Package: multipletR (via r-universe)

November 25, 2026

Type Package

Title Adaptive Detection of Human-Mouse Multiplets in PDX Single-Cell Data

Version 1.0.0

Description Detects human-mouse multiplets in patient-derived xenograft (PDX) single-cell RNA-seq data using an adaptive threshold method that does not assume a fixed species proportion. Takes a 10x CellRanger GEM classification file and returns the data with added multiplet classifications, with optional diagnostic plots and helpers to annotate a Seurat or SingleCellExperiment object with multiplet classifications or remove multiplets from it. The adaptive approach starts from a conservative central region of balanced droplets and expands its thresholds until the selected cells no longer resemble true multiplets, which handles the imbalanced species mixtures common in real PDX samples.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Imports diptest, ggplot2, methods, patchwork, scales, SingleCellExperiment, stats, utils

Suggests BiocStyle, Seurat, SummarizedExperiment, knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

biocViews SingleCell, RNASeq, Preprocessing, QualityControl, Transcriptomics, Classification

URL https://github.com/Alex05a/multipletR

BugReports https://github.com/Alex05a/multipletR/issues

Config/testthat/edition 3

Config/pak/sysreqs zlib1g-dev

Repository Bioconductor 3.24

Date/Publication 2026-11-23 22:56:59 UTC

RemoteUrl https://github.com/bioc/multipletR

RemoteRef RELEASE_3_24

RemoteSha bf170b7cb5da7f9398894dee32b7627e3daf334a

Contents

detect_multiplets 2
remove_multiplets 4

Index **6**

`detect_multiplets` *Detect human-mouse multiplets in PDX single-cell data*

Description

Reads a 10x CellRanger GEM classification file, applies the adaptive threshold method to identify human-mouse multiplets without assuming a fixed species proportion, and writes the input data back out with added classification columns. Optionally produces diagnostic plots.

Usage

```
detect_multiplets(
  fileIn,
  fileOut,
  T1 = 70,
  T2 = 30,
  T3 = 25,
  plotPercent = TRUE,
  plotTotalReads = FALSE,
  overlapDrop = 10,
  modeDiff = 0.9,
  verbose = FALSE
)
```

Arguments

`fileIn`	Path to the input GEM classification CSV (from 10x CellRanger). Expected columns: a barcode column, a human read-count column (GRCh38), a mouse read-count column (e.g. GRCm39 / mm10 / mm39 / mouse_reads), and a 10x call column.
`fileOut`	Path to write the output CSV. The output is the input data with added columns: our_classification (Multiplet / Singlet), pct_human, and pct_mouse.

T1	Upper percent-mouse threshold: the starting upper bound on percent mouse content. Cells above this are too mouse-heavy. Default 70.
T2	Lower percent-mouse threshold: the starting lower bound on percent mouse content (equivalently, the percent-human side). Cells below this are too human-heavy. Default 30.
T3	Total-reads threshold: the starting lower bound on total reads, given as a percentile of the total-reads range (e.g. 25 = 25th percentile). Default 25.
plotPercent	Logical, whether to draw the two percent plots (one for the original 10x classification, one for our classification), total reads vs percent mouse. Default TRUE.
plotTotalReads	Logical, whether to draw the total-reads plots (mouse reads vs human reads, colored by 10x and our classification). Default FALSE.
overlapDrop	Stop expanding a threshold when the overlap between the human and mouse read distributions drops by more than this percentage from its running peak. Lower values are more conservative. Default 10.
modeDiff	Stop expanding when the difference between the human and mouse distribution modes increases by more than this amount from the previous step. Default 0.9.
verbose	Logical, print progress messages including the final thresholds, multiplet count, and the step-by-step algorithm trace. Set FALSE to run silently. Default FALSE.

Details

The three starting thresholds T1, T2 and T3 follow the notation used in the manuscript. They define the conservative starting region that is then expanded adaptively; the defaults are the values we recommend and they rarely need to be changed.

Value

Invisibly, a data frame: the input data with added columns our_classification, pct_human, and pct_mouse. Also written to `fileOut`.

Examples

```
# Detect multiplets in the bundled example PDX dataset (sample PC65)
gem_file <- system.file("extdata", "PC65_gem_classification.csv",
  package = "multipletR"
)
res <- detect_multiplets(gem_file, tempfile(fileext = ".csv"),
  plotPercent = FALSE, plotTotalReads = FALSE
)
table(res$our_classification)
```

`remove_multiplets` *Annotate (and optionally remove) multiplets in a Seurat or SingleCellExperiment object*

Description

Adds multipletR's per-cell classification to a single-cell object's cell metadata. By default all cells are kept and only annotated (useful for visualizing where the multiplets fall, e.g. on a UMAP colored by `multipletR_class` before deciding whether to filter). Set `remove = TRUE` to drop the detected multiplets so the object is ready for downstream analysis. The object type (Seurat or `SingleCellExperiment`) is detected automatically.

Usage

```
remove_multiplets(
  x,
  multiplets,
  remove = FALSE,
  barcode_col = "barcode",
  verbose = TRUE
)
```

Arguments

`x`	A single-cell object whose cell names (`colnames`) are barcodes: either a Seurat object or a `SingleCellExperiment`. The object type is detected automatically.
`multiplets`	The data frame returned by `detect_multiplets` (must contain `barcode`, `our_classification`, `pct_human`, and `pct_mouse`).
`remove`	Logical. If `TRUE`, the detected multiplets are removed from the returned object. If `FALSE` (the default), all cells are kept and only the metadata is added.
`barcode_col`	Name of the barcode column in `multiplets`. Default "barcode".
`verbose`	Logical. If `TRUE` (the default), print a summary of the annotation and removal. Set `FALSE` to suppress these messages.

Details

The classification follows the same idea as doublet-detection tools such as DoubletFinder: each cell is labeled Human, Mouse, or Multiplet, and the percent human / percent mouse are added per cell.

Value

The input object with three added cell-metadata columns (`multipletR_class`, `multipletR_pct_human`, `multipletR_pct_mouse`). If `remove = TRUE`, the multiplet cells are also dropped.

Examples

```
# Detect multiplets in the bundled example dataset
gem_file <- system.file("extdata", "PC65_gem_classification.csv",
  package = "multipletR"
)
res <- detect_multiplets(gem_file, tempfile(fileext = ".csv"),
  plotPercent = FALSE, plotTotalReads = FALSE
)

# Build a minimal Seurat object whose cell names are the same barcodes
counts <- matrix(rpois(20 * nrow(res), 5),
  nrow = 20,
  dimnames = list(paste0("gene", 1:20), res$barcode)
)
seu <- Seurat::CreateSeuratObject(counts)

# Annotate only (keep all cells), e.g. to visualize before filtering
seu <- remove_multiplets(seu, res)
table(seu$multipletR_class)

# Or annotate and remove the multiplets in one step
seu_clean <- remove_multiplets(seu, res, remove = TRUE)
```

Index

detect_multiplets, 2, *4*

remove_multiplets, 4