

# Package ‘Maaslin2’

July 23, 2025

**Title** ``Multivariable Association Discovery in Population-scale  
Meta-omics Studies"

**Year** 2021

**Version** 1.22.0

**Depends** R ( $\geq$  3.6)

**Description** MaAsLin2 is comprehensive R package for efficiently determining multivariable association between clinical metadata and microbial meta'omic features. MaAsLin2 relies on general linear models to accommodate most modern epidemiological study designs, including cross-sectional and longitudinal, and offers a variety of data exploration, normalization, and transformation methods. MaAsLin2 is the next generation of MaAsLin.

**License** MIT + file LICENSE

**LazyData** false

**Imports** robustbase, biglm, pcaPP, edgeR, metagenomeSeq, pbapply, car,  
dplyr, vegan, chemometrics, ggplot2, pheatmap, logging,  
data.table, lmerTest, hash, optparse, grDevices, stats, utils,  
glmmTMB, MASS, cplm, pscl, lme4, tibble

**Suggests** knitr, testthat ( $\geq$  2.1.0), rmarkdown, markdown

**VignetteBuilder** knitr

**Collate** fit.R utility\_scripts.R viz.R Maaslin2.R

**URL** <http://huttenhower.sph.harvard.edu/maaslin2>

**biocViews** Metagenomics, Software, Microbiome, Normalization

**BugReports** <https://github.com/biobakery/maaslin2/issues>

**git\_url** <https://git.bioconductor.org/packages/Maaslin2>

**git\_branch** RELEASE\_3\_21

**git\_last\_commit** fcec85e

**git\_last\_commit\_date** 2025-04-15

**Repository** Bioconductor 3.21

**Date/Publication** 2025-07-23

**Author** Himel Mallick [aut],  
 Ali Rahnavard [aut],  
 Lauren McIver [aut, cre]

**Maintainer** Lauren McIver <lauren.j.mciver@gmail.com>

## Contents

|              |          |
|--------------|----------|
| Maaslin2     | 2        |
| <b>Index</b> | <b>5</b> |

---

|          |                                                                                                                                                                                                          |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Maaslin2 | <i>MaAsLin2 is the next generation of MaAsLin, a multivariable statistical framework for finding associations between clinical metadata and potentially high-dimensional microbial multi-omics data.</i> |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

---

## Description

MaAsLin2 finds associations between microbiome meta-omics features and complex metadata in population-scale epidemiological studies. The software includes multiple analysis methods (including support for multiple covariates and repeated measures), filtering, normalization, and transform options to customize analysis for your specific study.

## Usage

```
Maaslin2(
  input_data,
  input_metadata,
  output,
  min_abundance = 0.0,
  min_prevalence = 0.1,
  min_variance = 0.0,
  normalization = "TSS",
  transform = "LOG",
  analysis_method = "LM",
  max_significance = 0.25,
  random_effects = NULL,
  fixed_effects = NULL,
  correction = "BH",
  standardize = TRUE,
  cores = 1,
  plot_heatmap = TRUE,
  heatmap_first_n = 50,
  plot_scatter = TRUE,
  max_pngs = 10,
  save_scatter = FALSE,
  save_models = FALSE,
```

```

    reference = NULL
  )

```

## Arguments

|                               |                                                                                                                                                                     |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>input_data</code>       | The tab-delimited input file of features.                                                                                                                           |
| <code>input_metadata</code>   | The tab-delimited input file of metadata.                                                                                                                           |
| <code>output</code>           | The output folder to write results.                                                                                                                                 |
| <code>min_abundance</code>    | The minimum abundance for each feature.                                                                                                                             |
| <code>min_prevalence</code>   | The minimum percent of samples for which a feature is detected at minimum abundance.                                                                                |
| <code>min_variance</code>     | Keep features with variance greater than.                                                                                                                           |
| <code>max_significance</code> | The q-value threshold for significance.                                                                                                                             |
| <code>normalization</code>    | The normalization method to apply.                                                                                                                                  |
| <code>transform</code>        | The transform to apply.                                                                                                                                             |
| <code>analysis_method</code>  | The analysis method to apply.                                                                                                                                       |
| <code>random_effects</code>   | The random effects for the model, comma-delimited for multiple effects.                                                                                             |
| <code>fixed_effects</code>    | The fixed effects for the model, comma-delimited for multiple effects.                                                                                              |
| <code>correction</code>       | The correction method for computing the q-value.                                                                                                                    |
| <code>standardize</code>      | Apply z-score so continuous metadata are on the same scale.                                                                                                         |
| <code>plot_heatmap</code>     | Generate a heatmap for the significant associations.                                                                                                                |
| <code>heatmap_first_n</code>  | In heatmap, plot top N features with significant associations.                                                                                                      |
| <code>plot_scatter</code>     | Generate scatter plots for the significant associations.                                                                                                            |
| <code>max_pngs</code>         | Set the maximum number of scatter plots for significant associations to save as png files.                                                                          |
| <code>save_scatter</code>     | Save all scatter plot ggplot objects to an RData file.                                                                                                              |
| <code>cores</code>            | The number of R processes to run in parallel.                                                                                                                       |
| <code>save_models</code>      | Return the full model outputs and save to an RData file.                                                                                                            |
| <code>reference</code>        | The factor to use as a reference for a variable with more than two levels provided as a string of 'variable,reference' semi-colon delimited for multiple variables. |

## Value

List containing the results from applying the model.

## Author(s)

Himel Mallick<himel.stat.iitk@gmail.com>,  
 Ali Rahnavard<gholamali.rahnavaard@gmail.com>,  
 Maintainers: Lauren McIver<lauren.j.mciver@gmail.com>,

**Examples**

```
input_data <- system.file(
  'extdata', 'HMP2_taxonomy.tsv', package="Maaslin2")
input_metadata <- system.file(
  'extdata', 'HMP2_metadata.tsv', package="Maaslin2")
fit_data <- Maaslin2(
  input_data, input_metadata, 'demo_output', transform = "AST",
  fixed_effects = c('diagnosis', 'dysbiosisnonIBD', 'dysbiosisUC', 'dysbiosisCD', 'antibiotics', 'age'),
  random_effects = c('site', 'subject'),
  normalization = 'NONE',
  reference = 'diagnosis,nonIBD',
  standardize = FALSE)
```

# Index

Maaslin2, [2](#)