Package: BatchSVG (via r-universe)

November 13, 2026

Type Package

Title Identify Batch-Biased Spatially Variable Genes

Version 1.4.0

Description BatchSVG is a method to identify batch-biased spatially variable genes (SVGs) in spatial transcriptomics data. The batch variable can be defined as sample, donor sex, or other batch effects of interest. The BatchSVG method is based on the binomial deviance model (Townes et al, 2019).

License Artistic-2.0

Imports scry, dplyr, stats, rlang, cowplot, ggrepel, ggplot2, RColorBrewer, scales, SummarizedExperiment

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 4.5.0)

Suggests testthat (>= 3.0.0), knitr, rmarkdown, BiocStyle, spatialLIBD

VignetteBuilder knitr

biocViews Spatial, Transcriptomics, BatchEffect, QualityControl

URL https://github.com/christinehou11/BatchSVG, https://christinehou11.github.io/BatchSVG

BugReports https://github.com/christinehou11/BatchSVG/issues

Config/testthat/edition 3

LazyData false

Config/pak/sysreqs zlib1g-dev

Repository Bioconductor 3.23

Date/Publication 2026-04-28 13:04:54 UTC

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

RemoteRef RELEASE_3_23

RemoteSha 402c3bd7e37792e0763d951076f063789d3ee88b

Contents

biasDetect 2
featureSelect 3
list_batch_df 5
svg_nSD 5

Index **7**

`biasDetect` *Biased Genes Identification*

Description

Function to identify the bias genes based on user-selected threshold of number of standard deviation in relative change in deviance and rank difference.

Usage

```
biasDetect(
  list_batch_df,
  threshold = c("both", "dev", "rank"),
  nSD_dev = NULL,
  nSD_rank = NULL,
  plot_point_size = 3,
  plot_point_shape = 16,
  plot_text_size = 3,
  plot_palette = "YlOrRd"
)
```

Arguments

`list_batch_df`	`list` : The list of data frame(s) generated from `featureSelection()` function. The length of the data frame list should be at least one.
`threshold`	A character string specifying the filtering criterion. Must be one of: • `"dev"`: Filters genes based on the deviance threshold only. • `"rank"`: Filters genes based on the rank threshold only. • `"both"`: Filters genes based on either the deviance or rank threshold. Default is "both".
`nSD_dev`	`integer`: A numeric vector specifying the number of standard deviation (nSD) for each batch when analyzing the relative change in deviance. The order of values must correspond to the order of batches in `list_batch_df`. Required if `threshold` is "dev" or "both". If a single value is provided, it is applied to all batches; otherwise, it must have the same length as `list_batch_df`.

nSD_rank vector: A numeric vector specifying the number of standard deviation (nSD) for each batch when analyzing rank differences. The order of values must correspond to the order of batches in list_batch_df. Required if threshold is "rank" or "both". If a single value is provided, it is applied to all batches; otherwise, it must have the same length as list_batch_df.

plot_point_size vector: A numeric vector specifying point sizes in plots. If asingle value is provided, it is applied to all batches.

plot_point_shape vector: A numeric vector specifying point shapes in plots. If a single value is provided, it is applied to all batches.

plot_text_size vector: A numeric vector specifying text label size in plots. Default is 3.

plot_palette vector: A character string vector specifying the color palette for plots. Default is "YlOrRd".

Value

A named list where each element corresponds to a batch and contains:

- "Plot": A diagnostic plot (either deviance, rank, or both).
- "Table": A filtered data frame containing outlier genes based on the specified threshold.

Examples

```
# use the result generated from featureSelect()
data(list_batch_df)
biaGenes <- biasDetect(list_batch_df = list_batch_df, threshold = "both",
   nSD_dev = 3, nSD_rank = 3)
```

---

featureSelect *Feature selection*

Description

The function computes batch-adjusted deviance values, ranks the genes accordingly, and quantifies batch effects in terms of standard deviations from the mean difference. The list follows the order of batch effects provided in batch_effects.

Usage

```
featureSelect(input, batch_effects = NULL, VGs = NULL, verbose = TRUE)
```

Arguments

input	A SpatialExperiment object containing spatial transcriptomics data..
batch_effects	A character vector specifying column names in colData(input) that indicate batch effects. Must match existing column names.
VGs	A character vector specifying the variable genes (VGs) to be analyzed. Only genes present in this vector will be retained for feature selection.
verbose	Logical (TRUE or FALSE). Default is TRUE. If TRUE, progress messages will be printed; If FALSE, messages will be suppressed.

Value

A named list where each element corresponds to a batch effect. Each batch contains a data frame with the following columns:

- **"gene_id"**: Gene identifier.
- **"gene_name"**: Gene name.
- **"dev_default"**: Deviance score without batch correction.
- **"dev_"**: Deviance score with batch correction.
- **"rank_default"**: Rank of the gene based on deviance without batch correction.
- **"rank_"**: Rank of the gene based on deviance with batch correction.
- **"d_diff"**: Relative change in deviance between default and batch-corrected models.
- **"nSD_dev_"**: number of standard deviation of relative change in deviance for the batch.
- **"r_diff"**: Rank difference between default and batch-corrected models.
- **"nSD_rank_"**: number of standard deviation of rank difference for the batch.

Examples

```
library(spatialLIBD)
spatialLIBD_spe <- fetch_data(type = "spe")
libd_svg <- read.csv(
    system.file("extdata","libd-all_nnSVG_p-05-features-df.csv",
              package = "BatchSVG"),
    row.names = 1, check.names = FALSE)

list_batch_df <- featureSelect(input = spatialLIBD_spe,
   batch_effects = "subject", VGs = libd_svg$gene_id)
```

list_batch_df *List of data frames of nSD(s) on the batch of subject*

Description

The `list_batch_df` dataset contains results from `featureSelect()` applied to the spatial transcriptomics data from the `spatialLIBD` package.

Usage

```
data(list_batch_df)
```

Format

A named list of data frames, where each element corresponds to a batch effect:

- **"gene_id"**: Gene identifier.
- **"gene_name"**: Gene name.
- **"dev_default"**: Deviance score without batch correction.
- **"dev_(batch name)"**: Deviance score with batch correction.
- **"rank_default"**: Rank of the gene based on deviance without batch correction.
- **"rank_(batch name)"**: Rank of the gene based on deviance with batch correction.
- **"d_diff"**: Relative change in deviance between default and batch-corrected models.
- **"nSD_dev_(batch name)"**: number of standard deviation of relative change in deviance for the batch.
- **"r_diff"**: Rank difference between default and batch-corrected models.
- **"nSD_rank_(batch name)"**: number of standard deviation of rank difference for the batch.

Source

https://github.com/christinehou11/BatchSVG/blob/main/inst/scripts/make-list_batch_df.R

svg_nSD *SVGs Plots*

Description

This function generates visualizations to assess the impact of batch effects on spatially variable genes (SVGs) by analyzing changes in deviance and rank. The function bins the deviations into normalized standard deviation (nSD) intervals and creates histograms and scatter plots to illustrate the distribution of batch effects.

Usage

```
svg_nSD(list_batch_df, sd_interval_dev, sd_interval_rank)
```

Arguments

list_batch_df	A named list of data frames, where each data frame corresponds to a batch effect and contains columns with deviance and rank differences.
sd_interval_dev	A numeric vector specifying the interval widths for standard deviation bins for each batch when analyzing the relative change in deviance. The order of values must correspond to the order of batches in list_batch_df . If a single value is provided, it is applied to all batches; otherwise, it must have the same length as list_batch_df.
sd_interval_rank	vector: A numeric vector specifying the interval widths for standard deviation bins when analyzing rank differences. The order of values must correspond to the order of batches in list_batch_df. If a single value is provided, it is applied to all batches; otherwise, it must have the same length as list_batch_df.

Value

A combined ggplot object containing:

- **Deviance Plots**:
  - Histogram of deviance differences across SVGs, colored by nSD intervals.
  - Scatter plot comparing deviance values before and after batch correction.
- **Rank Plots**:
  - Histogram of rank differences across SVGs, colored by nSD intervals.
  - Scatter plot comparing ranks before and after batch correction.

The function arranges plots for each batch in a grid format for easy comparison.

Examples

```
# use the result generated from featureSelect()
data(list_batch_df)
plots <- svg_nSD(list_batch_df = list_batch_df,
   sd_interval_dev = 3, sd_interval_rank = 3)
```

Index

∗ **datasets**
list_batch_df, 5

biasDetect, 2

featureSelect, 3

list_batch_df, 5

svg_nSD, 5