

Package: MicEco (via r-universe)

November 6, 2024

Title Various functions for microbial community data

Version 0.9.19

Maintainer Jakob Russel <russel2620@gmail.com>

Description Collection of functions for microbiome analyses. E.g. fitting neutral models and standardized effect sizes of phylogenetic beta diversities, and much more.

Depends R (>= 3.2.5)

Imports stats, utils, phyloseq, foreach, doSNOW, picante, vegan, snow, bbmle, Hmisc, abind, reshape2, eulerr, pheatmap

License GPL (>= 3) | file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.0

Remotes bioc::release/phyloseq

Repository <https://metagenlab.r-universe.dev>

RemoteUrl <https://github.com/Russel88/MicEco>

RemoteRef HEAD

RemoteSha 91f8e6f5d67e0bfd018dc3b46da3994b1eadeb46

Contents

adonis_OmegaSq	2
clr	3
comdist.par	3
comdistnt.par	4
community_rrna	5
gg_13_5_16S	5
neutral.fit	6
neutral.rand	6
proportionality	7
ps_euler	8

ps_pheatmap	9
ps_prune	10
ps_refactor	10
ps_tax_clean	11
ps_venn	11
rarefy_rrna	12
rcurve	13
ses.comdist	14
ses.comdist2	15
ses.comdistnt	17
ses.comdistnt2	18
ses.mntd.par	20
ses.mpd.par	21
ses.permtest	23
ses.UniFrac	24
UniFrac.multi	25
WdS.test	26

Index	27
--------------	-----------

adonis_OmegaSq	<i>Calculate (partial) Omega-squared (effect-size calculation) for PERMANOVA and add it to the input object</i>
----------------	---

Description

Calculate (partial) Omega-squared (effect-size calculation) for PERMANOVA and add it to the input object

Usage

```
adonis_OmegaSq(adonisOutput, partial = TRUE)
```

Arguments

adonisOutput	An adonis object
partial	Should partial omega-squared be calculated (sample size adjusted). Default TRUE

Value

Original adonis object with the (partial) Omega-squared values added

clr	<i>CLR transformation of community matrix, with multiplicative zero replacement</i>
-----	---

Description

This function performs a CLR transformation on the input. Prior to transformation it does a multiplicative zero replacement, which adds a pseudocount, but corrects the non-zero abundances such that log-ratios between non-zero elements are unchanged after correction.

Usage

```
clr(mat, delta = 1)
```

Arguments

mat	A community matrix with features as rows
delta	The pseudocount to exchange zeroes with. Zero-correction is multiplicative such that the log-ratios between any entirely non-zero features will not be affected by the pseudocount.

Value

CLR transformed community matrix

comdist.par	<i>Inter-community mean pairwise distance</i>
-------------	---

Description

Parallel calculation of MPD (mean pairwise distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
comdist.par(comm, dis, abundance.weighted = FALSE, cores = 1, progress = TRUE)
```

Arguments

comm	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
cores	How many cores should be used for parallel computing
progress	Show a progress bar

Value

A distance matrix of MPD values

comdistnt.par	<i>Inter-community mean nearest taxon distance</i>
---------------	--

Description

Parallel calculation of MNTD (mean nearest taxon distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
comdistnt.par(
  comm,
  dis,
  abundance.weighted = FALSE,
  exclude.conspecifics = FALSE,
  cores = 1,
  progress = TRUE
)
```

Arguments

comm	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
exclude.conspecifics	Should conspecific taxa in different communities be exclude from MNTD calculations? (default = FALSE)
cores	How many cores should be used for parallel computing'
progress	Show a progress bar

Value

A distance matrix of MNTD values

community_rrna	<i>Calculate the community/sample-wise mean 16S rRNA copy numbers</i>
----------------	---

Description

Calculate the average 16S rRNA copy number for each sample in an OTU table

Usage

```
community_rrna(x, copy.database = "v13.5", weighted = TRUE)
```

Arguments

x	A phyloseq object OR an OTU-table with taxa as rows and OTU names as row-names. OTUs should be picked against the Greengenes v13.5 database, unless a another copy number database is provided.
copy.database	What Greengenes database version was used to find OTUs. Atm only "v13.5" is available. Alternatively, A dataframe with two variables: "ID" is the OTU id matched by rownames in x and "Copy" is the copy number.
weighted	Logical. Should the average copy number be weighted by the relative abundances of the OTUs? If not, it is advised to rarefy the otu-table first if there are large differences in sample reads.

Value

A dataframe with an average copy number for each sample

gg_13_5_16S	<i>16S copy numbers for GreenGenes 13_5 database OTUs</i>
-------------	---

Description

Precalculated for the PICRUSt pipeline (https://picrust.github.io/picrust/picrust_precalculated_files.html)

neutral.fit	<i>Fit Sloan et al. (2006) Neutral Model</i>
-------------	--

Description

Fit neutral model developed by Sloan et al. (2006, Environ Microbiol 8(4):732-740) and implemented by Burns et al. (2015, ISME J 10(3):655-664).

Usage

```
neutral.fit(otu)
```

Arguments

otu	An OTU-table with taxa as columns and samples as rows.
-----	--

Value

A list of length two; first element contains fit statistics, the second element contains predictions.

neutral.rand	<i>Fit Sloan et al. (2006) Neutral Model several times</i>
--------------	--

Description

Fit neutral model developed by Sloan et al. (2006, Environ Microbiol 8(4):732-740) and implemented by Burns et al. (2015, ISME J 10(3):655-664) several times on randomly picked samples and with 16S rRNA gene copy number corrected rarefaction.

Usage

```
neutral.rand(  
  data,  
  n = NULL,  
  s = NULL,  
  rRNA = NULL,  
  rn = NULL,  
  cores = 1,  
  naming = NULL  
)
```

Arguments

data	A phyloseq object
n	Integer. Number of times to repeat analysis
s	Integer. Number of random samples to for each repetition.
rRNA	What Greengenes database version was used to find OTUs. Atm only "v13.5" is available. Alternatively, A dataframe with two variables: "ID" is the OTU id matched by names in x and "Copy" is the copy number.
rn	Integer. Number of reads to sample for rarefaction
cores	Integer. Number of cores to use for parallel computing.
naming	Optional. A list for naming the output e.g. list(Time="1 Week",Type="Gut")

Value

A list of length two; first element contains fit statistics, the second element contains predictions.

proportionality	<i>Calculate proportionality between all pairs of OTUs, for network analysis</i>
-----------------	--

Description

Calculate proportionality as proposed by Lovell et al. 2016 Proportionality: a valid alternative to correlation for relative data

Usage

```
proportionality(x, delta = 1)
```

Arguments

x	An otu-table with OTUs as rows OR a phyloseq object
delta	The pseudocount to exchange zeroes with. Zero-correction is multiplicative such that the proportionality between any entirely non-zero OTUs will not be affected by the pseudocount.

Value

The network association matrix with proportionality values

ps_euler	<i>Make Euler diagram of shared taxa (ASVs, OTUs) across sample groups</i>
----------	--

Description

Make Euler diagram of shared taxa (ASVs, OTUs) across sample groups from a phyloseq object. Overlap can be weighted by relative abundance

Usage

```
ps_euler(
  ps,
  group,
  fraction = 0,
  shape = "circle",
  weight = FALSE,
  relative = TRUE,
  plot = TRUE,
  ...
)
```

Arguments

ps	A phyloseq object
group	The grouping factor. Should match variable in sample_data(ps)
fraction	The fraction (0 to 1) of samples in a group in which the taxa should be present to be included in the count.
shape	Shape of the plot "circle" or "ellipse"?
weight	If TRUE, the overlaps are weighted by abundance
relative	Should abundances be made relative
plot	If TRUE return a plot, if FALSE return a list with shared and unique taxa
...	Additional arguments

Details

Any further arguments to this function are passed to the plot.euler function from the eulerr package. This can be used to change colors, fonts, and other graphical parameters. For example: ps_euler(phy, "Time", quantities = list(type=c("percent","counts"), font = 2), labels = list(cex = 2), col = "red", fill = c("red","blue","green"))

Value

An euler plot

ps_pheatmap

Make pheatmap directly from a phyloseq object

Description

Make pheatmap directly from a phyloseq object, including agglomeration and filtering. Default color map is viridis with absent taxa as black.

Usage

```
ps_pheatmap(
  ps,
  annot_samp = NULL,
  annot_taxa = NULL,
  relative = TRUE,
  log10 = TRUE,
  tax_agg = NULL,
  order_taxa = TRUE,
  min_samples = 1,
  min_reads = 1,
  min_abundance = 0,
  label_rank = NULL,
  color = c("black", viridis::viridis(10)),
  ...
)
```

Arguments

ps	A phyloseq object
annot_samp	Sample variables to annotate with
annot_taxa	Taxa variables to annotate with
relative	If TRUE, abundances total sum scaled. Default TRUE
log10	If TRUE, log10 transform abundances. Default TRUE
tax_agg	Taxa rank to agglomerate. Default NULL
order_taxa	If TRUE, taxa are ordered from most to least abundant. Default TRUE
min_samples	Minimum number of samples the features should be present in. Default 1
min_reads	Minimum number of total reads the features should have. Default 1
min_abundance	Minimum mean relative abundance features should have. Default 0
label_rank	Taxa rank to label the taxa. If NULL will label by taxa_names(ps). Default NULL
color	Color palette. Default viridis with absent
...	Additional arguments to pheatmap function

Value

A heatmap

ps_prune	<i>Prune taxa from phyloseq object by their prevalence or abundance</i>
----------	---

Description

Prune taxa from phyloseq object by their prevalence or abundance

Usage

```
ps_prune(data, min.samples = 0, min.reads = 0, min.abundance = 0)
```

Arguments

data	phyloseq object.
min.samples	Minimum number of samples the features should be present in. Default 0
min.reads	Minimum number of total reads the features should have. Default 0
min.abundance	Minimum mean relative abundance features should have. Default 0

Value

Similar to input, but with features not reaching the criteria given grouped as "Others"

ps_refactor	<i>Relevel the Sample variable in a psmelted phyloseq object</i>
-------------	--

Description

Relevel the Sample variable in a psmelted phyloseq object, such that similar samples are plotted together with ggplot barcharts

Usage

```
ps_refactor(psmelted, ...)
```

Arguments

psmelted	A phyloseq object melted into a data.frame with psmelt
...	Arguments passed to hclust

ps_tax_clean	<i>Clean tax_table such that NAs are replaced with names of the most specific known taxonomy prefixed with the rank.</i>
--------------	--

Description

Example: An OTU is annotated as Proteobacteria in the Phylum rank, but NAs in the more specific ranks. All annotations below Phylum for this OTU will be replaced with Phylum_Proteobacteria.

Usage

```
ps_tax_clean(data)
```

Arguments

data phyloseq or tax_table object.

Value

Similar to input, but with NAs replaced in the tax_table

ps_venn	<i>Make Venn diagram of shared taxa (ASVs, OTUs) across sample groups</i>
---------	---

Description

Make Venn diagram of shared taxa (ASVs, OTUs) across sample groups from a phyloseq object. Overlap can be weighted by relative abundance #' Any further arguments to this function are passed to the plot.venn function from the eulerr package. This can be used to change colors, fonts, and other graphical parameters. For example: ps_venn(phy, "Time", quantities = list(type=c("percent", "counts"), font = 2), labels = list(cex = 2), col = "red", fill = c("red", "blue", "green"))

Usage

```
ps_venn(
  ps,
  group,
  fraction = 0,
  weight = FALSE,
  relative = TRUE,
  plot = TRUE,
  ...
)
```

Arguments

ps	A phyloseq object
group	The grouping factor. Should match variable in sample_data(ps)
fraction	The fraction (0 to 1) of samples in a group in which the taxa should be present to be included in the count.
weight	If TRUE, the overlaps are weighted by abundance
relative	Should abundances be made relative
plot	If TRUE return a plot, if FALSE return a list with shared and unique taxa
...	Additional arguments

Value

An venn plot

rarefy_rrna	<i>Rarefy and normalize based on 16S rRNA copy numbers</i>
-------------	--

Description

Rarefy an OTU-table with the probability of the inverse 16S rRNA copy numbers: The result is a normalized AND copy number corrected OTU-table.

Usage

```
rarefy_rrna(
  x,
  reads,
  copy.database = "v13.5",
  seed = NULL,
  trim = FALSE,
  replace = FALSE
)
```

```
rarefy_rrna.matrix(x, reads, copy.database, seed = NULL, trim, replace)
```

```
rarefy_rrna.phyloseq(x, reads, copy.database, seed = NULL, trim, replace)
```

Arguments

x	A phyloseq object OR an OTU-table with taxa as columns and OTU names as colnames. OTUs should be picked against the Greengenes v13.5 database, unless another copy number database is provided.
reads	Number of reads to sample.

copy.database	What Greengenes database version was used to find OTUs. Atm only "v13.5" is available. Alternatively, A dataframe with two variables: "ID" is the OTU id matched by names in x and "Copy" is the copy number.
seed	Random seed for sampling.
trim	Should samples with less than the set amount of reads be trimmed away?
replace	Should reads be sampled with replacement? Default FALSE. If FALSE, rRNA copy numbers will have small effect if the rarefied depth is close to the actual depth. If TRUE, some taxa can end up with more reads in the rarefied matrix that they had in the input.

Value

A rarefied otu-table

rcurve	<i>Rarefaction curve on phyloseq object Theoretical richness with vegan's rarefy function</i>
--------	---

Description

Rarefaction curve on phyloseq object Theoretical richness with vegan's rarefy function

Usage

```
rcurve(physeq, subsamp = 10^c(1:5), trim = TRUE, add_sample_data = TRUE)
```

Arguments

physeq	phyloseq object
subsamp	Vector of number of reads to subsample
trim	Remove richness estimations from subsamples larger than the library size
add_sample_data	Add sample data to data.frame

 ses.comdist

Standardized effect size of inter-community MPD (betaMPD, betaNRI)

Description

Standardized effect size of MPD (mean pairwise distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
ses.comdist(
  samp,
  dis,
  null.model = c("taxa.labels", "richness", "frequency", "sample.pool",
    "phylogeny.pool", "independentswap", "trialswap"),
  abundance.weighted = FALSE,
  runs = 999,
  iterations = 1000,
  cores = 1
)
```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
null.model	Null model to use (see Details section for description)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
runs	Number of randomizations
iterations	Number of iterations to use for each randomization (for independent swap and trial null models)
cores	Number of cores to use for parallel computing

Details

Currently implemented null models (arguments to null.model):

- taxa.labels - Shuffle distance matrix labels (across all taxa included in distance matrix)
- richness - Randomize community data matrix abundances within samples (maintains sample species richness)
- frequency - Randomize community data matrix abundances within species (maintains species occurrence frequency)
- sample.pool - Randomize community data matrix by drawing species from pool of species occurring in at least one community (sample pool) with equal probability

- phylogeny.pool- Randomize community data matrix by drawing species from pool of species occurring in the distance matrix (phylogeny pool) with equal probability
- independentswap - Randomize community data matrix with the independent swap algorithm (Gotelli 2000) maintaining species occurrence frequency and sample species richness
- trialswap - Randomize community data matrix with the trial-swap algorithm (Miklos & Podani 2004) maintaining species occurrence frequency and sample species richness

Value

A list of results:

- ntaxa - Number of taxa in community
- comdist.obs - Observed mpd between communities
- comdist.rand.mean - Mean mpd between null communities
- comdist.rand.sd - Standard deviation of mpd between null communities
- comdist.obs.rank - Rank of observed mpd vs. null mpd
- comdist.obs.z - Standardized effect size of mpd vs. null mpd (= (comdist.obs - comdist.rand.mean) / comdist.rand.sd, equivalent to -betaNRI)
- comdist.obs.p - P-value (quantile) of observed mpd vs. null communities (= comdist.obs.rank / runs + 1)
- runs - Number of randomizations

ses.comdist2

Standardized effect size of inter-community MPD (betaMPD, betaNRI)

Description

Standardized effect size of MPD (mean pairwise distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
ses.comdist2(
  samp,
  dis,
  method = "quasiswap",
  fixedmar = "both",
  shuffle = "both",
  strata = NULL,
  mtype = "count",
  burnin = 0,
  thin = 1,
  abundance.weighted = FALSE,
  runs = 999,
  cores = 1
)
```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
method	Character for method used for the swap algorithm ("swap", "tswap", "quasiswap", "backtrack"). If NULL no swap algorithm is applied (uses permatfull from vegan). If mtype="count" the "quasiswap", "swap", "swsh" and "abuswap" methods are available (see details).
fixedmar	Character, stating which of the row/column sums should be preserved ("none", "rows", "columns", "both").
shuffle	Character, indicating whether individuals ("ind"), samples ("samp") or both ("both") should be shuffled, see details.
strata	Numeric vector or factor with length same as nrow(m) for grouping rows within strata for restricted permutations. Unique values or levels are used.
mtype	Matrix data type, either "count" for count data, or "prab" for presence-absence type incidence data.
burnin	Number of null communities discarded before proper analysis in sequential ("swap", "tswap") methods.
thin	Number of discarded permuted matrices between two evaluations in sequential ("swap", "tswap") methods.
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
runs	Number of randomizations
cores	Number of cores to use for parallel computing

Details

See permat (vegan) for detailed options

Value

A list of results:

- ntaxa - Number of taxa in community
- comdist.obs - Observed mpd between communities
- comdist.rand.mean - Mean mpd between null communities
- comdist.rand.sd - Standard deviation of mpd between null communities
- comdist.obs.rank - Rank of observed mpd vs. null mpd
- comdist.obs.z - Standardized effect size of mpd vs. null mpd (= (comdist.obs - comdist.rand.mean) / comdist.rand.sd, equivalent to -betaNRI)
- comdist.obs.p - P-value (quantile) of observed mpd vs. null communities (= comdist.obs.rank / runs + 1)
- runs - Number of randomizations

ses.comdistnt	<i>Standardized effect size of inter-community MNTD (betaMNTD, betaNTI)</i>
---------------	---

Description

Standardized effect size of MNTD (mean nearest taxon distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
ses.comdistnt(
  samp,
  dis,
  null.model = c("taxa.labels", "richness", "frequency", "sample.pool",
    "phylogeny.pool", "independentswap", "trialswap"),
  abundance.weighted = FALSE,
  exclude.conspecifics = FALSE,
  runs = 999,
  iterations = 1000,
  cores = 1
)
```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
null.model	Null model to use (see Details section for description)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
exclude.conspecifics	Should conspecific taxa in different communities be exclude from MNTD calculations? (default = FALSE)
runs	Number of randomizations
iterations	Number of iterations to use for each randomization (for independent swap and trial null models)
cores	Number of cores to use for parallel computing

Details

Currently implemented null models (arguments to null.model):

- taxa.labels - Shuffle distance matrix labels (across all taxa included in distance matrix)
- richness - Randomize community data matrix abundances within samples (maintains sample species richness)

- frequency - Randomize community data matrix abundances within species (maintains species occurrence frequency)
- sample.pool - Randomize community data matrix by drawing species from pool of species occurring in at least one community (sample pool) with equal probability
- phylogeny.pool - Randomize community data matrix by drawing species from pool of species occurring in the distance matrix (phylogeny pool) with equal probability
- independentswap - Randomize community data matrix with the independent swap algorithm (Gotelli 2000) maintaining species occurrence frequency and sample species richness
- trialsswap - Randomize community data matrix with the trial-swap algorithm (Miklos & Podani 2004) maintaining species occurrence frequency and sample species richness

Value

A list of results:

- ntaxa - Number of taxa in community
- comdistnt.obs - Observed mntd between communities
- comdistnt.rand.mean - Mean mntd between null communities
- comdistnt.rand.sd - Standard deviation of mntd between null communities
- comdistnt.obs.rank - Rank of observed mntd vs. null mntd
- comdistnt.obs.z - Standardized effect size of mntd vs. null mntd (= (comdistnt.obs - comdistnt.rand.mean) / comdistnt.rand.sd, equivalent to -betaNTI)
- comdistnt.obs.p - P-value (quantile) of observed mntd vs. null communities (= comdistnt.obs.rank / runs + 1)
- runs - Number of randomizations

ses.comdistnt2	<i>Standardized effect size of inter-community MNTD (betaMNTD, betaNTI)</i>
----------------	---

Description

Standardized effect size of MNTD (mean nearest taxon distance) separating taxa in two communities, a measure of phylogenetic beta diversity

Usage

```
ses.comdistnt2(
  samp,
  dis,
  method = "quasiswap",
  fixedmar = "both",
  shuffle = "both",
  strata = NULL,
```

```

mtype = "count",
burnin = 0,
thin = 1,
abundance.weighted = FALSE,
exclude.conspecific = FALSE,
runs = 999,
cores = 1
)

```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
method	Character for method used for the swap algorithm ("swap", "tswap", "quasiswap", "backtrack"). If NULL no swap algorithm is applied (uses permatfull from vegan). If mtype="count" the "quasiswap", "swap", "swsh" and "abuswap" methods are available (see details).
fixedmar	Character, stating which of the row/column sums should be preserved ("none", "rows", "columns", "both").
shuffle	Character, indicating whether individuals ("ind"), samples ("samp") or both ("both") should be shuffled, see details.
strata	Numeric vector or factor with length same as nrow(m) for grouping rows within strata for restricted permutations. Unique values or levels are used.
mtype	Matrix data type, either "count" for count data, or "prab" for presence-absence type incidence data.
burnin	Number of null communities discarded before proper analysis in sequential ("swap", "tswap") methods.
thin	Number of discarded permuted matrices between two evaluations in sequential ("swap", "tswap") methods.
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
exclude.conspecific	Should conspecific taxa in different communities be exclude from MNTD calculations? (default = FALSE)
runs	Number of randomizations
cores	Number of cores to use for parallel computing

Details

See permat (vegan) for detailed options

Value

A list of results:

- ntaxa - Number of taxa in community
- comdistnt.obs - Observed mntd between communities
- comdistnt.rand.mean - Mean mntd between null communities
- comdistnt.rand.sd - Standard deviation of mntd between null communities
- comdistnt.obs.rank - Rank of observed mntd vs. null mntd
- comdistnt.obs.z - Standardized effect size of mntd vs. null mntd (= (comdistnt.obs - comdistnt.rand.mean) / comdistnt.rand.sd, equivalent to -betaNTI)
- comdistnt.obs.p - P-value (quantile) of observed mntd vs. null communities (= comdistnt.obs.rank / runs + 1)
- runs - Number of randomizations

 ses.mntd.par

Standardized effect size of MNTD

Description

Parallel calculation of standardized effect size of mean nearest taxon distances in communities. When used with a phylogenetic distance matrix, equivalent to -1 times the Nearest Taxon Index (NTI).

Usage

```
ses.mntd.par(
  samp,
  dis,
  null.model = c("taxa.labels", "richness", "frequency", "sample.pool",
    "phylogeny.pool", "independentswap", "trialswap"),
  abundance.weighted = FALSE,
  runs = 999,
  iterations = 1000,
  cores = 1
)
```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
null.model	Null model to use (see Details section for description)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
runs	Number of randomizations
iterations	Number of iterations to use for each randomization (for independent swap and trial null models)
cores	Number of cores to use for parallel computing

Details

Faster than ses.mntd from *picante* when there are many samples and taxa.

Currently implemented null models (arguments to `null.model`):

- `taxa.labels` - Shuffle distance matrix labels (across all taxa included in distance matrix)
- `richness` - Randomize community data matrix abundances within samples (maintains sample species richness)
- `frequency` - Randomize community data matrix abundances within species (maintains species occurrence frequency)
- `sample.pool` - Randomize community data matrix by drawing species from pool of species occurring in at least one community (sample pool) with equal probability
- `phylogeny.pool` - Randomize community data matrix by drawing species from pool of species occurring in the distance matrix (phylogeny pool) with equal probability
- `independentswap` - Randomize community data matrix with the independent swap algorithm (Gotelli 2000) maintaining species occurrence frequency and sample species richness
- `trialswap` - Randomize community data matrix with the trial-swap algorithm (Miklos & Podani 2004) maintaining species occurrence frequency and sample species richness

Value

A data frame of results for each community

- `ntaxa` Number of taxa in community
- `mntd.obs` Observed mntd in community
- `mntd.rand.mean` Mean mntd in null communities
- `mntd.rand.sd` Standard deviation of mntd in null communities
- `mntd.obs.rank` Rank of observed mntd vs. null communities
- `mntd.obs.z` Standardized effect size of mntd vs. null communities (= $(\text{mntd.obs} - \text{mntd.rand.mean}) / \text{mntd.rand.sd}$, equivalent to -NRI)
- `mntd.obs.pP-value` (quantile) of observed mntd vs. null communities (= $\text{mntd.obs.rank} / \text{runs} + 1$)
- `runs` Number of randomizations

ses.mpd.par

Standardized effect size of MPD

Description

Parallel calculation of standardized effect size of mean pairwise distances in communities. When used with a phylogenetic distance matrix, equivalent to -1 times the Nearest Relative Index (NRI).

Usage

```
ses.mpd.par(
  samp,
  dis,
  null.model = c("taxa.labels", "richness", "frequency", "sample.pool",
    "phylogeny.pool", "independentswap", "trialswap"),
  abundance.weighted = FALSE,
  runs = 999,
  iterations = 1000,
  cores = 1
)
```

Arguments

samp	Community data matrix with samples as rows
dis	Distance matrix (generally a phylogenetic distance matrix)
null.model	Null model to use (see Details section for description)
abundance.weighted	Should mean nearest taxon distances for each species be weighted by species abundance? (default = FALSE)
runs	Number of randomizations
iterations	Number of iterations to use for each randomization (for independent swap and trial null models)
cores	Number of cores to use for parallel computing

Details

Faster than `ses.mpd` from `picante` when there are many samples and taxa.

Currently implemented null models (arguments to `null.model`):

- `taxa.labels` - Shuffle distance matrix labels (across all taxa included in distance matrix)
- `richness` - Randomize community data matrix abundances within samples (maintains sample species richness)
- `frequency` - Randomize community data matrix abundances within species (maintains species occurrence frequency)
- `sample.pool` - Randomize community data matrix by drawing species from pool of species occurring in at least one community (sample pool) with equal probability
- `phylogeny.pool` - Randomize community data matrix by drawing species from pool of species occurring in the distance matrix (phylogeny pool) with equal probability
- `independentswap` - Randomize community data matrix with the independent swap algorithm (Gotelli 2000) maintaining species occurrence frequency and sample species richness
- `trialswap` - Randomize community data matrix with the trial-swap algorithm (Miklos & Podani 2004) maintaining species occurrence frequency and sample species richness

Value

A data frame of results for each community

- ntaxaNumber of taxa in community
- mpd.obsObserved mpd in community
- mpd.rand.meanMean mpd in null communities
- mpd.rand.sdStandard deviation of mpd in null communities
- mpd.obs.rankRank of observed mpd vs. null communities
- mpd.obs.zStandardized effect size of mpd vs. null communities ($= (\text{mpd.obs} - \text{mpd.rand.mean}) / \text{mpd.rand.sd}$, equivalent to -NRI)
- mpd.obs.pP-value (quantile) of observed mpd vs. null communities ($= \text{mpd.obs.rank} / \text{runs} + 1$)
- runsNumber of randomizations

ses.permtest

Permutation test of z-matrix

Description

Permutation test of z-matrix from ses.comdist(2), ses.comdistnt and ses.UniFrac. Test if means within or between groups is higher or lower than the overall mean of the matrix

Usage

```
ses.permtest(zmat, sampleGroups, R = 10000)
```

Arguments

zmat	Symmetric matrix with z-values
sampleGroups	Vector with the grouping of the samples in the same order as in zmat
R	Number of permutations

Value

A dataframe with p-values (pval), fdr corrected p-values (pval.adj), and average z-values (avg) for each group and combination of groups.

ses.UniFrac *Standardized effect size of unifrac*

Description

Standardized effect size of unifrac

Usage

```
ses.UniFrac(
  physeq,
  method = "taxa.labels",
  fixedmar = "both",
  shuffle = "both",
  strata = NULL,
  mtype = "count",
  burnin = 0,
  thin = 1,
  weighted = TRUE,
  normalized = TRUE,
  runs = 99,
  cores = 1
)
```

Arguments

physeq	phyloseq-class, containing at minimum a phylogenetic tree and otu table
method	"taxa.labels" shuffles labels in phylogenetic tree (Ignores fixedmar, shuffle, strata, mtype). If NULL then no swap algorithm is applied (i.e. uses permatfull from vegan). Else the method used for the swap algorithm ("swap", "tswap", "quasiswap", "backtrack"). If mtype="count" the "quasiswap", "swap", "swsh" and "abuswap" methods are available (see details).
fixedmar	Character, stating which of the row/column sums should be preserved ("none", "rows", "columns", "both").
shuffle	Character, indicating whether individuals ("ind"), samples ("samp") or both ("both") should be shuffled.
strata	Numeric vector or factor for grouping samples within strata for restricted permutations. Unique values or levels are used.
mtype	Matrix data type, either "count" for count data, or "prab" for presence-absence type incidence data.
burnin	Number of null communities discarded before proper analysis in sequential ("swap", "tswap") methods.
thin	Number of discarded permuted matrices between two evaluations in sequential ("swap", "tswap") methods.
weighted	Should unifrac be weighted by species abundance? (default = TRUE)

normalized	(Optional). Logical. Should the output be normalized such that values range from 0 to 1 independent of branch length values? Default is TRUE. Note that (unweighted) UniFrac is always normalized by total branch-length, and so this value is ignored when weighted == FALSE.
runs	Number of randomizations
cores	Number of cores to use for UniFrac of null communities. Default is 1

Details

See `permat` (vegan) for detailed options on permutation

Value

A list of results:

- `unifrac.obs` - Observed unifrac between communities
- `unifrac.rand.mean` - Mean unifrac between null communities
- `unifrac.rand.sd` - Standard deviation of unifrac between null communities
- `unifrac.obs.rank` - Rank of observed unifrac vs. null unifrac
- `unifrac.obs.z` - Standardized effect size of unifrac vs. null unifrac (= $(\text{unifrac.obs} - \text{unifrac.rand.mean}) / \text{unifrac.rand.sd}$)
- `unifrac.obs.p` - P-value (quantile) of observed unifrac vs. null communities (= $\text{unifrac.obs.rank} / \text{runs} + 1$)

UniFrac.multi

Run UniFrac multiple times in parallel and take the average

Description

With unrooted phylogenies UniFrac sets the root randomly on the tree. The position of the root affects the results. The function runs UniFrac multiple times, with different roots, and takes the average to smooth potential bias.

Usage

```
UniFrac.multi(physeq, R = 100, seed = 42, cores = 1, ...)
```

Arguments

physeq	Phyloseq object. Required
R	Number of times to repeat calculation. Default 100
seed	Random seed for reproducibility. Default 42
cores	Number of cores to use for parallel computing. Default 1 aka not parallel
...	Additional arguments passed to the UniFrac function

Value

A distance object with the average UniFrac distances

WdS.test	<i>Permutation test of Wd* - robust distance-based multivariate analysis of variance</i>
----------	--

Description

Robust distance-based multivariate analysis of variance (<https://doi.org/10.1186/s40168-019-0659-9>)

Usage

```
WdS.test(dm, f, nrep = 999, strata = NULL)
```

Arguments

dm	Distance matrix
f	Factor
nrep	Number of permutations
strata	Factor for permuting in strata

Details

This code is taken from <https://github.com/alekseyenko/WdStar/>

Value

A list with a p-value, test statistic, and number of permutation

Index

- * **MNTD**
 - comdistnt.par, 4
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
 - ses.mntd.par, 20
- * **MPD**
 - comdist.par, 3
 - ses.comdist, 14
 - ses.comdist2, 15
 - ses.mpd.par, 21
- * **NRI**
 - ses.mpd.par, 21
- * **NTI**
 - ses.mntd.par, 20
- * **association**
 - proportionality, 7
- * **bMNTD**
 - comdistnt.par, 4
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
- * **bMPD**
 - comdist.par, 3
 - ses.comdist, 14
 - ses.comdist2, 15
- * **bNRI**
 - comdist.par, 3
 - ses.comdist, 14
 - ses.comdist2, 15
- * **bNTI**
 - comdistnt.par, 4
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
- * **betaMNTD**
 - comdistnt.par, 4
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
- * **betaMPD**
 - comdist.par, 3
 - ses.comdist, 14
 - ses.comdist2, 15
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
- * **betaNRI**
 - comdist.par, 3
 - ses.comdist, 14
 - ses.comdist2, 15
- * **betaNTI**
 - comdistnt.par, 4
 - ses.comdistnt, 17
 - ses.comdistnt2, 18
- * **clr**
 - clr, 3
- * **data**
 - gg_13_5_16S, 5
- * **diagram**
 - ps_euler, 8
 - ps_venn, 11
- * **distance**
 - WdS.test, 26
- * **euler**
 - ps_euler, 8
- * **network**
 - proportionality, 7
- * **neutral**
 - neutral.fit, 6
 - neutral.rand, 6
- * **normalize**
 - rarefy_rrna, 12
- * **pseudocount**
 - clr, 3
- * **rarefy**
 - rarefy_rrna, 12
- * **rrna**
 - community_rrna, 5
 - rarefy_rrna, 12
- * **ses**
 - ses.comdist, 14
 - ses.comdist2, 15
 - ses.comdistnt, 17
 - ses.comdistnt2, 18

- ses.mntd.par, 20
- ses.mpd.par, 21
- ses.UniFrac, 24
- * **unifrac**
 - ses.UniFrac, 24
- * **venn**
 - ps_venn, 11
- adonis_OmegaSq, 2
- clr, 3
- comdist.par, 3
- comdistnt.par, 4
- community_rrna, 5
- gg_13_5_16S, 5
- neutral.fit, 6
- neutral.rand, 6
- proportionality, 7
- ps_euler, 8
- ps_pheatmap, 9
- ps_prune, 10
- ps_refactor, 10
- ps_tax_clean, 11
- ps_venn, 11
- rarefy_rrna, 12
- rcurve, 13
- ses.comdist, 14
- ses.comdist2, 15
- ses.comdistnt, 17
- ses.comdistnt2, 18
- ses.mntd.par, 20
- ses.mpd.par, 21
- ses.permtest, 23
- ses.UniFrac, 24
- UniFrac.multi, 25
- WdS.test, 26