

# Package ‘lefser’

October 10, 2021

**Type** Package

**Title** R implementation of the LefSE method for microbiome biomarker discovery

**Description** lefser is an implementation in R of the popular ``LDA Effect Size (LEfSe)'' method for microbiome biomarker discovery. It uses the Kruskal-Wallis test, Wilcoxon-Rank Sum test, and Linear Discriminant Analysis to find biomarkers of groups and sub-groups.

**Version** 1.2.0

**License** Artistic-2.0

**LazyData** true

**Depends** SummarizedExperiment, R (>= 4.0.0)

**Imports** coin, MASS, ggplot2, stats, methods

**Suggests** knitr, rmarkdown, curatedMetagenomicData, BiocStyle, testthat, pkgdown, covr, withr

**Encoding** UTF-8

**BugReports** <https://github.com/waldronlab/lefser/issues>

**URL** <https://github.com/waldronlab/lefser>

**VignetteBuilder** knitr

**biocViews** Software, Sequencing, DifferentialExpression, Microbiome, StatisticalMethod, Classification

**RoxygenNote** 7.1.1

**git\_url** <https://git.bioconductor.org/packages/lefser>

**git\_branch** RELEASE\_3\_13

**git\_last\_commit** 6976537

**git\_last\_commit\_date** 2021-05-19

**Date/Publication** 2021-10-10

**Author** Asya Khleborodova [cre, aut],  
Ludwig Geistlinger [ctb],  
Marcel Ramos [ctb] (<<https://orcid.org/0000-0002-3242-0582>>),  
Levi Waldron [ctb]

**Maintainer** Asya Khleborodova <[asya.bioconductor@gmail.com](mailto:asya.bioconductor@gmail.com)>

R topics documented:

|              |          |
|--------------|----------|
| lefser       | 2        |
| lefserPlot   | 3        |
| zeller14     | 4        |
| <b>Index</b> | <b>5</b> |

---

|        |                                             |
|--------|---------------------------------------------|
| lefser | <i>R implementation of the LEfSe method</i> |
|--------|---------------------------------------------|

---

Description

Perform a LEfSe analysis: the function carries out differential analysis between two sample groups for multiple microorganisms and uses linear discriminant analysis to establish their effect sizes. Subclass information for each class can be incorporated into the analysis (see examples). Microorganisms with large differences between two sample groups are identified as biomarkers.

Usage

```
lefser(  
  expr,  
  kruskal.threshold = 0.05,  
  wilcox.threshold = 0.05,  
  lda.threshold = 2,  
  groupCol = "GROUP",  
  blockCol = NULL,  
  assay = 1L,  
  trim.names = FALSE  
)
```

Arguments

|                   |                                                                                                                                                            |
|-------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| expr              | A <a href="#">SummarizedExperiment</a> with expression data.                                                                                               |
| kruskal.threshold | numeric(1) The p-value for the Kruskal-Wallis Rank Sum Test (default 0.05).                                                                                |
| wilcox.threshold  | numeric(1) The p-value for the Wilcoxon Rank-Sum Test when 'blockCol' is present (default 0.05).                                                           |
| lda.threshold     | numeric(1) The effect size threshold (default 2.0).                                                                                                        |
| groupCol          | character(1) Column name in 'colData(expr)' indicating groups, usually a factor with two levels (e.g., 'c("cases", "controls")'; default "GROUP").         |
| blockCol          | character(1) Optional column name in 'colData(expr)' indicating the blocks, usually a factor with two levels (e.g., 'c("adult", "senior")'; default NULL). |
| assay             | The i-th assay matrix in the 'SummarizedExperiment' ('expr'; default 1).                                                                                   |
| trim.names        | If 'TRUE' extracts the most specific taxonomic rank of organism.                                                                                           |

**Value**

The function returns a dataframe with two columns, which are names of microorganisms and their LDA scores.

**Examples**

```
# (1) Using classes only
data(zeller14)
# exclude 'adenoma'
zeller14 <- zeller14[, zeller14$study_condition != "adenoma"]
res_group <- lefser(zeller14, groupCol = "study_condition")
head(res_group)

# (2) Using classes and subclasses
data(zeller14)
# exclude 'adenoma'
zeller14 <- zeller14[, zeller14$study_condition != "adenoma"]
res_block <- lefser(
  zeller14, groupCol = "study_condition", blockCol = "age_category"
)
head(res_block)
```

---

lefserPlot

---

*Plots results from 'lefser' function*


---

**Description**

'lefserPlot' function displays effect sizes for differentially expressed microorganisms and whether they are more abundant in '0' or '1' sample group.

**Usage**

```
lefserPlot(df, colors = c("red", "forestgreen"), trim.names = TRUE)
```

**Arguments**

|            |                                                                                                                  |
|------------|------------------------------------------------------------------------------------------------------------------|
| df         | Data frame produced by 'lefser'.                                                                                 |
| colors     | character(2) The two colors corresponding to class 0 and 1, respectively. Defaults to 'c("red", "forestgreen")'. |
| trim.names | If 'TRUE' extracts the most specific taxonomic rank of organism.                                                 |

**Value**

Function returns plot of effect size scores produced by 'lefser'. Positive scores represent microorganisms with that are more abundant in class '1'. Negative scores represent microorganisms with that are more abundant in class '0'.

**Examples**

```
example("lefser")  
lefserPlot(res_group)
```

---

zeller14

*Example dataset for lefser*

---

**Description**

The ZellerG\_2014 dataset contains microbiome count data for CRC patients and controls. It was for curatedMetagenomicData using the script in the package directory "data-raw".

**Usage**

```
zeller14
```

**Format**

A SummarizedExperiment with 1585 features, 199 samples

**study\_condition** adenoma, control, CRC

**age\_category** adult, senoir

**Source**

<https://pubmed.ncbi.nlm.nih.gov/25432777/>

# Index

## \* **datasets**

zeller14, [4](#)

lefser, [2](#)

lefserPlot, [3](#)

SummarizedExperiment, [2](#)

zeller14, [4](#)