

Package ‘adiv’

August 24, 2018

Type Package

Title Analysis of Diversity

Version 1.2

Date 2018-08-23

Author Sandrine Pavoine

Maintainer Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

Description Includes functions, data sets and examples for the calculation of various indices of biodiversity including species, functional and phylogenetic diversity. Part of the indices are expressed in terms of equivalent numbers of species. It also provides ways to partition biodiversity across spatial or temporal scales (alpha, beta, gamma diversities). In addition to the quantification of biodiversity, ordination approaches are available which rely on diversity indices and allow the detailed identification of species, functional or phylogenetic differences between communities.

Depends R (>= 3.5.0), ade4, adegraphics, adephylo, ape, phylobase

Imports cluster, methods, rgl

Suggests phylosignal

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2018-08-24 13:44:27 UTC

R topics documented:

adiv-package	3
abgevodivparam	3
aptree	7
batcomm	10
betastat	11
birdData	12
CF	14
crossdpcoa	15

decdiv	19
dissABC	22
dissRicotta	25
distinctDis	26
distinctTopo	27
distinctTree	29
distinctUltra	30
distMS	31
divparam	32
dsimcom	34
EH	39
EqRao	40
eveparam	45
evoCA	47
evodiss	49
evodiss_family	51
evodivparam	54
evoNSCA	57
evopca	59
K	62
optimEH	64
orisaved	65
PADDis	66
QE	68
qHdiv	71
randEH	73
rare_Rao	74
RDMCCP16	76
Rentropy	77
rlqESLTP	78
rockfish	81
RP15EI	83
RP15JVS	84
RPP16EE	85
speciesdiv	85
specieseve	87
twoHmax	88
uniqueness	91

Description

Package adiv focuses on the analysis of biodiversity

Details

Package adiv is dedicated to biodiversity indices used in ecology and conservation biology (see e.g. functions `divparam`, `optimEH`, `QE`). It also treats the concepts of species' originality (e.g. `distinctDis`, `distinctTree`, `distinctUltra`) and of dissimilarities between species and communities (e.g. `dissABC`, `dissRicotta`, `dsimcom`). Focus is placed on species, functional and phylogenetic diversity and on the apportionment of biodiversity across spatial and temporal scales (see for notably functions `EqRao`, `EqRS`, `EqRSintra`). Ordination methods are also available to depict in detail biodiversity patterns in space and/or time (e.g. `crossdpcoa`, `evoCA`, `evoNSCA`, `evopca`).

Author(s)

Sandrine Pavoine

Maintainer: Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

Description

Function `abgevodivparam` calculates alpha, beta and gamma components of phylogenetic diversity using parametric indices derived from Tsallis and Hill compositional indices. Alpha is for within-community diversity, beta for between-community diversity and gamma for the diversity of all combined communities.

Usage

```
abgevodivparam(phyl, comm, w = c("evoab", "even", "speciesab"),
method = c("hillCJC", "hillR", "tsallis"), q = 2,
option = c("multiplicative", "additive", "proportional",
"C", "U", "V", "S", "Renyi"), tol = 1e-08)

## S3 method for class 'abgevodivparam'
plot(x, legend = TRUE,
legendposi = "topright", type = "b",
col = if (is.numeric(x)) NULL else 1:nrow(x$div),
lty = if (is.numeric(x)) NULL else rep(1, nrow(x$div)),
pch = if (is.numeric(x)) NULL else 1:nrow(x$div),
ylim1 = range(x$div[c("Alpha", "Gamma"), ]), ylim2 = NULL, ...)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>), or <code>hclust</code> .
<code>comm</code>	a data frame or a matrix typically with communities as rows, species as columns and an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
<code>w</code>	either a numeric vector giving weights for communities (same order as in <code>comm</code>), or a code: one of <code>"even"</code> , <code>"evoab"</code> , and <code>"speciesab"</code> . If several codes are given, only the first one is used. See details.
<code>method</code>	a character with one of the following codes: <code>"tsallis"</code> , <code>"hillR"</code> , or <code>"hillCJC"</code> . See details.
<code>q</code>	a vector with nonnegative value(s) for parameter <code>q</code> . See details.
<code>option</code>	a character code: either <code>"multiplicative"</code> , <code>"additive"</code> or <code>"proportional"</code> . If several codes are given, only the first one is used. Only, with <code>method="hillCJC"</code> , other options are possible: <code>"C"</code> , <code>"U"</code> , <code>"V"</code> , <code>"S"</code> , <code>"Renyi"</code> . See details.
<code>tol</code>	numeric tolerance threshold: values between <code>-tol</code> and <code>tol</code> are considered equal to zero.
<code>x</code>	an object of class <code>abgevodivparam</code> obtained with function <code>abgevodivparam</code> .
<code>legend</code>	a logical. If <code>TRUE</code> a legend is given with the colour, the type of line (etc.) used to define the diversity curve of each diversity level (<code>gamma</code> , <code>alpha</code> , <code>beta</code>).
<code>legendposi</code>	a character that gives the position of the legend to be passed to function <code>legend</code> of the base of R.
<code>type</code>	a character to be passed to the graphic parameter <code>type</code> of functions <code>plot</code> and <code>lines</code> used to draw the diversity curve of each diversity level (<code>gamma</code> , <code>alpha</code> , <code>beta</code>).
<code>col</code>	vector of colours to be passed to the graphic parameter <code>col</code> of functions <code>plot</code> and <code>lines</code> to define the colour of the diversity curve of each diversity level (<code>gamma</code> , <code>alpha</code> , <code>beta</code> , in that order).
<code>lty</code>	vector of type of line (plain, broken etc.) to be passed to the graphic parameter <code>lty</code> of functions <code>plot</code> and <code>lines</code> used to draw the diversity curve of each diversity level (<code>gamma</code> , <code>alpha</code> , <code>beta</code> , in that order).
<code>pch</code>	type of point (open circle, close circle, square etc.) to be passed to the graphic parameter <code>pch</code> of functions <code>plot</code> and <code>lines</code> used to draw the diversity level (<code>gamma</code> , <code>alpha</code> , <code>beta</code> , in that order).
<code>ylim1</code>	a vector with two numerics indicating the range to be used to display <code>alpha</code> and <code>gamma</code> diversity.
<code>ylim2</code>	a vector with two numerics indicating the range to be used to display <code>beta</code> diversity.
<code>...</code>	other parameters can be added and passed to the functions <code>plot</code> and <code>lines</code> used to draw the graphic.

Details

Consider a phylogenetic tree T , b_T the set of branches in T , k a branch, L_k the length of branch k , j a community ($j=1, \dots, m$), a_{jk} the abundance associated with branch k in community j (sum of abundance of all species descending from the branch). q is the parameter that increases with the importance given to abundant species compared to rare species in diversity.

The methods available are: `tsallis` (decomposition of Tsallis diversity into alpha, beta, gamma components adapted here to PD-diversity):

$${}^q\gamma_{evoTsallis} = \left[1 - \sum_{k \in b_T} L_k \left(\sum_{j=1}^m w_j \frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right] / (q - 1)$$

$${}^q\alpha_{evoTsallis} = \sum_{j=1}^m w_j \left[1 - \sum_{k \in b_T} L_k \left(\frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right] / (q - 1)$$

`hillR` (Routledge decomposition of Hill diversity into alpha, beta, gamma components adapted here to PD-diversity):

$${}^q\gamma_{evoHill} = \left[\sum_{k \in b_T} L_k \left(\sum_{j=1}^m w_j \frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right]^{1/(1-q)}$$

$${}^q\alpha_{evoHill-R} = \left[\sum_{j=1}^m w_j \sum_{k \in b_T} L_k \left(\frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right]^{1/(1-q)}$$

`hillCJC` (Chiu et al. (2014) decomposition of PD diversity into alpha, beta, gamma components, see Supplementary material Appendix 2 in Pavoine (2016) for a justification of the formulas):

$${}^q\gamma_{evoHill} = \left[\sum_{k \in b_T} L_k \left(\sum_{j=1}^m w_j \frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right]^{1/(1-q)}$$

$${}^q\alpha_{evoHill-CJC} = \frac{1}{m} \left[\sum_{k \in b_T} L_k \sum_{j=1}^m (w_j)^q \left(\frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right]^{1/(1-q)}$$

Then option "additive" calculates β diversity as $\gamma - \alpha$. Option "proportional" calculates β as $(\gamma - \alpha)/\gamma$. Option "multiplicative" calculates β diversity as γ/α . Only for method="hillCJC", options "C", "U", "V", "S", use the multiplicative option and also calculate one of the transformations introduced by Chiu et al. (2014): indices $1 - C_{qm}$, $1 - U_{qm}$, $1 - V_{qm}$, and $1 - S_{qm}$, respectively. "Renyi" is the ${}^q\text{evo}D_{\text{Renyi}}$ index introduced in Pavoine (2016), see also Supplementary material Appendix 1 in Pavoine (2016).

The weights of the sites (parameter w) can be "even" (even weights), "evoab" (proportional to the summed abundances of all evolutionary units), or "speciesab" (proportional to the summed abundances of all species). Note that if the phylogenetic tree is ultrametric (the distance from any species to the root is constant), then options "evoab" and "speciesab" are equivalent.

Value

If only one value of q is given, a vector with alpha, beta, and gamma diversities. If more than one value of q is given, a list of two objects is returned:

q the vector of values for q
 div a data frame with alpha, beta, gamma calculated for all values of q

Only if `option = "C"`, `"U"`, `"V"`, or `"S"`, the index $1 - C_{qm}$ (for `"C"`), $1 - U_{qm}$ (for `"U"`), $1 - V_{qm}$ (for `"V"`) or $1 - S_{qm}$ (for `"S"`) is also provided in the `div` data frame under the name `"transformed.beta"`.

The function `plot.abgevodivparam` returns a graphic.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- The methodologies and scripts were presented in
 Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.
 using earlier work by
 Havrda, M., Charvat F. (1967) Quantification method of classification processes: concept of structural alpha- entropy. *Kybernetik*, **3**, 30–35
 Hill, M.O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
 Routledge, R.D. (1979) Diversity indices: which ones are admissible? *Journal of Theoretical Biology*, **76**, 503–515.
 Rao, C.R. (1986) Rao's axiomatization of diversity measures. In: Kotz S, Johnson NL, editors. *Encyclopedia of Statistical Sciences*. New York: Wiley and Sons. pp. 614–617.
 Chiu, C.-H., Jost, L., Chao, A. (2014) Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, **84**, 21–44.

See Also

[evodiss](#), [divparam](#)

Examples

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
ab <- batcomm$ab[,phy$tip.label]
abgevodivparam(phy, ab)
plot(abgevodivparam(phy, ab))
abgevodivparam(phy, ab, q=0:4)
plot(abgevodivparam(phy, ab, q=0:4))
```

aptree

Hierarchical Partitioning of Evolutionary and Ecological Patterns in the Organization of Phylogenetically-Structured Species Assemblages

Description

apd performs Hardy (2008)'s test for phylogenetic structure in species abundance distribution;

aptree apportions the diversity (according to index I_α by Pavoine et al. 2009) within one or several communities between evolutionary periods;

plot.aptree displays the phylogenetic tree with vertical lines at each speciation event (limits of the evolutionary periods), the first period starts at the tips and the last one ends at the root node; the phylogenetic tree is pruned retaining only the species present in at least one of the observed communities;

abgaptree provides the apportionment of alpha, beta and gamma diversities between evolutionary periods, according to index I_α by Pavoine et al. 2009;

rtestaptree performs the test of phylogenetic signal in the differences between communities at each evolutionary periods;

plot.rtestaptree displays the phylogenetic tree with vertical lines at each speciation event (limits of the evolutionary periods; see above); a different colour is used for periods where the dissimilarities between communities are different from that expected at random;

tecAptree provides technical information for the apportionment of diversity between evolutionary periods;

pIa calculates the index I_α by Pavoine et al. 2009 within each community.

Usage

```
apd(phy1, comm, wcom = c("even", "speciesab"), nrep = 99,
    alter = "two-sided", tol = 1e-08)

aptree(phy1, comm, exponent = 2, tol = 1e-08)

## S3 method for class 'aptree'
plot(x, col.line = 'blue', ...)

abgaptree(phy1, comm, exponent = 2,
           wcom = c("even", "speciesab"), tol = 1e-08)

rtestaptree(phy1, comm, nrep = 99, alter = "two-sided",
            exponent = 2, wcom = c("even", "speciesab"), tol = 1e-08)

## S3 method for class 'rtestaptree'
plot(x, col.line = c("blue", "red"),
     alpha = 0.05, ...)
```

```
tecAptree(phy1, v = NULL, tol = 1e-08)
```

```
pIa(phy1, comm, exponent = 2, tol = 1e-08)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase), or hclust.
comm	a data frame or a matrix typically with communities as rows, species as columns and presence/absence or an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips. In function aptree, comm can be a vector that provides the presence/absence or an index of abundance within a single community.
wcom	a numeric vector that gives the weight attributed to the community. The weights must be positive and their sum equals 1.
nrep	a numeric that gives the number of permutations.
alter	a character specifying the alternative hypothesis; it must be one of "greater", "less" or "two-sided".
tol	a numeric. If the absolute value of a statistic is less than tol, this statistic is considered equal to zero.
exponent	a numeric that gives the value of parameter a in index I_a .
x	in plot.aptree, x is an object inheriting class aptree obtained with function aptree. In plot.rtestaptree, x is an object inheriting class rtestaptree obtained with function rtestaptree.
col.line	in plot.aptree, col.line is a character which attributes a colour to the vertical lines placed at each speciation event and defining the evolutionary periods. In plot.rtestaptree, it is a vector with two characters. These characters give colours to the vertical lines placed at each speciation event. The first colour is used when the differences between communities at the evolutionary period are not significantly different from random; the second colour is used when they are.
alpha	a numeric: the nominal alpha level for significance (the p-values calculated with function rtestaptree are compared to alpha to determine the result of the test).
...	further arguments passed to or from other methods.
v	either NULL or a vector that provides the presence/absence or an index of abundance of species within a single community.

Details

The approaches developed in these functions rely on a parametric index of phylogenetic diversity named I_a . The parameter a controls the importance given to rare versus abundant species in communities. Index I_a generalizes Rao's quadratic entropy (QE) applied to phylogenetic distances between species (when $a=2$) and Faith's Phylogenetic Diversity index (PD) (when $a=0$). When a tends towards 1, the index is a generalization of the Shannon index of diversity applied to phylogenetic data in addition to abundance data. In Pavoine et al. (2009), we developed this index and demonstrated how it can be used to partition diversity simultaneously across evolutionary periods

in the phylogeny and across spatial (e.g. local communities in a region) and/or time units (e.g. a community investigated yearly).

Value

The function `apd` returns an object of class `randtest` with the results of the test. (see function [randtest](#) in package `ade4`)

The function `aptree` returns a data frame with the evolutionary periods as rows, the communities as columns and the diversity values as entries.

The function `plot.aptree` returns a graph.

The function `abgaptree` returns a data frame with the evolutionary periods as rows, alpha diversity, beta diversity and gamma diversity as columns and the diversity values as entries.

The function `rtestaptree` returns an object of class `krandtest` with the results of the permutation tests. (see function [krandtest](#) in package `ade4`)

The function `plot.rtestaptree` returns a graph.

The function `tecAptree` returns a list. If `v` is `NULL`, the values of the list are:

<code>h</code>	the height at which each period ends;
<code>p.length</code>	period length;
<code>ngroups</code>	number of monophyletic groups per evolutionary period;
<code>list</code>	list of the species per monophyletic group at each evolutionary period;
<code>call</code>	original call.

If `v` contains a vector of presence/absence or abundance, the following object is added in the output:

<code>relab</code>	the relative abundance (sum of species' presences or abundances depending on <code>v</code>) of each monophyletic group at each evolutionary period.
--------------------	---

The function `pIa` returns a data frame with communities as rows and only one column. This column gives, for each community, the value taken by index I_a of phylogenetic diversity developed by Pavoine et al. (2009).

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

with contributions of Stephane Dray.

References

Pavoine, S., Love, M., Bonsall, M.B. (2009) Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: *Sebastes*) in the Southern California Bight. *Ecology Letters*, **12**, 898–908.

See Also

See Also [QE](#)

Examples

```

data(rockfish)
phy <- read.tree(text=rockfish$tre)
ABG <- abgaptree(phy, rockfish$fau, wcom="speciesab")
colSums(ABG)
A <- aptree(phy, rockfish$fau)
colSums(A)
plot(A)
P <- pIa(phy, rockfish$fau)
P
T <- apd(phy, rockfish$fau)
plot(T)
#R <- rtestaptree(phy, rockfish$fau, nrep=999, wcom="speciesab")
#plot(R)
TA <- tecAptree(phy)
TA$h

```

batcomm

Bat Abundance and Phylogeny Along a Disturbance Gradient in a Neotropical Rainforest

Description

The data were collected by Medellín et al. (2000) on bats in four habitats in the Selva Lacandona of Chiapas, Mexico. The phylogeny is Fritz et al. (2009) phylogeny pruned for retaining only the species present in Medellín et al. data set.

Usage

```
data("batcomm")
```

Format

batcomm is a list of two components:

ab, a data frame with species as rows, habitats as columns and abundance of species in habitats as entries.

tre, a phylogenetic tree in newick format for all bat species in ab.

Details

In ab, four habitat types were analyzed: F=rainforest; P=cacao plantations; O=oldfields; C=cornfields (Pavoine 2016).

The resolution of the bat phylogeny is uncertain especially at the older nodes (Pavoine 2016).

Source

www.oikosjournal.org/appendix/oik-03262

References

- Medellin, R. Equihua M., Amin, M.A. (2000) Bat diversity and abundance as indicators of disturbance in Neotropical rainforest. *Conservation Biology*, **14**, 1666–1675.
- Fritz, S.A., Bininda-Emonds, O.R.P., Purvis, A. (2009) Geographic variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, **12**, 538–549.
- Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.

Examples

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
plot(phy)
ab <- batcomm$ab[,phy$tip.label]
plot(abgevodivparam(phy, ab, q=0:4))
```

betastat	<i>Multiple-Site Dissimilarity Measure for Species Presence/Absence Data</i>
----------	--

Description

Functions `betastatjac` and `betastatsor` calculate multiple-site dissimilarity (beta diversity). The first one is derived from Jaccard coefficient of similarity and the second from Sorensen coefficient. These proposed dissimilarity indices can be additively partitioned into species nestedness and turnover.

Usage

```
betastatjac(comm)

betastatsor(comm)
```

Arguments

comm	A data frame typically with communities as rows, species as columns and presence/absence (1/0) as entries.
------	--

Value

The two functions return a vector of 4 values:

beta	Ricotta and Pavoine (2015) β^+ relative measure of additive beta diversity (multiple-site dissimilarity)
betaT	Ricotta and Pavoine (2015) β_T contribution of species turnover to multiple-site dissimilarity

betaN	Ricotta and Pavoine (2015) β_N contribution of species nestedness to multiple-site dissimilarity
sim	Ricotta and Pavoine (2015) $\bar{\beta}^x$ relative measure of multiple-site similarity

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Ricotta, C. and Pavoine, S. (2015) A multiple-site dissimilarity measure for species presence/absence data and its relationship with nestedness and turnover. *Ecological Indicators*, **54**,203–206.

Examples

```
data(RP15EI)
# Scripts used in Figure 1 of Ricotta and Pavoine (2015)
betastatjac(RP15EI$M1)
betastatjac(RP15EI$M2)
betastatjac(RP15EI$M3)
betastatjac(RP15EI$M4)

#see also
betastatsor(RP15EI$M1)
betastatsor(RP15EI$M2)
betastatsor(RP15EI$M3)
betastatsor(RP15EI$M4)
```

birdData

Avian Communities along Successional Forest Gradients

Description

The density of bird species in 3 locations in the Mediterranean region and 2 locations in the central Europe region were collected by Blondel and Farre (1988) along a gradient of habitats. The densities and a phylogenetic tree are provided for all observed species.

Usage

```
data("birdData")
```

Format

birdData is a list of 4 objects:

fau, a data frame with communities as rows, species as columns and densities as entries. The names of the communities start with the first three letters of the location and ends with the seral stage number. See details.

tre, a character. It contains the phylogeny in a newick format.

facA, a factor. It indicates which location each community belongs to.

facB, a factor. It indicates which seral stage each community belongs to.

Details

The locations are Alg = Algeria, Bur = Burgundy, Cor = Corsica Island, Pol = Poland, Pro = Provence.

Three are in the Mediterranean region (Provence, southern France; Corsica Island, southern France; and north east Algeria) and two in the central European region (Burgundy, central France; and Poland).

In each location, a habitat gradient has been conventionally divided into six seral stages (intermediate stages found in forest ecosystems advancing towards their climax stage after a disturbance event) in such a way that all five selected habitat gradients match one another reasonably well in terms of the number, patterns and overall structure of habitats: from low bushy vegetation, less than 1 m height (stage 1), to forests with trees at least 20 m high (stage 6).

The composite phylogenetic tree was obtained based on Davis' supertree: a strict consensus of 2000 trees (see details in Text S5 of Pavoine et al. 2013).

Source

Dataset S1 in Pavoine et al. (2013)

References

Blondel, J. and Farre, H. (1988) The convergent trajectories of bird communities along ecological successions in European forests. *OEcologia (Berlin)*, **75**, 83–93.

Davis, K.E. (2008) *Reweaving the tapestry: a supertree of birds* [PhD thesis]. Glasgow, U.K.: University of Glasgow.

Pavoine, S., Blondel, J., Dufour, A.-B., Gasc, A., Bonsall, M.B. (2013) A new technique for analysing interacting factors affecting biodiversity patterns: crossed-DPCoA. *PloS One*, **8**, e54530.

Examples

```
data(birdData)
phy <- read.tree(text=birdData$tre)
phydis <- sqrt(distTips(phy, method="nNodes")+1)

fau <- birdData$fau[, phy$tip.label]
facA <- birdData$facA
facB <- birdData$facB

cd_mainB <- crossdpcoa_maineffect(fau, facB, facA, phydis, w=rep(1/30, 30), scannf = FALSE)
s.label(cd_mainB$12)
```

Description

The function `CFprop` calculates the matrix *CF* of intra- (diagonal) and inter-specific (off-diagonal) similarities as defined in the main text of Pavoine and Izsak (2014), and matrix *CwF* as defined in Appendix S3 of Pavoine and Izsak (2014) for weighting functional attributes. The function `multiCFprop` calculates matrices *CwmF1*, *CwmF2*, *CwmF3* when several functional traits are considered (Appendix S3 of Pavoine and Izsak 2014). Traits and the attributes of the traits can be weighted. These two functions consider functional traits expressed as proportion (compositional) vectors. The functions `CFbinary` and `multiCFbinary` are the equivalents of `CFprop` and `multiCFprop` when traits are expressed as binary vectors as shown in Appendix S3 of Pavoine and Izsak (2014).

Usage

```
CFbinary(df, wA = rep(1, ncol(df)))
```

```
multiCFbinary(Ktab, w.attributes = lapply(Ktab, function(x) rep(1, ncol(x))),
w.traits = rep(1/length(Ktab), length(Ktab)),
labels = rownames(Ktab[[1]]), solution = c(2, 1))
```

```
CFprop(df, wA = rep(1, ncol(df)))
```

```
multiCFprop(Ktab, w.attributes = lapply(Ktab, function(x) rep(1, ncol(x))),
w.traits = rep(1/length(Ktab), length(Ktab)),
labels = rownames(Ktab[[1]]), solution = c(2, 1))
```

Arguments

<code>df</code>	a data frame or a matrix with species (or any entities of interest) as rows, functional attributes as columns, and proportions (for <code>CFprop</code>) or 0/1 values (for <code>CFbinary</code>) as entries. The sum of the columns will be standardized to equal 1 in function <code>CFprop</code> but not in <code>CFbinary</code> .
<code>wA</code>	a vector of weights that should be given to the attributes (same order as the columns of <code>df</code>).
<code>Ktab</code>	a list of data frames, each of which represent a trait. For a given trait, the data frame should have species (or any entities of interest) as rows, functional attributes as columns, and proportions (for <code>multiCFprop</code>) or 0/1 values (for <code>multiCFbinary</code>) as entries. The sum of the columns will be standardized to equal 1 in function <code>multiCFprop</code> but not in <code>multiCFbinary</code>
<code>w.attributes</code>	a list of weights that should be given to the attributes of each trait. Traits should be in the same order as they appear in the list of tables <code>Ktab</code> . The attributes of a trait should be ordered as the columns of the corresponding data frame in <code>Ktab</code> .

w.traits	a vector of weights that should be given to the traits (same order as the tables of Ktab).
labels	a vector of characters that gives the names of the species (or the other entities of interest; same order as the rows of all tables in Ktab).
solution	a numeric value (either 1 or 2) that indicates which equations are used to summarize the information given by several traits among the 2 approaches given in Appendix S3 of Pavoine and Izsak (2014) page 9. If a vector is given, only the first value of the vector is considered.

Value

A matrix with nonnegative values

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S. and Izsak, J. (2014) New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, **5**, 165–172.

See Also

[qHdiv](#)

crossdpcoa

Crossed-DPCoA

Description

Crossed-DPCoA typically analyze the phylogenetic or functional compositions of communities according to two factors affecting the communities (e.g. space and time; habitat and region)

The function `crossdpcoa_maineffect` obtains the space of DPCoA (the double principal coordinate analysis) where species are placed according to their functional traits or phylogeny and communities are placed at the center of their species. Next, levels of each factor are placed at the center of their communities. The function `crossdpcoa_maineffect` determines the principal axes of the positions of the levels of one of the factors in this space and project species' points on these principal axes. The main effect of the factor named `facA` is analysed by this process (Pavoine et al. 2013).

The function `crossdpcoa_version1` performs version 1 of the crossed DPCoA in Pavoine et al. (2013) where the effect of factor `facA` on the diversity of communities, given factor `facB`, is analysed.

The function `crossdpcoa_version2` performs version 2 of the crossed DPCoA in Pavoine et al. (2013) where the effect of factor `facA` on the diversity of communities, given factor `facB`, is also analysed.

Usage

```
crossdpcoa_maineffect(df, facA, facB, dis = NULL,
  scannf = TRUE, nf = 2, w = c("classic", "independence"),
  tol = 1e-07)
```

```
crossdpcoa_version1(df, facA, facB, dis = NULL,
  scannf = TRUE, nf = 2, w = c("classic", "independence"),
  tol = 1e-07)
```

```
crossdpcoa_version2(df, facA, facB, dis = NULL,
  scannf = TRUE, nf = 2, w = c("classic", "independence"),
  tol = 1e-07)
```

Arguments

df	a data frame or a matrix of 0/1 or nonnegative values (e.g. a communities x species data frame or matrix)
facA	a factor with the same length as the number of rows in df
facB	another factor with the same length as the number of rows in df
dis	an object of class <code>dist</code> that contains the distances among species (e.g. functional or phylogenetic distances). If <code>NULL</code> equidistances are used among species. The distances must have Euclidean properties. Distances are integrated in the Euclidean Diversity Index (Champely and Chessel 2002), which corresponds to a particular formulation of Rao (1982) quadratic entropy. For example the diversity within a community i is $EDI = \sum_{k=1}^S \sum_{l=1}^S p_{k i} p_{l i} \frac{d_{kl}^2}{2}$, where S is the number of species, $p_{k i}$ is the relative abundance of species k within community i ; d_{kl} is the (phylogenetic or functional) dissimilarity between species k and l .
scannf	a logical value indicating whether the screeplot (with eigenvalues) should be displayed.
nf	if <code>scannf</code> is <code>FALSE</code> , an integer indicating the number of kept axes.
w	either a character or a numeric vector of positive values that indicates how the rows of <code>df</code> (e.g. the communities) should be weighted. If <code>w="classic"</code> , the weights are defined from the sum of the values in each row (e.g. sum of all species abundances within a community). If <code>w="independence"</code> , then the weight attributed to a row of <code>df</code> (e.g. a community) is the product of the weight attributed to a level of factor A with the weight attributed to a level of factor B. If a vector of characters is given, only the first one is retained. If numeric, values in <code>w</code> must be in the same order as the rows of <code>df</code> (see Pavoine et al. 2013 for details on the definition of these weights).
tol	a numeric tolerance threshold: a value between <code>-tol</code> and <code>tol</code> is considered as null.

Value

The functions `crossdpcoa_maineffect`, `crossdpcoa_version1` and `crossdpcoa_version2` return a list containing the following information used for computing the crossed-DPCoA:

l1	coordinates of the columns of df (e.g. the species).
l2	coordinates of the levels of factor A.
l3	(for functions crossdpcoa_version1 and crossdpcoa_version2 only) coordinates of the rows of df (e.g. the communities).
eig	the eigenvalues.
lX	the weights attributed to the columns of df (e.g. species).
lA	the weights attributed to the levels of factor A.
lB	the weights attributed to the levels of factor B.
lC	the weights attributed to the rows of df (e.g. communities).
div	a numeric vector with the apportionment of Rao's quadratic diversity (APQE).
call	the call function.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S., Blondel, J., Dufour, A.-B., Gasc, A., Bonsall, M.B. (2013) A new technique for analysing interacting factors affecting biodiversity patterns: crossed-DPCoA. *PLoS One*, **8**, e54530.

Examples

```

data(birdData)
phy <- read.tree(text=birdData$tre)
phydis <- sqrt(distTips(phy, method="nNodes")+1)

fau <- birdData$fau[, phy$tip.label]
facA <- birdData$facA
facB <- birdData$facB

#Here factor B is put first to analyze
#the main effect of the strata:
cd_mainB <- crossdpcoa_maineffect(fau, facB, facA, phydis, w=rep(1/30, 30), scannf = FALSE)
barplot(cd_mainB$eig)
cd_mainB$eig[1:2]/sum(cd_mainB$eig)

#Positions of the levels of factor B on its principal axes:
s.label(cd_mainB$l2)
# The "d" value on graphs indicates the length of the edge of a grid cell (scale of the graphic).

#The coordinates of the species on the same axes
#can be displayed in front of the phylogeny:
mainB11.4d <- phylo4d(phy, as.matrix(cd_mainB$l1))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(mainB11.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE)
par(mar=parmar)

```

```

#If factor A is put first, the analysis focus
#on the main effect of the region:
cd_mainA <- crossdpcoa_maineffect(fau, facA, facB, phydis, w=rep(1/30, 30), scannf = FALSE)
barplot(cd_mainA$eig)
cd_mainA$eig[1:2]/sum(cd_mainA$eig)

#Positions of the levels of factor A on its principal axes:
s.label(cd_mainA$l2)
# The "d" value on graphs indicates the length of the edge of a grid cell (scale of the graphic).

#The coordinates of the species on the same axes
# can be displayed in front of the phylogeny:
mainA1.4d <- phylo4d(phy, as.matrix(cd_mainA$l1))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(mainA1.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE)
par(mar=parmar)

#Crossed DPCoA Version 1
cd_v1 <- crossdpcoa_version1(fau, facA, facB, phydis, w=rep(1/30, 30), scannf = FALSE)
#Proportion of SS(A) expressed by the two first axes:
cd_v1$eig[1:2]/sum(cd_v1$eig)
#To view the positions of the locations on the first two axes, write:
s.label(cd_v1$l2)
#To view the positions of all communities on the first two axes, write:
s.label(cd_v1$l3)
#To view the positions of the species on the first two axes in front of the phylogeny, write:
v1l1.4d <- phylo4d(phy, as.matrix(cd_v1$l1))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(v1l1.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE)
par(mar=parmar)

#Crossed DPCoA Version 2
#Crossed DPCoA version 2 can now be performed as follows:
cd_v2 <- crossdpcoa_version2(fau, facA, facB, phydis, w=rep(1/30, 30), scannf = FALSE)
#Proportion of variation among levels of factor A
#in the subspace orthogonal to the principal axes of B
#expressed by the two first axes:
cd_v2$eig[1:2]/sum(cd_v2$eig)
#To view the positions of the locations on the first two axes, write:
s.label(cd_v2$l2)
#To view the positions of all communities on the first two axes, write:
s.label(cd_v2$l3)
#To view the positions of the species on the first two axes in front of the phylogeny, write:
v2l1.4d <- phylo4d(phy, as.matrix(cd_v2$l1))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(v2l1.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE)
par(mar=parmar)

```

decdiv	<i>Decomposition of trait-based diversity along the nodes of a phylogenetic tree</i>
--------	--

Description

The function `decdiv` calculates trait-based differences between the lineages that descend from a node of a phylogenetic tree in one or several communities (using presence/absence or abundance data).

The function `plot.decdiv` plots the result of function `decdiv` for one of the communities.

The function `rtestdecdiv` tests, for one community (with presence/absence or abundance data), if the representation of trait diversity on the phylogenetic tree highlights a nonrandom pattern.

Usage

```
decdiv(phy1, comm, dis = NULL, tol = 1e-08, option = 1:5,
       formula = c("QE", "EDI"))

## S3 method for class 'decdiv'
plot(x, ncom = 1, col = "black",
     csize = 1, legend = TRUE, ...)

rtestdecdiv(phy1, vecab, dis = NULL, tol = 1e-08,
            option = 1:5, formula = c("QE", "EDI"),
            vranking = c("complexity", "droot"),
            ties.method = c("average", "first", "last", "random",
                           "max", "min"), statistic = 1:3, optiontest = NULL, nrep = 99)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>) or <code>hclust</code> with species as tips. To ease the interpretation of the results, I advise you to add labels to the nodes of the phylogeny. For example, if your phylogeny is named <code>'tree'</code> and is of class <code>phylo</code> , you can name the nodes with the following command: <code>tree\$node.label <- paste("n", 1:tree\$Nnode, sep="")</code> . Use <code>plot(tree, show.node.label=TRUE)</code> to see the result. If it is of class <code>phylo4</code> , you can use: <code>nodeLabels(tree) <- paste("n", 1:nNodes(tree), sep="")</code> .
<code>comm</code>	a vector with species presence/absence or an index of abundance as entries, or a data frame or a matrix typically with communities as rows, species as columns and presence/absence or an index of abundance as entries. Species names in <code>comm</code> must be similar as in <code>phy1</code> . Even if <code>comm</code> is a vector, each entry needs to have a species name (use function name to attributes names to a vector).
<code>dis</code>	either <code>NULL</code> or an object of class <code>dist</code> that contains the trait-based distances among species. If <code>NULL</code> , the Gini-Simpson index is used.
<code>tol</code>	a tolerance threshold (a value between <code>-tol</code> and <code>tol</code> is considered equal to zero)

option	a numeric (either 1, 2, 3, 4 or 5) indicating the option to use to calculate the contribution of each node of the phylogenetic tree to trait-based diversity. See details. If several values are given, the function retains only the first one.
formula	either "QE" (default) or "EDI". See details. If several values are given, the function retains only the first one.
x	an object of class <code>decdiv</code>
ncom	if <code>comm</code> is a matrix, the number of its row corresponding to the focal community for which the results of <code>decdiv</code> will be plotted. Ignored if <code>comm</code> is a vector.
col	the color of circles displayed at each node.
csize	a positive numeric giving the scale for plotting the circle at each node. 1 is the default size; if zero, no circle is drawn.
legend	a logical indicating whether the legend for the circle size needs to be displayed.
...	further arguments that can be specified to the internal use of function <code>plot.phylo</code> .
vecab	a numeric vector giving the presence/absence(1/0) or abundance(non-negative value) of species in a community.
vranking	a character with 2 possible values: either "complexity" for a ranking according to the complexity of the subtree rooted on each interior node (see Pavoine et al. 2010 for an explanation), or "droot" for ordering interior nodes by the distance between them and the root node of the tree.
ties.method	a character to be passed to function <code>rank</code> of the base of R. It can be one of "average", "first", "last", "random", "max", "min".
statistic	a numeric value or a vector of numeric values. Possible values are 1, 2, or 3. They correspond to the three statistics S1, S2 and S3, respectively, developed by Pavoine et al. (eqs. 5 to 7 in Pavoine et al. 2010).
optiontest	a vector of character strings specifying the alternative hypothesis of each test, which must be one of "greater", "less" or "two-sided". If null, then <code>statistic=1</code> is associated with "greater" and <code>statistic=2</code> and <code>=3</code> with "two-sided". See function <code>link{as.randtest}</code> of package <code>ade4</code> for details on these alternatives. The length of the vector given to <code>optiontest</code> must be equal to that given to <code>statistic</code> .
nrep	the number of permutations to be done in each permutation test.

Details

The function `decdiv` relies on Rao's (1982) quadratic entropy (QE) to measure the trait-based diversity of a set of species. Two formulas for QE have been introduced in the literature one is the original formula by Rao (1982) (which corresponds to `formula = "QE"`) and the other one introduced by Champely and Chessel (2002), named Euclidean Diversity Index (which corresponds to `formula = "EDI"`). See function [QE](#) for more details.

In function `decdiv`, each node has a weight proportional to the summed relative abundance of its descending species (or to the relative number of descending species if presence/absence data are used).

With `option = 1`, the function `decdiv` apportions trait-based diversity across the nodes of a phylogenetic tree using the algorithm defined in Pavoine et al. (2010). In that case the value at a given

node is equal to the node weight times a measure of beta trait-based diversity between the lineages that descend from the node. The sum of all values attributed to the nodes of a phylogeny is equal to the total trait-based diversity of the species (tips of the phylogeny). In case of dichotomic trees, only two lineages descend from a given node. Here I consider a more general case where more than two lineages may descend from a node (polytomy). The beta trait-based diversity among the lineages that descend from a node is measured here as the average trait-based dissimilarity between any two of these descending lineages. With `option = 1`, the trait-based dissimilarity between two lineages is measured by Rao's DISC index (gamma diversity [average trait-based dissimilarity between any two species descending from the node] - alpha diversity [average trait-based dissimilarity between any two species descending from one of the lineages branched to the node]).

In the present version of function `decdiv`, I have added other options. Options 2 and 3 code different ways of measuring trait-based differences between lineages, standardized between 0 and 1: With `option = 2`, the formula is $(\text{gamma} - \text{alpha}) / (1 - \text{alpha}) * M / (M - 1)$, where gamma and alpha are defined above (for `option = 1`) and M is the number of lineages descending from the node.

With `option = 3`, the formula is $(\text{gamma} - \text{alpha}) / (1 - \text{gamma}) / (M - 1)$, where gamma and alpha are defined above (for `option = 1`) and M is the number of lineages descending from the node.

Options 4 and 5 decompose the result given by `option = 1`. `option = 4` returns gamma minus alpha (a measure of beta functional diversity between the lineages that descend from a node). `option = 5` returns node weights the summed relative abundance of its descending species (or to the relative number of descending species if presence/absence data are used).

Values for `option` different from 1 needs that values in `dis` (the trait-based dissimilarities between species) are bounded between 0 and 1 if `formula = QE` or `sqrt(2)` if `formula = EDI`. If they are not bounded, the function `decdiv` will bound them using the maximum observed value in `dis`.

The argument `ties.method` in `rtestdecdiv` allows you to explicitly take into account potential ties when ranking nodes according to their complexity or their distance to root.

Value

Function `decdiv` returns a matrix with nodes of the phylogenetic tree as rows and the decomposition of trait-based diversity in communities as columns. An attributes of this matrix is the phylogenetic tree (of class `phylo` with specified names for nodes. If the nodes of `phy1` had no names, the function `decdiv` automatically attributed names.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Champely, S. and Chessel, D. (2002) Measuring biological diversity using Euclidean metrics. *Environmental and Ecological Statistics*, **9**, 167–177.
- Pavoine, S., Baguette, M., Bonsall, M.B. (2010) Decomposition of trait diversity among the nodes of a phylogenetic tree. *Ecological Monographs*, **80**, 485–507.
- Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.

Examples

```

data(ungulates)
ung.phy <- read.tree(text=ungulates$tre)
ung.phy$node.label
plot(ung.phy, show.node.label=TRUE)
# Regarding traits, we log-transformed the first three traits
# measuring volumes
# and we standardized all the traits (mean=0; variance=1).
tab <- cbind.data.frame(afbw = log(ungulates$tab$afbw),
mnw = log(ungulates$tab$mnw), fnw = log(ungulates$tab$fnw),
ls = unguates$tab$ls)
ung.tab0 <- data.frame(scalewt(tab))
ung.tab0 <- data.frame(scalewt(log(ungulates$tab)))
ung.pres <- rep(1, nrow(ung.tab0))
names(ung.pres) <- rownames(ung.tab0)
ung.dec1 <- decdiv(ung.phy, ung.pres, dist(ung.tab0),
option=1, formula = "EDI")
plot.decdiv(ung.dec1)

ung.dec2 <- decdiv(ung.phy, ung.pres, dist(ung.tab0),
option=2, formula = "EDI")
plot.decdiv(ung.dec2)

ung.dec3 <- decdiv(ung.phy, ung.pres, dist(ung.tab0),
option=3, formula = "EDI")
plot.decdiv(ung.dec3)

ung.dec4 <- decdiv(ung.phy, ung.pres, dist(ung.tab0),
option=4, formula = "EDI")
plot.decdiv(ung.dec4)

ung.dec5 <- decdiv(ung.phy, ung.pres, dist(ung.tab0),
option=5, formula = "EDI")
plot.decdiv(ung.dec5)

```

dissABC

Phylogenetic and Functional Similarity between Communities

Description

Coefficients of similarity between communities that rely on the presence/absence of species are generally based on various combinations of the matching/mismatching components of the classical 2 x 2 contingency table. Three of these components are: a=the number of species shared by the two communities; b=the number of species in the first community that are not in the second; c=the number of species in the second community that are not in the first. These coefficients are extended in dissABC to include phylogenetic or functional information on species (Ricotta and Pavoine 2015).

Usage

```
dissABC(comm, dis, option = 1:4, method = c("J", "S", "O", "K", "SS"))
```

Arguments

comm	a data frame or a matrix typically with communities as rows, species as columns and relative abundance or absolute abundance as entries. Column labels (species names) should be assigned as in the object <code>dis</code> .
dis	a matrix (or data frame) of (phylogenetic or functional) dissimilarities among species rescaled in the range [0, 1] or an object of class <code>dist</code> containing these dissimilarities [obtained by functions like <code>vegdist</code> in package <code>vegan</code> , <code>gowdis</code> in package <code>FD</code> , or <code>dist.ktab</code> in package <code>ade4</code> for functional dissimilarities, or functions like <code>cophenetic.phylo</code> in package <code>ape</code> or <code>distTips</code> in package <code>ade4phylo</code> for phylogenetic dissimilarities]. Note that dissimilarities among species need first to be rescaled in the range [0, 1]. If the dissimilarities are outside the range 0-1, a warning message is displayed.
option	a numeric, either 1, 2, 3, or 4 (if several characters are given only the first one is considered). See details.
method	a character, either "J", "S", "O", "K", or "SS" (if several characters are given only the first one is considered). See details.

Details

To obtain the dissimilarities among plots, one needs to choose the equations to be used for the (phylogenetic or functional) components A, B, and C thanks to parameter `option` and the way the parameters will be combined, thanks to parameter `method`.

Let $\mathbf{D} = (d_{ij})$ a matrix of (functional, morphological or phylogenetic) dissimilarities between pairs of species with $d_{ij} = d_{ji}$ and $d_{ii} = 0$. If the dissimilarity coefficient d is in the range [0, 1], it is possible to define a corresponding similarity coefficient s as the complement of d : $s = 1 - d$. Let x_{ik} the abundance of species i in community k . $S(kh)$ is the number of species in the pooled communities k and h (i.e. the species for which $\min\{x_{ik}, x_{ih}\} > 0$). The (absolute) abundance of species similar to i in plot k is

$$Z_{ik} = \sum_{j=1}^{S(kh)} x_{jk} s_{ij}$$

If `option=1`, equations 6-8 of the main text of Ricotta and Pavoine (2015) are used for calculating components A, B, C:

$$A = \sum_{i=1}^{S(kh)} \min\{Z_{ik}, Z_{ih}\}$$

$$B = \sum_{i=1}^{S(kh)} (\max\{Z_{ik}, Z_{ih}\} - Z_{ih})$$

$$C = \sum_{i=1}^{S(kh)} (\max\{Z_{ik}, Z_{ih}\} - Z_{ik})$$

If `option=2`, equations A1-A3 from Appendix S1 of Ricotta and Pavoine (2015) are used.

If `option=3`, equations A5-A7 from Appendix S1 of Ricotta and Pavoine (2015) are used.

If option=4, equations A10-A12 from Appendix S1 of Ricotta and Pavoine (2015) are used.

If method="J"=the Jaccard index is used:

$$\frac{A}{A + B + C}$$

If method="S"=the Sorensen index is used:

$$\frac{2A}{2A + B + C}$$

If method="O"=the Ochiai index is used:

$$A/(\sqrt{A+B}\sqrt{A+C})$$

If method="K", the Kulczynski index is used:

$$\frac{1}{2} \left(\frac{A}{A+B} + \frac{A}{A+C} \right)$$

If method="SS", the Sokal-Sneath index is used:

$$\frac{A}{A + 2B + 2C}$$

Value

Function dissABC returns a matrix with the values of the proposed ISR-based similarities among communities.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Ricotta, C. and Pavoine, S. (2015) Measuring similarity among plots including similarity among species: an extension of traditional approaches. *Journal of Vegetation Science*, **26**, 1061–1067.

See Also

See Also [discomQE](#), [evodiss](#)

Examples

```
data(RP15JVS)
dissABC(RP15JVS$ab, RP15JVS$D1, method="J", option=1)
J <- as.matrix(dissABC(RP15JVS$ab, RP15JVS$D1, method="J", option=1))[, 1]
SS <- as.matrix(dissABC(RP15JVS$ab, RP15JVS$D1, method="SS", option=1))[, 1]
S <- as.matrix(dissABC(RP15JVS$ab, RP15JVS$D1, method="S", option=1))[, 1]
O <- as.matrix(dissABC(RP15JVS$ab, RP15JVS$D1, method="O", option=1))[, 1]
K <- as.matrix(dissABC(RP15JVS$ab, RP15JVS$D1, method="K", option=1))[, 1]
```



```

plot(1:9, J,
     xlab="Number of the plots which plot 1 is compared to",
     ylab="Similarity", type="b", ylim=c(0,1), pch=18)
lines(1:9, SS, type="b", pch=15)
lines(1:9, S, type="b", pch=17)
lines(1:9, O, type="b", pch=12)
lines(1:9, K, type="b", pch=1)
legend("bottomleft",
       c("Jaccard", "Sokal-Sneath", "Sorensen", "Ochiai", "Kulczynski"),
       pch=c(18,15,17,12,1), lty=1)

```

dissRicotta

Plot-to-Plot Functional or Phylogenetic Dissimilarity

Description

The function calculates plot-to-plot functional or phylogenetic dissimilarity based on index symmetry (Index D_{AB} in Ricotta et al. 2015)

Usage

```
dissRicotta(comm, dis)
```

Arguments

comm	a matrix of the relative or absolute abundance of species in communities. Columns are species and communities are rows. Column labels (species names) should be assigned as in dis
dis	a matrix of (functional or phylogenetic) dissimilarities rescaled in the range [0, 1] or an object of class <code>dist</code> [obtained by functions like <code>vegdist</code> in package <code>vegan</code> , <code>gowdis</code> in package <code>FD</code> , or <code>dist.ktab</code> in package <code>ade4</code> for functional dissimilarities, or functions like <code>cophenetic.phylo</code> in package <code>ape</code> or <code>distTips</code> in package <code>adephylo</code> for phylogenetic dissimilarities].

Value

The function returns a semi-matrix of class `dist` with the values of the proposed dissimilarities for each pair of plots. Note that dissimilarities among species need first to be rescaled in the range [0, 1]. If the dissimilarities are outside the range 0-1 (as it is usually the case in phylogenetic studies for instance), a warning message is displayed

Author(s)

Giovanni Bacaro and Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Ricotta, C., Bacaro, G., Pavoine, S. (2015) A cautionary note on some phylogenetic dissimilarity measures. *Journal of Plant Ecology*, **8**, 12–16.

Examples

```
# Phylogenetic tree
s<-"test(((v:20,w:20):10,(x:20,y:20):10):15,z:45):5;"
plot(test <- read.tree(text=s))

# Phylogenetic distances among species
tdist <- cophenetic(test)/100

# Matrix of abundances of the species in four communities;
# communities A and C are identical;
# communities B and D are identical;
comm <- t(data.frame(A = rep(0.2, 5),
  B = c(0.1, 0.2, 0.2, 0, 0.5), C = rep(0.2, 5),
  D = c(0.1, 0.2, 0.2, 0, 0.5), row.names = letters[22:26]))

# Index DAB
dissRicotta(comm, tdist)
```

distinctDis

Dissimilarity-based Species' Originality

Description

The function calculates three indices of species' originality.

Usage

```
distinctDis(dis, method = 1:3, standardized = FALSE)
```

Arguments

dis	an object of class <code>dist</code> containing pair-wise (functional or phylogenetic) dissimilarities between species.
method	a numeric or a vector of numerics. Possible values are 1, 2 and 3. 1 is for Pavoine et al. (2017) index Rb; 2 is for AV, the average dissimilarity between a species and all others in a set (Eiswerth and Haney 1992; Ricotta 2004); 3 is for FV, the average dissimilarity between a species and any other (including the focal species itself) (Schmera et al. 2009).
standardized	a logical. If TRUE, the vector of originalities is divided by its sum (transforming absolute originalities into relative originalities).

Value

A data frame with species as rows and originality indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Eiswerth, M.E. and Haney, J.C. (1992) Allocating conservation expenditures: accounting for inter-species genetic distinctiveness. *Ecological Economics*, **5**, 235–249.
- Ricotta, C. (2004) A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. *Diversity and Distributions*, **10**, 143–146.
- Schmera, D., Podani, J., Eros, T. (2009) Measuring the contribution of community members to functional diversity. *Oikos*, **118**, 961–971.
- Pavoine, S., Bonsall, M.B., Dupaix, A., Jacob, U., Ricotta, C. (2017) From phylogenetic to functional originality: guide through indices and new developments. *Ecological Indicators*, **82**, 196–205.

See Also

See Also [distinctTopo](#), [distinctTree](#), [distinctUltra](#)

Examples

```
e <- rlnorm(10)
e <- sort(e)
names(e) <- paste("s", 1:10, sep="")
d <- dist(e)
barplot(e)

D <- distinctDis(d, standardized = TRUE)

par(mfrow=c(3,1))
plot(e, D[,1], xlab="trait", ylab="Rb")
plot(e, D[,2], xlab="trait", ylab="AV")
plot(e, D[,3], xlab="trait", ylab="FV")

plot(D[,1], D[,2], xlab="Rb", ylab="AV")
plot(D[,1], D[,3], xlab="Rb", ylab="FV")
plot(D[,2], D[,3], xlab="AV", ylab="FV")
par(mfrow=c(1,1))
```

distinctTopo

Topology-based Species' Originality

Description

The function calculates three indices of species' originality that rely on the topology of (phylogenetic) trees.

Usage

```
distinctTopo(phy1, method = 1:3, standardized = FALSE)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase), or hclust.
method	a numeric or a vector of numerics. Possible values are 1, 2 and 3. 1 is for Vane-Wright et al. index (Vane-Wright et al. 1991) (noted VW); 2 is for May index (May 1990) (noted M); 3 is for Pavoine et al. index of originality derived from Abouheif (1999) (Pavoine et al. 2008) (noted A).
standardized	a logical. If FALSE, the minimum score is scaled to 1 as in May (1990). If TRUE, the vector of originalities is divided by its sum (transforming absolute originalities into relative originalities).

Value

A data frame with species as rows and originality indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>
with contributions of Stephane Dray

References

- May, R.M. (1990) Taxonomy as destiny. *Nature*, **347**, 129–130.
- Vane-Wright, R.I., Humphries, C.J., Williams, P.H. (1991) What to protect? Systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.
- Abouheif, E. (1999) A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research*, **1**, 895–909.
- Pavoine, S., Ollier, S., Pontier, D., Chessel, D. (2008) Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theoretical Population Biology*, **73**, 79–91.

See Also

See Also [distinctDis](#), [distinctTree](#), [distinctUltra](#)

Examples

```
data(carni70, package = "adephylo")
tre <- read.tree(text=carni70$tre)
U <- distinctTopo(tre, standardized = TRUE)
U.4d <- phylo4d(tre, as.matrix(U))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(U.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE, cex.symbol=2)
par(mar=parmar)

## Not run:
if(requireNamespace("phylosignal", quietly = TRUE)) {
dotplot.phylo4d(U.4d, center = FALSE, scale = FALSE)
```

```

}

## End(Not run)

```

distinctTree	<i>Tree-based Species' Originality</i>
--------------	--

Description

The function calculates two indices of species' originality that rely on the structure and branch lengths of (phylogenetic) trees.

Usage

```
distinctTree(phy1, method = c("ED", "ES"), standardized = FALSE)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust. Polytomies are allowed.
method	a character a vector of characters. Possible values are "ED" and "ES". "ED" is for the evolutionary distinctiveness, also named fair-proportion, index (Redding 2003; Isaac et al. 2007); "ES" is for the Equal-Splits index (Redding and Mooers 2006).
standardized	a logical. If TRUE, the vector of originalities is divided by its sum (transforming absolute originalities into relative originalities).

Value

A data frame with species as rows and originality indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Redding, D.W. (2003) *Incorporating genetic distinctness and reserve occupancy into a conservation prioritisation approach*. Master thesis: University of East Anglia, Norwich.

Redding, D.W., Mooers, A.O. (2006) Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, **20**, 1670–1678.

Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C., Baillie, J.E. (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE*, **2**, e296.

See Also

See Also [distinctDis](#), [distinctTopo](#), [distinctUltra](#)

Examples

```

data(carni70, package = "adephylo")
tre <- read.tree(text=carni70$tre)
U <- distinctTree(tre, standardize = TRUE)
U.4d <- phylo4d(tre, as.matrix(U))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(U.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE, cex.symbol=2)
par(mar=parmar)

## Not run:
if(requireNamespace("phylosignal", quietly = TRUE)) {
dotplot.phylo4d(U.4d, center = FALSE, scale = FALSE)
}

## End(Not run)

```

distinctUltra

Ultrametric Tree-based Species' Originality

Description

The function calculates two indices of species' originality that rely on the structure and branch lengths of ultrametric (phylogenetic) trees.

Usage

```
distinctUltra(phy1, method = 1:2)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase), or hclust, with ultrametric properties.
method	a numeric or a vector of numerics. Possible values are 1, and 2. 1 is for the Pavoine et al. QE-based (also named Qb) index (Pavoine et al. 2005); 2 is for Pavoine et al. 2H-based index (which could also be named more shortly as 2Hb) (vector that maximizes index 2H , Pavoine and Izsak 2014).

Value

A data frame with species as rows and originality indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>
with contributions of Stephane Dray

References

- Pavoine, S., Ollier, S., Dufour, A.B. (2005) Is the originality of a species measurable? *Ecology Letters*, **8**, 579–586.
- Pavoine, S. and Izsak, J. (2014) New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, **5**, 165–172.

See Also

See Also [distinctDis](#), [distinctTopo](#), [distinctTree](#)

Examples

```
data(carni70, package = "adephylo")
tre <- read.tree(text=carni70$tre)
U <- distinctUltra(tre)
U.4d <- phylo4d(tre, as.matrix(U))
parmar <- par()$mar
par(mar=rep(.1,4))
table.phylo4d(U.4d, show.node=FALSE, symbol="squares", center=FALSE, scale=FALSE, cex.symbol=2)
par(mar=parmar)

## Not run:
if(requireNamespace("phylosignal", quietly = TRUE)) {
  dotplot.phylo4d(U.4d, center = FALSE, scale = FALSE)
}

## End(Not run)
```

distMS	<i>Marczewski-Steinhaus Coefficient</i>
--------	---

Description

The function calculates the Marczewski-Steinhaus coefficient of dissimilarity between pairs of entities (e.g. communities)

Usage

```
distMS(comm, diag = FALSE, upper = FALSE)
```

Arguments

- | | |
|-------|---|
| comm | a data frame or a matrix of nonnegative values (e.g. abundance of species (columns) within communities (rows) to obtain dissimilarities between communities). |
| diag | a logical value indicating whether the diagonal of the distance matrix should be printed by function <code>print.dist</code> . |
| upper | a logical value indicating whether the upper triangle of the distance matrix should be printed by function <code>print.dist</code> . |

Value

an object of class `dist`

Note

This function is a modification of function `dist.quant` from library `ade4` where other dissimilarity coefficients can be found.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Orloci, L. (1978) *Multivariate Analysis in Vegetation Research*. The Hague: Junk.
- Legendre, P. and Legendre, L. (1998) *Numerical Ecology*. Amsterdam: Elsevier.
- Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B.E., Pavoine, S. (2016). Measuring the functional redundancy of biological communities: a quantitative guide. *Methods in Ecology and Evolution*, **7**, 1386–1395.

Examples

```
data(birdData)
distMS(birdData$fau)
```

divparam
Parametric Indices of Species Diversity

Description

The function `divparam` calculates parametric diversity indices. The parameter controls the relative importance given to rare versus abundant species in a community. The function `plot.divparam` plots the results of function `divparam`.

Usage

```
divparam(comm, method = c("hill", "tsallis", "renyi"), q = 2, tol = 1e-08)
```

```
## S3 method for class 'divparam'
plot(x, legend = TRUE,
     legendposi = "topright", type = "b",
     col = if (is.numeric(x)) NULL
     else sample(colors(distinct = TRUE), nrow(x$div)),
     lty = if (is.numeric(x)) NULL else rep(1, nrow(x$div)),
     pch = if (is.numeric(x)) NULL else rep(19, nrow(x$div)),
     ...)
```


Arguments

comm	a data frame or a matrix typically with communities as rows, species as columns and abundance as entry.
method	a character: either "hill" for the Hill numbers (Hill 1973), "tsallis" for the Tsallis or HCDT entropy (Harvda and Charvat 1967; Daroczy 1970; Tsallis 1988), or "renyi" for Renyi's entropy (Renyi 1960).
q	a positive numeric or a vector of positive numerics that give values for the q parameter.
tol	numeric tolerance threshold: values between -tol and tol are considered equal to zero.
x	an object of class divparam obtained with function divparam.
legend	a logical. If TRUE a legend is given with the colour, the type of line (etc.) used to define the diversity curve of each community.
legendposi	a character or a numeric that gives the position of the legend to be passed to function legend of the base of R.
type	a character to be passed to the graphic parameter type of functions plot and lines used to draw the diversity curve of each community.
col	colours to be passed to the graphic parameter col of functions plot and lines to define the colour of the diversity curve of each community.
lty	type of line (plain, broken etc.) to be passed to the graphic parameter lty of functions plot and lines used to draw the diversity curve of each community.
pch	type of point (open circle, close circle, square etc.) to be passed to the graphic parameter pch of functions plot and lines used to draw the diversity curve of each community.
...	other parameters can be added and passed to the functions plot and lines used to draw the graphic.

Value

Function `divparam`, if only one value of `q` is given, returns a vector with communities' diversities. If more than one value of `q` is given, a list of two objects is returned:

<code>q</code>	the vector of values for <code>q</code>
<code>div</code>	a data frame with communities' diversities calculated for all values of <code>q</code>

The function `plot.divparam` returns a graphic.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Daroczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.
- Havrda, M., Charvat, F. (1967) Quantification method of classification processes: concept of structural alpha-entropy. *Kybernetika*, **3**, 30–35.
- Hill, M.O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
- Renyi, A. (1960) On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547–561.
- Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 480–487.

Examples

```
data(batcomm)
ab <- batcomm$ab
plot(divparam(ab))
plot(divparam(ab, q=0:4))
```

dsimcom

Functional or Phylogenetic Similarity between Species and Communities

Description

The R function `dsimcom` calculates Pavoine and Ricotta (2014) coefficients SSokal-Sneath, SJaccard, SSorensen, SOchiai, and Sbeta of similarities among communities or their complements (1-S, dissimilarities).

Function `SQ` calculates Pavoine and Ricotta (2014) index S_Q (similarities between communities) and its complements $1 - S_Q$ (dissimilarities between communities).

The functions `dsimTax` and `dsimTree` calculate pair-wise taxonomic and phylogenetic (dis)similarities between species, respectively. `dsimTree` can also be used to calculate pair-wise functional (dis)similarities between species if a functional dendrogram is used to describe species.

Function `dsimFun` calculates pair-wise functional (dis)similarities between species.

Usage

```
dsimcom(comm, Sigma = NULL, method = 1:5,
option = c("relative", "absolute"),
type = c("similarity", "dissimilarity"))
```

```
SQ(comm, Sigma = NULL, type = c("similarity", "dissimilarity"))
```

```
dsimTaxo(tax, method = c(1, 2, 3, 4, 5),
type = c("similarity", "dissimilarity"))
```

```
dsimTree(phy1, method = c(1, 2, 3, 4, 5), rootedge = NULL,
type = c("similarity", "dissimilarity"))
```

```
dsimFun(df, vartype = c("Q", "N", "M", "P"), method = 1:5,
type = c("similarity", "dissimilarity"))
```

Arguments

comm	a data frame or matrix with communities as rows, species as columns and non-negative values as entries.
Sigma	matrix of similarities among species (species as rows and columns in the same order as in comm; values in Sigma are bounded between 0 and 1).
method	an integer (1, 2, 3, 4, 5) indicating which basic coefficient should be used: Sokal-Sneath, Jaccard, Sorensen, Ochiai, beta, respectively.
option	a character. If option = "relative", the rows of comm are standardized into proportions that sum to 1. If option = "absolute", raw values are retained in comm.
tax	an object of class <code>taxo</code> (of package <code>ade4</code>).
type	a character. If type = "similarity", the functions <code>dsimcom</code> and <code>SQ</code> calculates similarities (S) between communities and the functions <code>dsimTaxo</code> , <code>dsimTree</code> and <code>dsimFun</code> calculates similarities between species. If type = "dissimilarity", the functions <code>dsimcom</code> and <code>SQ</code> calculates dissimilarities (1-S) between communities and the functions <code>dsimTaxo</code> , <code>dsimTree</code> and <code>dsimFun</code> calculates dissimilarities between species.
phy1	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>), or <code>hclust</code> .
rootedge	a numeric equal to the length of the branch at the nearest common ancestor of all species (here referred to as the root node). This branch is thus anterior to the root. It is ignored if <code>rootedge</code> is <code>NULL</code> .
df	either a data frame, a matrix or an object of class <code>ktab</code> of package <code>ade4</code> . If class <code>ktab</code> is used, each data frame must contain a single type of traits (see parameter <code>vartype</code>).
vartype	a vector of characters indicating the type of traits used in each data frame of <code>df</code> : "Q" for quantitative; "N" for nominal; "M" for multichoice; "P" for traits that can be expressed as proportions. Values in <code>type</code> must be in the same order as data frames in <code>df</code> (one value per table).

Details

Formulas for the indices are given in Pavoine and Ricotta (2014) main text and appendixes.

Value

If type = "similarities", a matrix of similarities between communities. If type = "dissimilarities", an object of class `dist` with the dissimilarities between communities.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S., Ricotta, C. (2014) Functional and phylogenetic similarity among communities. *Methods in Ecology and Evolution*, **5**, 666–675.

Examples

```
data(macroloire, package="ade4")

Ssokalsneath <- dsimcom(t(macroloire$fau), method=1, option=c("relative"))
Sjaccard <- dsimcom(t(macroloire$fau), method=2, option=c("relative"))
Ssorensen <- dsimcom(t(macroloire$fau), method=3, option=c("relative"))
Sochiai <- dsimcom(t(macroloire$fau), method=4, option=c("relative"))
Sbeta <- dsimcom(t(macroloire$fau), method=5, option=c("relative"))

SQUNIF <- SQ(t(macroloire$fau))

# The taxonomy is contained in macroloire$taxo

s_species_sokalsneath_taxo <- dsimTaxo(macroloire$taxo, method=1) # Using formula  $a/(a+2b+2c)$ 
# (see notations below)
s_species_jaccard_taxo <- dsimTaxo(macroloire$taxo, method=2) # Using formula  $a/(a+b+c)$ 
s_species_sorensen_taxo <- dsimTaxo(macroloire$taxo, method=3) # Using formula  $2a/(2a+b+c)$ 
s_species_ochiai_taxo <- dsimTaxo(macroloire$taxo, method=4) # Using  $a/\sqrt{(a+b)(a+c)}$ 
s_species_beta_taxo <- dsimTaxo(macroloire$taxo, method=5) # Using  $4a/(4a+b+c)$ 

# To check that these matrices of taxonomic similarities among species are p.s.d.
# we have to verify that their eigenvalues are all nonnegative:
all(eigen(s_species_sokalsneath_taxo)$val>-(1e-8))
all(eigen(s_species_jaccard_taxo)$val>-(1e-8))
all(eigen(s_species_sorensen_taxo)$val>-(1e-8))
all(eigen(s_species_ochiai_taxo)$val>-(1e-8))
all(eigen(s_species_beta_taxo)$val>-(1e-8))

# Compositions of the communities
m <- t(macroloire$fau[rownames(s_species_sokalsneath_taxo), ])
# Taxonomic similarities among species
s <- s_species_sokalsneath_taxo

Ssokalsneath_taxo <- dsimcom(m, s, method = 1)
SQwith_species_sokalsneath_taxo <- SQ(m, s)

# Compositions of the communities
m <- t(macroloire$fau[rownames(s_species_jaccard_taxo), ])
# Taxonomic similarities among species
s <- s_species_jaccard_taxo
Sjaccard_taxo <- dsimcom(m, s, method = 2)
SQwith_species_jaccard_taxo <- SQ(m, s)
```

```

# Compositions of the communities
m <- t(macroloire$fau[rownames(s_species_sorensen_taxo), ])
# Taxonomic similarities among species
s <- s_species_sorensen_taxo

Ssorensen_taxo <- dsimcom(m, s, method = 3)
SQwith_species_sorensen_taxo <- SQ(m, s)

# Compositions of the communities
m <- t(macroloire$fau[rownames(s_species_ochiai_taxo), ])
# Taxonomic similarities among species
s <- s_species_ochiai_taxo

Sochiai_taxo <- dsimcom(m, s, method = 4)
SQwith_species_ochiai_taxo <- SQ(m, s)

# Compositions of the communities
m <- t(macroloire$fau[rownames(s_species_beta_taxo), ])
# Taxonomic similarities among species
s <- s_species_beta_taxo

Sbeta_taxo <- dsimcom(m, s, method = 5)
SQwith_species_beta_taxo <- SQ(m, s)

# The matrix named feed below contains feeding attributes as rows,
# species as columns, and affinities (proportions) as entries.
feed <- macroloire$traits[ ,-(1:4)]

# Feeding habits comprise seven categories: engulfers, shredders, scrapers,
# deposit-feeders, active filter-feeders, passive filter-feeders and piercers, in this order.

# Functional similarities among species are computed as indicated in the main text
s_species_sokalsneath_feed <- dsimFun(feed, vartype = "P", method=1)
s_species_jaccard_feed <- dsimFun(feed, vartype = "P", method=2)
s_species_sorensen_feed <- dsimFun(feed, vartype = "P", method=3)
s_species_ochiai_feed <- dsimFun(feed, vartype = "P", method=4)
s_species_beta_feed <- dsimFun(feed, vartype = "P", method=5)

all(eigen(s_species_sokalsneath_feed)$val>-(1e-8))
all(eigen(s_species_jaccard_feed)$val>-(1e-8))
all(eigen(s_species_sorensen_feed)$val>-(1e-8))
all(eigen(s_species_ochiai_feed)$val>-(1e-8))
all(eigen(s_species_beta_feed)$val>-(1e-8))

Ssokalsneath_feed <- dsimcom(t(macroloire$fau), s_species_sokalsneath_feed, method=1)
SQwith_species_sokalsneath_feed <- SQ(t(macroloire$fau), s_species_sokalsneath_feed)

Sjaccard_feed <- dsimcom(t(macroloire$fau), s_species_jaccard_feed, method=2)
SQwith_species_jaccard_feed <- SQ(t(macroloire$fau), s_species_jaccard_feed)

Ssorensen_feed <- dsimcom(t(macroloire$fau), s_species_sorensen_feed, method=3)

```

```

SQwith_species_sorensen_feed <- SQ(t(macroloire$fau), s_species_sorensen_feed)

Sochiai_feed <- dsimcom(t(macroloire$fau), s_species_ochiai_feed, method=4)
SQwith_species_ochiai_feed <- SQ(t(macroloire$fau), s_species_ochiai_feed)

Sbeta_feed <- dsimcom(t(macroloire$fau), s_species_sorensen_feed, method=5)

all(eigen(Ssokalsneath_feed)$val>-(1e-8))
all(eigen(Sjaccard_feed)$val>-(1e-8))
all(eigen(Ssorensen_feed)$val>-(1e-8))
all(eigen(Sochiai_feed)$val>-(1e-8))
all(eigen(Sbeta_feed)$val>-(1e-8))

par(mfrow=c(3, 5))
plot(SQUNIF, Ssokalsneath, xlim=c(0,1), ylim=c(0,1), asp=1)
segments(0, 0, 1,1)
plot(SQUNIF, Sjaccard, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQUNIF, Ssorensen, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQUNIF, Sochiai, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQUNIF, Sbeta, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)

plot(SQwith_species_sokalsneath_taxo, Ssokalsneath_taxo, xlim=c(0,1), ylim=c(0,1), asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_jaccard_taxo, Sjaccard_taxo, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_sorensen_taxo, Ssorensen_taxo, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_ochiai_taxo, Sochiai_taxo, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_beta_taxo, Sbeta_taxo, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)

plot(SQwith_species_sokalsneath_feed, Ssokalsneath_feed, xlim=c(0,1), ylim=c(0,1), asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_jaccard_feed, Sjaccard_feed, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_sorensen_feed, Ssorensen_feed, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_ochiai_feed, Sochiai_feed, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)
plot(SQwith_species_sorensen_feed, Sbeta_feed, xlim=c(0,1), ylim=c(0,1) , asp=1)
segments(0, 0, 1,1)

par(mfrow=c(1,1))

```

Description

Computes the sum of branch lengths on a phylogenetic tree.

Usage

```
EH(phy1, select = NULL)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust.
select	a vector containing the numbers of the leaves (species) which must be considered in the computation of Phylogenetic Diversity (PD) (or merely sum of branch lengths on the tree). This parameter allows the calculation of PD for a subset of species (including the branch between the subtree and the most ancient node of the full tree).

Value

Function EH returns a real value

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr> with contributions of Stephane Dray

References

- Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.
- Nee, S. and May, R.M. (1997) Extinction and the loss of evolutionary history. *Science*, **278**, 692–694.

See Also

See Also [optimEH](#), [randEH](#)

Examples

```
data(carni70, package = "adephylo")
tre <- read.tree(text = carni70$tre)
EH(tre)
EH(tre, select=c("Mustela.nigripes", "Mustela.frenata", "Puma.concolor"))
EH(tre, select=c(1,68,70))
```

Description

Eight functions are available.

EqRS performs the first decomposition of diversity developed in Pavoine et al. (2016) and `rtestEqRS` can be used for the associated permutation tests.

EqRSintra performs the second decomposition of diversity developed in Pavoine et al. (2016) and `rtestEqRSintra` can be used for the associated permutation tests.

EqRao performs the third decomposition of diversity introduced in Pavoine et al. (2016) and `rtestEqRao` can be used for the associated permutation tests.

`wapqe` performs the apportionment of quadratic entropy (Rao 1986) and `rtestwapqe` associated permutation tests (Pavoine et al. 2016).

Usage

```
EqRS(comm, dis = NULL, structures = NULL,
      option = c("eq", "normed1", "normed2"),
      formula = c("QE", "EDI"), tol = 1e-08)
```

```
EqRSintra(comm, dis = NULL, structures = NULL,
           option = c("eq", "normed1", "normed2"),
           formula = c("QE", "EDI"), tol = 1e-08,
           metmean = c("harmonic", "arithmetic"))
```

```
EqRao(comm, dis = NULL, structures = NULL,
       option = c("eq", "normed1", "normed2"),
       formula = c("QE", "EDI"), wopt = c("even", "speciesab"),
       tol = 1e-08, metmean = c("harmonic", "arithmetic"))
```

```
wapqe(comm, dis = NULL, structures = NULL,
       formula = c("QE", "EDI"), wopt = c("even", "speciesab"),
       tol = 1e-08)
```

```
rtestEqRS(comm, dis = NULL, structures = NULL,
          formula = c("QE", "EDI"),
          option = c("normed1", "normed2", "eq"), level = 1, nrep = 99,
          alter = c("greater", "less", "two-sided"), tol = 1e-08)
```

```
rtestEqRSintra(comm, dis = NULL, structures = NULL,
               formula = c("QE", "EDI"),
               option = c("normed1", "normed2", "eq"), level = 1,
               nrep = 99, alter = c("greater", "less", "two-sided"),
               tol = 1e-08, metmean = c("harmonic", "arithmetic"))
```



```
rtestEqRao(comm, dis = NULL, structures = NULL,
formula = c("QE", "EDI"),
option = c("normed1", "normed2", "eq"),
wopt = c("even", "speciesab"), level = 1, nrep = 99,
alter = c("greater", "less", "two-sided"), tol = 1e-08,
metmean = c("harmonic", "arithmetic"))
```

```
rtestwapqe(comm, dis = NULL, structures = NULL,
formula = c("QE", "EDI"), wopt = c("even", "speciesab"),
level = 1, nrep = 99, alter = c("greater", "less", "two-sided"),
tol = 1e-08)
```

Arguments

comm	a data frame or a matrix with communities as rows and species as columns. Entries are abundances of species within sites.
dis	either NULL or an object of class <code>dist</code> that contains the (functional or phylogenetic) distances among species. If NULL, the Gini-Simpson index is used.
structures	either NULL or a data frame that contains, in the j th row and the k th column, the name of the group of level k to which the j th community belongs. Communities in <code>structures</code> should be in the same order as in <code>comm</code> . See details.
option	"eq", "normed1" or "normed2" (if a vector with several codes is given, only the first one is considered). "eq" means that the diversity components are given in terms of equivalent number (E) of species, sites, regions etc. "normed1" means that the normed components of diversity will be returned with formula $(1 - 1/E) / (1 - 1/E_{\max})$; "normed2" means that the normed components of diversity will be returned with formula $(E - 1) / (E_{\max} - 1)$. For E_{α} and E_{γ} , $E_{\max}=S$ (the number of species in the data set).
formula	either "QE" (default) or "EDI". See details.
wopt	either "even" (default), "speciesab", or a numeric vector. If $w="speciesab"$, then the communities will be weighted by their sum of species' abundances. If $w="even"$, the communities will be evenly weighted within the factors defined by the parameter <code>structures</code> .
tol	a tolerance threshold (a value between $-tol$ and tol is considered equal to zero)
metmean	a character string: either "arithmetic" or "harmonic" (default).
level	a number. The number is discarded if the parameter <code>structures</code> is set to NULL. If <code>structures</code> is different from NULL then 1 means test for differences among communities, within the levels of the first factor given in parameter <code>structures</code> (column 1), 2 means test for differences among levels of the first factor given in parameter <code>structures</code> (column 1) but within levels of the second factor given in parameter <code>structures</code> (column 2) (if available), etc. <code>level</code> should thus be between 1 and 1+the number of columns in <code>structures</code> . See details for concrete examples.
nrep	the number of permutations.
alter	a character string specifying the alternative hypothesis, must be one of "greater" (default), "less" or "two-sided".

Details

If formula = "QE", the definition of the quadratic entropy is:

$$QE(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} d_{kl}$$

where $\mathbf{p}_i = (p_{1|i}, \dots, p_{k|i}, \dots, p_{S|i})$ is the vector of relative species abundance within community i ; S is the number of species; $\mathbf{D} = (d_{kl})$ is the matrix of (phylogenetic or functional) dissimilarities among species, and d_{kl} is the (phylogenetic or functional) dissimilarity between species k and l .

If formula = "EDI", the definition of the quadratic entropy is:

$$EDI(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} \frac{d_{kl}^2}{2}$$

EDI stands for the Euclidean Diversity Index of Champely and Chessel (2002). For example, applying EqRao with the distances dis=mydis and formula=QE corresponds to applying it with dis=sqrt(2*mydis) and formula=EDI.

Parameter level permits to choose the component of beta diversity that will be tested for significance.

Examples are given below:

Scenario 1: Imagine that structures is NULL, then we only have a set of communities, species within them and a matrix of (functional or phylogenetic) dissimilarities among species. Then there is only one level of beta diversity which represents the average dissimilarities among communities. The functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe will always, in that case, test for the significance of the dissimilarities among communities (whatever the value given to parameter level).

Scenario 2: Imagine that structures is a data frame with sites as rows and only one column representing how sites are distributed among regions. Then, there are two levels of beta diversity: beta1 diversity represents the (functional or phylogenetic) dissimilarities among sites within regions; and beta2 diversity represents the (functional or phylogenetic) dissimilarities among regions. If level = 1 then functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe will test for the significance of the dissimilarities among sites within regions (beta1 diversity); in contrast, if level = 2 functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe will test for the significance of the dissimilarities among regions (beta2 diversity). As there is only one column in parameter structures and thus only two levels of diversity, level cannot be higher than 2.

Scenario 3: Imagine now that parameter comm contains quadrats as rows and species as columns and that structures is a data frame with quadrats as rows and three columns representing how quadrats are organized in three nested factors: hamlets, towns, and counties. Then there would be four levels of beta diversity: beta1 diversity represents the (functional or phylogenetic) dissimilarities among quadrats within hamlets, towns, and counties; beta2 diversity represents the dissimilarities among hamlets within towns and counties; beta3 diversity represents the dissimilarities among towns within counties; and beta4 diversity represents the dissimilarities among counties. If level = 1 then functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe will test for the significance of the dissimilarities among quadrats within hamlets, towns, and counties (beta1 diversity); if level = 2 functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe will test for the significance of the dissimilarities among hamlets within towns and regions (beta2 diversity); if level = 3 functions rtestEqRS, rtestEqRSintra, rtestEqRao, and rtestwapqe

will test for the significance of the dissimilarities among towns within regions (beta3 diversity); if level = 4 functions `rtestEqRS`, `rtestEqRSintra`, `rtestEqRao`, and `rtestwapqe` will test for the significance of the dissimilarities among regions (beta4 diversity). As there are only three columns in parameter structures and thus only four levels of diversity, level cannot be higher than 4. Test for level=1 is performed by permuting the abundances of each species across quadrats but within hamlets, towns, and counties. Test for level=2 is performed by permuting the quadrats among hamlets within towns and counties. Test for level=3 is performed by permuting the hamlets among towns within counties. Test for level=4 is performed by permuting the towns among counties.

Other permutation schemes will be added in the future.

Value

The functions `EqRS`, `EqRSintra`, `EqRao`, `wapqe` return a data frame with each component of the selected diversity decomposition. The functions `rtestEqRS`, `rtestEqRSintra`, `rtestEqRao`, and `rtestwapqe` return a list of class `randtest` or `krandtest` (classes of package `ade4`).

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Pavoine, S., Marcon, E., Ricotta, C. (2016) "Equivalent numbers" for species, phylogenetic, or functional diversity in a nested hierarchy of multiple scales. *Methods in Ecology and Evolution*, **7**, 1152–1163.
- Champely, S. and Chessel, D. (2002) Measuring biological diversity using Euclidean metrics. *Environmental and Ecological Statistics*, **9**, 167–177.
- Rao, C.R. (1986) Rao's axiomatization of diversity measures. In: *Encyclopedia of Statistical Sciences*. Vol. 7 (eds S. Kotz and N.L. Johnson), pp. 614–617. New York: Wiley and Sons.

Examples

```
data(macroloire, package="ade4")

# Taxonomic dissimilarities among species:
dTaxo <- dist.taxo(macroloire$taxo)^2/2
dTaxo <- dTaxo/max(dTaxo)
# Size-based dissimilarities among species
dSize <- dist.prop(macroloire$traits[,1:4], method = 2)
# Dissimilarities among species in terms of feeding categories
dFeed <- dist.prop(macroloire$traits[,5:11], method = 2)
# Dissimilarities among species in terms of
# both size and feeding categories
dSF <- (dSize+dFeed)/2

# Table with sites as rows (stations),
# species as columns and abundances as entries
ab <- as.data.frame(t(macroloire$fau))
# Table with sites as rows and one column only.
```

```

# Entries indicate the geological region associated
# with each site
stru <- macroloire$envir["Morphoregion"]

EqRao(ab, , stru, option="eq")
EqRao(ab, dTaxo, stru, formula = "QE", option="eq")
EqRao(ab, dSize, stru, formula = "QE", option="eq")
EqRao(ab, dFeed, stru, formula = "QE", option="eq")
EqRao(ab, dSF, stru, formula = "QE", option="eq")

EqRao(ab, , stru, option="normed2")
EqRao(ab, dTaxo, stru, formula = "QE", option="normed2")
EqRao(ab, dSize, stru, formula = "QE", option="normed2")
EqRao(ab, dFeed, stru, formula = "QE", option="normed2")
EqRao(ab, dSF, stru, formula = "QE", option="normed2")

## Not run:
# Tests for dissimilarities among sites within regions:

rb1_GS <- rtestEqRao(ab, , stru, level=1, nrep=999, option="normed2")
rb1_GS
plot(rb1_GS)
rb1_Taxo <- rtestEqRao(ab, dTaxo, stru, formula = "QE",
level=1, nrep=999, option="normed2")
rb1_Taxo
plot(rb1_Taxo)
rb1_Size <- rtestEqRao(ab, dSize, stru, formula = "QE",
level=1, nrep=999, option="normed2")
rb1_Size
plot(rb1_Size)
rb1_Feed <- rtestEqRao(ab, dFeed, stru, formula = "QE",
level=1, nrep=999, option="normed2")
rb1_Feed
plot(rb1_Feed)
rb1_SF <- rtestEqRao(ab, dSF, stru, formula = "QE",
level=1, nrep=999, option="normed2")
rb1_SF
plot(rb1_SF)

# Tests for dissimilarities among regions:

r2_GS <- rtestEqRao(ab, , stru, level=2, nrep=999, option="normed2")
r2_GS
plot(r2_GS)
r2_Taxo <- rtestEqRao(ab, dTaxo, stru, formula = "QE",
level=2, nrep=999, option="normed2")
r2_Taxo
plot(r2_Taxo)
r2_Size <- rtestEqRao(ab, dSize, stru, formula = "QE",
level=2, nrep=999, option="normed2", w="even")
r2_Size
plot(r2_Size)
r2_Feed <- rtestEqRao(ab, dFeed, stru, formula = "QE",

```

```

level=2, nrep=999, option="normed2", w="even")
r2_Feed
plot(r2_Feed)
r2_SF <- rtestEqRao(ab, dSF, stru, formula = "QE",
level=2, nrep=999, option="normed2", w="even")
r2_SF
plot(r2_SF)

## End(Not run)

```

eveparam

Parametric Indices of Species Evenness

Description

The function `eveparam` calculates parametric evenness indices. The parameter controls the relative importance given to rare versus abundant species in a community. The function `plot.eveparam` plots the results of function `eveparam`.

Usage

```
eveparam(comm, method = c("hill", "tsallis", "renyi"), q = 2, tol = 1e-08)
```

```

## S3 method for class 'eveparam'
plot(x, legend = TRUE,
legendposi = "topright", type = "b",
col = if (is.numeric(x)) NULL
else sample(colors(distinct = TRUE), nrow(x$eve)),
lty = if (is.numeric(x)) NULL else rep(1, nrow(x$eve)),
pch = if (is.numeric(x)) NULL else rep(19, nrow(x$eve)), ...)

```

Arguments

<code>comm</code>	a data frame or a matrix typically with communities as rows, species as columns and abundance as entry.
<code>method</code>	a character: either "hill" for the Hill numbers (Hill 1973), "tsallis" for the Tsallis or HCDDT entropy (Harvda and Charvat 1967; Daroczy 1970; Tsallis 1988), or "renyi" for Renyi's entropy (Renyi 1960). These indices are divided by the value they would have if all species had even abundances.
<code>q</code>	a positive numeric or a vector of positive numerics that give values for the q parameter.
<code>tol</code>	numeric tolerance threshold: values between -tol and tol are considered equal to zero.
<code>x</code>	an object of class <code>eveparam</code> obtained with function <code>eveparam</code> .
<code>legend</code>	a logical. If TRUE a legend is given with the colour, the type of line (etc.) used to define the evenness curve of each community.

legendposi	a character or a numeric that gives the position of the legend to be passed to function legend of the base of R.
type	a character to be passed to the graphic parameter type of functions plot and lines used to draw the evenness curve of each community.
col	colours to be passed to the graphic parameter col of functions plot and lines to define the colour of the evenness curve of each community.
lty	type of line (plain, broken etc.) to be passed to the graphic parameter lty of functions plot and lines used to draw the evenness curve of each community.
pch	type of point (open circle, close circle, square etc.) to be passed to the graphic parameter pch of functions plot and lines used to draw the evenness curve of each community.
...	other parameters can be added and passed to the functions plot and lines used to draw the graphic.

Value

Function eveparam, if only one value of q is given, returns a vector with communities' evenness. If more than one value of q is given, a list of two objects is returned:

q	the vector of values for q
div	a data frame with communities' evenness calculated for all values of q

The function plot.eveparam returns a graphic.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Daroczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.
- Havrda, M., Charvat, F. (1967) Quantification method of classification processes: concept of structural alpha-entropy. *Kybernetika*, **3**, 30–35.
- Hill, M.O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
- Renyi, A. (1960) On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547–561.
- Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 480–487.

Examples

```
data(batcomm)
ab <- batcomm$ab
plot(eveparam(ab))
plot(eveparam(ab, q=0:4))
```

Description

The function `evoCA` performs the evolutionary correspondence analysis (evoCA) (Pavoine 2016): an adaptation of the correspondence analysis (CA) to analyse the distributions of lineages among sites and, simultaneously, to analyse the phylogenetic composition of sites. The function `plot.evoCA` displays the phylogeny on the factorial maps of evoCA.

Usage

```
evoCA(phy1, comm, scannf = TRUE, nf = 2, abundance = TRUE)
```

```
## S3 method for class 'evoCA'
plot(x, xaxis = 1, yaxis = 2, ...)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>) or <code>hclust</code> . To ease the interpretation of factorial maps, it is advised to add labels to the nodes of the phylogeny. For example, if your phylogeny is named <code>'tree'</code> and is of class <code>phylo</code> , you can name the nodes with the following command: <code>tree\$node.label <- paste("n", 1:tree\$Nnode, sep="")</code> . If it is of class <code>phylo4</code> , you can use: <code>nodeLabels(tree) <- paste("n", 1:nNodes(tree), sep="")</code> . Use <code>plot(tree)</code> to see the result.
<code>comm</code>	a data frame or a matrix typically with communities (or sites, plots, etc.) as rows, species as columns and presence/absence or an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
<code>scannf</code>	a logical value indicating whether the screeplot (eigenvalues) should be displayed for choosing the number of axes to be kept.
<code>nf</code>	if <code>scannf</code> is <code>FALSE</code> , an integer indicating the number of axes to be kept.
<code>abundance</code>	a logical value, if <code>FALSE</code> , only the presences/absences of the evolutionary units in sites are considered.
<code>x</code>	an object of class <code>evoCA</code>
<code>xaxis</code>	the number of the evoCA axis chosen for the x-axis of the 3d plot
<code>yaxis</code>	the number of the evoCA axis chosen for the y-axis of the 3d plot
<code>...</code>	other parameters can be added and passed to the function <code>plot3d</code> of package <code>rgl</code> .

Value

`evoCA` returns an object of class `evoCA` and of class `dudi` (see package `ade4`, `?dudi`). Graphical tools are associated with class `dudi` in packages `ade4` and `adegraphics` (see example section below and `?scatter.dudi`).

The object returned contains the following components:

tab	a data frame with n rows and p columns, with communities as rows and nodes of the phylogeny as columns; the entries of the data frame evaluate the degree of dependence (values that depart from zero)/independence(close-to-zero values) between the occurrence in a community and the position in the phylogeny
cw	weights attributed to the nodes of the phylogeny, a vector with n components
lw	weights attributed to the communities, a vector with p components
eig	eigenvalues, a vector with min(n,p) components
rank	integer, number of axes
nf	integer, number of kept axes
c1	normed scores for the nodes of the phylogeny, data frame with p rows and nf columns
l1	normed scores for the communities, data frame with n rows and nf columns
co	scores for the nodes of the phylogeny, data frame with p rows and nf columns
li	scores for the communities, data frame with n rows and nf columns
call	the original call

If X is an object of class evoCA, then `attributes(X)$phy` contains the phylogenetic tree (of class phylo) with names for internal nodes.

`plot.evoCA` returns a dynamics 3-dimensional plot

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.

See Also

See Also [evoNSCA](#), [evopcachord](#), [evopcahellinger](#), [evodiss](#)

Examples

```
data(batcomm)
ab <- batcomm$ab
phy <- read.tree(text=batcomm$tre)
plot(phy, show.node=TRUE)
evoCAbat <- evoCA(phy, ab, scan=FALSE, nf=2)
evoCAbat$eig/sum(evoCAbat$eig)
s.label(evoCAbat$li)
s.label(evoCAbat$co)
s.arrow(evoCAbat$co)
inertia.dudi(evoCAbat, row=TRUE)$row.abs
inertia.dudi(evoCAbat, col=TRUE)$col.abs
```



```

evoCAbat <- evoCA(phy, ab, scan=FALSE, nf=3) ## All axes are now retained
# The Euclidean (canonical) distances among habitat points on the evoCA space is
dist(evoCAbat$li)
# which is equal to evoDchi2:
evodiss(phy, ab, "chi2")

plot(evoCAbat)

plot(evoCAbat, xaxis=1, yaxis=3)

```

evodiss

Pair-wise Phylogenetic Dissimilarities between Communities

Description

The function calculates PD-dissimilarity indices described and/or discussed in Pavoine (2016). Part of them are parametric indices.

Usage

```

evodiss(phy1, comm, method = NULL, q = NULL,
w = c("evoab", "even", "speciesab"), diag = FALSE,
upper = FALSE, tol = 1e-08)

```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust.
comm	a data frame typically with communities as rows, species as columns and an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
method	one of the following character codes: "Minkowski", "Euclidean", "Manhattan", "Chord", "ScaledCanberra", "Divergence", "BC", "MH", "LG", "Hellinger", "chi2", "Hill", "Renyi", "C", "U", "S". See Details.
q	a vector with nonnegative value(s) for parameter q only used with method="Minkowski", "Hill", "Renyi" (parametric indices). q controls the importance given to rare compared to abundant evolutionary units. See Details.
w	either a numeric vector giving weights for communities (same order as in comm), or a character among "even", "evoab", and "speciesab". If several characters are given, only the first one is used. w is used only if method="Hill", "Renyi", "C", "U", or "S" is chosen. See details.
diag	parameter (boolean) passed to function as.dist (R base).
upper	parameter (boolean) passed to function as.dist (R base).
tol	numeric tolerance threshold: values between -tol and tol are considered equal to zero.

Details

The indices available are (formulas can be found in Supplementary material Appendix 1 of Pavoine 2016):

"Minkowski": ${}^q\text{evo}D_{\text{Minkowski}}$

"Euclidean": $\text{evo}D_{\text{Euclidean}}$

"Manhattan": $\text{evo}D_{\text{Manhattan}}$

Chord: $\text{evo}D_{\text{Chord}}$

ScaledCanberra: $\text{evo}D_{\text{ScaledCanberra}}$

Divergence: $\text{evo}D_{\text{Divergence}}$

BC: $\text{evo}D_{\text{Bray-Curtis}}$

MH: $\text{evo}D_{\text{Morisita-Horn}}$

LG: $\text{evo}D_{\text{Profile}}$

Hellinger: $\text{evo}D_{\text{Hellinger}}$

chi2: $\text{evo}D_{\text{Chi2}}$

Hill: $1 - \bar{V}_{q2}$

Renyi: ${}^q\text{evo}D_{\text{Renyi}}$

C: $1 - \bar{C}_{q2}$

U: $1 - \bar{U}_{q2}$

S: $1 - \bar{S}_{q2}$

The weights of the communities (parameter w) can be "even" (even weights, i.e. relative abundances are considered for evolutionary units), "evoab" (proportional to the summed abundances of all evolutionary units, i.e. absolute abundances are considered for evolutionary units), or "speciesab" (proportional to the summed abundances of all species). Note that if the phylogenetic tree is ultrametric (the distance from any species to the root is constant), then options "evoab" and "speciesab" are equivalent.

Value

An object of class `dist` containing the PD-dissimilarities (phylogenetic dissimilarities) between communities.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

The methodologies are presented in

Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.

They gather, in a common framework, indices introduced in

Chiu, C.-H., Jost, L., Chao, A. (2014) Phylogenetic beta diversity, similarity and differentiation measures based on Hill numbers. *Ecological Monographs*, **84**, 21–44.

and earlier work extended here in a phylogenetic context and reviewed in

Legendre, P. and De Caceres, M. (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, **16**, 951–963.

See Also

[evodiss_family](#)

Examples

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
ab <- batcomm$ab[,phy$tip.label]

evodiss(phy, ab, "LG")
evodiss(phy, ab, "Hellinger")
evodiss(phy, ab, "Chord")

evodiss(phy, ab, "Hill", q=2)
evodiss(phy, ab, "Hill", q=2, w="even")
```

evodiss_family	<i>A Family of Indices Dedicated to Pair-wise Phylogenetic Dissimilarities between Communities</i>
----------------	--

Description

The function `evodiss_family` was written thanks to function `dist.binary` of package `ade4`. Function `dist.binary` calculates specific compositional distances. The new function here replaces species with evolutionary units. It calculates Nipperess et al. (2010) parameters a , b , c , d (with incidence data), or A , B , C , D (with abundance data) and then use these parameters to compute pair-wise phylogenetic dissimilarities between communities.

The graphical function `evodiss_ternaryplot` displays Nipperess et al. (2010) parameters a , b , c (with incidence data), or A , B , C (with abundance data) on a ternary plot (see Koleff et al. 2003).

Usage

```
evodiss_family(phyl, comm, method = NULL, abundance = TRUE,
squareroot = TRUE, diag = FALSE, upper = FALSE, tol = 1e-08)
```

```
evodiss_ternaryplot(phyl, comm, abundance = TRUE,
tol = 1e-08, ...)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust.
comm	a data frame typically with communities as rows, species as columns and presence/absence or an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
method	a number between 1 and 10. If NULL, the choice is made with a console message. See details.
abundance	a logical indicating whether abundance data (if TRUE) or incidence data (presence/absence if FALSE) must be used.
squareroot	a logical. First a similarity index (S) is calculated (see details). Then if squareroot=TRUE, the dissimilarity is calculated as $\sqrt{1 - S}$. If squareroot=FALSE, it is calculated as $1 - S$. The square root is useful to obtain Euclidean properties for the PD-dissimilarity indices.
diag	parameter (boolean) passed to function as.dist (R base).
upper	parameter (boolean) passed to function as.dist (R base).
tol	numeric tolerance threshold: values between -tol and tol are considered equal to zero.
...	other parameters can be added and passed to the function triangle.plot of the ade4 package (labels cannot be changed and are defined thanks to the row names of comm).

Details

The function was written thanks to function `dist.binary` of package `ade4`. Function `dist.binary` calculates specific compositional distances. The new function here replaces species with evolutionary units. It calculates Nipperess et al. (2010) parameters a, b, c, d (with incidence data), or A, B, C, D (with abundance data). Then, the parameters are combined thanks to one out of ten methods.

1: Jaccard index (1901) $S3$ coefficient of Gower and Legendre (1986) $s1 = a / (a+b+c)$.

2: Simple matching coefficient of Sokal and Michener (1958) $S4$ coefficient of Gower and Legendre (1986) $s2 = (a+d) / (a+b+c+d)$.

3: Sokal and Sneath(1963) $S5$ coefficient of Gower and Legendre (1986) $s3 = a / (a + 2(b + c))$.

4: Rogers and Tanimoto (1960) $S6$ coefficient of Gower and Legendre (1986) $s4 = (a + d) / (a + 2(b + c) + d)$.

5: Dice (1945) or Sorensen (1948) $S7$ coefficient of Gower and Legendre (1986) $s5 = 2a / (2a + b + c)$.

6: Hamann coefficient $S9$ index of Gower and Legendre (1986) $s6 = (a - (b + c) + d) / (a + b + c + d)$.

7: Ochiai (1957) $S12$ coefficient of Gower and Legendre (1986) $s7 = a / \sqrt{(a + b)(a + c)}$.

8: Sokal and Sneath (1963) $S13$ coefficient of Gower and Legendre (1986) $s8 = ad / \sqrt{(a + b)(a + c)(d + b)(d + c)}$.

9: Phi of Pearson $S14$ coefficient of Gower and Legendre (1986) $s9 = (ad - bc) / \sqrt{(a + b)(a + c)(d + b)(d + c)}$.

10: $S2$ coefficient of Gower and Legendre (1986) $s10 = a / (a + b + c + d)$.

Value

Function `evodiss_family` returns an object of class `dist` containing the PD-dissimilarities (phylogenetic dissimilarities) between communities.

Function `evodiss_ternaryplot` returns a graph.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

The methodologies are presented in Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.

They gather in a common framework and extend earlier work introduced in Koleff, P., Gaston, K.J., Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.

Nipperess, D.A., Faith, D.P., Barton, K. (2010) Resemblance in phylogenetic diversity among ecological assemblages. *Journal of Vegetation Science*, **21**, 809–820.

Dice, L.R. (1945) Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302.

Gower, J.C., Legendre, P. (1986) Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, **3**, 5–48.

Jaccard, P. (1901) Etude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Societe Vaudoise des Sciences Naturelles*, **37**, 547–579.

Ochiai, A. (1957) Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring regions. *Bulletin of the Japanese Society of Scientific Fisheries*, **22**, 526–530.

Rogers, J.S. and Tanimoto, T.T. (1960) A computer program for classifying plants. *Science*, **132**, 1115–1118.

Sokal, R.R. and Michener, C.D. (1958) A Statistical Method for Evaluating Systematic Relationships. *The University of Kansas Science Bulletin*, **38**, 1409–1438.

Sokal, R.R. and Sneath, P.H.A. (1963) *Principles of numerical taxonomy*. San Francisco: W. H. Freeman.

Sorensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. *Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter*, **5**, 1–34.

See Also

[evodiss](#)

Examples

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
ab <- batcomm$ab[,phy$tip.label]
```

```

# PD-dissimilarity indices that use Nipperess et al. (2010)
# parameters can be obtained thanks to function evodiss_family.
# For example, with incidence data,
# indices evoDJaccard, evoDSorensen, and evoD0chiai
# (supplementary Appendix 1 in Pavoine 2016)
# can be obtained as follows:

evodiss_family(phy, ab, method=1, abundance=FALSE) # Jaccard
evodiss_family(phy, ab, method=5, abundance=FALSE) # Sorensen
evodiss_family(phy, ab, method=7, abundance=FALSE) # Ochiai

# With abundance data, indices evoDTJ, evoDTS, evoDTO
# (supplementary Appendix 1 in Pavoine 2016)
# can be obtained as follows:

evodiss_family(phy, ab, method=1) # evoDTJ
evodiss_family(phy, ab, method=5) # evoDTS
evodiss_family(phy, ab, method=7) # evoDTO

# Ternary plots can be obtained for Nipperess et al. (2010)
# parameters a, b, c (incidence data)
# (see Supplementary material Appendix 4 in Pavoine 2016):

evodiss_ternaryplot(phy, ab, abundance = FALSE)

# and for Nipperess et al. (2010) parameters A, B, C
# (abundance data):

evodiss_ternaryplot(phy, ab, abundance = TRUE)

# The ternary plots can be adjusted thanks
# to parameters of function triangle.plot (package ade4).
# For example, full triangles can be obtained as follows
# (previous graphs were zoomed on the smallest principal
# equilateral triangle that contained the points,
# as indicated by the embedded close grey triangle
# at the left-hand corner of ternary plot given above):

evodiss_ternaryplot(phy, ab, abundance = FALSE, adjust=FALSE, showposition=FALSE)
# Incidence data

evodiss_ternaryplot(phy, ab, abundance = TRUE, adjust=FALSE, showposition=FALSE)
# abundance data

```

Description

Function `evodivparam` calculates phylogenetic diversity in communities using parametric indices derived from Tsallis and Hill compositional indices. It can also be applied to functional trees rather

than phylogenies, to calculate a functional diversity. The function `plot.evodivparam` plots the results of function `evodivparam`.

Usage

```
evodivparam(phy1, comm,
  method = c("hill", "tsallis", "renyi"),
  q = 2, tol = 1e-08)

## S3 method for class 'evodivparam'
plot(x, legend = TRUE,
  legendposi = "topright", type="b",
  col = if(is.numeric(x)) NULL
  else sample(colors(distinct = TRUE), nrow(x$div)),
  lty = if(is.numeric(x)) NULL else rep(1, nrow(x$div)),
  pch = if(is.numeric(x)) NULL else rep(19, nrow(x$div)),
  ...)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>), or <code>hclust</code> .
<code>comm</code>	a data frame or a matrix typically with communities as rows, species as columns and abundance as entry. Species should be labeled as in the phylogenetic tree where they are the tips.
<code>method</code>	a character: either "hill" for the Hill numbers (Hill 1973), "tsallis" for the Tsallis or HCDT entropy (Harvda and Charvat 1967; Daroczy 1970; Tsallis 1988), or "renyi" for Renyi's entropy (Renyi 1960). If several value are given, only the first one is considered. See details.
<code>q</code>	a vector with nonnegative value(s) for parameter <code>q</code> . See details.
<code>tol</code>	numeric tolerance threshold: values between <code>-tol</code> and <code>tol</code> are considered equal to zero.
<code>x</code>	an object of class <code>evodivparam</code> obtained with function <code>evodivparam</code> .
<code>legend</code>	a logical. If <code>TRUE</code> a legend is given with the colour, the type of line (etc.) used to define the diversity curve of each community.
<code>legendposi</code>	a character that gives the position of the legend to be passed to function <code>legend</code> of the base of R.
<code>type</code>	a character to be passed to the graphic parameter <code>type</code> of functions <code>plot</code> and <code>lines</code> used to draw the diversity curve of each community.
<code>col</code>	vector of colours to be passed to the graphic parameter <code>col</code> of functions <code>plot</code> and <code>lines</code> to define the colour of the diversity curve of each community.
<code>lty</code>	vector of type of line (plain, broken etc.) to be passed to the graphic parameter <code>lty</code> of functions <code>plot</code> and <code>lines</code> used to draw the diversity curve of each community.

pch type of point (open circle, close circle, square etc.) to be passed to the graphic parameter pch of functions plot and lines used to draw the diversity level of each community.

... other parameters can be added and passed to the functions plot and lines used to draw the graphic.

Details

Consider a phylogenetic tree T , b_T the set of branches in T , k a branch, L_k the length of branch k , j a community ($j=1, \dots, m$), a_{jk} the abundance associated with branch k in community j (sum of abundance of all species descending from the branch). q is the parameter that increases with the importance given to abundant species compared to rare species in diversity.

The methods available to calculate the phylogenetic diversity in community j are: tsallis:

$${}^q\text{evoTsallis}_j = \left[1 - \sum_{k \in b_T} L_k \left(\frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right] / (q - 1)$$

hill:

$${}^q\text{evoHill}_j = \left[\sum_{k \in b_T} L_k \left(\frac{a_{jk}}{\sum_{k \in b_T} L_k a_{jk}} \right)^q \right]^{1/(1-q)}$$

renyi:

$${}^q\text{evoRenyi}_j = \log({}^q\text{evoHill}_j)$$

Value

If only one value of q is given, a vector with the phylogenetic diversity of each community is returned. If more than one value of q is given, a list of two objects is returned:

q the vector of values for q

div a data frame with the phylogenetic diversity of each community calculated for all values of q

The function `plot.evodivparam` returns a graphic.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

The methodologies and scripts were developed using earlier work by:

Chao, A., Chiu, C.-H., Jost, L. (2010) Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society London Series B*, **365**, 3599–3609.

Daroczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.

Havrda, M., Charvat F. (1967) Quantification method of classification processes: concept of structural alpha- entropy. *Kybernetik*, **3**, 30–35.

Hill, M.O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.

Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.

Renyi, A. (1960) On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547–561.

Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 480–487.

See Also

[abgevodivparam](#), [divparam](#)

Examples

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
ab <- batcomm$ab[, phy$tip.label]
plot(evodivparam(phy, ab))
plot(evodivparam(phy, ab, q=seq(0, 10, length=20)))
```

evoNSCA

Evolutionary Non-Symmetric Correspondence Analysis

Description

The function performs evoNSCA (Pavoine 2016): an adaptation of the non-symmetric correspondence analysis (NSCA) (see e.g. Kroonenberg and Lombardo 1999) to analyse the distributions of lineages among sites and, simultaneously, to analyse the phylogenetic composition of sites.

Usage

```
evoNSCA(phy1, comm, scannf = TRUE, nf = 2, abundance = TRUE)
```

Arguments

phy1 an object inheriting the class `phylo` (see package `ape`), `phylo4` (see package `phylobase`) or `hclust`. To ease the interpretation of factorial maps, it is advised to add labels to the nodes of the phylogeny. For example, if your phylogeny is named `'tree'` and is of class `phylo`, you can name the nodes with the following command: `tree$node.label <- paste("n", 1:tree$Nnode, sep="")`. If it is of class `phylo4`, you can use: `nodeLabels(tree) <- paste("n", 1:nNodes(tree), sep="")`. Use `plot(tree)` to see the result.

comm	a data frame or a matrix typically with communities as rows, species as columns and presence/absence or an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
scannf	a logical value indicating whether the screeplot (eigenvalues) should be displayed for choosing the number of axes to be kept.
nf	if scannf is FALSE, an integer indicating the number of axes to be kept
abundance	a logical value, if FALSE, only the presences/absences of the evolutionary units in sites are considered.

Value

evoNSCA returns an object of class evoNSCA and of class dudi (see package ade4, ?dudi). Graphical tools are associated with class dudi in packages ade4 and adegraphics (see example section below and ?scatter.dudi).

The object returned contains the following components:

tab	a data frame with n rows and p columns, with communities as rows and nodes of the phylogeny as columns. Internal data frame used by the algorithm.
cw	weights attributed to the nodes of the phylogeny, a vector with n components
lw	weights attributed to the communities, a vector with p components
eig	eigenvalues, a vector with min(n,p) components
rank	integer, number of axes
nf	integer, number of kept axes
c1	normed scores for the nodes of the phylogeny, data frame with p rows and nf columns
l1	normed scores for the communities, data frame with n rows and nf columns
co	scores for the nodes of the phylogeny, data frame with p rows and nf columns
li	scores for the communities, data frame with n rows and nf columns
call	the original call

If X is an object of class evoNSCA, then attributes(X)\$phy contains the phylogenetic tree (of class phylo) with names for internal nodes.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.
- Kroonenberg PM, Lombardo R (1999) Nonsymmetric correspondence analysis: a tool for analysing contingency tables with a dependence structure. *Multivariate Behavioral Research*, **34**, 367–396.

See Also

See Also [evoCA](#), [evopcachord](#), [evopcahellinger](#), [evodiss](#)

Examples

```
data(batcomm)
ab <- batcomm$ab
phy <- read.tree(text=batcomm$tre)
plot(phy, show.node=TRUE)
evoNSCAbat <- evoNSCA(phy, ab, scan=FALSE, nf=2)
evoNSCAbat$eig/sum(evoNSCAbat$eig)
s.label(evoNSCAbat$li)
s.label(evoNSCAbat$co)
s.arrow(evoNSCAbat$co)
inertia.dudi(evoNSCAbat, row=TRUE)$row.abs
inertia.dudi(evoNSCAbat, col=TRUE)$col.abs

evoNSCAbat <- evoNSCA(phy, ab, scan=FALSE, nf=3) ## All axes are now retained
# The Euclidean (canonical) distances among habitat points on the evoNSCA space is
dist(evoNSCAbat$li)
# which is equal to evoDprofile (see function \code{evodiss}):
evodiss(phy, ab, "LG")
```

 evopca

Evolutionary Principal Component analyses

Description

The functions `evopcachord` and `evopcahellinger` perform two adaptations of Principal Component Analysis (PCA) for the analysis of phylogenetic diversity patterns across species communities: the evolutionary principal component analysis based on Chord distance (`evoPCACHord`) and the evolutionary principal component analysis based on Hellinger distance (`evoPCAHellinger`) (Pavoine 2016).

Usage

```
evopcachord(phy1, comm, option = c("centred", "decentred"),
w = c("evoab", "even", "speciesab"), scannf = TRUE,
nf = 2, abundance = TRUE)

evopcahellinger(phy1, comm, option = c("centred", "decentred"),
w = c("evoab", "even", "speciesab"), scannf = TRUE,
nf = 2, abundance = TRUE)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust. To ease the interpretation of factorial maps, it is advised to add labels to the nodes of the phylogeny. For example, If your phylogeny is named 'tree' and is of class phylo, you can name the nodes with the following command: <code>tree\$node.label <- paste("n", 1:tree\$Nnode, sep="")</code> . If it is of class phylo4, you can use: <code>nodeLabels(tree) <- paste("n", 1:nNodes(tree), sep="")</code> . Use <code>plot(tree)</code> to see the result.
comm	a data frame or a matrix typically with communities as rows, species as columns and presence/absence or an index of abundance as entries. Species should be labeled as in the phylogenetic tree where they are the tips.
option	a character code: either "centered" or "decentered". If several codes are given, only the first one is used. See details.
w	character code: either "evoab", "even", or "speciesab". If several codes are given, only the first one is used. See details.
scannf	a logical value indicating whether the screeplot (eigenvalues) should be displayed for choosing the number of axes to be kept.
nf	if scannf is FALSE, an integer indicating the number of axes to be kept.
abundance	a logical value, if FALSE, only the presences/absences of the evolutionary units in communities are considered.

Details

Let L_k be the length of branch k in the phylogenetic tree (out of K branches); a_{jk} the sum of abundances, in community j , for all species descending from branch k ; w_j a positive weight attributed to community j (the definition for w_j is flexible with the only requirement that $\sum_{j=1}^m w_j = 1$); $a_{j+} = \sum_{k=1}^K L_k a_{jk}$; $a_{+k} = \sum_{j=1}^m a_{jk}$; $a_{++} = \sum_{j=1}^m \sum_{k=1}^K L_k a_{jk}$.

The weights of the communities (parameter w) can be "even" (even weights, i.e. relative abundances are considered for evolutionary units), "evoab" (proportional to the summed abundances of all evolutionary units, i.e. absolute abundances are considered for evolutionary units), or "speciesab" (proportional to the summed abundances of all species). Note that if the phylogenetic tree is ultrametric (the distance from any species to the root is constant), then options "evoab" and "speciesab" are equivalent.

In `evopcahellinger`, as recommended by Rao (1995), the vector used to centre matrix $(\sqrt{a_{jk}/a_{j+}})_{j,k}$ in PCA can be defined as

$$\left(\sum_{j=1}^m w_j \sqrt{a_{jk}/a_{j+}} \right)_k$$

(ordinary weighted mean, option "centered") or as

$$\left(\sqrt{\frac{\sum_{j=1}^m w_j a_{jk}}{\sum_{k=1}^K L_k \sum_{j=1}^m w_j a_{jk}}} \right)_k$$

(option "decentered"); see Pavoine (2016) for an introduction of all ordination approaches.

Similarly, in `evopcachord`, the vector used to centre matrix $(a_{jk}/\sqrt{\sum_{k=1}^K a_{jk}^2})_{j,k}$ can be defined as

$$\left(\sum_{j=1}^m w_j a_{jk} / \sqrt{\sum_{k=1}^K a_{jk}^2} \right)_k$$

(ordinary weighted mean, option "centered") or as

$$\left(\frac{\sum_{j=1}^m w_j a_{jk}}{\sqrt{\sum_{k=1}^K L_k (\sum_{j=1}^m w_j a_{jk})^2}} \right)_k$$

(option "decentered").

Value

`evopcachord` (`evopcahellinger`) returns an object of class `evopca` and of class `dudi` (see package `ade4`, `?dudi`). Graphical tools are associated with class `dudi` in packages `ade4` and `adegraphics` (see example section below and `?scatter.dudi`).

The object returned contains the following components:

<code>tab</code>	a data frame with <code>n</code> rows and <code>p</code> columns, with communities as rows and nodes of the phylogeny as columns. Internal data frame used by the algorithm.
<code>cw</code>	weights attributed to the nodes of the phylogeny, a vector with <code>n</code> components
<code>lw</code>	weights attributed to the communities, a vector with <code>p</code> components
<code>eig</code>	eigenvalues, a vector with <code>min(n,p)</code> components
<code>rank</code>	integer, number of axes
<code>nf</code>	integer, number of kept axes
<code>c1</code>	normed scores for the nodes of the phylogeny, data frame with <code>p</code> rows and <code>nf</code> columns
<code>l1</code>	normed scores for the communities, data frame with <code>n</code> rows and <code>nf</code> columns
<code>co</code>	scores for the nodes of the phylogeny, data frame with <code>p</code> rows and <code>nf</code> columns
<code>li</code>	scores for the communities, data frame with <code>n</code> rows and <code>nf</code> columns
<code>call</code>	the original call

If `X` is an object of class `evopca`, then `attributes(X)$phy` contains the phylogenetic tree (of class `phylo`) with names for internal nodes.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, **125**, 1719–1732.
- Rao, C.R. (1995) A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Questio*, **19**, 23–63.

See Also

See Also [evoCA](#), [evoNSCA](#), [evodiss](#)

Examples

```

data(batcomm)
ab <- batcomm$ab
phy <- read.tree(text=batcomm$tre)
plot(phy, show.node=TRUE)

evopcaHbat <- evopcahellinger(phy, ab, scan=FALSE, nf=3)
dist(evopcaHbat$li)
evodiss(phy, ab, "Hellinger")

evopcaHbat$eig/sum(evopcaHbat$eig)
s.label(evopcaHbat$li)
s.label(evopcaHbat$co)
s.arrow(evopcaHbat$co)
inertia.dudi(evopcaHbat, row=TRUE)$row.abs
inertia.dudi(evopcaHbat, col=TRUE)$col.abs

evopcaCbat <- evopcachord(phy, ab, scan=FALSE, nf=3)
dist(evopcaCbat$li)
evodiss(phy, ab, "Chord")

evopcaCbat$eig/sum(evopcaCbat$eig)
s.label(evopcaCbat$li)
s.label(evopcaCbat$co)
s.arrow(evopcaCbat$co)
inertia.dudi(evopcaCbat, row=TRUE)$row.abs
inertia.dudi(evopcaCbat, col=TRUE)$col.abs

```

Description

The function `K` calculates Blomberg et al. (2003) statistic K and performs a permutation test where species identities are maintained in the phylogeny while the trait values of species are randomly shuffled (permuted) (Pavoine and Ricotta 2013).

The function `Kstar` calculates Blomberg et al. (2003) statistic K^* and performs a permutation test where species identities are maintained in the phylogeny while the trait values of species are randomly shuffled (permuted) (Pavoine and Ricotta 2013).

The function `Kw` calculates Pavoine and Ricotta (2013) statistic K_w and performs a permutation test where species identities are maintained in the phylogeny while the trait values of species are randomly shuffled (permuted).

Usage

```
K(phy1, trait, nrep = 999, alter = c("greater", "less", "two-sided"))
Kstar(phy1, trait, nrep = 999, alter = c("greater", "less", "two-sided"))
Kw(phy1, trait, nrep = 999, alter = c("greater", "less", "two-sided"))
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>) or <code>hclust</code> .
<code>trait</code>	a vector with the trait value for each species (tip) in the phylogenetic tree. Trait values for species must be in the same order as species in the phylogenetic tree.
<code>nrep</code>	a numeric: the number of randomizations
<code>alter</code>	a character specifying the alternative hypothesis; must be one of "greater" (phylogenetic signal; the default), "less" or "two-sided". If several values are given, only the first one is retained.

Value

Each function returns an object of class `randtest` with the results of the permutation tests. (see function `randtest` in package `ade4`)

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Blomberg, S.P., Garland, T., Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, **57**, 717–745.

Pavoine, S., Ricotta, C. (2013) Testing for phylogenetic signal in biological traits: the ubiquity of cross-product statistics. *Evolution*, **67**, 828–840.

See Also

See Also [Kstar](#), [Kw](#)

Examples

```
data(rockfish)
phy <- read.tree(text=rockfish$tre)
K(phy, rockfish$traits[phy$tip.label, 1])
Kstar(phy, rockfish$traits[phy$tip.label, 1])
Kw(phy, rockfish$traits[phy$tip.label, 1])
```

 optimEH

Nee and May's Optimizing Process

Description

The function `optimEH` performs Nee and May's optimizing scheme. When branch lengths in an ultrametric phylogenetic tree are expressed as divergence times, the total sum of branch lengths in that tree expresses the amount of evolutionary history. Nee and May's algorithm optimizes the amount of evolutionary history preserved if only k species out of n were to be saved. The $k-1$ closest-to-root nodes are selected, which defines k clades; one species from each clade is picked. At this last step, we decide to select the most original species of each of the k clades.

Usage

```
optimEH(phy1, nbofsp, tol = 1e-08, give.list = TRUE)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>) or <code>hclust</code> .
<code>nbofsp</code>	an integer indicating the number of species saved (k).
<code>tol</code>	a tolerance threshold for null values (a value less than <code>tol</code> in absolute terms is considered as zero).
<code>give.list</code>	a logical indicating whether a list of optimizing species should be provided. If <code>give.list = TRUE</code> , <code>optimEH</code> provides the list of the k species that optimize the amount of evolutionary history preserved and are the most original species in their clades. If <code>give.list = FALSE</code> , <code>optimEH</code> returns directly the real value giving the amount of evolutionary history preserved.

Value

If `give.list=FALSE`, function `optimEH` returns a numeric.

If `give.list=TRUE`, function `optimEH` returns a list containing:

<code>value</code>	a real value providing the amount of evolutionary history preserved.
<code>selected.sp</code>	a data frame containing the list of the k species which optimize the amount of evolutionary history preserved and are the most original species in their clades.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

with contributions of Stephane Dray

References

Pavoine, S., Ollier, S. and Dufour, A.-B. (2005) Is the originality of a species measurable? *Ecology Letters*, **8**, 579–586.

See Also

See Also [randEH](#)

Examples

```
data(carni70, package="adephylo")
tre <- read.tree(text=carni70$tre)
EH(tre)
optimEH(tre, 10)
```

orisaved

Maximal or Minimal Amount of Originality Saved under Optimal Conditions

Description

Function `orisaved` computes the maximal or minimal amount of originality saved over all combinations of species optimizing the amount of evolutionary history preserved. The originality of a species is measured with the QE-based index (Pavoine et al. 2005).

Usage

```
orisaved(phy1, rate = 0.1, method = 1:2)
```

Arguments

<code>phy1</code>	an object inheriting the class <code>phylo</code> (see package <code>ape</code>), <code>phylo4</code> (see package <code>phylobase</code>) or <code>hclust</code> .
<code>rate</code>	a real value (between 0 and 1) indicating how many species will be saved for each calculation. For example, if the total number of species is 70 and <code>'rate = 0.1'</code> then the calculations will be done at a rate of 10% i.e. for 0 (= 0 %), 7 (= 10 %), 14 (= 20 %), 21 (= 30 %), ..., 63 (= 90 %) and 70(= 100 %) species saved. If <code>'rate = 0.5'</code> then the calculations will be done for only 0 (= 0 %), 35 (= 50 %) and 70(= 100 %) species saved.
<code>method</code>	an integer either 1 or 2 (see details).

Details

1 = maximum amount of originality saved 2 = minimum amount of originality saved

Value

Returns a numeric vector.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>
with contributions of Stephane Dray

References

Pavoine, S., Ollier, S. and Dufour, A.-B. (2005) Is the originality of a species measurable? *Ecology Letters*, **8**, 579–586.

See Also

See Also [optimEH](#), [randEH](#)

Examples

```
data(carni70, package="adephylo")
tre <- read.tree(text=carni70$tre)
tmax <- orisaved(tre, rate = 1 / 70, method = 1)
tmin <- orisaved(tre, rate = 1 / 70, method = 2)
plot(c(0, 1:70), tmax, xlab = "nb of species saved", ylab = "Originality saved", type = "l")
lines(c(0, 1:70), tmin, lty = 2)
```

PADDis

Functional Dissimilarity Measures for Presence-Absence Data

Description

Functions PADDis, DJac and Jac calculate the dissimilarity coefficients introduced in Ricotta et al. (2016). These dissimilarity coefficients use traditional mismatching components a, b and c of the 2 x 2 contingency table expressed as to include functional or phylogenetic differences among species and noted A, B, C. Components B and C represent the functional or phylogenetic uniqueness of community X compared with community Y and vice versa. Component A represent the functional or phylogenetic similarities between communities X and Y.

Usage

```
PADDis(comm, dis, method = NULL, diag = FALSE, upper = FALSE)
```

```
DJac(comm, dis, diag = FALSE, upper = FALSE)
```

```
Jac(comm, diag = FALSE, upper = FALSE)
```

Arguments

comm	a matrix or a data frame with communities (or plots, assemblages, etc.) as rows and species as columns containing the incidence (0/1) of all species in the communities.
dis	an object of class <code>dist</code> containing the (functional) dissimilarities among species
method	an integer between 0 and 5. If <code>NULL</code> the choice is made with a console message. See details.
diag	a logical value indicating whether the diagonal of the distance matrix should be printed by function <code>print.dist</code>

upper a logical value indicating whether the upper triangle of the distance matrix should be printed by function `print.dist`

Details

In PADDIS, dissimilarities among communities are calculated with the following formulas:

Generalized Jaccard dissimilarity, with `method = 1`

$$\frac{B + C}{a + b + c}$$

Generalized Sorensen dissimilarity, with `method = 2`

$$\frac{B + C}{2a + b + c}$$

Generalized Sokal and Sneath dissimilarity, with `method = 3`

$$\frac{2(B + C)}{a + 2(b + c)}$$

Generalized Ochiai dissimilarity, with `method = 4`

$$\frac{\sqrt{A + B}\sqrt{A + C} - A}{\sqrt{a + b}\sqrt{a + c}}$$

Generalized Simpson dissimilarity, with `method = 5`

$$\frac{\min\{B + C\}}{a + \min\{b + c\}}$$

DJac and Jac use the additive decomposition of the Jaccard index into turnover and richness difference. DJac takes into account the (functional or phylogenetic) dissimilarities among species while Jac does not.

Value

In function PADDIS, if `method=0`, then the function PADDIS returns 6 matrices corresponding to the a, b, c, A, B, and C values per pair of communities. Otherwise, it returns an object of class `dist` corresponding to the dissimilarities among communities.

Functions DJac and Jac return a list of three objects of class `dist`:

J	for the full dissimilarities between communities
JRepl	for the turnover component of the dissimilarities
JRich	for the component of difference in richness

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Ricotta, C., Podani, J., Pavoine, S. (2016) A family of functional dissimilarity measures for presence and absence data. *Ecology and Evolution*, **6**, 5383–5389

Examples

```
data(RPP16EE)
Com <- RPP16EE$Com
Dis <- as.dist(RPP16EE$Dis)
J <- Jac(Com)
DJ <- DJac(Com, Dis)

plot(c(as.matrix(DJ$J)[1,]), ylab="Dissimilarity",
      xlab="Plot-to-plot comparison", pch=15, type="b",
      ylim=c(0,1), main="Jaccard")

lines(c(as.matrix(J$J)[1,]), type="b", pch=18)

legend("bottomright", legend=c("P/A scores", "functional data"),
      pch=c(15,18), lty=1)

plot(c(as.matrix(DJ$JRepl)[1,]), ylab="Dissimilarity",
      xlab="Plot-to-plot comparison", pch=15, type="b",
      ylim=c(0,1), main="Species replacement")

lines(c(as.matrix(J$JRepl)[1,]), type="b", pch=18)

legend("bottomright", legend=c("P/A scores", "functional data"),
      pch=c(15,18), lty=1)

#Use the following instruction to obtain all components:

PADDis(Com, Dis, method=0)
```

QE

Quadratic Entropy

Description

Function QE calculates Rao's quadratic entropy within communities

Function discomQE calculates Rao's dissimilarities between communities

Usage

```
QE(comm, dis = NULL, formula = c("QE", "EDI"), scale = FALSE)
```

```
discomQE(comm, dis = NULL, structures = NULL, formula = c("QE", "EDI"))
```

Arguments

<code>comm</code>	a data frame or a matrix with communities as rows and species as columns. Entries are abundances of species within sites
<code>dis</code>	either NULL or an object of class <code>dist</code> that contains the (functional or phylogenetic) distances among species. If NULL the Gini-Simpson index is used.
<code>formula</code>	either "QE" (default) or "EDI". See details.
<code>scale</code>	a logical value indicating whether or not the diversity coefficient should be scaled by its maximal value over all species abundance distributions.
<code>structures</code>	either NULL or a data frame that contains, in the j th row and the k th column, the name of the group of level k to which the j th community belongs. Communities in <code>structures</code> should be in the same order as in <code>comm</code> .

Details

If `formula = "QE"`, the definition of the quadratic entropy is:

$$QE(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} d_{kl}$$

where $\mathbf{p}_i = (p_{1|i}, \dots, p_{k|i}, \dots, p_{S|i})$ is the vector of relative species abundance within community i ; S is the number of species; $\mathbf{D} = (d_{kl})$ is the matrix of (phylogenetic or functional) dissimilarities among species, and d_{kl} is the (phylogenetic or functional) dissimilarity between species k and l . For the calculations of dissimilarities between communities see the description of the apportionment of quadratic entropy in Pavoine et al. (2016) and references therein.

If `formula = "EDI"`, the definition of the quadratic entropy is:

$$EDI(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} \frac{d_{kl}^2}{2}$$

EDI stands for the Euclidean Diversity Index of Champely and Chessel (2002) (equation 3 in Pavoine et al. 2004). If EDI is used, the dissimilarities calculated by `discomQE` are obtained from the same scheme (equation 4 in Pavoine et al. 2004).

Value

Function `QE` returns a data frame with communities as rows and the diversity within communities as columns.

If `structures` is NULL, function `discomQE` returns an object of class `dist`. Otherwise it returns a list of objects of class `dist`.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Gini, C. (1912) *Variabilita e mutabilita*. Universite di Cagliari III, Parte II.
- Simpson, E.H. (1949) Measurement of diversity. *Nature*, **163**, 688.
- Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.
- Champely, S. and Chessel, D. (2002) Measuring biological diversity using Euclidean metrics. *Environmental and Ecological Statistics*, **9**, 167–177.
- Pavoine, S., Dufour, A.B., Chessel, D. (2004) From dissimilarities among species to dissimilarities among communities: a double principal coordinate analysis. *Journal of Theoretical Biology*, **228**, 523–537.
- Pavoine, S., Marcon, E., Ricotta, C. (2016) "Equivalent numbers" for species, phylogenetic, or functional diversity in a nested hierarchy of multiple scales. *Methods in Ecology and Evolution*, **7**, 1152–1163.

Examples

```

data(ecomor, package="ade4")
dtaxo <- dist.taxo(ecomor$taxo)
QE(t(ecomor$habitat), dtaxo, formula="EDI")
QE(t(ecomor$habitat), dtaxo^2/2, formula="QE")
table.value(as.matrix(discomQE(t(ecomor$habitat), dtaxo, formula="EDI")))
EDIcom <- discomQE(t(ecomor$habitat), dtaxo, formula="EDI")
QEcom <- discomQE(t(ecomor$habitat), dtaxo^2/2, formula="QE")
QEcom
EDIcom^2/2

bird.QE <- QE(t(ecomor$habitat), dtaxo, formula="EDI")
dotchart(bird.QE$diversity, labels = rownames(bird.QE))

data(humDNAm, package="ade4")
QE(t(humDNAm$samples), humDNAm$distances/2, formula="QE")
QE(t(humDNAm$samples), sqrt(humDNAm$distances), formula="EDI")
QEhumDNA.dist <- discomQE(t(humDNAm$samples),
humDNAm$distances/2, humDNAm$structures)
is.euclid(QEhumDNA.dist$communities)
is.euclid(QEhumDNA.dist$regions)

EDIhumDNA.dist <- discomQE(t(humDNAm$samples),
sqrt(humDNAm$distances), humDNAm$structures, formula="EDI")
is.euclid(EDIhumDNA.dist$communities)
is.euclid(EDIhumDNA.dist$regions)

QEhumDNA.dist$communities
EDIhumDNA.dist$communities^2/2

hum.QE <- QE(t(humDNAm$samples), humDNAm$distances/2, formula="QE")
dotchart(hum.QE$diversity, labels = rownames(hum.QE))

```

qHdiv	<i>Biodiversity Measure that Includes Consistent Interspecific and Intraspecific Components</i>
-------	---

Description

Function qHdiv calculates the parametric diversity index developed by Pavoine and Izsak (2014)

Usage

```
qHdiv(comm, C, q = 2)
```

Arguments

comm	a data frame or a matrix typically with communities as rows, species as columns and an index of abundance as entries.
C	a matrix that contains measures of the chosen intraspecific components (as defined in Pavoine and IZsak 2014) on the diagonal and measures of interspecific components off diagonal. These interspecific components reflect functional or phylogenetic similarities among species.
q	a positive numeric for the value of the parameter q.

Value

a vector with the diversity in each of the communities (same order as in comm.)

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S., Izsak, J. (2014) New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, **5**, 165–172.

See Also

See Also [twoHmax](#), [CFprop](#), [CFbinary](#)

Examples

```
tre <- "((((((sA:4,sB:1):1,sC:3):2,((sD:2,sE:1):1,sF:1):2):1,sG:7):1,sH:1):3,(sI:2,sJ:1):2):0;"
#The number of tips is kept in parameter n:
n<-10
# Next we need to obtain matrix CP.
library(ape)
phyape <- read.tree(text = tre)
plot(phyape)
```

```

CP <- vcv(phyape)
WP <- diag(diag(CP))
# With this particular illustration, a maximizing vector,
# for 2H used with CP, that does not contain any zero can be found.
# This maximizing vector can thus be obtained directly,
# instead of being estimated.
# Two equivalent equations have been given to obtain
# the maximizing vector in Appendix S1 of
# Pavoine and Izsak (2014).
# We use the first one below:

Pmax<-(solve(CP^2)%*%diag(CP))/sum(solve(CP^2)%*%diag(CP))
Pmax

# The second equation equivalently provides:

Z <- ((diag(1/sqrt(diag(CP))))%*%CP%*(diag(1/sqrt(diag(CP)))))^2
Pmax<-(solve(WP)%*%solve(Z)%*%rep(1,n))/sum(solve(WP)%*%solve(Z))
Pmax

# Applied to our case study, the function twoHmax
# provides good approximations:

twoHmax(CP)

# Redundancy among variables:
data(rhone, package="ade4")
V <- rhone$tab
# First consider the covariances among the variables:
C <- cov(V)
# A vector that maximizes 2H applied to C is estimated
# as follows:
pmax_covariances <- twoHmax(C)$vector
dotchart(as.matrix(pmax_covariances))

# If we apply 2H only to the diagonal matrix with the variances
# of the variables, the vector that maximizes 2H is:
W <- diag(diag(C))
rownames(W)<-colnames(W)<-rownames(C)
pmax_variances <- twoHmax(W)$vector
dotchart(as.matrix(pmax_variances))

# If C contains the correlations among variables,
# a vector that maximizes 2H applied to C is estimated
# as follows:
C <- cor(V)
pmax_correlations <- twoHmax(C)$vector
dotchart(as.matrix(pmax_correlations))

# By attributing equal weights to the variables,
# 2H applied to the correlation matrix measures
# the number of effective variables:
# from 0 if all variables are completely correlated

```



```

# with each other to n if they are not correlated.
# Similarly, by attributing equal weights to the variables,
# 2H applied to the covariance matrix measures
# the effective amount of variation:
# from 0 if all variables are completely correlated
# with each other to n if they are not correlated
# and have similar variances.
C <- cor(V)
equalproportions <- cbind.data.frame(rep(1/ncol(C), ncol(C)))
names(equalproportions) <- "equalprop"
qHdiv(equalproportions, C)

#Even if the data set contains 15 variables,
# the effective number of variables is lower:
C <- cov(V)
equalproportions <- cbind.data.frame(rep(1/ncol(C), ncol(C)))
names(equalproportions) <- "equalprop"
qHdiv(equalproportions, C)

# When considering the covariances, instead of the correlations,
# the effective number of variables is even lower:
# indicating also an imbalance in the variances
# of the variables.

```

randEH	<i>Amount of Evolutionary History Preserved when Random Species are Saved</i>
--------	---

Description

When branch lengths in an ultrametric phylogenetic tree are expressed as divergence times, the total sum of branch lengths in that tree expresses the amount of evolutionary history. The function randEH calculates the amount of evolutionary history preserved when k random species out of n original species are saved.

Usage

```
randEH(phy1, nbofsp, nrep = 10)
```

Arguments

phy1	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust.
nbofsp	an integer indicating the number of species saved (k).
nrep	an integer indicating the number of random sampling.

Value

Returns a numeric vector with the amount of evolutionary history preserved by each random drawing of the k species to be saved.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>
with contributions of Stephane Dray

References

- Nee, S. and May, R.M. (1997) Extinction and the loss of evolutionary history. *Science*, **278**, 692–694.
- Pavoine, S., Ollier, S. and Dufour, A.-B. (2005) Is the originality of a species measurable? *Ecology Letters*, **8**, 579–586.

See Also

See Also [optimEH](#)

Examples

```
data(carni70, package = "adephylo")
tre <- read.tree(text=carni70$tre)
EH(tre)
R <- randEH(tre, 10, nrep=1000)
hist(R)
```

rare_Rao

Functional Rarefaction for Species Abundance Data

Description

The function `Rare_Rao` performs distance-based rarefaction curves using species abundance data. It finds the expected functional diversity (if functional distances between species are used) as a function of the sampling effort. Two approaches are available: an analytical solution, a resampling approach.

Usage

```
rare_Rao(comm, dis, sim = TRUE, resampling = 999, formula = c("QE", "EDI"))
```

Arguments

- | | |
|-------------------|---|
| <code>comm</code> | a data frame or a matrix with samples as rows, species as columns, and abundance or frequencies as entries. |
| <code>dis</code> | an object of class <code>dist</code> containing pairwise distances among species. The distance matrix should be squared Euclidean or simply Euclidean (see definition and details in Appendix S1 of Ricotta et al. (2012)). Alternatively, if <code>dis</code> is <code>NULL</code> , the Gini-Simpson index is used. |
| <code>sim</code> | a logical; if <code>TRUE</code> , the resampling approach is used; if <code>FALSE</code> , the analytical solution is given. |

resampling	number of times data are resampled to calculate the mean functional rarefaction curve (used if sim=TRUE)
formula	Either "QE" (default) or "EDI". See details.

Details

If formula = "QE", the definition of the quadratic entropy is:

$$QE(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} d_{kl}$$

where $\mathbf{p}_i = (p_{1|i}, \dots, p_{k|i}, \dots, p_{S|i})$ is the vector of relative species abundance within sample i ; S is the number of species; $\mathbf{D} = (d_{kl})$ is the matrix of (phylogenetic or functional) dissimilarities among species, and d_{kl} is the (phylogenetic or functional) dissimilarity between species k and l .

If formula = "EDI", the definition of the quadratic entropy is:

$$EDI(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S p_{k|i} p_{l|i} \frac{d_{kl}^2}{2}$$

EDI stands for the Euclidean Diversity Index of Champely and Chessel (2002) (equation 3 in Pavoine et al. 2004).

Value

If sim = TRUE, the function returns a data frame containing the Expected Rao Quadratic Entropy (column 'ExpRao'), the limits of the 95% Confidence Interval (columns 'LeftIC' and 'RightIC') for each subsample dimension (M) out of the total set of samples (N). If sim = FALSE, the function returns a data frame containing the analytical solution for the Expected Rao Quadratic Entropy (column 'ExpRao') for each subsample dimension (M) out of the total set of samples (N).

Author(s)

Giovanni Bacaro and Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Ricotta, C., Pavoine, S., Bacaro, G., Acosta, A. (2012) Functional rarefaction for species abundance data. *Methods in Ecology and Evolution*, **3**, 519–525.
- Champely, S. and Chessel, D. (2002) Measuring biological diversity using Euclidean metrics. *Environmental and Ecological Statistics*, **9**, 167–177.
- Pavoine, S., Dufour, A.B., Chessel, D. (2004) From dissimilarities among species to dissimilarities among communities: a double principal coordinate analysis. *Journal of Theoretical Biology*, **228**, 523–537.

See Also

See Also [QE](#)

Examples

```

data(aviurba, package="ade4")
distances<-dist.ktab(ktab.list.df(list(aviurba$traits)), type = "N")
# The distances should be squared Euclidean;
# note that Euclidean distances can be used
# as they also are squared Euclidean.

abundances<- aviurba$fau
rare_Rao(abundances, distances, sim = TRUE, resampling = 100)
rare_Rao(abundances, distances, sim = FALSE)

```

RDMCCP16

Theoretical Data Set used in Ricotta et al. (2016)

Description

Hypothetical communities C1-C9 composed of nine species S1-S9 with varying abundances divided into three groups of three species (S1-S3, S4-S6 and S7-S9) (say, legumes, herbs and forbs). For simplicity, all species within the same group are functionally identical to each other, while two species belonging to different groups are always maximally dissimilar (Ricotta et al. 2016).

Usage

```
data("RDMCCP16")
```

Format

A list of two objects:

ab, a matrix with communities as rows, species as columns and abundance values as entries.

dis, an object of class `dist` that contains theoretical dissimilarities between species.

Source

Table 1 in Ricotta et al. (2016)

References

Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B.E., Pavoine, S. (2016). Measuring the functional redundancy of biological communities: A quantitative guide. *Methods in Ecology and Evolution*, **7**, 1386–1395.

Examples

```

data(RDMCCP16)
uniqueness(RDMCCP16$ab, as.dist(RDMCCP16$dis))

```

 Rentropy

Pavoine et al. (2017) functional or phylogenetic entropy

Description

Function Rentropy calculates Pavoine et al. (2017) functional or phylogenetic R entropy within communities; this index is closely related to Rao's quadratic entropy

Usage

```
Rentropy(comm, dis = NULL, scale = FALSE)
```

Arguments

comm	a data frame or a matrix with communities as rows and species as columns. Entries are abundances of species within sites
dis	either NULL or an object of class <code>dist</code> that contains the (functional or phylogenetic) distances among species. If NULL species are said equidistant with a distance of 1 between any two species and a distance of 0 between a species and itself.
scale	a logical value indicating whether or not the diversity coefficient should be scaled by its maximal value over all species abundance distributions.

Details

The definition of the R entropy is:

$$R(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^S \sum_{l=1}^S \sqrt{p_{k|i}} \sqrt{p_{l|i}} d_{kl}$$

where $\mathbf{p}_i = (p_{1|i}, \dots, p_{k|i}, \dots, p_{S|i})$ is the vector of relative species abundance within community i ; S is the number of species; $\mathbf{D} = (d_{kl})$ is the matrix of (phylogenetic or functional) dissimilarities among species, and d_{kl} is the (phylogenetic or functional) dissimilarity between species k and l .

Value

Function Rentropy returns a data frame with communities as rows and the R entropy within communities as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S., Bonsall, M.B., Dupaix, A., Jacob, U., Ricotta, C. (2017) From phylogenetic to functional originality: guide through indices and new developments. *Ecological Indicators*, **82**, 196–205.

Examples

```
data(ecomor, package="ade4")
dtaxo <- dist.taxo(ecomor$taxo)
bird.R <- Rentropy(t(ecomor$habitat), dtaxo^2/2)
dotchart(bird.R$diversity, labels = rownames(bird.R))
```

```
data(batcomm)
phy <- read.tree(text=batcomm$tre)
dphy <- as.dist(cophenetic(phy))/2
ab <- batcomm$ab[, phy$tip.label]
bat.R <- Rentropy(ab, dphy)
dotchart(bat.R$diversity, labels = rownames(bat.R))
```

rlqESLTP

Linking Patterns in Phylogeny, Traits, Abiotic Variables and Space

Description

An extension of the RLQ approach to identify potential environmental filters in species' traits in a phylogenetic and spatial context.

Usage

```
rlqESLTP(dudiE, dudiS, dudiL, dudiT, dudiP, ...)

## S3 method for class 'rlqESLTP'
plot(x, which = NULL, phyl = NULL,
xy = NULL, traits = NULL, env = NULL, type = NULL, ax = 1, ...)

## S3 method for class 'rlqESLTP'
summary(object, ...)
```

Arguments

dudiE	an object of class dudi of package ade4 (e.g. a principal component applied to environmental variables or a principal coordinate analysis applied to environmental distances between plots). Plots must be weighted as in dudiL. dudiE can be set to NULL if dudiS is not.
dudiS	an object of class dudi of package ade4 (e.g. an eigenvector decomposition applied to a spatial neighbourhood matrix (see Pavoine et al. 2011 for examples)). Plots must be weighted as in dudiL. dudiS can be set to NULL if dudiE is not.
dudiL	an object of class dudi obtained with function dudi.coa of ade4 applied to a plot (rows) by species (columns) data frame. dudiL cannot be set to NULL.
dudiT	an object of class dudi of package ade4 (e.g. a principal component analysis applied to functional traits of species or a principal coordinate analysis applied to functional distances between species). Species must be weighted as in dudiL. dudiT can be set to NULL if dudiP is not.

dudiP	an object of class dudi of package ade4 (e.g. a principal coordinate analysis applied to phylogenetic distances between species). Species must be weighted as in dudiL. dudiP can be set to NULL if dudiT is not.
x	an object of class rlqESTP (obtained with function rlqESTP).
object	an object of class rlqESTP (obtained with function rlqESTP).
which	a character that might be "E" (in that case parameters env and type must be given); "S" (in that case parameter xy must be given); "T" (in that case parameters traits and type must be given); or "P" (in that case parameter phyl must be given).
phyl	an object inheriting the class phylo (see package ape), phylo4 (see package phylobase) or hclust.
xy	a data frame with two columns giving the coordinates of the plots (longitude, latitude in that order).
traits	a list of data frames for the traits, each data frame contains variables from a single statistical type (see parameter type below). NAs are allowed. However, in that case, the calculations depend, for each trait, on the subset of available data only.
env	a list of data frames for the environmental variables, each data frame contains variables from a single statistical type (see parameter type below). NAs are allowed. However, in that case, the calculations depend, for each variable, on the subset of available data only.
type	a vector that provides the type of each table to be analysed in env (if environmental variables should be displayed) or traits (if the focus is on species' traits). The possible types are "Q" (quantitative, numeric), "O" (ordinal), "N" (nominal), "D" (dichotomous), "F" (fuzzy, or expressed as a proportion), "B" (multichoice nominal variables, coded by binary columns), "C" (circular). Values in type must be in the same order as in env or traits.
ax	a numeric indicating the axis of interest.
...	further arguments passed to or from other methods.

Value

Function `rlqESLTP` returns an object of classes `rlqESLTP`, `rlq` and `dudi`. It is a list of 26 objects:

tab	a data frame. Crossed Table (CT): crossing the columns of the merged trait and phylogenetic table with those of the merged environmental and spatial table.
cw	a vector of numerics. Weights attributed to the columns of the merged trait and phylogenetic table.
lw	a vector of numerics. weights attributed to the columns of the merged environmental and spatial table.
eig	a vector of numerics. The vector of eigenvalues.
rank	an integer. The total number of axes in the analysis.
nf	a numeric. The number of axes kept.
c1	a data frame. Principal axes. Normed scores for the columns of merged trait and phylogenetic table.

co	a data frame. Scores for the columns of merged trait and phylogenetic table.
l1	a data frame. Principal axes. Normed scores for the columns of merged environmental and spatial table.
li	a data frame. Scores for the columns of merged environmental and spatial table.
call	call
lQ	data frame. Scores for the species.
mQ	data frame. Normed scores for the species.
lR	data frame. Scores of the plots.
mR	data frame. Normed scores for the plots.
aQ	data frame. Correlations trait/phylogeny axes / coinertia axes.
aR	data frame. Correlations environment/space axes / coinertia axes.
lR_givenE	data frame. Contributions of environmental information to plot scores.
lR_givenS	data frame. Contributions of spatial information to plot scores.
lQ_givenT	data frame. Contributions of trait information to species scores.
lQ_givenP	data frame. Contributions of phylogenetic information to species scores.
row.w	a vector. Weights attributed to plots.
col.w	a vector. Weights attributed to species.
dudiL	object of class dudi (dudiL).
dudiR	object of class dudi (merging information on environment and space).
dudiQ	object of class dudi (merging information on traits and phylogeny).

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Pavoine, S., Vela, E., Gachet, S., de Belair, G., Bonsall, M.B. (2011) Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *Journal of Ecology*, **99**, 165–175.

Examples

```
data(mafragh, package="ade4")
xy <- mafragh$xy
#The object that defines the neighbourhood between plots is in
mneig <- mafragh$mneig
#mneig is an object of class neig
s.label(mafragh$xy, nb = mafragh$nb, paxes.draw = FALSE)

#The environmental variables are in
env <- mafragh$env
names(env)
#The abundance data are in
flo <- mafragh$flo
```



```

# Adjustment of the writing of species names
names <- gsub(" ", "_", mafragh$spenames[colnames(flo), 1])
for(i in 1:26){
names <- gsub(LETTERS[i], letters[i], names)
}
names <- gsub("alisma_plantago", "alisma_plantago_aquatica", names)
colnames(flo) <- names
#The data on traits are in
traits <- lapply(mafragh$traits, function(x) x[colnames(flo), , drop=FALSE])
#The phylogenetic tree is in
phy <- read.tree(text=mafragh$tre)
plot(phy)

#Traits are separated by statistical type. The object traits is a list of data frames.
tabBinary <- prep.binary(traits$tabBinary, c(3, 4))
tabQuantitative <- traits$tabQuantitative
tabCircular <- prep.circular(traits$tabCircular, 1, 12)
tabOrdinal <- traits$tabOrdinal

#Analyses:
coaflo <- dudi.coa(flo, scan = FALSE, nf = 55)
vecspa <- scores.neig(mneig)
pcaspa <- dudi.pca(vecspa, coaflo$lw, scan = FALSE, nf = 96)
#We first removed environmental variables
env <- env[-(8:10)]
env[4:8] <- log(env[4:8])
pcaenv <- dudi.pca(env, coaflo$lw, scale = FALSE, scan = FALSE, nf = 8)
disT <- dist.ktab(ktab.list.df(list(tabBinary,
tabOrdinal[c("Spikiness", "Hairy leaves")])), c("B", "O"))
# The definition of the functional distances between species
# were slightly different in Pavoine et al. (2011).

pcotraits <- dudi.pco(disT, coaflo$cw, full = TRUE)
pcophy <- dudi.pco(sqrt(as.dist(cophenetic(phy)[names(flo), names(flo)])/2), coaflo$cw, full = TRUE)

rlqmix <- rlqESLTP(pcaenv, pcaspa, coaflo, pcotraits, pcophy, scan = FALSE, nf = 2)
barplot(rlqmix$eig)
rlqmix$eig[1]/sum(rlqmix$eig)
plot(rlqmix, xy=xy, ax=1, wh="S")
plot(rlqmix, phy=phy, ax=1, wh="P")
plot(rlqmix, traits=tabBinary, ax=1, type="B", wh="T")
plot(rlqmix, traits=tabOrdinal[2:3], ax=1, type="O", wh="T")
plot(rlqmix, env=env, ax=1, type="Q", wh="E")

```

Description

The data set was analyzed in Pavoine et al. (2009) to determine whether the observed decrease in the number of individuals and species in the Southern California Bight through years due to fishing activities was accompanied by changes in the phylogenetic structure of the community.

Usage

```
data("rockfish")
```

Format

The format is a list of 3 objects:

tre, a character: the phylogenetic tree in newick format.

fau, a data frame with years as rows, species as columns (genus: *Sebastes*), and abundance as entries.

traits, a data frame with species as rows (same as in fau and two variables as columns: MaxSize, the maximum body size; and Vu1, an index of species' vulnerability).

Details

The abundance data were obtained from the Marine Recreational Fishery Statistics Survey (MRFSS). We considered the compositions of rockfish assemblages caught by party and charter boats with hooks and lines from 1980-1986, 1993-1994, 1996, 1998-2007. The phylogenetic tree was obtained from Hyde and Vetter (2007).

Source

Appendixes of Pavoine et al. (2009).

References

Pavoine, S., Love, M., Bonsall, M. (2009) Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: *Sebastes*) in the Southern California Bight. **Ecology Letters**, **12**, 898–908.

Hyde, J.R. and Vetter, R.D. (2007). The origin, evolution, and diversification of rockfishes of the genus *Sebastes* (Cuvier). *Molecular Phylogenetics and Evolution*, **44**, 790–811.

Examples

```
data(rockfish)
phy <- read.tree(text=rockfish$tre)
plot(phy)
```

RP15EI	<i>Theoretical Data Set used in Ricotta and Pavoine (2015) in Ecological Indicators</i>
--------	---

Description

The data set was used in Ricotta and Pavoine (2015) to show how the coefficients of multi-site dissimilarity that they developed can be applied to data.

Usage

```
data("RP15EI")
```

Format

The format is a list of 4 data frames. Each data frame gives the presence (1) and absence (0) of 10 species (columns, S1-S10) in 8 plots (rows P1-P8). These data are characterized by different levels of nestedness and turnover in species compositions between sites.

M1: Intermediate nestedness and turnover.

M2: Perfectly nested pattern.

M3: Very high species turnover.

M4: Random configuration.

Source

Figure 1 of Ricotta and Pavoine (2015)

References

Ricotta, C. and Pavoine, S. (2015) A multiple-site dissimilarity measure for species presence/absence data and its relationship with nestedness and turnover. *Ecological Indicators*, **54**,203–206.

Examples

```
data(RP15EI)  
betastatjac(RP15EI$M1)
```

RP15JVS

Theoretical Data Set used in Ricotta and Pavoine (2015) in Journal of Vegetation Science

Description

The data set was used in Ricotta and Pavoine (2015) to illustrate the relationships between the coefficients of similarity between communities that they developed. These data represent an artificial ecological gradient.

Usage

```
data("RP15JVS")
```

Format

The format is a list of four objects:

ab, a data frame with communities as rows, species as columns, and number of individuals as entries.

D1, a data frame of pairwise dissimilarities between species. In D1, dissimilarities reflect the species ecological differences along the artificial gradient of table ab: interspecies dissimilarities were set roughly proportional to the distance between the locations of the species optima and to the difference between their optimal abundances (see Ricotta and Pavoine 2015 for details).

D2, a data frame of pairwise dissimilarities between species. In D2, dissimilarities were randomly assigned using an even distribution.

D3, a data frame of pairwise dissimilarities between species. In D3, dissimilarities were uniformly set to $d_{ij} = 2/3$.

Source

Appendixes of Ricotta and Pavoine (2015)

References

Ricotta, C. and Pavoine, S. (2015) Measuring similarity among plots including similarity among species: An extension of traditional approaches. *Journal of Vegetation Science*, **26**, 1061–1067

Examples

```
data(RP15JVS)
dissABC(RP15JVS$ab, RP15JVS$D1, method="J", option=1)
```

RPP16EE	<i>Theoretical Data Set used in Ricotta et al. (2016) in Ecology and Evolution</i>
---------	--

Description

The data set was used in Ricotta et al. (2016) to show how the coefficients of plot-to-plot dissimilarity that they developed can be applied to data.

Usage

```
data("RPP16EE")
```

Format

The format is a list of 2 objects:

Com, a data frame. Artificial data table composed of 15 species (S1-S15) (columns) and 9 plots (P1-P9) (rows).

Dis, a data frame with the artificial dissimilarities between species.

Source

Appendixes 1 and 2 of Ricotta et al. (2016)

References

Ricotta, C., Podani, J., Pavoine, S. (2016) A family of functional dissimilarity measures for presence and absence data. *Ecology and Evolution*, **6**, 5383–5389.

Examples

```
data(RPP16EE)
table.value(RPP16EE$Com)
Jac(RPP16EE$Com)
```

speciesdiv	<i>Indices of Species Diversity</i>
------------	-------------------------------------

Description

The function speciesdiv calculates diversity indices that rely on relative or absolute species abundance.

Usage

```
speciesdiv(comm, method = "full", tol = 1e-8)
```

Arguments

comm	a data frame or a matrix typically with communities as rows, species as columns and presence/absence or abundance as entry. Note that with presence/absence (0/1) data, only species richness will be calculated correctly.
method	a character or a vector of characters: one or several of "richness", "GiniSimpson", "Simpson", "Shannon", "Margalef", "Menhinick", "McIntosh", "full". See details.
tol	a tolerance threshold (a value between -tol and tol is considered equal to zero)

Details

Let S_i be the number of species in community i , n_{ij} be the absolute abundance of species j in community i , N_i the sum of all species abundance in community i ($N_i = \sum_j n_{ij}$; the sum of row i in comm), p_{ij} the relative abundance of species j in community i ($p_{ij} = n_{ij}/N_i$). If method="richness", the diversity index is the number of species. If method="GiniSimpson", the diversity index is that of Gini (1912) and Simpson (1949): $1 - \sum_j p_{ij}^2$. If method="Simpson", the diversity index is (Simpson 1949): $1/\sum_j p_{ij}^2$. If method="Shannon", the diversity index is that of Shannon (1948) with neperian logarithm: $-\sum_j p_{ij} \log(p_{ij})$. If method="Margalef", the diversity index is that of Margalef (1972): $(S_i - 1)/\log(N_i)$. If method="Menhinick", the diversity index is that of Menhinick (): $S_i/\sqrt{N_i}$. If method="McIntosh", the diversity index is that of McIntosh (1967): $(N_i - \sqrt{\sum_j n_{ij}^2})/(N_i - \sqrt{N_i})$. If one of the characters is "full", then all indices are calculated.

Value

Function speciesdiv returns a matrix with communities as rows and the diversity indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Gini, C. (1912) Variabilita e mutabilita. Studi economicoaguridici delle facoltta di giurizprudenza dell, Universite di Cagliari III, Parte II.
- Magurran, A.E. (2004) Measuring biological diversity. Blackwell Publishing, Oxford, U.K.
- Margalef, R. (1972) Homage to Evelyn Hutchinson, or why is there an upper limit to diversity? *Transactions of the Connecticut Academy of Arts and Sciences*, **44**, 211–235.
- McIntosh, R.P. (1967) An index of diversity and the relation of certain conepts to diversity. *Ecology*, **48**, 392–404.
- Shannon, C.E. (1948) A mathematical theory of communication. *Bell System technical journal*, **27**, 379–423, 623–656.
- Simpson, E.H. (1949) Measurement of diversity. *Nature*, **163**, 688.

Examples

```
data(batcomm)
ab <- batcomm$ab
speciesdiv(ab)
```

specieeve

*Indices of Species Evenness***Description**

The function `specieeve` calculates evenness indices that rely on relative or absolute species abundance.

Usage

```
specieeve(comm, method = "full", tol = 1e-8)
```

Arguments

<code>comm</code>	a data frame or a matrix typically with communities as rows, species as columns and abundance as entry.
<code>method</code>	a character or a vector of characters: one or several of "GiniSimpson", "Simpson", "Shannon", "Heip", "McIntosh", "Carmargo", "SmithWilson", "full". See details.
<code>tol</code>	a tolerance threshold (a value between <code>-tol</code> and <code>tol</code> is considered equal to zero)

Details

Let S_i be the number of species in community i , n_{ij} be the absolute abundance of species j in community i , N_i the sum of all species abundance in community i ($N_i = \sum_j n_{ij}$; the sum of row i in `comm`), p_{ij} the relative abundance of species j in community i ($p_{ij} = n_{ij}/N_i$). If `method="GiniSimpson"`, the evenness index is that associated with Gini (1912) and Simpson (1949) diversity index: $(1 - \sum_j p_{ij}^2) * S_i / (S_i - 1)$. If `method="Simpson"`, the evenness index is (Simpson 1949; Magurran 2004): $(1 / \sum_j p_{ij}^2) / S_i$. If `method="Shannon"`, the evenness index is that associated with Shannon (1948) diversity index with neperian logarithm: $(-\sum_j p_{ij} \log(p_{ij})) / \ln(S_i)$. If `method="Heip"`, the evenness index is that of Heip (1974) (Magurran 2004): $[\exp(-\sum_j p_{ij} \log(p_{ij})) - 1] / (S_i - 1)$. If `method="McIntosh"`, the evenness index is that of Pielou (1975) associated with McIntosh (1967) index of diversity: $(N_i - \sqrt{\sum_j n_{ij}^2}) / (N_i - N_i / \sqrt{S_i})$. If `method="SmithWilson"`, the Smith and Wilson (1996) evenness index is calculated (Magurran 2004): $1 - [(2/\pi)(\arctan(\sum_{j=1}^{S_i} (\ln n_{ij} - \sum_{k=1}^{S_i} \ln n_{ik} / S_i)^2 / S_i))]$. If one of the characters is "full", then all indices are calculated.

Value

Function `specieeve` returns a matrix with communities as rows and the evenness indices as columns.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

- Gini, C. (1912) *Variabilita e mutabilita*. Studi economicoaguridici delle facoltta di giurizprudenza dell, Universite di Cagliari III, Parte II.
- Heip, C. (1974) A new index measuring evenness. *Journal of the Marine Biological Association UK*, **54**, 555–557.
- Magurran, A.E. (2004) *Measuring biological diversity*. Oxford, UK: Blackwell Publishing.
- McIntosh, R.P. (1967) An index of diversity and the relation of certain conepts to diversity. *Ecology*, **48**, 392–404.
- Pielou, E.C. (1975) *Ecological diversity*. New York: Wiley InterScience.
- Shannon, C.E. (1948) A mathematical theory of communication. *Bell System technical journal*, **27**, 379–423, 623–656.
- Simpson, E.H. (1949) Measurement of diversity. *Nature*, **163**, 688.
- Smith, B. and Wilson, J.B. (1996) A consumer’s guide to evenness measures. *Oikos*, **76**, 70–82.

Examples

```
data(batcomm)
ab <- batcomm$ab
specieeve(ab)
```

twoHmax

Maximization of the Diversity Index TwoH

Description

The R function twoHmax maximizes function 2H (Pavoine and Izsak 2014) for a given matrix **C** of (functional or phylogenetic) similarities between species. It is based on function divcmax of the ade4 package in R. As function divcmax of package ade4, function twoHmax uses an optimization technique based on Rosen’s projection gradient algorithm and is verified using the Kuhn-Tucker conditions.

Usage

```
twoHmax(C, epsilon = 1e-08, smooth = TRUE, comment = FALSE)
```


Arguments

C	a matrix that contains measures of the chosen intraspecific components (see functions CFprop, multiCFprop, CFbinary, multiCFbinary) on the diagonal and measures of interspecific components off diagonal. These interspecific components reflect functional or phylogenetic similarities among species.
epsilon	a numeric tolerance threshold: a frequency is non null if it is higher than epsilon.
smooth	a logical value: if TRUE, the estimated positive proportions are checked and re-estimated with exact matrix products.
comment	a logical value indicating whether or not comments on the optimization technique should be printed.

Value

	A list of two objects
value	the maximum value of index 2H (see function qHdiv).
vector	a data frame with the vector pmax that maximizes index 2H (see function qHdiv).

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

The code is a modification of function divcmax of package ade4 written by Stephane Champely.

References

Pavoine, S. and Izsak, J. (2014) New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, **5**, 165–172.

See Also

See Also [qHdiv](#), [CFprop](#), [CFbinary](#)

Examples

```
tre <- "((((((sA:4,sB:1):1,sC:3):2,((sD:2,sE:1):1,sF:1):2):1,sG:7):1,sH:1):3,(sI:2,sJ:1):2):0;"
#The number of tips is kept in parameter n:
n<-10
# Next we need to obtain matrix CP.
phyape <- read.tree(text = tre)
plot(phyape)
CP <- vcv(phyape)
WP <- diag(diag(CP))
# With this particular illustration, a maximizing vector,
# for 2H used with CP, that does not contain any zero can be found.
# This maximizing vector can thus be obtained directly,
# instead of being estimated. Two equivalent equations
# have been given to obtain the maximizing vector in
# Appendix S1 of Pavoine and Izsak (2014).
# We use the first one below
```

```

Pmax<-(solve(CP^2)%*%diag(CP))/sum(solve(CP^2)%*%diag(CP))
Pmax

# The second equation equivalently provides
Z <- ((diag(1/sqrt(diag(CP))))%*%CP%*(diag(1/sqrt(diag(CP)))))^2
Pmax<-(solve(WP)%*%solve(Z)%*%rep(1,n))/sum(solve(WP)%*%solve(Z))
Pmax

# Applied to our case study, the function twoHmax provides good approximations
twoHmax(CP)

# Redundancy among variables:
data(rhone, package="ade4")
V <- rhone$tab
# First consider the covariances among the variables:
C <- cov(V)
# A vector that maximizes 2H applied to C is estimated as follows:
pmax_covariances <- twoHmax(C)$vector
dotchart(as.matrix(pmax_covariances))

# If we apply 2H only to the diagonal matrix with the variances
# of the variables, the vector that maximizes 2H is:
W <- diag(diag(C))
rownames(W)<-colnames(W)<-rownames(C)
pmax_variances <- twoHmax(W)$vector
dotchart(as.matrix(pmax_variances))

# If C contains the correlations among variables,
# a vector that maximizes 2H applied to C is estimated as follows:
C <- cor(V)
pmax_correlations <- twoHmax(C)$vector
dotchart(as.matrix(pmax_correlations))

# By attributing equal weights to the variables,
# 2H applied to the correlation matrix measures
# the number of effective variables:
# from 0 if all variables are completely correlated
# with each other to n if they are not correlated.
# Similarly, by attributing equal weights to the variables,
# 2H applied to the covariance matrix measures
# the effective amount of variation:
# from 0 if all variables are completely correlated
# with each other to n if they are not correlated
# and have similar variances.
C <- cor(V)
equalproportions <- cbind.data.frame(rep(1/ncol(C), ncol(C)))
names(equalproportions) <- "equalprop"
qHdiv(equalproportions, C)

#Even if the data set contains 15 variables,
# the effective number of variables is lower.

```

```

C <- cov(V)
equalproportions <- cbind.data.frame(rep(1/ncol(C), ncol(C)))
names(equalproportions) <- "equalprop"
qHdiv(equalproportions, C)

# When considering the covariances among species,
# instead of the correlations, the effective number
# of variables is even lower: indicating also an imbalance
# in the variances of the variables.

```

uniqueness	<i>Functional Uniqueness and Functional Redundancy of Biological Communities</i>
------------	--

Description

Hypothetical communities C1-C9 composed of nine species S1-S9 with varying abundances divided into three groups of three species each displayed with different shades (S1-S3, S4-S6 and S7-S9). All species within the same group are functionally identical to each other, while two species belonging to different groups are always maximally dissimilar.

Usage

```
uniqueness(comm, dis, tol = 1e-08, abundance = TRUE)
```

Arguments

comm	a matrix or a data frame containing the abundance or incidence (0/1) of species in communities (or plots). Columns are species and communities are rows.
dis	an object of class <code>dist</code> containing the functional distances among species
tol	a tolerance threshold (a value between <code>-tol</code> and <code>tol</code> is considered as null)
abundance	a logical. If <code>TRUE</code> , abundance data are used when available; if <code>FALSE</code> , incidence (0/1) data are used.

Value

The function `uniqueness` returns a list of three data frames:

`kbar`: this first data frame gives values for Ricotta et al. coefficient \bar{K}_i 's per species (rows) and community (columns).

`V`: this second data frame gives values for Ricotta et al. coefficient V_i 's per species (rows) and community (columns).

`red`: this third data set gives values for Ricotta et al. coefficients N (species richness), Q (quadratic diversity), D (Simpson diversity), and U (uniqueness) per community; in this third data frame, coefficients are columns and communities are rows, the coefficients are thus calculated per community only.

Author(s)

Sandrine Pavoine <sandrine.pavoine@mnhn.fr>

References

Ricotta, C, de Bello, F, Moretti, M, Caccianiga, M, Cerabolini, BE, Pavoine, S (2016). Measuring the functional redundancy of biological communities: A quantitative guide. *Methods in Ecology and Evolution*, **7**, 1386–1395.

See Also

See Also [QE](#)

Examples

```
data(RDMCCP16)
uniqueness(RDMCCP16$ab, as.dist(RDMCCP16$dis))
```

Index

*Topic **datasets**

batcomm, [10](#)
birdData, [12](#)
RDMCCP16, [76](#)
rockfish, [81](#)
RP15EI, [83](#)
RP15JVS, [84](#)
RPP16EE, [85](#)

*Topic **models**

abgevodivparam, [3](#)
aptree, [7](#)
betastat, [11](#)
CF, [14](#)
crossdpcoa, [15](#)
decdiv, [19](#)
dissABC, [22](#)
dissRicotta, [25](#)
distinctDis, [26](#)
distinctTopo, [27](#)
distinctTree, [29](#)
distinctUltra, [30](#)
distMS, [31](#)
divparam, [32](#)
dsimcom, [34](#)
EH, [39](#)
EqRao, [40](#)
eveparam, [45](#)
evoCA, [47](#)
evodiss, [49](#)
evodiss_family, [51](#)
evodivparam, [54](#)
evoNSCA, [57](#)
evopca, [59](#)
K, [62](#)
optimEH, [64](#)
orisaved, [65](#)
PADDis, [66](#)
QE, [68](#)
qHdiv, [71](#)

randEH, [73](#)
rare_Rao, [74](#)
Rentropy, [77](#)
rlqESLTP, [78](#)
speciesdiv, [85](#)
specieseve, [87](#)
twoHmax, [88](#)
uniqueness, [91](#)

abgaptree (aptree), [7](#)
abgevodivparam, [3](#), [57](#)
adiv (adiv-package), [3](#)
adiv-package, [3](#)
apd (aptree), [7](#)
aptree, [7](#)

batcomm, [10](#)
betastat, [11](#)
betastatjac (betastat), [11](#)
betastatsor (betastat), [11](#)
birdData, [12](#)

CF, [14](#)
CFbinary, [71](#), [89](#)
CFbinary (CF), [14](#)
CFprop, [71](#), [89](#)
CFprop (CF), [14](#)
crossdpcoa, [15](#)
crossdpcoa_maineffect (crossdpcoa), [15](#)
crossdpcoa_version1 (crossdpcoa), [15](#)
crossdpcoa_version2 (crossdpcoa), [15](#)

decdiv, [19](#)
discomQE, [24](#)
discomQE (QE), [68](#)
dissABC, [22](#)
dissRicotta, [25](#)
dist.binary, [51](#)
distinctDis, [26](#), [28](#), [29](#), [31](#)
distinctTopo, [27](#), [27](#), [29](#), [31](#)

- distinctTree, [27](#), [28](#), [29](#), [31](#)
- distinctUltra, [27–29](#), [30](#)
- distMS, [31](#)
- divparam, [6](#), [32](#), [57](#)
- DJac (PADDis), [66](#)
- dsimcom, [34](#)
- dsimFun (dsimcom), [34](#)
- dsimTaxo (dsimcom), [34](#)
- dsimTree (dsimcom), [34](#)
- EH, [39](#)
- EqRao, [40](#)
- EqRS (EqRao), [40](#)
- EqRSintra (EqRao), [40](#)
- eveparam, [45](#)
- evoCA, [47](#), [59](#), [62](#)
- evodiss, [6](#), [24](#), [48](#), [49](#), [53](#), [59](#), [62](#)
- evodiss_family, [51](#), [51](#)
- evodiss_ternaryplot (evodiss_family), [51](#)
- evodivparam, [54](#)
- evoNSCA, [48](#), [57](#), [62](#)
- evopca, [59](#)
- evopcachord, [48](#), [59](#)
- evopcachord (evopca), [59](#)
- evopcahellinger, [48](#), [59](#)
- evopcahellinger (evopca), [59](#)
- Jac (PADDis), [66](#)
- K, [62](#)
- krandtest, [9](#)
- Kstar, [63](#)
- Kstar (K), [62](#)
- Kw, [63](#)
- Kw (K), [62](#)
- multiCFbinary (CF), [14](#)
- multiCFprop (CF), [14](#)
- optimEH, [39](#), [64](#), [66](#), [74](#)
- orisaved, [65](#)
- PADDis, [66](#)
- pIa (aptree), [7](#)
- plot.abgevodivparam (abgevodivparam), [3](#)
- plot.aptree (aptree), [7](#)
- plot.decdiv (decdiv), [19](#)
- plot.divparam (divparam), [32](#)
- plot.eveparam (eveparam), [45](#)
- plot.evoCA (evoCA), [47](#)
- plot.evodivparam (evodivparam), [54](#)
- plot.rlqESLTP (rlqESLTP), [78](#)
- plot.rtestaptree (aptree), [7](#)
- QE, [9](#), [20](#), [68](#), [75](#), [92](#)
- qHdiv, [15](#), [71](#), [89](#)
- randEH, [39](#), [65](#), [66](#), [73](#)
- randtest, [9](#), [63](#)
- rare_Rao, [74](#)
- RDMCCP16, [76](#)
- Rentropy, [77](#)
- rlqESLTP, [78](#)
- rockfish, [81](#)
- RP15EI, [83](#)
- RP15JVS, [84](#)
- RPP16EE, [85](#)
- rtestaptree (aptree), [7](#)
- rtestdecdiv (decdiv), [19](#)
- rtestEqRao (EqRao), [40](#)
- rtestEqRS (EqRao), [40](#)
- rtestEqRSintra (EqRao), [40](#)
- rtestwapqe (EqRao), [40](#)
- speciesdiv, [85](#)
- specieseve, [87](#)
- SQ (dsimcom), [34](#)
- summary.rlqESLTP (rlqESLTP), [78](#)
- tecAptree (aptree), [7](#)
- twoHmax, [71](#), [88](#)
- uniqueness, [91](#)
- wapqe (EqRao), [40](#)