adephylo: exploratory analyses for the phylogenetic comparative method

Thibaut Jombart and Stéphane Dray

February 21, 2015

Contents

1 Introduction

This document describes the adephylo package for the R software. adephylo aims at implementing exploratory methods for the analysis of phylogenetic comparative data, i.e. biological traits measured for taxa whose phylogeny is also provided. This package extends and replaces implementation of phylogeny-related methods in the ade4 package http://pbil.univ-lyon1.fr/ADE-4/home.php?lang=eng.

Procedures implemented in adephylo rely on exploratory data analysis. They include data visualization and manipulation, tests for phylogenetic

autocorrelation, multivariate analysis, computation of phylogenetic proximities and distances, and modelling phylogenetic signal using orthonormal bases.

These methods can be used to visualize, test, remove or investigate the phylogenetic signal in comparative data. The purpose of this document is to provide a general view of the main functionalities of `adephylo`, and to show how this package can be used along with `ape`, `phylobase` and `ade4` to analyse comparative data.

2 First steps

2.1 Data representation: why we are not reinventing the weel

Data representation can be defined as the way data are stored in a software (R, in our case). Technically, data representation is defined by classes of objects that contain the information. In the case of phylogeny and comparative data, very efficient data representation are already defined in other packages. Hence, it makes much more sense to use directly objects from these classes.

Phylogenies are best represented in Emmanuel Paradis's `ape` package (http://ape.mpl.ird.fr/), as the class `phylo`. As `ape` is by far the largest package dedicated to phylogeny, using the `phylo` class assures a good interoperability of data. This class is defined in an online document: http://ape.mpl.ird.fr/misc/FormatTreeR_28July2008.pdf.

However, data that are to be analyzed in `adephylo` do not only contain trees, but also traits associated to the tips of a tree. The package `phylobase` (http://r-forge.r-project.org/projects/phylobase/) is a collaborative effort designed to handling such data. Its representation of phylogenies slightly differs from that of `ape`; the class `phylo4` was originally an extension of the `phylo` class into formal (S4) class, but it has now evolved into something more original. The S4 class `phylo4d` ('d' for 'data') can be used to store a tree and data associated to tips, internal nodes, or even edges of a tree. Classes of `phylobase` are described in a vignette of the package, accessible by typing:

```
> vignette("phylobase")
```

As trees and comparative data are already handled by `ape` and `phylobase`, no particular data format shall be defined in `adephylo`. In particular, we are no longer using `phylog` objects, which were used to represent phylogenies in `ade4` in a very *ad hoc* way, without much compatibility with other packages. This class is now deprecated, but all previous functionalities available for `phylog` objects have been re-implemented and – in some cases – improved in `adephylo`.

2.2 Installing the package

What is tricky here is that a vignette is basically available once the package is installed. Assuming you got this document before installing the package, here are some clues about installing `adephylo`.

First of all, `adephylo` depends on other packages, being `methods`, `ape`, `phylobase`, and `ade4`. These dependencies are mandatory, that is, you actually need to have these packages installed before using `adephylo`. Also, it is better to make sure you are using the latest versions of these packages. This can be achieved using the `update.packages` command, or by installing devel versions from R-Forge (http://r-forge.r-project.org/). In all cases, the latest version of `adephylo` can be found from http://r-forge.r-project.org/R/?group_id=303.

We load *adephylo*, alongside some useful packages:

```
> library(ape)
> library(phylobase)
> library(ade4)
> library(adephylo)
> search()
```

```
 [1] ".GlobalEnv"        "package:adephylo"  "package:ade4"
 [4] "package:phylobase" "package:grid"      "package:ape"
 [7] "package:stats"     "package:graphics"  "package:grDevices"
[10] "package:utils"     "package:datasets"  "package:methods"
[13] "Autoloads"         "package:base"
```

Note that possibly conflicting, deprecated functions or datasets from `ade4` are masked by `adephylo`. In case the converse would occur (i.e. deprecated function masking a function of `adephylo`), one can refer to the 'good' version of a function by adding the prefix `adephylo::` to the function. Hence, it is possible to coerce the version of a masked function, using a kludge like:

```
> cat("\n=== Old - deprecated- version ===\n")
```

```

=== Old - deprecated- version ===
```

```
> orthogram <- ade4::orthogram
> args(orthogram)
```

```
function (x, orthobas = NULL, neig = NULL, phylog = NULL, nrepet = 999,
    posinega = 0, tol = 1e-07, na.action = c("fail", "mean"),
    cdot = 1.5, cfont.main = 1.5, lwd = 2, nclass, high.scores = 0,
    alter = c("greater", "less", "two-sided"))
NULL
```

```
> cat("\n=== New version === \n")
```

```

=== New version ===
```

```
> orthogram <- adephylo::orthogram
> args(orthogram)


function (x, tre = NULL, orthobas = NULL, prox = NULL, nrepet = 999,
    posinega = 0, tol = 1e-07, cdot = 1.5, cfont.main = 1.5,
    lwd = 2, nclass, high.scores = 0, alter = c("greater", "less",
        "two-sided"))
NULL
```

Luckily, this should not be required as long as one is not playing with loading and unloading `ade4` once `adephylo` is loaded.

2.3 Getting started

All the material of the package is summarized in a manpage accessible by typing:

```
> ?adephylo
```

The html version of this manpage may be preferred to browse easily the content of `adephylo`; this is accessible by typing:

```
> help("adephylo", package="adephylo", html=TRUE)
```

To revert help back to text mode, simply type:

```
> options(htmlhelp = FALSE)
```

2.4 Putting data into shape

While this is not the purpose of this document to go through the details of `phylo`, `phylo4` and `phylo4d` objects, we shall show briefly how these objects can be obtained.

2.4.1 Making a `phylo` object

The simplest way of turning a tree into a `phylo` object is using ape's function `read.tree`. This function reads a tree with the Newick (or 'parentetic') format, from a file (default, argument `file`) of from a character string (argument `text`).

```
> data(ungulates)
> ungulates$tre


[1] "((Antilocapra_americana,((Gorgon_taurinus,Oryx_leucoryx)W1,(Taurotragus_livingstoni,Tautragus_oryx)W2,(Gazel


> myTree <- read.tree(text=ungulates$tre)
> myTree


Phylogenetic tree with 18 tips and 13 internal nodes.

Tip labels:
        Antilocapra_americana, Gorgon_taurinus, Oryx_leucoryx, Taurotragus_livingstoni, Tautragus_oryx, Gazella_t
Node labels:
        Root, W11, W10, W1, W2, W7, ...

Rooted; no branch lengths.
```

```
> plot(myTree, main="ape's plotting of a tree")
```

It is easy to convert `ade4`'s `phylog` objects to a `phylo`, as `phylog` objects store the Newick format of the tree in the `$tre` component.

Note that `phylo` trees can also be constructed from alignements (see `read.GenBank`, `read.dna`, `dist.dna`, `nj`, `bionj`, and `mlphylo`, all in `ape`), or even simulated (for instance, see `rtree`).

Also note that, if needed, conversion can be done back and forward with `phylo4` trees:

```
> temp <- as(myTree, "phylo4")
> class(temp)
```

```
[1] "phylo4"
attr(,"package")
[1] "phylobase"
```

```
> temp <- as(temp, "phylo")
> class(temp)
```

```
[1] "phylo"
```

```
> all.equal(temp, myTree)
```

```
[1] TRUE
```

2.4.2 Making a `phylo4d` object

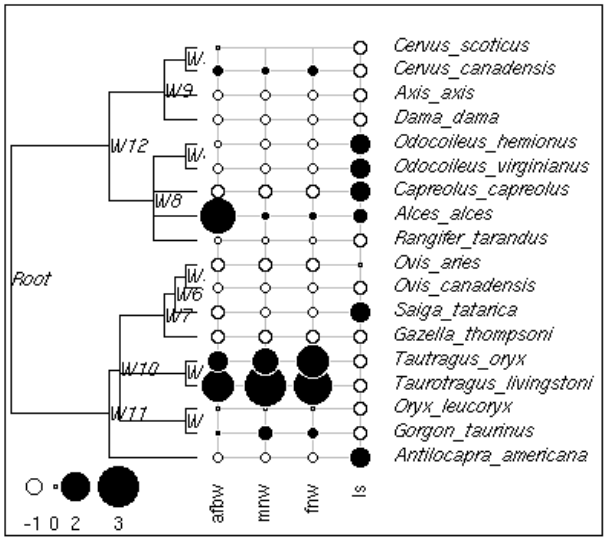
`phylo4d` objects are S4 objects, and are thus created in a particular way. These objects can be obtained in two ways, by reading a Nexus file containing tree and data information, or by 'assembling' a tree and data provided for tips, nodes, or edges.

Nexus files containing both tree and data can be read by `phylobase`'s function `readNexus` (see corresponding manpage for more information). The other way of creating a `phylo4d` object is using the constructor, also named `phylo4d`. This is a function that takes two arguments: a tree (`phylo` or `phylo4` format) and a `data.frame` containing data, for tips by default (see `?phylo4d` for more information). Here is an example:

```
> ung <- phylo4d(myTree, ungulates$tab)
> class(ung)
```

```
[1] "phylo4d"
attr(,"package")
[1] "phylobase"
```

```
> table.phylo4d(ung)
```



Data are stored inside the `@data` slot of the object. They can be accessed using the function `tdata`:

```
> x <- tdata(ung, type="tip")
> head(x)
```

```
                        afbw   mnw   fnw  ls
Antilocapra_americana    50586  3832  3908 1.8
Gorgon_taurinus         165000 18600 14500 1.0
Oryx_leucoryx            87700  6840  6490 1.0
Taurotragus_livingstoni 405000 36300 28500 1.0
Tautragus_oryx          316000 26800 24700 1.0
Gazella_thompsoni        21300  2500  2500 1.0
```

3 Exploratory data analysis

3.1 Quantifying and testing phylogenetic signal

In this document, the terms 'phylogenetic signal' and 'phylogenetic autocorrelation' are used interchangeably. They refer to the fact that values of life-history traits or ecological features are not independent in closely related taxa. Several procedures are implemented by `adephylo` to measure and test phylogenetic autocorrelation.

3.1.1 Moran's I

The function `moran.idx` computes Moran's I, the most widely-used autocorrelation measure. It can also provide additionnal information (argument `addInfo`), being the null value of I (i.e., the expected value in absence of phylogenetic autocorrelation), and the range of variation of I. It requires the degree of relatedness of tips on the phylogeny to be modelled by a matrix of phylogenetic proximities. Such a matrix can be obtained using different methods implemented by the function `proxTips`.

```
> W <- proxTips(myTree, met="Abouheif")
> moran.idx(tdata(ung, type="tip")$afbw, W)
```

```
[1] 0.1132682
```

```
> moran.idx(tdata(ung, type="tip")[,1], W, addInfo=TRUE)
```

```
[1] 0.1132682
attr(,"I0")
[1] -0.05882353
attr(,"Imin")
[1] -0.5217391
attr(,"Imax")
[1] 1.000699
```

From here, it is quite straightforward to build a non-parametric test based on Moran's I. For instance (taken from `?moran.idx`):

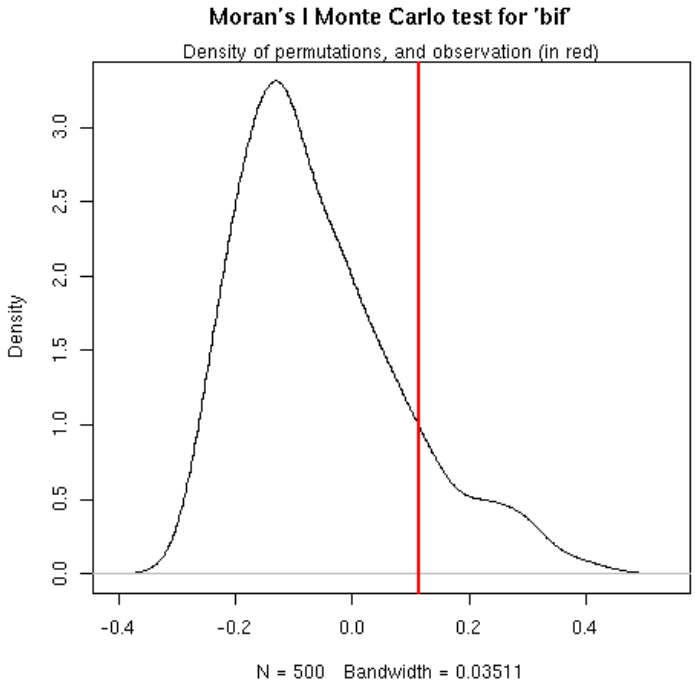
```
> afbw <- tdata(ung, type="tip")$afbw
> sim <- replicate(499, moran.idx(sample(afbw), W)) # permutations
> sim <- c(moran.idx(afbw, W), sim)
> cat("\n=== p-value (right-tail) === \n")
```

```
=== p-value (right-tail) ===
```

```
> pval <- mean(sim>=sim[1])
> pval
```

```
[1] 0.124
```

```
> plot(density(sim), main="Moran's I Monte Carlo test for 'bif'") # plot
> mtext("Density of permutations, and observation (in red)")
> abline(v=sim[1], col="red", lwd=3)
>
```



Here, `afbw` is likely not phylogenetically autocorrelated.

3.1.2 Abouheif's test

The test of Abouheif (see reference in `?abouheif.moran`) is designed to test the existence of phylogenetic signal. In fact, it has been shown that this test amounts to a Moran's I test with a particular proximity matrix (again, see references in the manpage). The implementation in `abouheif.moran` proposes different phylogenetic proximities, using by default the original one.

The function can be used on different objects; in particular, it can be used with a `phylo4d` object. In such case, all traits inside the object are tested. The returned object is a `krandtest`, a class of object defined by `ade4` to store multiple Monte Carlo tests. Here is an example using the ungulates dataset:

```
> ung.abTests <- abouheif.moran(ung)
> ung.abTests


class: krandtest
Monte-Carlo tests
Call: as.krandtest(sim = matrix(res$result, ncol = nvar, byrow = TRUE),
    obs = res$obs, alter = alter, names = test.names)

Number of tests:   4

Adjustment method for multiple comparisons:   none
Permutation number:   999
  Test       Obs  Std.Obs   Alter Pvalue
1 afbw 0.1653920 1.231621 greater  0.118
```

```
2  mnw 0.3681410 2.735761 greater  0.017
3  fnw 0.3843272 2.800180 greater  0.019
4   ls 0.3002425 1.968999 greater  0.036

other elements: adj.method call
```

```
> plot(ung.abTests)
```

In this case, it seems that all variables but `afbm` are phylogenetically structured.

Note that other proximities than those proposed in `abouheif.moran` can be used: on has just to pass the appropriate proximity matrix to the function (argument `W`). For instance, we would like to use the correlation corresponding to a Brownian motion as a measure of phylogenetic proximity.

First, we must estimate branch lengths, as our tree does not have any (ideally, we would already have a tree with meaningful branch lengths):

```
> hasEdgeLength(ung)
```

```
[1] FALSE
```

```
> myTree.withBrLe <- compute.brlen(myTree)
```

Now, we can use ape's function `vcv.phylo` to compute the matrix of phylogenetic proximities, and use this matrix in Abouheif's test:

```
> myProx <- vcv.phylo(myTree.withBrLe)
> abouheif.moran(ung, W=myProx)
```

```
class: krandtest
Monte-Carlo tests
Call: as.krandtest(sim = matrix(res$result, ncol = nvar, byrow = TRUE),
    obs = res$obs, alter = alter, names = test.names)

Number of tests:   4

Adjustment method for multiple comparisons:   none
Permutation number:   999
  Test        Obs    Std.Obs   Alter Pvalue
1 afbw 0.09113999 -0.2793974 greater  0.489
2  mnw 0.17453316  0.9330601 greater  0.160
3  fnw 0.16952884  0.8362817 greater  0.169
4   ls 0.15842344  0.5284066 greater  0.165

other elements: adj.method call
```

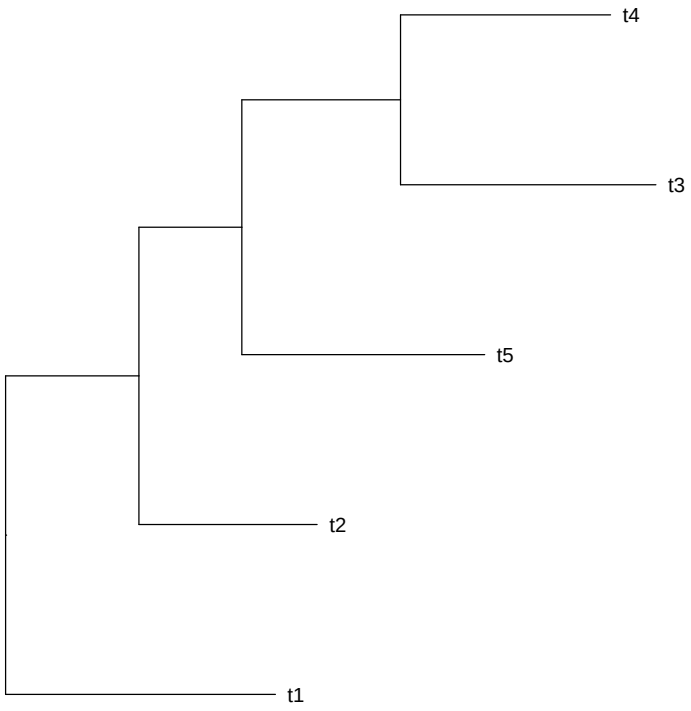
In the present case, traits no longer appear as phylogenetically autocorrelated. Several explanation can be proposed: the procedure for estimating branch length may not be appropriate in this case, or the Brownian motion may fail to describe the evolution of the traits under study for this set of taxa.

3.1.3 Phylogenetic decomposition of trait variation

The phylogenetic decomposition of the variation of a trait proposed by Ollier et al. (2005, see references in `?orthogram`) is implemented by the function `orthogram`. This function replaces the former, deprecated version from `ade4`.

The idea behind the method is to model different levels of variation on a phylogeny. Basically, these levels can be obtained from dummy vectors indicating which tip descends from a given node. A partition of tips can then be obtained for each node. This job is achieved by the function `treePart`. Here is an example using a small simulated tree:

```
> x <- as(rtree(5),"phylo4")
> plot(x,show.n=TRUE)
```

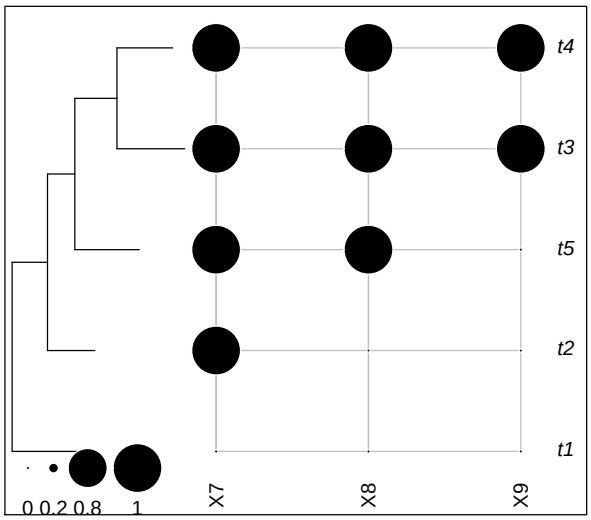



```
> x.part <- treePart(x)
> x.part
```

```
   X7 X8 X9
t1  0  0  0
t2  1  0  0
t5  1  1  0
t3  1  1  1
t4  1  1  1
```

The obtained partition can also be plotted:

```
> temp <- phylo4d(x, x.part)
> table.phylo4d(temp, cent=FALSE, scale=FALSE)
```



What we would like to do is assess where the variation of a trait is structured on the phylogeny; to do so, we could use these dummy vectors as regressors and see how variation is distributed among these vectors. However, these dummy vectors cannot be used as regressors because they are linearly dependent. The orthogram circumvents this issue by transforming and selecting dummy vectors into a new set of variables that are orthonormal. The obtained orthonormal basis can be used to decompose the variation of the trait. Even if not necessary to get an orthogram, this basis can be obtained from `treePart`:

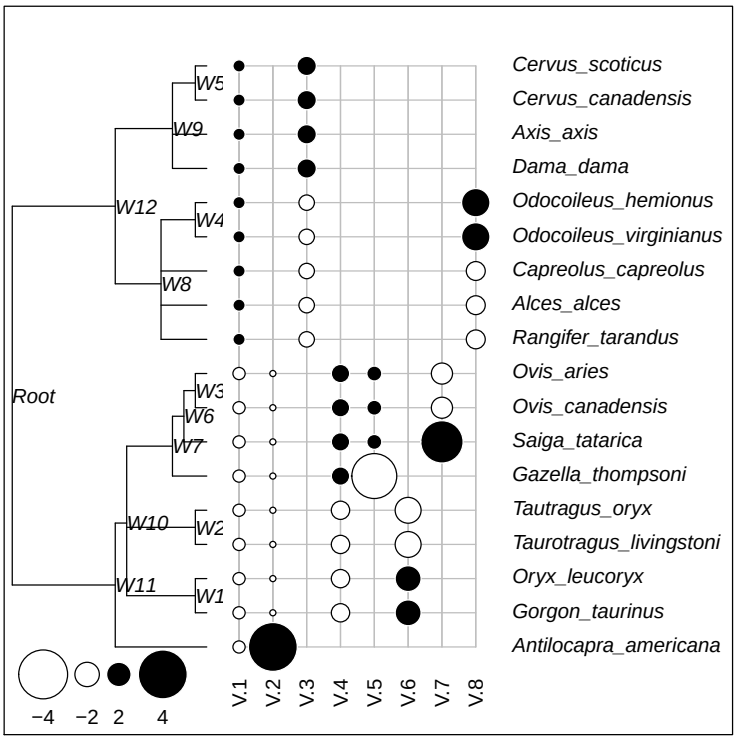
```
> args(treePart)
```

```
function (x, result = c("dummy", "orthobasis"))
NULL
```

```
> temp <- phylo4d(x, treePart(x, result="orthobasis") )
```

And here are the first 8 vectors of the orthonormal basis for the ungulate dataset:

```
> temp <- phylo4d(myTree, treePart(myTree, result="orthobasis") )
> par(mar=rep(.1,4))
> table.phylo4d(temp, repVar=1:8, ratio.tree=.3)
```



The decomposition of variance achieved by projecting a trait onto this orthonormal basis gives rise to several test statistics, that are performed by the function `orthogram`. Like the `abouheif.moran` function, `orthogram` outputs a `krandtest` object:

```
> afbw.ortgTest <- orthogram(afbw, myTree)
> afbw.ortgTest


class: krandtest
Monte-Carlo tests
Call: orthogram(x = afbw, tre = myTree)

Number of tests:   4

Adjustment method for multiple comparisons:   none
Permutation number:   999
   Test       Obs    Std.Obs   Alter Pvalue
1 R2Max 0.3298815  0.8937201 greater  0.163
2 SkR2k 8.2600870 -0.4654565 greater  0.699
3  Dmax 0.2066299  0.2828508 greater  0.333
4   SCE 0.1797097 -0.5641744 greater  0.653

other elements: adj.method call
```

Here again, `afbw` does not seem to be phylogenetically structured.

3.2 Modelling phylogenetic signal

3.2.1 Using orthonormal bases

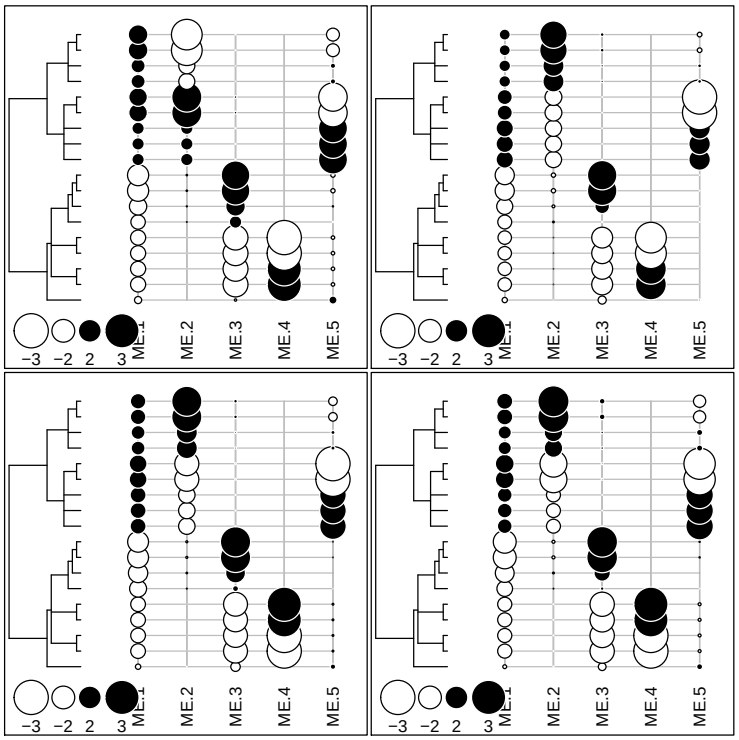
The previous section describing the orthogram has shown that testing phylogenetic signal underlies a model of phylogenetic structure. In the case of the orthogram, several tests are based on the decomposition of the variance of a trait onto an orthonormal basis describing tree topology. In fact, it is possible to extend this principle to any orthonormal basis modelling phylogenetic topology. Another example of such bases is offered by Moran's eigenvectors, which can be used to model different observable phylogenetic structures (see references in `me.phylo`).

Moran's phylogenetic eigenvectors are implemented by the function `me.phylo` (also nicknamed `orthobasis.phylo`). The returned object is a data.frame with the class `orthobasis` defined in `ade4`; columns of this object are Moran's eigenvectors. An `orthobasis` can be coerced to a regular `data.frame` or to a matrix using `as.data.frame` and `as.matrix`.

```
> me.phylo(myTree.withBrLe)


Orthonormal basis: data.frame with 18 rows and 17 columns
--------------------------------------
Columns are an orthonormal basis of 1n-orthogonal for
the inner product defined by the weights attribute
--------------------------------------
names = ME 1 ... ME 17
row.names = Antilocapra_americana ... Cervus_scoticus
weights = 0.05555556 ... 0.05555556
values = 0.5804085 ... -0.3981598
class = orthobasis data.frame
call =me.phylo(x = myTree.withBrLe)
```

Moran's eigenvectors are constructed from a matrix of phylogenetic proximities between tips. Any proximity can be used (argument `prox`); the 5 proximities implemented by the `proxTips` function are available by default, giving rise to different orthobases:

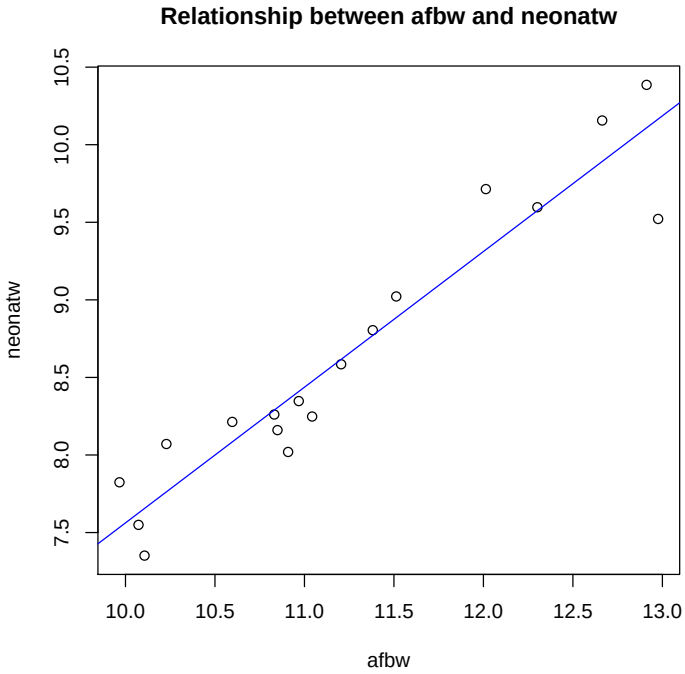
```
> ung.listBas <- list()
> ung.listBas[[1]] <- phylo4d(myTree, as.data.frame(me.phylo(myTree.withBrLe, method="patristic")))
> ung.listBas[[2]] <- phylo4d(myTree, as.data.frame(me.phylo(myTree, method="nNodes")))
> ung.listBas[[3]]<- phylo4d(myTree, as.data.frame(me.phylo(myTree, method="Abouheif")))
> ung.listBas[[4]] <- phylo4d(myTree, as.data.frame(me.phylo(myTree, method="sumDD")))
> par(mar=rep(.1,4), mfrow=c(2,2))
> invisible(lapply(ung.listBas, table.phylo4d, repVar=1:5, cex.sym=.7, show.tip.label=FALSE, show.node=FALSE))
```



In this case, the first Moran's eigenvectors are essentially similar. In other cases, however, the orthobases built from different proximities can be quite different.

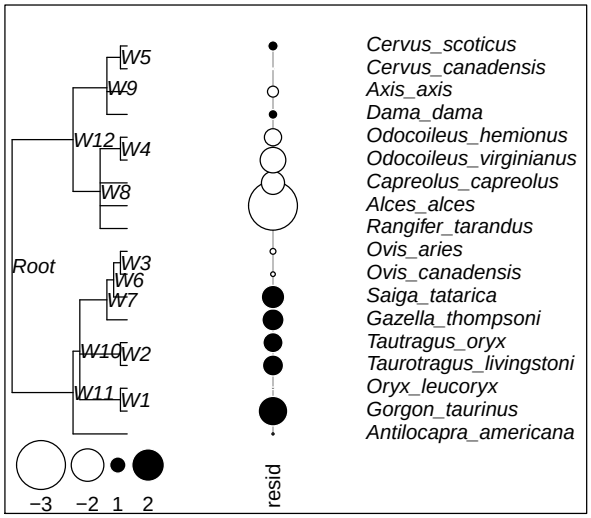
One of the interests of Moran's eigenvectors in phylogeny is to account for phylogenetic autocorrelation in a linear model. This can be achieved using the appropriate eigenvector as covariate. Here is an example when studying the link of two traits in ungulate dataset.

```
> afbw <- log(ungulates$tab[,1])
> neonatw <- log((ungulates$tab[,2]+ungulates$tab[,3])/2)
> names(afbw) <- myTree$tip.label
> names(neonatw) <- myTree$tip.label
> plot(afbw, neonatw, main="Relationship between afbw and neonatw")
> lm1 <- lm(neonatw~afbw)
> abline(lm1, col="blue")
> anova(lm1)


Analysis of Variance Table

Response: neonatw
          Df  Sum Sq Mean Sq F value    Pr(>F)
afbw       1 12.1625 12.1625  159.43 9.81e-10 ***
Residuals 16  1.2206  0.0763
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```



Are the residuals of this model independent?

```
> resid <- residuals(lm1)
> names(resid) <- myTree$tip.label
> temp <- phylo4d(myTree,data.frame(resid))
> abouheif.moran(temp)


class: krandtest
Monte-Carlo tests
Call: as.krandtest(sim = matrix(res$result, ncol = nvar, byrow = TRUE),
    obs = res$obs, alter = alter, names = test.names)

Number of tests:   1

Adjustment method for multiple comparisons:   none
Permutation number:   999
   Test       Obs Std.Obs   Alter Pvalue
1 resid 0.4566173 3.30083 greater  0.001

other elements: adj.method call


> table.phylo4d(temp)
```



No, residuals are clearly not independent, as they exhibit strong phylogenetic autocorrelation. In this case, autocorrelation can be removed by using the first Moran's eigenvector as a covariate. In general, the appropriate eigenvector(s) can be chosen by usual variable-selection approaches, like the forward selection, or using a selection based on the existence of autocorrelation in the residuals.

```
> myBasis <- me.phylo(myTree, method="Abouheif")
> lm2 <- lm(neonatw~myBasis[,1] + afbw)
> resid <- residuals(lm2)
> names(resid) <- myTree$tip.label
> temp <- phylo4d(myTree,data.frame(resid))
> abouheif.moran(temp)


class: krandtest
Monte-Carlo tests
Call: as.krandtest(sim = matrix(res$result, ncol = nvar, byrow = TRUE),
    obs = res$obs, alter = alter, names = test.names)

Number of tests:   1

Adjustment method for multiple comparisons:   none
Permutation number:   999
   Test       Obs  Std.Obs   Alter Pvalue
1 resid 0.1805854 1.291656 greater  0.102

other elements: adj.method call


> anova(lm2)
```

```
Analysis of Variance Table

Response: neonatw
             Df  Sum Sq Mean Sq  F value    Pr(>F)
myBasis[, 1]  1  0.1630  0.1630   3.1444   0.09649 .
afbw          1 12.4427 12.4427 240.0840 1.227e-10 ***
Residuals    15  0.7774  0.0518
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The link between the two variables is still very statistically significant, but this time the model is not invalid because of non-independence of residuals.

3.2.2 Autoregressive models

Autoregressive models can also be used to remove phylogenetic autocorrelation from residuals. This approach implies the use of a phylogenetically lagged vector, for some or all of the variates of a model (see references in `?proxTips`). The lag vector of a trait x, denoted $\tilde{x}$, is computed as:

$$\tilde{x} = Wx$$

where W is a matrix of phylogenetic proximities, as returned by `proxTips`. Hence, one can use an autoregressive approach to remove phylogenetic autocorrelation quite simply. We here re-use the example from the previous section:

```
> W <- proxTips(myTree, method="Abouheif", sym=FALSE)
> lagNeonatw <- W %*% neonatw
> lm3 <- lm(neonatw ~ lagNeonatw + afbw)
> resid <- residuals(lm3)
> abouheif.moran(resid,W)


class: krandtest
Monte-Carlo tests
Call: as.krandtest(sim = matrix(res$result, ncol = nvar, byrow = TRUE),
    obs = res$obs, alter = alter, names = test.names)

Number of tests:   1

Adjustment method for multiple comparisons:   none
Permutation number:   999
  Test       Obs  Std.Obs   Alter Pvalue
1    x 0.1658522 1.530926 greater  0.069

other elements: adj.method call
```

Here, this most simple autoregressive model may not be sufficient to account for all phylogenetic signal; yet, phylogenetic autocorrelation is no longer detected at the usual threshold $\alpha = 0.05$.

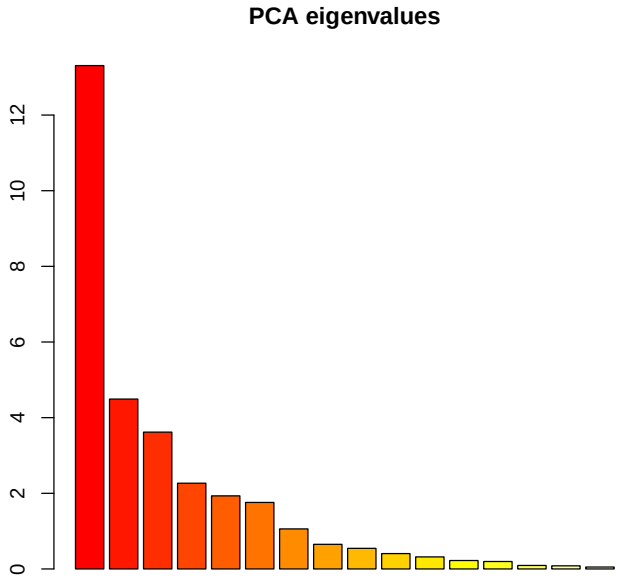
3.3 Using multivariate analyses

Multivariate analyses can be used to identify the main biodemographic strategies in a large set of traits. This could be the topic of an entire book. Such application is not particular to `adephylo`, but some practices are made easier by the package, used together with `ade4`. We here provide a simple example, using

the `maples` dataset. This dataset contains a tree and a set of 31 quantitative traits (see `?maples`).

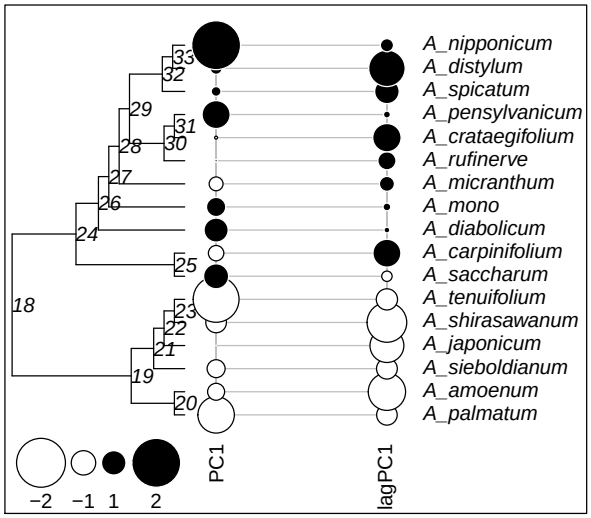
First of all, we seek a summary of the variability in traits using a principal component analysis. Missing data are replaced by mean values, so they are placed at the origin of the axes (the 'non-informative' point).

```
> f1 <- function(x){
+     m <- mean(x,na.rm=TRUE)
+     x[is.na(x)] <- m
+     return(x)
+ }
> data(maples)
> traits <- apply(maples$tab, 2, f1)
> pca1 <- dudi.pca(traits, scannf=FALSE, nf=1)
> barplot(pca1$eig, main="PCA eigenvalues", col=heat.colors(16))
```



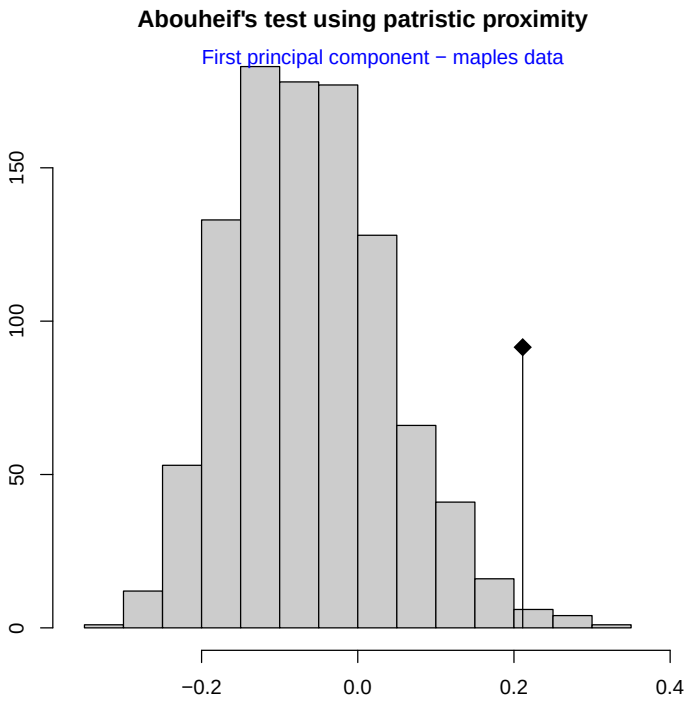
One axis shall be retained. Does this axis reflect a phylogenetic structure? We can represent this principal component onto the phylogeny. In some cases, positive autocorrelation can be better perceived by examining the lag vector (see previous section on autoregressive models) instead of the original vector. Here, we shall plot both the retained principal component, and its lag vector:

```
> tre <- read.tree(text=maples$tre)
> W <- proxTips(tre)
> myComp <- data.frame(PC1=pca1$li[,1], lagPC1=W %*% pca1$li[,1])
> myComp.4d <- phylo4d(tre, myComp)
> nodeLabels(myComp.4d) <- names(nodeLabels(myComp.4d))
> table.phylo4d(myComp.4d)
```



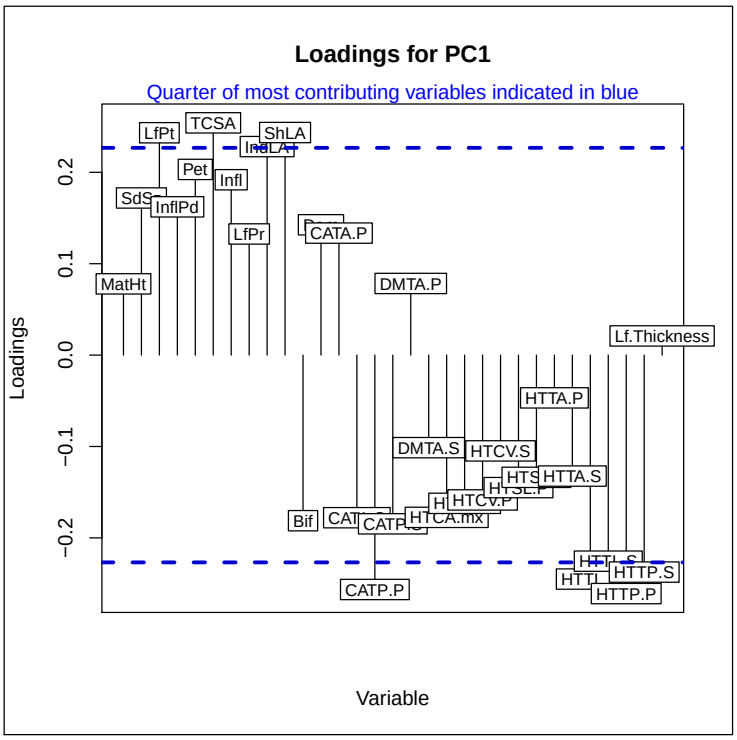
It is quite clear that the main component of diversity among taxa separates descendants from node 19 from descendants of node 24. Phylogenetic autocorrelation can be checked in 'PC1' (note that testing it in the lag vector would be circulary, as the lag vector already otimizes positive autocorrelation), for instance using Abouheif's test:

```
> myTest <- abouheif.moran(myComp[,1], W=W)
> plot(myTest, main="Abouheif's test using patristic proximity")
> mtext("First principal component - maples data", col="blue", line=1)
```



To dig further into the interpretation of this structure, one can have a look at the loadings of the traits, to see to which biological traits these opposed life histories correspond:

```
> ldgs <- pca1$c1[,1]
> plot(ldgs, type="h", xlab="Variable", xaxt="n", ylab="Loadings")
> s.label(cbind(1:31, ldgs), lab=colnames(traits), add.p=TRUE, clab=.8)
> temp <- abs(ldgs)
> thres <- quantile(temp, .75)
> abline(h=thres * c(-1,1), lty=2, col="blue3", lwd=3)
> title("Loadings for PC1")
> mtext("Quarter of most contributing variables indicated in blue", col="blue")
```



As a reminder, species with a large black symbol would be on the top of this graph, while species with a large white symbol would lie on the bottom.