Package ‘PCPS’

February 19, 2015

Type Package
Title Principal Coordinates of Phylogenetic Structure
Version 1.0.2
Date January 13, 2015
Author Vanderlei Julio Debastiani
Maintainer Vanderlei Julio Debastiani
        <vanderleidebastiani@yahoo.com.br>
Depends SYNCSA, vegan
Imports ape, picante, phylobase, plotrix
Description Set of functions for analysis of Principal Coordinates of Phylogenetic Structure (PCPS).
License GPL-2
NeedsCompilation no
Repository CRAN
Date/Publication 2015-01-13 14:49:22

R topics documented:

  define.clade .................................................. 2
  pcoa.sig .................................................. 3
  pcps .................................................. 5
  pcps.curve .................................................. 6
  pcpps.sig .................................................. 8
  self.belonging ........................................... 10

Index 12
Define clade

Description

Function to define groups (clades) in a phylogenetic tree.

Usage

```r
define.clade(tree, threshold, time, method = c("threshold", "time"))
```

Arguments

- `tree`: Phylogenetic tree.
- `threshold`: A threshold value to form the groups.
- `time`: A cutting height (age) to form the groups.
- `method`: Method to define the clades, "threshold" or "time".

Details

In the method `threshold` the total length of phylogenetic tree is used as cutting factor. If `threshold` is near to zero the cutting is near the root, if `threshold` near to one cutting is near the tips.

The phylogenetic tree must contain the node labels for the function work. Use the `makeNodeLabel` for defining node labels in a flexible way.

Value

- `clades`: Tips and their clades.
- `height`: The cutting height.

Author(s)

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>

See Also

- `makeNodeLabel`

Examples

```r
require(ape)
tree<-makeNodeLabel(rcoal(20))
clades<-define.clade(tree, threshold = 0.8, method = "threshold")
clades
plot.phylo(tree, show.node.label = TRUE)
abline(v = clades$height)
```
**pcoa.sig**

**Significant dimensions in principal coordinate analysis**

**Description**

Function for determine the number of significant dimensions in principal coordinate analysis (PCoA).

**Usage**

```r
pcoa.sig(data, dist = "gower", correction = "none", squareroot = FALSE,
          n.start = NULL, axis = 6, iterations = 1000)
```

## S3 method for class 'pcoasig'

- `print(x, ...)`
- `summary(object, ...)`
- `scores(x, choices = c(1, 2), ...)`

**Arguments**

- `data` Community data matrix.
- `dist` Dissimilarity index, as accepted by `vegdist` (Default dist="gower").
- `correction` Correction methods for negative eigenvalues, as accepted by `pcoa`: "lingoes" and "cailliez" (Default correction="none").
- `squareroot` Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = FALSE).
- `axis` Maximum number of ordination principal axes to be monitored (Default axis=6).
- `n.start` Initial sample size. One sampling unit is added at each sampling step. If n.start = NULL initial sample size is equal to total sample size (Default n.start=NULL).
- `iterations` Number of permutations to assess significance (Default iterations=1000).
- `object, x` An object of class pcoasig.
- `choices` Axes for re-scaling. Choices must have length equal to two (Default choices = c(1,2)).
- `...` Other parameters for the respective functions.

**Details**

At each iteration step a bootstrap sample is subjected to PCoA ordination, the scores are submitted to a procrustean adjustment, and the correlation between observed and bootstrap ordination scores is computed. It compares such correlations to the same parameter generated in a parallel bootstrapped ordination of randomly permuted data. The number of axes in bootstrap or null PCoA with eigenvectors corresponding to positive eigenvalues may be smaller than the number of axes monitored, in this case, axes with values equal to 0 are created. The number of iterations with original values for each axis is shown in n.permut.bootstrap and n.permut.null.

The function scores.pcoasig re-scales the correlation values for biplot graphics.
**Value**

- PCoA: PCoA result, exactly as returned for the pcoa function.
- correlations: Correlations between axis and original data.
- mean.cor.null: Mean correlations, for axis, between null and reference scores.
- mean.cor.bootstrap: Mean correlations, for axis, between bootstrap and reference scores.
- cumulative.frequency: Cumulative frequency in which the null correlations were greater than the bootstrap correlation.
- n.permut.bootstrap: Number of iterations for each axis in bootstrap step.
- n.permut.null: Number of iterations for each axis in null step.
- probabilities: Probabilities for each axis.

**Note**

**Principal Component Analysis (PCA)**

You can use the same function to determine the number of significant dimensions in principal component analysis (PCA). For this, standardize each variable for zero mean and uni variance (function `decostand` and method `standardize`) and use euclidean distance as dissimilarity index.

**Interpretation**

If the higher dimension is significant, then all lower dimensions will also be significant.

**Author(s)**

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>

**References**


**See Also**

- `pcoa`
- `procrustes`

**Examples**

```r
data(flona)
res<-pcoa.sig(flona$community, axis = 6, dist = "bray", iterations = 100)
res
summary(res)
```
pcps

Principal Coordinates of Phylogenetic Structure

Description

Function to generate Principal Coordinates of Phylogenetic Structure (PCPS).

Usage

```r
pcps(comm, distNspp, method = "bray", squareroot = TRUE)
```

Arguments

- `comm`: Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
- `distNspp`: Matrix containing phylogenetic distances between species.
- `method`: Dissimilarity index, as accepted by `vegdist` (Default dist="bray").
- `squareroot`: Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).
- `object.x`: An object of class pcps.
- `choices`: Axes for re-scaling. Choices must have length equal to two (Default choices = c(1, 2)).
- `display`: Display text or points for the sampling units.
- `groups`: Factor giving the groups (Clades) for each species.
- `showlabel`: Label the groups by their names in the centroid of the object.
- `...`: Other parameters for the respective functions.

Details

The function obtains a matrix containing phylogeny-weighted species composition (matrix.p) and is submitted to principal coordinates analysis (PCoA). This method generates the principal coordinates of phylogenetic structure (PCPS) (Duarte, 2011).

The function scores.pcps re-scales the correlation values for biplot graphics. The function plot.pcps draws a simple biplot and represent clades as "spider" graphs (see ordispider).
Value

P          Phylogeny-weighted species composition matrix.
values     The eigenvalues, relative eigenvalues and cumulative relative eigenvalues.
vectors    The principal coordinates of phylogenetic structure (PCPS).
correlations Correlations between a PCPS axis and phylogenetically weighted species abundances or frequencies.

Note

IMPORTANT: The sequence species show up in the community data matrix MUST be the same as they show up in the phylogenetic distance matrix. See organize.synCs.

Author(s)

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>

References


See Also

matrix.p, wcmdscale

Examples

data(flona)
res<-pcps(flona$community, flona$phylo)
res
summary(res)
scores(res)
plot(res, display = "text", groups = c(rep("Clade-A", 30), rep("Clade-B", 29)))

pcps.curve   Curve of phylogenetic signal at metacommunity level

Description

The function estimate the phylogenetic signal at metacommunity level and draws a representation curve.
Usage

pcps.curve(comm, dist.spp, trait, method = "bray", squareroot = TRUE, 
null.model = TRUE, runs = 99, progressbar = FALSE)

## S3 method for class 'pcpscurve'
print(x, ...)

## S3 method for class 'pcpscurve'
plot(x, type = "b", errorbars = c("none", "sd", "se", "quantile"), 
probs = c(0.025, 0.975), col = "black", errbar.col = "black",
errbar.pch = 16, ...)

Arguments

comm Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
dist.spp Matrix containing phylogenetic distances between species.
trait Matrix data of species described by traits, with traits as columns and species as rows.
method Dissimilarity index, as accepted by vegdist (Default dist="bray").
squareroot Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).
null.model Logical argument (TRUE or FALSE) to specify if use null model to generate null curves (Default null.model = TRUE)
runs Number of randomizations.
progressbar Logical argument (TRUE or FALSE) to specify if display a progress bar on the R console (Default progressbar = FALSE).
x An object of class pcpscurve.
type Type of the plot to be drawn (Default type="b").
errorbars Plot type of error bars; none (none), standard deviation (sd), standard error (se) and quantiles corresponding to the given probabilities (quantile).
probs Numeric vector of probabilities used by quantile. (Default probs=c(0.025,0.975)).
col Plot color.
errbar.col Color of error bars.
errbar.pch Symbol to use of error bars.
... Further graphical parameters for points.

Details

The PCPS are used, in a sequential manner, as predictors in a linear regression to model the trait averages across the metacommunity. The curve is drawn as the percentage of cumulative eigenvalues in the abscissa and as the determination coefficient of regressions in the ordinate.

The null curves are generated via a null model that shuffles terminal tips across the phylogenetic tree, generates a set of random PCPS and recalculates the curves.
Value

- curve.obs: The cumulative PCPS eigenvalues and the coefficient of determination.
- curve.null: The cumulative PCPS eigenvalues and the coefficient of determination for each randomization.

Author(s)

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>

References


See Also

- matrix.p.pcps

Examples

data(flona)
res_curve<-pcps.curve(flona$community, flona$phylo, flona$trait[,1], method = "bray",
  squareroot = TRUE, null.model = TRUE, runs = 9, progressbar = FALSE)
res_curve
plot(res_curve, type = "b", errorbars = "se", col = "red",
  errbar.col = "black", errbar.pch = 21,pch = 19)

pcps.sig  Association between PCPS axes and environmental predictors

Description

Analyze the association between PCPS axes and environmental predictors via null model.

Usage

pcps.sig(comm, dist.spp, envir, method = "bray",
  squareroot = TRUE, formula, family = gaussian,
  runs = 999, AsFactors = NULL)

Arguments

- comm: Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
- dist.spp: Matrix containing phylogenetic distances between species.
- envir: Environmental variables for each community, with variables as columns and sampling units as rows.
method
Dissimilarity index, as accepted by `vegdist` (Default dist="bray").

squeroot
Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).

formula
An object of class `formula` quotation marks. See Details.

family
A description of the error distribution to be used in the model. See `family` (Default family = gaussian).

runs
Number of permutations for assessing significance.

AsFactors
Encode an environmental variable as factor. See Details.

Details
The function generates principal coordinates of phylogenetic structure (`pcps`) and run a generalized linear model (`glm`) with a gaussian error distribution (see `family`). The significance is obtained via two null models, one that shuffles sites across the environmental gradient and another that shuffles terminal tips(taxa) across the phylogenetic tree. The first null model (site shuffle) shuffles the site position across the environmental gradient and rerun the same model, generating a null F value. The second null model (taxa shuffle), shuffles terminal tips across the phylogenetic tree and generates a set of null PCPS. Each null PCPS is submitted to a procrustean adjustment (see `procrustes`), and the fitted values between observed PCPS and null PCPS is obtained. The adjusted null PCPS is used to rerun the model, generating another null F value. The observed F value is compared independently with both null sets of F values to generate a probability value of the original F value being generated merely by chance according to each null model.

The item formula is an expression of the form pcps.1 ~ model. The response term must be the pcps name, for example pcps.1, pcps.2, pcps.12.

The item AsFactors changes a environmental variable for the class `factor`. The sequence is the same that in the environmental data matrix. Use `c` to combine more that one variable.

Value

model
The model, an object of class glm.

Envir_class
The class of each variable in environmental data.

formula
The formula used.

f.obs
Observed F value in GLM.

p.site.shuffle
The p value for the site shuffle null model.

p.taxa.shuffle
The p value for the taxa shuffle null model.

Author(s)

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>

References

self.belonging

See Also

`matrix.p, pcps, procrustes, glm`

Examples

data(flona)
pcps.sig(flona$community, flona$phylo, flona$environment,
    formula = "pcps.1-alt", runs = 100)

---

self.belonging

Degree of self belonging of species

Description

Define the degree of self belonging of species.

Usage

self.belonging(dis, standardize = TRUE)

Arguments

dis Matrix containing distance between species.
standardize Logical argument (TRUE or FALSE) to specify if dis must be standardize in values into range 0 from 1 (Default standardize = TRUE).

Details

For the calculation of self-belonging of a set of species the dissimilarities between the species are transformed into similarities and used to define degrees of belonging to fuzzy sets (Pillar et al. 2009; Pillar & Duarte 2010). Every species among all species specifies a fuzzy set in relation to all other species, with a certain degree of belonging. The self-belonging of a given species i expresses its degree of belonging to the root node of the phylogenetic/functional tree, conditioned to the similarities between i and all other internal nodes connecting it to the root.

Value

The self-belonging for each species.

Author(s)

Vanderlei Júlio Debastiani <vanderleidebastiani@yahoo.com.br>
self.belonging

References


See Also

belonging

Examples

data(flona)
self.belonging(flona$phylo)
Index

*Topic PCPS
  define.clade, 2
  pcoa.sig, 3
  pcps, 5
  pcps.curve, 6
  pcps.sig, 8
  self.belonging, 10

belonging, 11
biplot, 3, 5
c, 9
define.clade, 2
factor, 9
family, 9
formula, 9
glm, 9, 10
makeNodeLabel, 2
matrix.p, 5, 6, 8, 10
ordispider, 5
organize.syncsa, 6
pcoa, 3, 4
pcoa.sig, 3
pcps, 5, 8–10
pcps.curve, 6
pcps.sig, 8
plot.pcps (pcps), 5
plot.pcpscurve (pcps.curve), 6
print.pcoasig (pcoa.sig), 3
print.pcps (pcps), 5
print.pcpscurve (pcps.curve), 6
procrustes, 4, 9, 10
quantile, 7
scores.pcoasig (pcoa.sig), 3
scores.pcps (pcps), 5
self.belonging, 10
summary.pcoasig (pcoa.sig), 3
summary.pcps (pcps), 5
vegdist, 3, 5, 7, 9
wcmdscale, 6