Package 'LOCOM2'

October 29, 2026

Type Package

Title A Logistic Regression Model for Testing Microbial Differential Abundance

Version 1.1

Date 2026-10-29

Maintainer Yi-Juan Hu <yijuanhu@bicmr.pku.edu.cn>

Description Testing differential abundance at individual taxa and in a whole microbial community. The tests are based on the log-ratio of relative abundances. The tests accommodate continuous, discrete (binary, categorical), and multivariate traits, and allow adjustment of confounders. For more details see He (2026) <doi:10.64898/2026.04.07.716976>. This package also includes the 'com2seq' function that combines two sequencing datasets (e.g., from 16S marker-gene and shotgun sequencing) for integrative analysis of differential abundance. For more details see <doi:10.1080/01621459.2025.2516205>.

License GPL (>= 2)

RoxygenNote 7.3.2

Depends R (>= 3.5.0)

Imports Rcpp (>= 1.0.11), stats, permute, parallel, BiocParallel, matrixStats, abind, car

LinkingTo Rcpp, RcppArmadillo

Suggests testthat, survival

URL https://github.com/yijuanhu/LOCOM2

BugReports https://github.com/yijuanhu/LOCOM2/issues

Encoding UTF-8

LazyData true

NeedsCompilation yes

Author Yi-Juan Hu [aut, cre]

Repository CRAN

Date/Publication 2026-10-29 12:50:02 UTC

Contents

com2seq 2
locom2 5
meta.1 7
meta.2 8
otu.table.1 8
otu.table.2 9
throat.meta.filter 9
throat.otu.table.filter 10
throat.otu.taxonomy 10

Index **12**

`com2seq` *com2seq: combining two sequencing datasets (e.g., from 16S marker-gene and shotgun metagenomic sequencing) for integrative analysis of differential abundance at both the genus and community levels*

Description

This function allows you to test (1). whether any genus is associated with the trait of interest with FDR control, based on the log ratio of relative abundances between pairs of taxa, and (2). whehter the whole community is associated with the trait (a global test). The tests accommodate continuous, discrete (binary, categorical), and multivariate traits, and allows adjustment of confounders.

Usage

```
com2seq(
  formula.1,
  formula.2,
  data.1 = NULL,
  data.2 = NULL,
  fdr.nominal = 0.2,
  seed = NULL,
  n.perm.max = NULL,
  n.rej.stop = 100,
  n.cores = 1,
  filter.thresh = 0.2,
  verbose = TRUE
)
```

Arguments

`formula.1` A formula in the form of `otu.table.1 | C1 ~ Y1`, where `otu.table.1` is the OTU table (or taxa count table) for the first sequencing dataset, `C1` contains confounding covariates to be adjusted for, and `Y1` is the trait of interest. In the OTU table, rows correspond to samples and columns correspond to OTUs; both

	row and column names are required for aligning two OTU tables. `C1` can be replaced by multiple (say `K`) covariates in the form of `C11 + C12 + ... + C1K`.
`formula.2`	A formula in the form of `otu.table.2 | C2 ~ Y2`, where `otu.table.2` is the OTU table (or taxa count table) for the second sequencing dataset, `C2` contains confounding covariates to be adjusted for, and `Y2` is the trait of interest. Same requirements for each component as for `formula.1`.
`data.1`	A data frame for metadata of the first sequencing dataset, with samples ordered in the same way as in `otu.table.1`.
`data.2`	A data frame for metadata of the second sequencing dataset, with samples ordered in the same way as in `otu.table.2`.
`fdr.nominal`	the nominal FDR. The default is 0.2.
`seed`	a user-supplied integer seed for the random number generator in the permutation procedure. The default is NULL, which means that an integer seed will be generated internally and randomly. In either case, the integer seed will be stored in the output object in case the user wants to reproduce the permutation replicates.
`n.perm.max`	the maximum number of permutations. The default is NULL, in which case `n.taxa * n.rej.stop * (1/fdr.nominal)` is used where `n.taxa` is the total number of taxa (that have non-zero counts in at least one sample).
`n.rej.stop`	the minimum number of rejections (i.e., the permutation test statistic exceeds the observed test statistic) to obtain before stopping the permutation procedure. The default is 100.
`n.cores`	the number of cores to be used for parallel computing. The default is 1.
`filter.thresh`	a real value between 0 and 1; OTUs present in fewer than `filter.thresh` * 100% samples are filtered out. The default is 0.2.
`verbose`	a logical value indicating whether to generate verbose output during the permutation process. Default is TRUE.

Details

This function uses a sequential stopping criterion (similar to that of Sandve et al. 2011) for the permutation procedure, which stops when all OTU-level tests have either reached the pre-specified number (default 100) of rejections or yielded a q-value (by the Benjamini-Hochberg [BH] procedure) that is below the nominal FDR level (default 0.2). The permutation procedure is always terminated if a pre-specified maximum number (see description of `n.perm.max`) of permutations have been generated. The global test is based on all permutation replicates when the procedure stops/terminates.

Value

A list consisting of

- effect.size.wc - effect size at each OTU, i.e., beta_j,1 - median_j'=1,...J (beta_j',1), based on the count weight
- p.taxa.wc - p-values by the OTU-specific tests based on the count weight
- q.taxa.wc - q-values (adjusted p-values by the BH procedure) by the OTU-specific tests based on the count weight

- detected.taxa.wc - detected taxa (using the column names of the OTU table) at the nominal FDR by the OTU-specific tests based on the count weight
- p.global.wc - p-value by the global test based on the count weight
- effect.size.we - effect size at each OTU, i.e., beta_j,1 - median_j'=1,...J(beta_j',1), based on the equal weight
- p.taxa.we - p-values by the OTU-specific tests based on the equal weight
- q.taxa.we - q-values (adjusted p-values by the BH procedure) by the OTU-specific tests based on the equal weight
- detected.taxa.we - detected taxa (using the column names of the OTU table) at the nominal FDR by the OTU-specific tests based on the equal weight
- p.global.we - p-value by the global test based on the equal weight
- p.taxa.omni - p-values by the OTU-specific omnibus tests
- q.taxa.omni - q-values (adjusted p-values by the BH procedure) by the OTU-specific omnibus tests
- detected.taxa.omni - detected taxa (using the column names of the OTU table) at the nominal FDR by the OTU-specific omnibus tests
- p.global.omni - p-value by the global omnibus test
- n.perm.completed - number of permutations completed
- ref.otu - the selected reference OTU
- seed - the seed used to generate the permutation replicates

Examples

```
library(LOCOM2)
data("otu.table.1")
data("otu.table.2")
data("meta.1")
data("meta.2")

# restricting to a subset of samples and OTUs for speed
sub.sam.1 <- c(31:60)
sub.sam.2 <- c(1:20, 51:60)
otu.table.sub.1 <- otu.table.1[sub.sam.1, 1:70]
otu.table.sub.2 <- otu.table.2[sub.sam.2, 1:70]
meta.sub.1 <- meta.1[sub.sam.1,]
meta.sub.2 <- meta.2[sub.sam.2,]

# running com2seq
res.com2seq <- com2seq(formula.1 = otu.table.sub.1 | C ~ as.factor(Y), data.1 = meta.sub.1,
                     formula.2 = otu.table.sub.2 | C ~ as.factor(Y), data.2 = meta.sub.2,
                       seed = 123, n.cores = 1, n.perm.max = 1000)
res.com2seq$p.global.omni
res.com2seq$detected.taxa.omni
```

`locom2` *A logistic regression model for testing differential abundance in compositional microbiome data (LOCOM2)*

Description

This function allows you to test (1). whether any OTU (or taxon) is associated with the trait of interest with FDR control, based on log ratios of relative abundances between pairs of OTUs, and (2). whether the whole community is associated with the trait (a global test), based on the harmonic mean method for combining individual p-values The tests accommodate continuous, discrete (binary or categorical), and multivariate traits, and allow adjustment for confounders.

Usage

```
locom2(
  formula,
  data = NULL,
  ref.otu = NULL,
  fdr.nominal = 0.1,
  filter = TRUE,
  permute = TRUE,
  n.perm.max = 1000,
  n.rej.stop = 100,
  n.cores = 1,
  seed = NULL,
  verbose = TRUE,
  Firth.thresh = 0.4
)
```

Arguments

`formula`	A formula in the form of `otu.table	C ~ Y`, where `otu.table` is the OTU table (or taxa count table), `C` contains confounding covariates to be adjusted for, and `Y` is the trait of interest. In the OTU table, rows correspond to samples and columns correspond to OTUs. `C` can be replaced by multiple (say `K`) covariates in the form of `C1 + C2 + ... + CK`.
`data`	A data frame for metadata.
`ref.otu`	A null OTU to be used as the reference. The default is NULL, in which case the most abundant OTU is selected internally as the reference. The user-supplied value should be a single character string specifying the corresponding column name in the OTU table, such as "Escherichia coli".
`fdr.nominal`	The nominal FDR level, with a default of 0.1.
`filter`	A logical value indicating whether to filter out rare OTUs. The default is TRUE, using a filtering threshold of min(0.1*n.sam, 10).
`permute`	A logical value indicating whether to perform permutation. The default is TRUE.

n.perm.max	The maximum number of permutations. The default is 1,000, used for the Wald-type test. At least 100 is required for performing the Wald-type test. The full permutation procedure as in LOCOM is performed when n.perm.max is set to NULL. In this case, the total number of permutations is determined as n.otu * n.rej.stop * (1/fdr.nominal), where n.otu is the number of OTUs (that have non-zero counts in at least one sample). The full permutation procedure adopts a sequential stopping criterion (similar to Sandve et al. 2011), which stops when all OTU-level tests have either reached the prespecified number of rejections (default 100) or yielded a q-value (by the Benjamini-Hochberg [BH] procedure) below the nominal FDR level (default 0.1).
n.rej.stop	The minimum number of rejections (i.e., instances where the permutation test statistic exceeds the observed test statistic) required before stopping the permutation procedure. The default is 100.
n.cores	The number of cores to be used for parallel computing. The default is 1.
seed	A user-supplied integer seed for the random number generator in the permutation procedure. The default is NULL, in which case an integer seed is generated internally at random. In either case, the seed is stored in the output object to enable reproducibility of the permutation replicates.
verbose	A logical value indicating whether to produce verbose output during the permutation process. The default is TRUE.
Firth.thresh	The threshold (between 0 and 1) of OTU prevalence for applying the Firth correction. The default is 0.4.

Details

This function extends LOCOM (Hu et al., 2022, PNAS) in the following ways: -. accommodating both relative abundance and read count data for OTUs; -. refining the weighting scheme in LOCOM to eliminate confounding by library size; -. incorporating a series of adjustments to ensure stable and reliable inference, even under extreme conditions such as rare OTUs and highly unbalanced case–control designs; -. replacing the computationally intensive permutation procedure with a Wald-type test (using a fixed 1,000 permutation replicates by default).

Value

A list consisting of

- p.otu.Wald - Wald p-values for OTU-specific tests
- q.otu.Wald - Wald q-values (adjusted p-values by BH) for OTU-specific tests
- detected.otu.Wald - OTUs detected by the Wald test at the nominal FDR level
- p.otu.perm - permutation p-values for OTU-specific tests
- q.otu.perm - permutation q-values (adjusted p-values by BH) for OTU-specific tests
- detected.otu.perm - OTUs detected by the permutation test at the nominal FDR level
- p.otu.asymptotic - asymptotic p-values for OTU-specific tests
- q.otu.asymptotic - asymptotic q-values (adjusted p-values by BH) for OTU-specific tests
- detected.otu.asymptotic - OTUs detected by the asymptotic test at the nominal FDR level

- beta - effect size at each OTU, defined as beta_j - median (beta_j') if the reference is internally selected and beta_j if the reference is user supplied
- beta.var - variance estimate for beta, calculated from 1000 permutations if available or asymptotic variance
- ref.otu - reference OTU, specified by the user or selected internally.
- p.global - p-value for the global test (not available in the asymptotic version). The global test is based on the harmonic mean of individual p-values, using all permutation replicates generated up to the point when the procedure terminates.
- n.perm.completed - number of permutations completed
- seed - the seed used to generate the permutation replicates

Examples

```
data("throat.otu.table.filter")
data("throat.meta.filter")
data("throat.otu.taxonomy")

Y <- ifelse(throat.meta.filter$SmokingStatus == "NonSmoker", 0, 1)
C <- ifelse(throat.meta.filter$Sex == "Male", 0, 1)

##################
# running LOCOM2
##################

## LOCOM2 (Wald), recommended for sample size greater than 100
## LOCOM2-P (permutation), recommended for sample size less than 100
## Recommend to use n.cores = 4 to speed up

res <- locom2(formula = throat.otu.table.filter | C ~ Y, seed = 123, n.perm.max = 100)
res$detected.otu.Wald
res$p.otu.Wald[res$detected.otu.Wald]
```

---

meta.1 *Metadata for samples in otu.table.1*

Description

Metadata for samples in otu.table.1

Usage

```
data("meta.1")
```

Format

A data frame with 80 observations on 2 variables.

Examples

```
data(meta.1)
```

`meta.2` *Metadata for samples in otu.table.2*

Description

Metadata for samples in otu.table.2

Usage

```
data("meta.2")
```

Format

A data frame with 60 observations on 2 variables.

Examples

```
data(meta.2)
```

`otu.table.1` *Simulated OTU count table*

Description

This table contains 80 subjects with 677 OTUs.

Usage

```
data("otu.table.1")
```

Format

A data frame with 80 observations on 677 variables.

Examples

```
data(otu.table.1)
```

otu.table.2 *Simulated OTU count table*

Description

This table contains 60 subjects with 695 OTUs.

Usage

```
data("otu.table.2")
```

Format

A data frame with 80 observations on 695 variables.

Examples

```
data(otu.table.2)
```

throat.meta.filter *Metadata of the throat microbiome samples*

Description

This data set includes samples from the microbiome of the nasopharynx and oropharynx on each side of the body. It were generated to study the effect of smoking on the microbiota of the upper respiratory tract in 57 individuals, after filtering out three individuals with antibiotic use.

Usage

```
data("throat.meta.filter")
```

Format

A data frame with 57 observations on 16 variables.

Source

Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, et al. (2010) Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers. PLoS ONE 5(12): e15216.

References

R package "GUniFrac"

Examples

```
data(throat.meta.filter)
```

`throat.otu.table.filter` *OTU count table from 16S sequencing of the throat microbiome samples*

Description

This data set contains 57 subjects, after filtering out three subjects with antibiotic use. Microbiome data were collected from right and left nasopharynx and oropharynx region to form an OTU table with 856 OTUs.

Usage

```
data("throat.otu.table.filter")
```

Format

A data frame with 57 observations on 856 variables.

Source

Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, et al. (2010) Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers. PLoS ONE 5(12): e15216.

References

R package "GUniFrac"

Examples

```
data(throat.otu.table.filter)
```

`throat.otu.taxonomy` *Taxonomy names for OTUs from 16S sequencing of the throat microbiome samples*

Description

This file contains 5683 taxonomy names.

Usage

```
data("throat.otu.taxonomy")
```

Format

A vector with 5683 taxonomy names

Source

Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, et al. (2010) Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers. PLoS ONE 5(12): e15216.

References

R package "GUniFrac"

Examples

```
data(throat.otu.taxonomy)
```

Index

∗ **datasets**
meta.1, 7
meta.2, 8
otu.table.1, 8
otu.table.2, 9
throat.meta.filter, 9
throat.otu.table.filter, 10
throat.otu.taxonomy, 10

com2seq, 2

locom2, 5

meta.1, 7
meta.2, 8

otu.table.1, 8
otu.table.2, 9

throat.meta.filter, 9
throat.otu.table.filter, 10
throat.otu.taxonomy, 10