heatmap_sample(
  lr_scores = lr_scores_eg,
  metadata = metadata_eg,
  score = "normalized",
  selected_sender = "Endothelial",
  selected_receiver = "Perivascular",
  selected_metadata = c("Sex", "Age_group", "IFN_type")
  )

print(p)
```

---

lr_db *Ligand-Receptor Pair Database*

Description

A comprehensive database of human ligand-receptor pairs with gene/protein identifiers and supporting evidence from literature. Data imported from `human_lr_pair.txt`. CellTalkDB: A manually curated database of ligand-receptor interactions in human and mouse

Usage

```
lr_db
```

Format

A data frame with 3,398 rows (pairs) and 10 columns:

lr_pair Character. Unique identifier for ligand-receptor pair, formatted as "LIGAND_RECEPTOR" (e.g., "SEMA3F_PLXNA3")

ligand_gene_symbol Character. Official HGNC symbol of the ligand gene (e.g., "SEMA3F")

receptor_gene_symbol Character. Official HGNC symbol of the receptor gene (e.g., "PLXNA3")

ligand_gene_id Integer. Entrez Gene ID of the ligand gene (NCBI identifier)

receptor_gene_id Integer. Entrez Gene ID of the receptor gene (NCBI identifier)

ligand_ensembl_protein_id Character. Ensembl protein ID of the ligand (e.g., "ENSP00000002829")

receptor_ensembl_protein_id Character. Ensembl protein ID of the receptor (e.g., "ENSP00000358696")

ligand_ensembl_gene_id Character. Ensembl gene ID of the ligand (e.g., "ENSG00000001617")

receptor_ensembl_gene_id Character. Ensembl gene ID of the receptor (e.g., "ENSG00000130827")

evidence Character. PubMed IDs (PMIDs) supporting the interaction, comma-separated (e.g., "15721238")

Source

Source from CellTalkDB (PMID: 33147626).

`lr_linear_model_discrete` *Compare Ligand-Receptor Interaction Scores with Group Variable using Linear Regression*

Description

Perform linear regression analysis to compare ligand-receptor (LR) interaction scores across groups, handling both continuous and binary group variables (ident1 vs ident2 or all others).

Usage

```
lr_linear_model_discrete(
  lr_scores,
  metadata,
  group_variable,
  ident1,
  ident2 = NULL,
  covariates = NULL,
  fdr_threshold = 0.05
)
```

Arguments

`lr_scores`	Data frame containing LR interaction scores per sample (data frame).
`metadata`	Data frame containing sample metadata (data frame).
`group_variable`	Column name in `metadata` to compare groups (categorical or continuous) (character).
`ident1`	If categorical, group to compare (coded as 1) (character).
`ident2`	Reference group or list of groups (coded as 0). If None, uses all others (character).
`covariates`	Optional list of covariate column names (character vector).
`fdr_threshold`	Significance cutoff for adjusted p-values (numeric, default: 0.05).

Value

Data frame with ligand, receptor, sender, receiver, coef, p-values, and adjusted p-values.

Examples

```
data(lr_scores_eg)
data(metadata_eg)

res <- lr_linear_model_discrete(
  lr_scores = lr_scores_eg,
  metadata = metadata_eg,
  group_variable = "IFN_type",
  ident1 = "high",
  covariates = c("Age_group", "Sex")
  )

head(res)
```

`lr_scores_eg` *Example for lr_scores*

Description

Example for lr_scores

Usage

```
lr_scores_eg
```

Format

An object of class `data.frame` with 187593 rows and 14 columns.

`matrix_object` *Example for matrix object*

Description

Example for matrix object

Usage

```
matrix_object
```

Format

An object of class `matrix` (inherits from `array`) with 223 rows and 144 columns.

`metadata_eg` *Example for metadata*

Description

Example for metadata

Usage

```
metadata_eg
```

Format

An object of class `data.frame` with 163 rows and 9 columns.

`one_step_all` *Analyze Ligand-Receptor Pair Correlations and Projection Scores (Across All Cell Types)*

Description

Performs integrated analysis of ligand-receptor (LR) pairs through two consecutive phases: (1) Filters LR pairs and analyzes correlations across all cell types; (2) Calculates projection scores based on regression models for valid pairs. Returns comprehensive results combining statistical metrics. This function supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names).

Usage

```
one_step_all(
  rna,
  lr_database,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  min_samples = 10,
  min_cell_ratio = 0.1,
  min_sample_ratio = 0.1,
  cor_method = "spearman",
  adjust_method = "BH",
  min_adjust_p = 0.05,
  min_cor = 0,
  min_r2 = 0,
  min_fstat = 0,
```

```
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

rna	A Seurat object or a matrix containing single-cell RNA expression data.
lr_database	A data frame of ligand-receptor pairs with columns "ligand_gene_symbol" and "receptor_gene_symbol".
sample_col	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).
cell_type_col	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
id_sep	Separator used in matrix column names to split sample and cell type (e.g., -- for "Cardiac–sample1"). Only used in Matrix mode.
min_cells	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.
min_samples	Minimum number of valid samples to proceed (numeric, default 10).
min_cell_ratio	Minimum ratio of cells expressing ligand and receptor genes in sender or receiver cells (numeric, default 0.1). Only used in Seurat mode.
min_sample_ratio	Minimum ratio of samples in which both the ligand and receptor genes must be expressed (numeric, default 0.1).
cor_method	Correlation method: "spearman" (default), "pearson", or "kendall".
adjust_method	P-value adjustment method (default "BH" for Benjamini-Hochberg). Options: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
min_adjust_p	Adjusted p-value threshold for significance (numeric, default 0.05).
min_cor	Minimum correlation coefficient threshold (numeric, default 0). Must be ≥ 0.
min_r2	Minimum R-squared threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
min_fstat	Minimum F-statistic threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
num_cores	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
verbose	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

Two data frames with columns:

ligand, receptor	Ligand and receptor gene symbols (res1/res2).
cor	Correlation coefficient (res1/res2).
p_val	Raw p-value (res1/res2).

`adjust.p` Adjusted p-value (res1/res2).

`sender`, `receiver` Sender and receiver cell types (res1/res2).

`slope` Slope of the linear regression model (res1/res2).

`intercept` Intercept of the linear regression model (res1/res2).

`r2` Coefficient of determination (R-squared) of the linear regression model (res1/res2).

`fstat` F-statistic of the linear regression model (res1/res2).

`sample` Sample identifier (res2).

`score` Projection score (raw co-expression intensity) (res2).

`normalized_score` Normalized score scaled between 0-1 (res2).

Returns `NULL` if:

- No cell types are found in the metadata.
- Insufficient samples or cells remain after filtering.
- No ligand-receptor pairs pass the filtering thresholds.
- One or both of the specified sender and receiver cell types are missing in the data.
- Fewer than two valid samples remain after filtering based on minimum cell number per sample.

Examples

```
data(matrix_object)
data(lr_db)

# Integrated analysis across all cell types
res_all <- one_step_all(
  rna = matrix_object,
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

if (!is.null(res_all)) {
  print(head(res_all$res1))
  print(head(res_all$res2))
}
```

one_step_single *Analyze Ligand-Receptor Pair Correlations and Projection Scores (Specified Sender and Receiver)*

Description

Performs integrated analysis of ligand-receptor (LR) pairs through two consecutive phases: (1) Filters LR pairs and analyzes correlations between specified cell types; (2) Calculates projection scores based on regression models for valid pairs. Returns comprehensive results combining statistical metrics. This function supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names).

Usage

```
one_step_single(
  rna,
  sender,
  receiver,
  lr_database = PopComm::lr_db,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  min_samples = 10,
  min_cell_ratio = 0.1,
  min_sample_ratio = 0.1,
  cor_method = "spearman",
  adjust_method = "BH",
  min_adjust_p = 0.05,
  min_cor = 0,
  min_r2 = 0,
  min_fstat = 0,
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

rna	A Seurat object or a matrix containing single-cell RNA expression data.
sender	Cell type designated as the ligand sender (character).
receiver	Cell type designated as the receptor receiver (character).
lr_database	A data frame of ligand-receptor pairs with columns "ligand_gene_symbol" and "receptor_gene_symbol".
sample_col	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).

cell_type_col	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
id_sep	Separator used in matrix column names to split sample and cell type (e.g., -- for "Cardiac–sample1"). Only used in Matrix mode.
min_cells	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.
min_samples	Minimum number of valid samples to proceed (numeric, default 10).
min_cell_ratio	Minimum ratio of cells expressing ligand and receptor genes in sender or receiver cells (numeric, default 0.1). Only used in Seurat mode.
min_sample_ratio	Minimum ratio of samples in which both the ligand and receptor genes must be expressed (numeric, default 0.1).
cor_method	Correlation method: "spearman" (default), "pearson", or "kendall".
adjust_method	P-value adjustment method (default "BH" for Benjamini-Hochberg). Options: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
min_adjust_p	Adjusted p-value threshold for significance (numeric, default 0.05).
min_cor	Minimum correlation coefficient threshold (numeric, default 0). Must be ≥ 0.
min_r2	Minimum R-squared threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
min_fstat	Minimum F-statistic threshold for the linear regression model (numeric, default 0). Must be ≥ 0.
num_cores	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
verbose	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

Two data frames with columns:

ligand, receptor	Ligand and receptor gene symbols (res1/res2).
cor	Correlation coefficient (res1/res2).
p_val	Raw p-value (res1/res2).
adjust.p	Adjusted p-value (res1/res2).
sender, receiver	Sender and receiver cell types (res1/res2).
slope	Slope of the linear regression model (res1/res2).
intercept	Intercept of the linear regression model (res1/res2).
r2	R-squared of the linear regression model (res1/res2).
fstat	F-statistic of the linear regression model (res1/res2).
sample	Sample identifier (res2).
score	Projection score (raw co-expression intensity) (res2).

normalized_score
: Normalized score scaled between 0-1 (res2).

Returns NULL if:

- No cell types are found in the metadata.
- Insufficient samples or cells remain after filtering.
- No ligand-receptor pairs pass the filtering thresholds.
- One or both of the specified sender and receiver cell types are missing in the data.
- Fewer than two valid samples remain after filtering based on minimum cell number per sample.

Examples

```
data(matrix_object)
data(lr_db)

# Integrated analysis with Perivascular -> Endothelial
res_single <- one_step_single(
  rna = matrix_object,
  sender = "Perivascular",
  receiver = "Endothelial",
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

if (!is.null(res_single)) {
  print(head(res_single$res1))
  print(head(res_single$res2))
}
```

---

pca_sample *Generate PCA of Ligand-Receptor Interaction Scores*

Description

This function performs principal component analysis (PCA) on ligand-receptor (LR) interaction scores across samples, and generates a scatter plot of the first two principal components. Optionally, sample metadata can be used to color the points.

Usage

```
pca_sample(
  lr_scores,
  metadata,
  selected_sender = NULL,
  selected_receiver = NULL,
  color_by = NULL,
  n_components = 2
)
```

Arguments

lr_scores	Data frame containing LR interaction scores per sample (data frame).
metadata	Data frame containing sample metadata (data frame).
selected_sender	Specific sender cell type to filter, default is None (use all) (character).
selected_receiver	Specific receiver cell type to filter, default is None (use all) (character).
color_by	metadata column name to color points in PCA plot (character).
n_components	Number of principal components to extract (numeric, default: 2).

Value

A list containing:

- plot - ggplot object of the PCA scatter plot
- df - data frame used for the PCA results

Examples

```
# PCA of LR Interaction Scores
data(lr_scores_eg)
data(metadata_eg)

res <- pca_sample(
  lr_scores = lr_scores_eg,
  metadata = metadata_eg,
  color_by = "IFN_type"
  )

print(res$plot)
head(res$df)
```

score_lr_all *Analyze Ligand-Receptor Projection Scores (Across All Cell Types)*

Description

This function calculates the ligand-receptor (LR) projection scores between all combinations of sender and receiver cell types, and it supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names). The projection score is computed based on linear regression models, measuring the normalized distance of each sample's LR expression from the origin of the regression line.

Usage

```
score_lr_all(
  rna,
  filtered_lr,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

rna	A Seurat object or a matrix containing single-cell RNA expression data.
filtered_lr	A data frame of ligand-receptor pairs from prior analysis (e.g., output of filter_lr_single). Must contain an "lr" column with pair identifiers in "Ligand_Receptor" format.
sample_col	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).
cell_type_col	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
id_sep	Separator used in matrix column names to split sample and cell type (e.g., -- for "Cardiac–sample1"). Only used in Matrix mode.
min_cells	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.
num_cores	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
verbose	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

A data frame with projection scores per sample and LR pair. Columns:

All input from filtered_lr
: Original columns provided by the user in filtered_lr.

sample
: Sample identifier.

score
: Projection score (raw co-expression intensity).

normalized_score
: Normalized score scaled between 0-1.

Rows are ordered by filtered_lr columns and descending score.

Returns NULL if:

- No cell types are found in the metadata.
- One or both of the specified sender and receiver cell types are missing in the data.
- Fewer than two valid samples remain after filtering based on minimum cell number per sample.

Examples

```
data(matrix_object)
data(lr_db)

# Analyzing ligand-receptor interactions between all cell types
result01a <- filter_lr_all(
  rna = matrix_object,
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

# Analyzing ligand-receptor projection scores between all cell types
result02a <- score_lr_all(
  rna = matrix_object,
  filtered_lr = result01a,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  num_cores = 1,
  verbose = TRUE
  )

if (!is.null(result02a)) {
print(head(result02a))
}
```

score_lr_single *Analyze Ligand-Receptor Projection Scores (Specified Sender and Receiver)*

Description

This function calculates the projection scores for ligand-receptor (LR) pairs between specified sender and receiver cell types, and it supports both Seurat objects and average expression matrices (matrix of gene expression data with cell types and samples as column names). The projection score is computed based on linear regression models, measuring the normalized distance of each sample's LR expression from the origin of the regression line.

Usage

```
score_lr_single(
  rna,
  sender,
  receiver,
  filtered_lr,
  sample_col,
  cell_type_col,
  id_sep,
  min_cells = 50,
  num_cores = 10,
  verbose = TRUE
)
```

Arguments

rna	A Seurat object or a matrix containing single-cell RNA expression data.
sender	Cell type designated as the ligand sender (character).
receiver	Cell type designated as the receptor receiver (character).
filtered_lr	A data frame of filtered ligand-receptor pairs from prior analysis (e.g., output of filter_lr_single). Must contain an "lr" column with pair identifiers in "Ligand_Receptor" format.
sample_col	Metadata column name (character) for sample identifiers in Seurat mode; Matrix mode uses column index (numeric).
cell_type_col	Metadata column name (character) for cell type in Seurat mode; Matrix mode uses column index (numeric).
id_sep	Separator used in matrix column names to split sample and cell type (e.g., -- for "Cardiac–sample1"). Only used in Matrix mode.
min_cells	Minimum number of cells per sample for both sender and receiver (numeric, default 50). Only used in Seurat mode.
num_cores	Number of CPU cores for parallel processing (numeric, default 10). Automatically capped at (system cores - 1).
verbose	Logical indicating whether to print progress messages (logical, default: TRUE).

Value

A data frame with projection scores per sample and LR pair. Columns:

`All input from filtered_lr`
: Original columns provided by the user in `filtered_lr`.

`sample`
: Sample identifier.

`score`
: Projection score (raw co-expression intensity).

`normalized_score`
: Normalized score scaled between 0-1.

Rows are ordered by `filtered_lr` columns and descending `score`.

Returns `NULL` if:

- No cell types are found in the metadata.
- One or both of the specified sender and receiver cell types are missing in the data.
- Fewer than two valid samples remain after filtering based on minimum cell number per sample.

Examples

```
data(matrix_object)
data(lr_db)

# Analyzing ligand-receptor interactions: Perivascular -> Endothelial
result01s <- filter_lr_single(
  rna = matrix_object,
  sender = "Perivascular",
  receiver = "Endothelial",
  lr_database = lr_db,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  min_samples = 10,
  min_sample_ratio = 0.1,
  min_adjust_p = 0.05,
  num_cores = 1,
  verbose = TRUE
  )

# Analyzing ligand-receptor projection scores: Perivascular -> Endothelial
result02s <- score_lr_single(
  rna = matrix_object,
  sender = "Perivascular",
  receiver = "Endothelial",
  filtered_lr = result01s,
  sample_col = 2,
  cell_type_col = 1,
  id_sep = "--",
  num_cores = 1,
  verbose = TRUE
```

```
  )

if (!is.null(result02s)) {
print(head(result02s))
}
```

Index