

# Package ‘PCPS’

May 24, 2018

**Type** Package

**Title** Principal Coordinates of Phylogenetic Structure

**Version** 1.0.6

**Date** 2018-05-24

**Author** Vanderlei Julio Debastiani

**Maintainer** Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

**Depends** SYNCSA (>= 1.3.3)

**Imports** ape, picante, phylobase, vegan, RcppArmadillo, stats,  
graphics, parallel

**Description** Set of functions for analysis of Principal Coordinates of Phylogenetic Structure (PCPS).

**License** GPL-2

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2018-05-24 13:16:12 UTC

## R topics documented:

define.clade . . . . .	2
matrix.p.sig . . . . .	3
pcoa.sig . . . . .	5
pcpc.curve.calc . . . . .	7
pcps . . . . .	10
self.belonging . . . . .	12
wcmdscale.org . . . . .	13

<b>Index</b>	<b>14</b>
--------------	-----------

---

define.clade	<i>Define clade</i>
--------------	---------------------

---

### Description

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

### Usage

```
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 (age).

### Author(s)

Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

### See Also

[makeNodeLabel](#)

### Examples

```
require(ape)
tree<-makeNodeLabel(rcoal(10))
clades<-define.clade(tree, threshold = 0.8, method = "threshold")
clades
plot.phylo(tree, show.node.label = TRUE)
abline(v = clades$height)
```

---

matrix.p.sig	<i>Association between phylogeny-weighted species composition and environmental predictors</i>
--------------	--

---

## Description

Analyses to relate an environmental gradient to the phylogenetic assembly of species across a meta-community by means of phylogenetic fuzzy weighting.

## Usage

```
matrix.p.sig(comm, phylodist, envir, analysis = c("adonis", "mantel"),
  method = "bray", squareroot = TRUE, method.envir = "euclidean",
  runs = 999, parallel = NULL, newClusters = TRUE, CL = NULL)
```

```
pcps.sig(comm, phylodist, envir, analysis = c("glm", "rda"),
  method = "bray", squareroot = TRUE, formula, family = stats::gaussian,
  AsFactors = NULL, pcps.choices = c(1, 2, 3, 4), runs = 999,
  parallel = NULL, newClusters = TRUE, CL = NULL)
```

## Arguments

comm	Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
phylodist	Matrix containing phylogenetic distances between species.
envir	Environmental variables for each community, with variables as columns and sampling units as rows.
analysis	Type of analysis. For the function pcps.sig <a href="#">glm</a> or <a href="#">rda</a> , for matrix.p.sig function <a href="#">adonis</a> or <a href="#">mantel</a> . See Details.
method	Dissimilarity index, as accepted by <a href="#">vegdist</a> (Default dist = "bray").
squareroot	Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).
method.envir	Resemblance index between communities based on environmental variables, as accepted by <a href="#">vegdist</a> used in Mantel analysis (Default method.envir = "euclidean")
runs	Number of permutations for assessing significance.
parallel	Number of parallel processes. Tip: use <a href="#">detectCores()</a> (Default parallel = NULL).
newClusters	Logical argument (TRUE or FALSE) to specify if make new parallel processes or use predefined socket cluster. Only if parallel is different of NULL (Default newClusters = TRUE).
CL	A predefined socket cluster done with parallel package.
formula	An object of class <a href="#">formula</a> quotation marks used in GLM analysis. See Details.
family	A description of the error distribution to be used in used in GLM analysis. See <a href="#">family</a> (Default family = gaussian).

AsFactors	Encode an environmental variable as factor used in GLM analysis. The sequence is the same that in the environmental data matrix. See Details.
pcps.choices	PCPS used in RDA analysis (Default pcps.choices = c(1, 2, 3, 4)).

## Details

Each metacommunity is submitted to phylogenetic fuzzy weighting, generating a matrix that describing the phylogeny-weighted species composition of the communities ([matrix.p](#)). The function `matrix.p.sig` test directly the association this matrix with the environmental predictors. The pairwise dissimilarities are submitted to Mantel test ([mantel](#)) or ADONIS test ([adonis](#)) to evaluate the influence of an environmental gradient on species dispersion across the communities. The function `pcps.sig` generates principal coordinates of phylogenetic structure ([pcps](#)) and use a single axis for run a generalized linear model (GLM, [glm](#)) or use set of axis for run a distance-based redundancy analysis (db-RDA, [rda](#)).

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 (or r value in Mantel test). The second null model (taxa shuffle), shuffles terminal tips across the phylogenetic tree and generates a null matrix containing phylogeny-weighted species composition and rerun the same model, generating another null F value. In the `pcps.sig` function are generate set of null PCPS and each null PCPS (or set of PCPS in RDA) 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 (or r value) is compared independently with both null sets of F values (or r value) 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, not the order in the formula. Use `c` to combine more that one variable.

## Value

<code>model</code>	The model, an object of class <code>glm</code> , <code>rda</code> , <code>adonis</code> or <code>mantel</code> .
<code>envir_class</code>	The class of each variable in environmental data in <code>glm</code> .
<code>formula</code>	The formula used in <code>glm</code> .
<code>statistic.obs</code>	Observed F value or r value.
<code>p.site.shuffle</code>	The p value for the site shuffle null model.
<code>p.taxa.shuffle</code>	The p value for the taxa shuffle null model.

## Author(s)

Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

**References**

Duarte, L.S. (2011). Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, 120, 208:215.

**See Also**

[matrix.p](#), [pcps](#), [procrustes](#), [glm](#), [rda](#), [adonis](#), [mantel](#)

**Examples**

```
data(flona)
pcps.sig(flona$community, flona$phylo, flona$environment, analysis = "glm",
         formula = "pcps.1~alt", runs = 99)
matrix.p.sig(flona$community, flona$phylo, flona$environment[, 2, drop = FALSE],
            analysis = "adonis", runs = 99)
```

---

pcoa.sig

*Significant dimensions in principal coordinate analysis*

---

**Description**

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

**Usage**

```
pcoa.sig(data, method = "gower", squareroot = FALSE, axis = 6,
         n.start = NULL, by = 1, iterations = 1000, parallel = NULL,
         newClusters = TRUE, CL = NULL)
```

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

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

```
## S3 method for class 'pcoasig'
summary(object, choices = c(1, 2), ...)
```

**Arguments**

data	Community data matrix.
method	Method for dissimilarity index, as accepted by <a href="#">vegdist</a> (Default method = "gower").

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. If n.start = NULL initial sample size is equal to total sample size (Default n.start = NULL).
by	Sampling unit is added at each sampling step (Default by = 1).
iterations	Number of permutations to assess significance (Default iterations = 1000).
parallel	Number of parallel processes. Tip: use detectCores() (Default parallel = NULL).
newClusters	Logical argument (TRUE or FALSE) to specify if make new parallel processes or use predefined socket cluster. Only if parallel is different of NULL (Default newClusters = TRUE).
CL	A predefined socket cluster done with parallel package.
x	An object of class pcoasig.
...	Other parameters for the respective functions.
object	An object of class pcoasig.
choices	Axes for re-scaling. Choices must have length equal to two (Default choices = c(1, 2)).

## 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

value	The eigenvalues, relative eigenvalues and cumulative relative eigenvalues..
vectors	The principal coordinates.
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.
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 Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

**References**

Pillar, V.D. (1999). The bootstrapped ordination reexamined. *Journal of Vegetation Science* 10, 895-902.

**See Also**

[pcoa](#), [procrustes](#)

**Examples**

```
data(flona)
res<-pcoa.sig(flona$community, method = "bray", squareroot = TRUE, axis = 6, iterations = 100)
res
summary(res)$scores
```

---

pcpc.curve.calc

*Curve of phylogenetic signal at metacommunity level*

---

**Description**

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

**Usage**

```
pcpc.curve.calc(values, vectors, mt)
```

```
pcps.curve(comm, phylodist, trait, method = "bray", squareroot = TRUE,
  ranks = TRUE, null.model.ts = FALSE, null.model.bm = FALSE, tree,
  runs = 99, progressbar = FALSE, parallel = NULL, newClusters = TRUE,
  CL = NULL)
```

```
## S3 method for class 'pcpscurve'
plot(x, draw.model = c("none", "ts", "bm"), type = "b",
     probs = c(0.025, 0.975), col = "black", model.col = "black", ...)

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

## S3 method for class 'pcpscurve'
summary(object, probs = c(0.025, 0.975), ...)
```

## Arguments

values	The eigenvalues, relative eigenvalues and cumulative relative eigenvalues returned by <a href="#">pcps</a> .
vectors	The principal coordinates of phylogenetic structure returned by <a href="#">pcps</a> .
mt	Matrix containing trait average at community level for one trait.
comm	Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
phyloDist	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 <a href="#">vegdist</a> (Default dist = "bray").
squareroot	Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).
ranks	Logical argument (TRUE or FALSE) to specify if ordinal variables are convert to ranks (Default ranks = TRUE).
null.model.ts	Logical argument (TRUE or FALSE) to specify if use null model that shuffles terminal tips across the phylogenetic tree to generate null curves. See details (Default null.model.ts = FALSE).
null.model.bm	Logical argument (TRUE or FALSE) to specify if use null model that simulate trait evolving under Brownian motion to generate null curves. See details (Default null.model.bm = FALSE).
tree	Phylogenetic tree, as phylo object.
runs	Number of randomizations.
progressbar	Logical argument (TRUE or FALSE) to specify if display a progress bar on the R console (Default progressbar = FALSE).
parallel	Number of parallel processes. Tip: use <a href="#">detectCores()</a> (Default parallel = NULL).
newClusters	Logical argument (TRUE or FALSE) to specify if make new parallel processes or use predefined socket cluster. Only if parallel is different of NULL (Default newClusters = TRUE).
CL	A predefined socket cluster done with parallel package.
x	An object of class <code>pcpscurve</code> .
draw.model	Type of null model to draw; none (none), taxa shuffle (ts), browian motion model (bm).



type	Type of the plot to be drawn (Default type = "b").
probs	Numeric vector of probabilities used by <a href="#">quantile</a> . (Default probs = c(0.025, 0.975)).
col	Plot color.
model.col	Color of lines of null models.
...	Further graphical parameters for points.
object	An object of class pcpscurve.

### 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.

Two null models are available. The first one (ts), the null curves are generated shuffling terminal tips across the phylogenetic tree, generates a set of random PCPS and recalculates the curves. The second (bm), the null curves are generated with simulate traits evolving under Brownian motion model.

### Value

curve.obs	The cumulative PCPS eigenvalues and the coefficient of determination.
curve.null.ts	The cumulative PCPS eigenvalues and the coefficient of determination for each randomization using the taxa shuffle null model.
curve.null.bm	The cumulative PCPS eigenvalues and the coefficient of determination for each randomization using the Brownian motion null model.

### Author(s)

Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

### References

Duarte, L.S. (2011). Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, 120, 208:215.

### See Also

[matrix.p](#), [pcps](#)

### Examples

```
data(flona)
res<-pcps.curve(flona$community, flona$phylo, flona$trait[,1,drop = FALSE],
               null.model.ts = TRUE, runs = 9)
res
summary(res)
plot(res, draw.model = "ts", type = "b", col = "red")
```

**Description**

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

**Usage**

```
pcps(comm, phylodist, method = "bray", squareroot = TRUE,
      correlations = TRUE)

## S3 method for class 'pcps'
plot(x, groups = NULL, choices = c(1, 2), display = "text",
     showlabel = TRUE, ...)

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

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

scores.pcps(x, choices = c(1, 2), ...)

## S3 method for class 'pcps'
summary(object, choices = c(1, 2), ...)
```

**Arguments**

comm	Community data, with species as columns and sampling units as rows. This matrix can contain either presence/absence or abundance data.
phylodist	Matrix containing phylogenetic distances between species.
method	Dissimilarity index, as accepted by <a href="#">vegdist</a> (Default dist="bray").
squareroot	Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index (Default squareroot = TRUE).
correlations	Logical argument (TRUE or FALSE) to specify if are calculated the correlations between each PCPS and each species in matrix P (Default correlations = TRUE).
x	An object of class pcps.
groups	Factor giving the groups (Clades) for each species (Default groups = NULL).
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, partial match to "text" or "points" (Default display = "text").
showlabel	Label the groups by their names in the centroid of the object.
...	Other parameters for the respective functions.
object	An object of class pcps.

## 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 summary or the function scores.pcps re-scales the correlation values for obtain the scores for [biplot](#) graphics. The function plot 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.
scores	Scores for biplot graphics.

## 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.syncsa](#).

## Author(s)

Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

## References

Duarte, L.S. (2011). Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, 120, 208:215.

## See Also

[matrix.p](#), [wcmdscale](#), [ordispider](#), [ordilabel](#)

## Examples

```
data(ADRS)
res<-pcps(ADRS$community, ADRS$phylo)
res
summary(res)
summary(res, choices = c(1, 2))$scores
plot(res, display = "text", groups = c(rep("Clade-A", 2), rep("Clade-B", 4)))
```

---

self.belonging	<i>Degree of self belonging of species</i>
----------------	--

---

**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 Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

**References**

Pillar, V.D.; Duarte, L.d.S. (2010). A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, 13, 587:596.

Pillar, V.D., Duarte, L.d.S., Sosinski, E.E. & Joner, F. (2009). Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science*, 20, 334:348.

**See Also**

[belonging](#)

## Examples

```
data(ADRS)
self.belonging(ADRS$phylo)
```

---

wcmdscale.org

*Internal function for organize the results of wcmdscale function*

---

## Description

Internal function for organize the results of [wcmdscale](#) function. The function computes dissimilarity indices using the function [vegdist](#) and perform Principal Coordinates Analysis (PCoA) using the function [wcmdscale](#). If data is of class `dist`, the function do not compute the dissimilarity indices.

## Usage

```
wcmdscale.org(data, method, squareroot, eig, correlations, ...)
```

## Arguments

<code>data</code>	Data matrix or dissimilarities of class <code>dist</code> .
<code>method</code>	Method for dissimilarity index, as accepted by <a href="#">vegdist</a> .
<code>squareroot</code>	Logical argument (TRUE or FALSE) to specify if use square root of dissimilarity index.
<code>eig</code>	Logical argument (TRUE or FALSE) to indicate if eigenvalues are returned.
<code>correlations</code>	Logical argument (TRUE or FALSE) to indicate if correlations between axis and original data are returned.
<code>...</code>	Other arguments passed to <code>wcmdscale</code> function.

## Value

<code>values</code>	The eigenvalues, relative eigenvalues and cumulative relative eigenvalues.
<code>vectors</code>	The principal coordinates.
<code>correlations</code>	Correlations between axis and original data.

## Author(s)

Vanderlei Julio Debastiani <[vanderleidebastiani@yahoo.com.br](mailto:vanderleidebastiani@yahoo.com.br)>

## See Also

[vegdist](#), [wcmdscale](#)

# Index

## \*Topic **PCPS**

- define.clade, 2
  - matrix.p.sig, 3
  - pcoa.sig, 5
  - pcpc.curve.calc, 7
  - pcps, 10
  - self.belonging, 12
  - wcmdscale.org, 13
- adonis, 3–5
- belonging, 12
- biplot, 6, 11
- c, 4
- define.clade, 2
- factor, 4
- family, 3
- formula, 3
- glm, 3–5
- makeNodeLabel, 2
- mantel, 3–5
- matrix.p, 4, 5, 9, 11
- matrix.p.sig, 3
- ordilabel, 11
- ordispider, 11
- organize.syncsa, 11
- pcoa, 7
- pcoa.sig, 5
- pcpc.curve.calc, 7
- pcps, 4, 5, 8, 9, 10
- pcps.curve (pcpc.curve.calc), 7
- pcps.sig (matrix.p.sig), 3
- plot.pcps (pcps), 10
- plot.pcpscurve (pcpc.curve.calc), 7
- print.pcoasig (pcoa.sig), 5
- print.pcps (pcps), 10
- print.pcpscurve (pcpc.curve.calc), 7
- print.summarypcoasig (pcoa.sig), 5
- print.summarypcps (pcps), 10
- procrustes, 4, 5, 7
- quantile, 9
- rda, 3–5
- scores.pcps (pcps), 10
- self.belonging, 12
- summary.pcoasig (pcoa.sig), 5
- summary.pcps (pcps), 10
- summary.pcpscurve (pcpc.curve.calc), 7
- vegdist, 3, 5, 8, 10, 13
- wcmdscale, 11, 13
- wcmdscale.org, 13