

Package ‘iNEXT.seq’

July 20, 2024

Type Package

Title Interpolation and Extrapolation for phylogenetic beta diversity and dissimilarity measure for genetic sequence data

Version 0.1.0

Author YaYun Teng & ChunYu Yeh

Maintainer ChunYu Yeh <881204ian@gmail.com>

Description iNEXT.seq provides the function to calculate beta diversity and dissimilarity measure for multiple assemblages.

Depends R (>= 4.0)

Imports tidyverse,
magrittr,
tidyr,
ggplot2,
abind,
ape,
ade4,
phytools,
phyclust,
tidytree,
colorRamps,
future.apply,
iNEXT.3D,
iNEXT.beta3D

Remotes AnneChao/iNEXT.beta3D

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

R topics documented:

esophagus	2
esophagus_tree	3
gghierPD	3
ggiNEXTseq	4
ggObsAsyPD	5
hierPD	5

iNEXTseq	6
ObsAsyPD	8
tongue_cheek	9
tongue_cheek_tree	9
wetland	10
wetland_mat	10
wetland_tree	11
Index	12

esophagus	<i>esophagus OTU count data</i>
-----------	---------------------------------

Description

This dataset is from package phyloseq and collected by Pei et al. (2004).

Usage

data(esophagus)

Format

esophagus is a list with four species-by-assemblage data.frames.

\$ esophagus_BC:

..\$ B: num [1:58] 50 0 2 0 0 ...

..\$ C: num [1:58] 19 2 13 2 0 ...

\$ esophagus_BD:

..\$ B: num [1:58] 50 0 2 0 0 ...

..\$ D: num [1:58] 5 0 0 1 1 ...

\$ esophagus_CD:

..\$ C: num [1:58] 19 2 13 2 0 ...

..\$ D: num [1:58] 5 0 0 1 1 ...

\$ esophagus_BCD:

..\$ B: num [1:58] 50 0 2 0 0 ...

..\$ C: num [1:58] 19 2 13 2 0 ...

..\$ D: num [1:58] 5 0 0 1 1 ...

References

Pei, Z., Bini, E. J., Yang, L., Zhou, M., Francois, F., & Blaser, M. J. (2004). Bacterial biota in the human distal esophagus. *Proceedings of the National Academy of Sciences of the United States of America*, 101(12), 4250-4255.

esophagus_tree	<i>phylogenetic tree for esophagus</i>
----------------	--

Description

This dataset includes a phylogenetic tree spanned by 58 OTU listed in the dataset esophagus. Use the `chronos` function in the `ape` package to convert to an ultrametric tree with a fixed tree height of 1.

Usage

```
data(esophagus_tree)
```

Format

`esophagus_tree` is a list with the following phylogenetic information:

A list of 4:

\$ edge: int [1:114, 1:2] 59 59 60 60 61 62 63 63 62 61 ...

\$ edge.length: num [1:114] 1 0.113 0.887 0.102 0.033 ...

\$ Nnode: int 57

\$ tip.label: chr [1:58] "OTU_59_8_22" "OTU_59_5_13" "OTU_59_8_12" ...

gghierPD	<i>ggplot2 extension for an hierPD object</i>
----------	---

Description

`gghierPD`: the `ggplot` extension for `hierPD` object to plot order q against to hierarchical phylogenetic diversity decomposition and dissimilarity measure.

Usage

```
gghierPD(output, type = "A")
```

Arguments

`output` the output from `hierPD`.

`type` selection of plot type :
 (type = "A") for alpha and gamma diversity;
 (type = "B") for beta diversity;
 (type = "D") for dissimilarity measure based on multiplicative decomposition.

Value

a figure for hierarchical phylogenetic diversity decomposition or dissimilarity measure.

Examples

```
data("wetland")
data("wetland_mat")
data("wetland_tree")
hier_output <- hierPD(wetland, mat = wetland_mat, q = seq(0, 2, 0.2), PDtree = wetland_tree)
gghierPD(hier_output, type = "A")
```

ggiNEXTseq

ggplot2 extension for an iNEXT.seq object

Description

ggiNEXTseq: the ggplot extension for [iNEXTseq](#) object to plot coverage- or sample-sized-based rarefaction/extrapolation curves for phylogenetic diversity decomposition and dissimilarity measure.

Usage

```
ggiNEXTseq(output, type = "B")
```

Arguments

output	the output from iNEXTseq.
type	(required only when base = "coverage"), selection of plot type : type = "B" for plotting the gamma, alpha, and beta diversity; type = "D" for plotting 4 turnover dissimilarities.

Value

a figure for phylogenetic diversity decomposition or dissimilarity measure.

Examples

```
data("esophagus")
data("esophagus_tree")
output <- iNEXTseq(esophagus[1], q = c(0,1,2), nboot = 10, PDtree = esophagus_tree)
ggiNEXTseq(output, type = "B")
```

ggObsAsyPD

ggplot2 extension for an ObsAsyPD object

Description

ggObsAsyPD: the ggplot extension for [ObsAsyPD](#) object to plot order q against to phylogenetic diversity decomposition and dissimilarity measure.

Usage

```
ggObsAsyPD(output, type = "B")
```

Arguments

output	the output from ObsAsyPD.
method	selection of plot type : type = "B" for plotting the gamma, alpha, and beta diversity; type = "D" for plotting 4 turnover dissimilarities.

Value

a figure for phylogenetic diversity decomposition or dissimilarity measure.

Examples

```
data("esophagus")
data("esophagus_tree")
ObsAsyPD_output <- ObsAsyPD(esophagus[1], q = seq(0, 2, 0.2), PDtree = esophagus_tree)
ggObsAsyPD(ObsAsyPD_output, type = "B")
```

hierPD

function to calculate hierarchical phylogenetic gamma, alpha, beta diversity and dissimilarity measure

Description

hierPD: function to calculate empirical estimates for hierarchical phylogenetic gamma, alpha, beta diversity and dissimilarity measure.

Usage

```
hierPD(
  data,
  mat,
  q = seq(0, 2, 0.2),
  weight = "size",
  nboot = 10,
  conf = 0.95,
  PDtree,
```

```

    type = "mle",
    decomposition = "relative"
  )

```

Arguments

<code>data</code>	data should be input as a <code>matrix/data.frame</code> (species by assemblages).
<code>mat</code>	hierarchical structure of data should be input as a <code>matrix</code> .
<code>q</code>	a numerical vector specifying the diversity orders. Default is <code>seq(0, 2, 0.2)</code> .
<code>weight</code>	(required only when <code>type = "mle"</code> and <code>decomposition = "relative"</code>) weight for relative decomposition empirical estimate. Select size-weighted (" <code>size</code> "), equal-weighted (" <code>equal</code> ") or a numerical vector for weight. Default is " <code>size</code> ".
<code>nboot</code>	a positive integer specifying the number of bootstrap replications when assessing sampling uncertainty and constructing confidence intervals. Bootstrap replications are generally time consuming. Enter 0 to skip the bootstrap procedures. Default is 10.
<code>conf</code>	a positive number < 1 specifying the level of confidence interval. Default is 0.95.
<code>PDtree</code>	a phylogenetic tree in Newick format for all observed species in the pooled assemblage.
<code>type</code>	estimate type: empirical (<code>type = "mle"</code>) or asymptotic estimate (<code>type = "est"</code>). Default is " <code>mle</code> ".
<code>decomposition</code>	decomposition type: relative (<code>decomposition = "relative"</code>) or absolute decomposition (<code>decomposition = "absolute"</code>). Default is " <code>relative</code> ".

Value

a data frames with hierarchical phylogenetic diversity (gamma, alpha, and beta) and four types dissimilarity measure.

Examples

```

data("wetland")
data("wetland_mat")
data("wetland_tree")
hier_output <- hierPD(wetland, mat = wetland_mat, q = seq(0, 2, 0.2), PDtree = wetland_tree)

```

iNEXTseq	<i>function to calculate phylogenetic gamma, alpha, beta diversity and dissimilarity measure</i>
----------	--

Description

iNEXTseq: function to calculate interpolation and extrapolation for phylogenetic gamma, alpha, beta diversity and dissimilarity measure.

Usage

```
iNEXTseq(
  data,
  q = c(0, 1, 2),
  base = "coverage",
  level = NULL,
  nboot = 10,
  conf = 0.95,
  PDtree = NULL,
  PDreftime = NULL
)
```

Arguments

data	OTU data can be input as a <code>matrix/data.frame</code> (species by assemblages), or a list of <code>matrices/data.frames</code> , each matrix represents species-by-assemblages abundance matrix.
q	a numerical vector specifying the diversity orders. Default is <code>c(0, 1, 2)</code> .
base	sample-sized-based rarefaction and extrapolation for gamma and alpha diversity (<code>base = "size"</code>) or coverage-based rarefaction and extrapolation for gamma, alpha and beta diversity (<code>base = "coverage"</code>). Default is <code>base = "coverage"</code> .
level	A numerical vector specifying the particular value of sample coverage (between 0 and 1 when <code>base = "coverage"</code>) or sample size (<code>base = "size"</code>). <code>level = 1</code> (<code>base = "coverage"</code>) means complete coverage (the corresponding diversity represents asymptotic diversity). If <code>base = "size"</code> and <code>level = NULL</code> , then this function computes the gamma and alpha diversity estimates up to double the reference sample size. If <code>base = "coverage"</code> and <code>level = NULL</code> , then this function computes the gamma and alpha diversity estimates up to one (for <code>q = 1, 2</code>) or up to the coverage of double the reference sample size (for <code>q = 0</code>); the corresponding beta diversity is computed up to the same maximum coverage as the alpha diversity.
nboot	a positive integer specifying the number of bootstrap replications when assessing sampling uncertainty and constructing confidence intervals. Bootstrap replications are generally time consuming. Enter 0 to skip the bootstrap procedures. Default is 10.
conf	a positive number < 1 specifying the level of confidence interval. Default is 0.95.
PDtree	a phylogenetic tree in Newick format for all observed species in the pooled assemblage.
PDreftime	a numerical value specifying reference time for PD. Default is NULL (i.e., the age of the root of PDtree).

Value

If `base = "coverage"`, return a list of seven data frames with three coverage-based diversity (gamma, alpha, and beta) and four types dissimilarity measure. If `base = "size"`, return a list of two data frames with two diversity (gamma and alpha).

Examples

```
data("esophagus")
data("esophagus_tree")
output <- iNEXTseq(esophagus[1], q = c(0,1,2), level = seq(0.5, 1, 0.05), nboot = 10,
  conf = 0.95, PDtree = esophagus_tree, PDreftime = NULL)
```

ObsAsyPD	<i>function to calculate observed or asymptotic phylogenetic gamma, alpha, beta diversity and dissimilarity of order q</i>
----------	--

Description

ObsAsyPD computes observed and asymptotic diversity of order q between 0 and 2 (in increments of 0.2) for phylogenetic gamma, alpha, beta diversity and dissimilarity; these values with different order q can be used to depict a q -profile in the ggObsAsyPD function.

Usage

```
ObsAsyPD(
  data,
  q = seq(0, 2, 0.2),
  weight = "size",
  nboot = 10,
  conf = 0.95,
  PDtree,
  type = "mle",
  decomposition = "relative"
)
```

Arguments

data	OTU data can be input as a matrix/data.frame (species by assemblages), or a list of matrices/data.frames, each matrix represents species-by-assemblages abundance matrix.
q	a numerical vector specifying the diversity orders. Default is seq(0, 2, 0.2).
weight	(required only when type = "mle" and decomposition = "relative") weight for relative decomposition empirical estimate. Select size-weighted ("size"), equal-weighted ("equal") or a numerical vector for weight. Default is "size".
nboot	a positive integer specifying the number of bootstrap replications when assessing sampling uncertainty and constructing confidence intervals. Bootstrap replications are generally time consuming. Enter 0 to skip the bootstrap procedures. Default is 10.
conf	a positive number < 1 specifying the level of confidence interval. Default is 0.95.
PDtree	a phylogenetic tree in Newick format for all observed species in the pooled assemblage.
type	estimate type: empirical (type = "mle") or asymptotic estimate (type = "est"). Default is "mle".

decomposition decomposition type: relative (decomposition = "relative") or absolute decomposition (decomposition = "absolute"). Default is "relative".

Value

a data frames with observed or asymptotic phylogenetic diversity (gamma, alpha, and beta) and four types dissimilarity measure for each dataset.

Examples

```
data("esophagus")
data("esophagus_tree")
ObsAsyPD_output <- ObsAsyPD(esophagus[1], q = seq(0, 2, 0.2), PDtree = esophagus_tree)
```

tongue_cheek	<i>tongue and cheek OTU count data</i>
--------------	--

Usage

```
data("tongue_cheek")
```

Format

A data frame with 2549 observations on the following 2 variables.

Cheek a numeric vector

Tongue a numeric vector

Examples

```
data(tongue_cheek)
```

tongue_cheek_tree	<i>phylogenetic for tongue and cheek data</i>
-------------------	---

Usage

```
data("tongue_cheek_tree")
```

Format

The format is:

List of 4

\$ edge : int [1:25400, 1:2] 12702 12703 12704 12704 12705 12706 12706 12707 12707 12705 ...

\$ edge.length: num [1:25400] 2.20e-03 9.91e-01 6.61e-03 7.87e-05 6.38e-03 ...

\$ Nnode : int 12700

\$ tip.label : chr [1:12701] "OTU_97.15099" "OTU_97.13686" "OTU_97.30326" "OTU_97.26112"

...

- attr(*, "class")= chr "phylo"

- attr(*, "order")= chr "cladewise"

Examples

```
data(tongue_cheek_tree)
```

wetland	<i>wetland OTU count data</i>
---------	-------------------------------

Description

This dataset is from package microeco and collected by An et al. (2019).

Usage

```
data(wetland)
```

Format

wetland is a species-by-assemblages data.frame.
 \$ NE: num 132 12 15 14 12 ...
 \$ NW: num 10 4 0 4 7 9 ...
 \$ NC: num 17 9 1 24 26 ...
 \$ YML: num 0 1 0 10 2 ...
 \$ SC: num 32 37 18 23 153 ...

References

An, J., Liu, C., Wang, Q., Yao, M., Rui, J., Zhang, S., & Li, X. (2019). Soil bacterial community structure in Chinese wetlands. *Geoderma*, 337, 290-299.

wetland_mat	<i>hierarchical structure matrix for wetland</i>
-------------	--

Description

This dataset includes a hierarchical structure matrix in the dataset wetland for calculate hierarchical diversity.

Usage

```
data(wetland_mat)
```

Format

wetland_mat is a 3 x 5 (number of hierarchical layers times number of assemblages) matrix.

wetland_tree	<i>phylogenetic tree for wetland</i>
--------------	--------------------------------------

Description

This dataset includes a phylogenetic tree spanned by 404 OTU listed in the dataset wetland. Use the `chronos` function in the `ape` package to convert to an ultrametric tree with a fixed tree height of 1.

Usage

```
data(wetland_tree)
```

Format

wetland_tree is a list with the following phylogenetic information:

A list of 4:

\$ edge: int [1:806, 1:2] 405 406 407 408 409 410 410 411 412 412 ...

\$ edge.length: num [1:806] 9.99e-16 3.86e-01 5.19e-02 2.13e-01 2.89e-02 ...

\$ Nnode: int 403

\$ tip.label: chr [1:404] "OTU_3829" "OTU_705" "OTU_3570" ...

Index

* datasets

- esophagus, [2](#)
- esophagus_tree, [3](#)
- tongue_cheek, [9](#)
- tongue_cheek_tree, [9](#)
- wetland, [10](#)
- wetland_mat, [10](#)
- wetland_tree, [11](#)

- esophagus, [2](#)
- esophagus_tree, [3](#)

- gghierPD, [3](#)
- ggiNEXTseq, [4](#)
- ggObsAsyPD, [5](#)

- hierPD, [3](#), [5](#)

- iNEXTseq, [4](#), [6](#)

- ObsAsyPD, [5](#), [8](#)

- tongue_cheek, [9](#)
- tongue_cheek_tree, [9](#)

- wetland, [10](#)
- wetland_mat, [10](#)
- wetland_tree, [11](#)