Package: scPharm (via r-universe)

April 25, 2026

Title Identification of Pharmacological Subpopulations of Single Cells for Precision Medicine in Cancers

Version 1.0.6

Description A computational framework for single-cell RNA-seq data that integrates pharmacogenomics profiles to uncover therapeutic heterogeneity within tumors at single-cell resolution. The tool prioritizes tailored drugs and provides insights into combination therapy regimens and drug toxicity in cancers.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

LazyData true

LazyDataCompression xz

URL https://github.com/Zaoqu-Liu/scPharm, https://zaoqu-liu.github.io/scPharm/

BugReports https://github.com/Zaoqu-Liu/scPharm/issues

Depends R (>= 4.1), Seurat (>= 4.0.1)

Imports CelliD, SeuratObject, fgsea (>= 1.16.0), Rcpp, mixtools (>= 2.0.0), dplyr (>= 1.1.2), sparseMatrixStats (>= 1.2.1), irlba (>= 2.3.5.1), tidyr (>= 1.3.0), methods, parallel, stats, utils, dlm, MCMCpack, parallelDist, cluster, matrixStats

LinkingTo Rcpp, RcppArmadillo

Suggests knitr, rmarkdown, ggplot2, patchwork, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/pak/sysreqs libcairo2-dev cmake libfontconfig1-dev libfreetype6-dev libfribidi-dev libglpk-dev make libharfbuzz-dev libicu-dev libjpeg-dev libpng-dev libtiff-dev libuv1-dev libwebp-dev libxml2-dev libssl-dev python3 zlib1g-dev

Repository https://zaoqu-liu.r-universe.dev

Date/Publication 2026-01-25 06:49:56 UTC

RemoteUrl https://github.com/Zaoqu-Liu/scPharm

RemoteRef main

RemoteSha 33a25f64d3748973d513bcdadb2c32a18c048064

Contents

`bulkdata` *Bulk RNA-seq Expression Data for Cancer Cell Lines*

Description

TPM-normalized gene expression profiles for tumor cell lines from the Cell Model Passports database.

Usage

```
bulkdata
```

Format

A data frame with 37,004 genes (rows) and 1,387 cell lines (columns). Row names are gene symbols; column names are cell line identifiers.

Source

Cell Model Passports https://cellmodelpassports.sanger.ac.uk/downloads

References

van der Meer D, et al. (2019). Cell Model Passports - a curated and standardised dataset of preclinical cancer models. Nucleic Acids Research.

drug_info *GDSC2 Drug Information*

Description

Drug metadata including targets and signaling pathways from the GDSC2 project.

Usage

```
drug_info
```

Format

A data frame with 295 drugs and 4 variables:

DRUG_ID GDSC drug identifier

DRUG_NAME Drug name

PUTATIVE_TARGET Known drug target(s)

PATHWAY_NAME Target signaling pathway

Source

GDSC https://www.cancerrxgene.org/downloads/drug_data

gdscdata *GDSC2 Pharmacogenomics Data*

Description

Drug sensitivity data (IC50 and AUC) for cancer cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) project.

Usage

```
gdscdata
```

Format

A data frame with 196,344 observations and 19 variables including:

DRUG_ID GDSC drug identifier

DRUG_NAME Drug name

CELL_LINE_NAME Cell line identifier

COSMIC_ID COSMIC cell line ID

SANGER_MODEL_ID Sanger model ID
TCGA_DESC TCGA cancer type classification
DATASET Dataset source
COMPANY_ID Company identifier
NLME_RESULT_ID NLME result ID
NLME_CURVE_ID NLME curve ID
LN_IC50 Natural log of IC50
AUC Area under the dose-response curve
RMSE Root mean square error
Z_SCORE Z-score
MAX_CONC Maximum concentration tested
MIN_CONC Minimum concentration tested
PUTATIVE_TARGET Known drug target
PATHWAY_NAME Target pathway
WEBRELEASE Web release version

Source

GDSC https://www.cancerrxgene.org/downloads/drug_data

References

Yang W, et al. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Research 41, D955-D961.

---

scPharmCombo *Identify Potential Drug Combinations*

Description

Identify potential drug combinations based on two strategies: (1) Compensation effects: drugs that target cells resistant to the primary drug (2) Booster effects: drugs that enhance sensitivity through different pathways

Usage

```
scPharmCombo(object, score, drug = NULL, topN = 1, drug_info = NULL)
```

Arguments

`object`	A Seurat object after running `scPharmIdentify`.
`score`	Output from `scPharmDr`.
`drug`	Name of the primary drug. If `NULL`, uses `topN` drugs.
`topN`	Number of top-ranked drugs to analyze. Default: 1.
`drug_info`	Drug information table. If `NULL`, uses built-in `scPharm::drug_info`.

Details

Compensation effects: Identifies drugs where cells resistant to the primary drug show sensitivity. This suggests the combination could overcome resistance.

Booster effects: Identifies drugs targeting different pathways that show high sensitivity in cells already sensitive to the primary drug. This suggests synergistic enhancement.

Value

A named list where each element corresponds to a primary drug and contains a data frame with:

DRUG_FIRST Primary drug name

DRUG_ID Combination drug ID

DRUG_NAME Combination drug name

Effect Combination effect score

Strategy "compensation effects" or "booster effects"

See Also

scPharmIdentify, scPharmDr

Examples

```
## Not run:
dr_scores <- scPharmDr(result)
combos <- scPharmCombo(result, dr_scores, topN = 3)

## End(Not run)
```

---

scPharmDr *Compute Drug Prioritization Score (Dr)*

Description

Calculate drug prioritization scores based on the ratio of sensitive and resistant tumor cell populations identified by scPharmIdentify.

Usage

```
scPharmDr(object)
```

Arguments

object	A Seurat object after running scPharmIdentify.

Details

The Dr score is calculated as:

$$Dr = S \times (1 - R)$$

where S is the proportion of sensitive cells and R is the proportion of resistant cells among tumor cells. Higher scores indicate better drug candidates for the patient.

Value

A data frame with the following columns:

DRUG_ID GDSC drug identifier

DRUG_NAME Drug name

SENSI_RATIO Proportion of sensitive tumor cells

RESIS_RATIO Proportion of resistant tumor cells

Dr Drug prioritization score

Rank Drug ranking (1 = best)

See Also

scPharmIdentify, scPharmCombo

Examples

```
## Not run:
# After running scPharmIdentify
dr_scores <- scPharmDr(result)
head(dr_scores)

## End(Not run)
```

---

scPharmDse *Predict Drug Side Effects (Dse)*

Description

Calculate drug side effect scores based on the sensitivity of adjacent (normal) cells to drugs. Higher scores indicate greater potential for off-target toxicity.

Usage

```
scPharmDse(object)
```

Arguments

object	A Seurat object after running scPharmIdentify with type="tissue".

Details

The Dse score represents the proportion of adjacent (non-tumor) cells that are classified as sensitive to each drug. Drugs with high Dse scores may cause more side effects by affecting normal cells.

Value

A data frame with the following columns:

DRUG_ID GDSC drug identifier

DRUG_NAME Drug name

Dse Side effect score (0-1, higher = more side effects)

See Also

scPharmIdentify, scPharmDr

Examples

```
## Not run:
# After running scPharmIdentify with type="tissue"
dse_scores <- scPharmDse(result)
head(dse_scores)

## End(Not run)
```

---

scPharmGenNullDist *Generate Null Distribution and Thresholds*

Description

Generate a null distribution from healthy tissue cells and calculate thresholds for classifying sensitive and resistant cells.

Usage

```
scPharmGenNullDist(
  object,
  cancer,
  nmcs = 50,
  nfeatures = 200,
  cores = 1,
  features = NULL,
  slot = "data",
  layer = NULL,
  assay = "RNA",
  bulkdata = NULL,
  gdscdata = NULL
)
```

Arguments

object	A Seurat object containing cells from healthy/normal tissue.
cancer	TCGA cancer type(s) for context. A character string or vector. Use "pan" for pan-cancer analysis.
nmcs	Number of MCA components. Default: 50.
nfeatures	Number of genes for cell identity signature. Default: 200.
cores	Number of CPU cores. Default: 1.
features	Character vector of gene names to use. If NULL, uses all.
slot	Slot for Seurat V4. Default: "data".
layer	Layer for Seurat V5. If NULL, uses slot value.
assay	Assay to use. Default: "RNA".
bulkdata	Bulk RNA-seq data. If NULL, uses built-in data.
gdscdata	GDSC data. If NULL, uses built-in data.

Details

This function computes NES distributions from normal cells and uses a two-component Gaussian mixture model to determine thresholds. The thresholds are calculated as mean +/- 1 standard deviation of each component.

Value

A list containing:

NullDist Numeric vector of NES values from normal cells

threshold_s Threshold for sensitive cells (NES < threshold_s)

threshold_r Threshold for resistant cells (NES > threshold_r)

See Also

scPharmIdentify

Examples

```
## Not run:
# Using healthy tissue cells
thresholds <- scPharmGenNullDist(healthy_seurat, cancer = "BRCA")
print(thresholds$threshold_s)
print(thresholds$threshold_r)

## End(Not run)
```

scPharmIdentify *Identify Pharmacological Cell Subpopulations*

Description

Classify single cells into drug-sensitive, drug-resistant, or other subpopulations based on pharmacogenomics profiles from the GDSC2 database.

Usage

```
scPharmIdentify(
  object,
  type,
  cancer,
  drug = NULL,
  nmcs = 50,
  nfeatures = 200,
  cores = 1,
  features = NULL,
  slot = "data",
  layer = NULL,
  assay = "RNA",
  threshold.s = -1.751302,
  threshold.r = 1.518551,
  tumor.cells = NULL,
  normal.cells = NULL,
  bulkdata = NULL,
  gdscdata = NULL
)
```

Arguments

object	A Seurat object containing single-cell RNA-seq data.
type	Data source type. Either "tissue" for tumor tissue samples or "cellline" for cell line samples. When "tissue", the function identifies tumor vs adjacent normal cells using CNV analysis.
cancer	TCGA cancer type(s). A character string or vector specifying cancer type(s) (e.g., "BRCA", c("LUAD", "LUSC")). Use "pan" for pan-cancer analysis.
drug	Drug name to analyze. If NULL (default), all drugs from GDSC2 project will be analyzed.
nmcs	Number of MCA components to compute. Default: 50.
nfeatures	Number of genes for cell identity signature. Default: 200.
cores	Number of CPU cores for parallel processing. Default: 1.
features	Character vector of gene names to use. If NULL, all features are used.
slot	Slot name for Seurat V4 data access. Default: "data".

layer	Layer name for Seurat V5 data access. If NULL, uses slot value.
assay	Assay to use. Default: "RNA".
threshold.s	Threshold for labeling sensitive cells. Cells with NES below this value are classified as sensitive. Default: -1.751302.
threshold.r	Threshold for labeling resistant cells. Cells with NES above this value are classified as resistant. Default: 1.518551.
tumor.cells	Character vector of known tumor cell barcodes. If provided when type="tissue", CNV-based detection is skipped.
normal.cells	Character vector of known normal cell barcodes. Used as reference for CNV analysis when provided.
bulkdata	Bulk RNA-seq data for cell lines. If NULL, uses built-in scPharm::bulkdata.
gdscdata	GDSC pharmacogenomics data. If NULL, uses built-in scPharm::gdscdata.

Details

The function performs the following steps:

1. For tissue samples: identifies tumor cells via CNV analysis (or uses provided annotations)
2. Computes Multiple Correspondence Analysis (MCA) for dimensionality reduction
3. Generates cell identity gene signatures
4. Correlates gene expression with drug response (AUC) across cell lines
5. Performs GSEA to score each cell's drug response profile
6. Classifies cells based on NES thresholds

Value

A Seurat object with pharmacological annotations added to metadata:

cell.label Cell type: "tumor" or "adjacent"

scPharm_label_DRUGID_DRUGNAME Drug response: "sensitive", "resistant", or "other"

scPharm_nes_DRUGID_DRUGNAME Normalized enrichment score for the drug

References

Tian P, Zheng J, et al. scPharm: identifying pharmacological subpopulations of single cells for precision medicine in cancers. 2023.

Examples

```
## Not run:
# Basic usage
result <- scPharmIdentify(seurat_obj, type = "tissue", cancer = "LUAD")

# With known tumor cells
result <- scPharmIdentify(seurat_obj, type = "tissue", cancer = "LUAD",
                          tumor.cells = tumor_barcodes)
```

```
# Specific drug analysis
result <- scPharmIdentify(seurat_obj, type = "cellline", cancer = "BRCA",
                          drug = "Erlotinib")

## End(Not run)
```

Index