

Simulate phylogenetic trees and calculate phylometrics.

Supplementary material in: Species and functional  
diversity accumulate differently in terrestrial mammals

*Brunno F. Oliveira<sup>1,2</sup>; Antonin Machac<sup>3,4,5</sup>; Gabriel C. Costa<sup>1</sup>; Thomas M. Brooks<sup>6</sup>; Ana D. Davidson<sup>7</sup>; Carlo Rondinini<sup>8</sup>; Catherine H. Graham<sup>2</sup>;*

*Global Ecology and Biogeography, 2016* / **1** Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal, Brasil. / **2** Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794, USA. / **3** Center for Macroecology, Evolution, and Climate, Natural History Museum of Denmark, Universitetsparken 15, DK 2100 Copenhagen, Denmark. / **4** Department of Ecology, Charles University, Vinicna 7, 128 44 Praha 2, Czech Republic. / **5** Center for Theoretical Study, Charles University and Academy of Sciences of the Czech Republic, Jilska 1, 110 00 Praha 1, Czech Republic. / **6** International Union for Conservation of Nature, 1196 Gland, Switzerland 7NatureServe, Arlington, VA, USA. / **8** Global Mammal Assessment Program, Department of Biology and Biotechnologies, Sapienza Universita di Roma, Rome, Italy. / **Corresponding author:** ([brunno.oliveira@me.com](mailto:brunno.oliveira@me.com))

## Contents

<b>1. Packages versions:</b>	<b>2</b>
<b>2. Load functions:</b>	<b>2</b>
<b>2. List of phylometrics:</b>	<b>2</b>
<b>4. Simulate phylogenetic trees:</b>	<b>3</b>
4.1. Manipulate richness: . . . . .	3
4.2. Manipulate time (age): . . . . .	4
4.3. Manipulate diversification rate: . . . . .	5
<b>5. Table of results:</b>	<b>6</b>

---

This document follows the principles of reproducible research (Peng, 2011). This document was generated in **R studio** with **kintr** package.

---

## 1. Packages versions:

We used R version 3.2.2 (2015-08-14) and the following packages:

```
library(TreeSim)
library(picante)
library(apTreeshape)
library(laser)
library(knitr)
```

## 2. Load functions:

```
### Function range data between zero and one
range01 <- function(x){(x-min(x))/(max(x)-min(x))}

### phylometrics function
source("phylometrics_function.R")
```

## 2. List of phylometrics:

Phylogenetic diversity (PD) Sum of all branch lengths in a tree. Faith (1992)

Mean phylogenetic distance (MPD) Mean of phylogenetic distances between all species pairs within a tree. Webb (2000)

Mean nearest taxon distance (MNTD) Mean phylogenetic distance between each species and its nearest neighbor in a tree. Webb (2000)

Phylogenetic species variability (PSV) Expected variance among species in a neutrally evolving trait. Helmus et al. (2007)

Mean root distance (MRD) The mean number of nodes that separating species from the root of their tree. Alternatively, when time trees are available MRD can be calculated as the mean distance between each species nodes and the root of its phylogenetic tree (MRD.time). Kerr and Currie (1999); Hawkins et al. (2012)

Species evolutionary distinctiveness (ED) After dividing each branch length by the number of species subtending that branch, the obtained values are summed across all branches from which a species descends. Redding and Mooers (2006); Isaac et al. (2007)

Species-level diversification rate (DivRate) Inverse of the evolutionary distinctness (ED). Species rapidly diversifying will have short edge lengths shared among many species. Jetz et al. (2012), Kembel et al. (2010)

Relative branch length (RBL) Approximates time between speciation events across a tree by calculating the mean ((stem age – crown age) / crown age). Marin and Blair, In prep.

Diversification rates (DR) Species accumulation through time (log(clade diversity)/clade age). Magallon and Sanderson (2001); Nee (2001); Phillimore et al. (2006)

Gamma statistics (GAM) The distribution of branching events throughout the tree. This phylometric distinguished between trees with relatively long inter-nodal distances towards the tips (tippy trees) and trees with relatively longer inter-nodal distances towards the root of the phylogeny (stemmy trees). Pybus and Harvey (2000)

Colless Measures the degree of tree imbalance. Here, we calculated standardized values of Colless according to the Yule and PDA models. Colless (1982)

## 4. Simulate phylogenetic trees:

### 4.1. Manipulate richness:

```
### set parameters
# number of simulations
N <- 5000
# number of taxa (tips) * Taken from the observed distribution of species
# richness values considering 1x1 degree resolution worldwide gridded data
# for mammals.
n <- rnorm(10000,mean=100,sd=38)
n = n+(-1*min(n))
# We consider that it is better to have at least 10 species to calculate phymetrics.
n = n +10

# Simulate a tree in parameters to the actual mammalian tree (Hedges et al., 2015)
trxs <- sim.bd.taxa.age(n=5000, numbsim=1, lambda=0.2, mu=0.14, age=180, mrca=T)
trxs <- trxs[[1]]

# Species' evolutionary distinctiveness: Calculate for the complete tree then
# use it to subset the resulting vector
spp.ED <- evol.distinct(trxs, type = c("fair.proportion"),
                        scale = TRUE, use.branch.lengths = TRUE)

# Species' ages: Calculate for the complete tree then use it to subset the
# resulting vector
spp.ages.c<-data.frame(matrix(data=NA,nrow=length(trxs$tip.label),ncol=2))
colnames(spp.ages.c)<-c("Species","age")
max<-max(branching.times(trxs))
for (j in 1:length(trxs$tip.label)){
  Spp<-trxs$tip.label[j]
  spp.ages.c$age[j] <- trxs$edge.length[which.edge(trxs,Spp)]
  spp.ages.c$Species[j]<-Spp
}

# Sample the tree to obtain assemblages with different richness
for(i in 1:N) { #cat("\r",i,"of", N)

  k<- as.integer(sample(n,1,replace = F)) #n species

  trx<-drop.tip(trxs,sample(1:5000,5000-k)) # sample tips to obtain
  # assemblages of n tips

  subED<-subset(spp.ED,Species%in%trx$tip.label)
  subAge<-subset(spp.ages.c,Species%in%trx$tip.label)

  run <- i

  ### Manipulated parameter
  par <- length(trx$tip.label)

  ### Species ages (Sp.Ages2) Subset the resulting vector
```

```

Sp.Ages2 <- mean(subAge[,2])

### Species evolutionary distinctiveness (ED2): Subset the resulting vector
ED2 <- mean(subED[,2])

### Species-level diversification rate (DivRate): Subset the resulting vector
DivRate2 <- mean(1/subED[,2])

### calculate phylometrics
phylo.res <- phylometrics(trx)

### store results
if(i==1){
  model1 <- c(run=run, par=par, phylo.res, ED2=ED2, DivRate2=DivRate2,
             Sp.Ages2=Sp.Ages2)
}
else{
  model1 <- rbind(model1,
                  c(run=run, par=par, phylo.res, ED2=ED2,
                    DivRate2=DivRate2, Sp.Ages2=Sp.Ages2))
}
}

model1 <- data.frame(model1)

```

## 4.2. Manipulate time (age):

```

# set parameters
ages <- rnorm(10000, 100, 20)
N <- 5000 # number of simulation
n <- 100 # number of taxa (tips)

# sample ages
tsamp<-sample(ages,N,replace = F)

# Simulate N trees under a uniform birth-death process
for(i in 1:N){ #cat("\r",i,"of", N)
  time<-tsamp[i]
  trx <- sim.bd.taxa.age(n=100, numbsim=1, lambda=log(n/2)/time, mu=0, age=time, mrca=TRUE)
  trx <- trx[[1]]

  subED<-subset(spp.ED,Species%in%trx$tip.label)
  subAge<-subset(spp.ages.c,Species%in%trx$tip.label)

  run <- i

  ### Manipulated parameter
  par <- time

  ### Species ages (Sp.Ages2) Subset the resulting vector
  Sp.Ages2 <- NA
}

```

```

### Species evolutionary distinctiveness (ED2): Subset the resulting vector
ED2 <- NA

### Species-level diversification rate (DivRate): Subset the resulting vector
DivRate2 <- NA

### calculate phylometrics
phylo.res <- phylometrics(trx)

### store results
if(i==1){
  model2 <- c(run=run, par=par, phylo.res, ED2=as.numeric(phylo.res["ED"]),
              DivRate2=as.numeric(phylo.res["DivRate"]),
              Sp.Ages2=as.numeric(phylo.res["Sp.Ages"]))
}
else{
  model2 <- rbind(model2,
                  c(run=run, par=par, phylo.res, ED2=as.numeric(phylo.res["ED"]),
                    DivRate2=as.numeric(phylo.res["DivRate"]),
                    Sp.Ages2=as.numeric(phylo.res["Sp.Ages"])))
}
}

model2 <- data.frame(model2)

```

### 4.3. Manipulate diversification rate:

```

# set parameters
lamb = rexp(10000, rate=2)
lamb = range01(lamb)
lamb = lamb+0.0001
N <- 5000 # number of simulation
n <- 100 # number of taxa (tips)

#simulate div values
tsamp=sample(lamb,N,replace = F)

# Simulate N trees under a uniform birth-death process

for(i in 1:N){ #cat("\r",i,"of", N)

  ts<-tsamp[i]
  trx <- sim.bd.taxa.age(n=100, numbsim=1, lambda=ts, mu=0.14, age=40, mrca=TRUE)
  trx <- trx[[1]]

  subED<-subset(spp.ED,Species%in%trx$tip.label)
  subAge<-subset(spp.ages.c,Species%in%trx$tip.label)

  run <- i

  ### Manipulated parameter
  par <- ts

```

```

### Species ages (Sp.Ages2): Subset the resulting vector
Sp.Ages2 <- NA

### Species evolutionary distinctiveness (ED2): Subset the resulting vector
ED2 <- NA

### Species-level diversification rate (DivRate): Subset the resulting vector
DivRate2 <- NA

### calculate phylometrics
phylo.res <- phylometrics(trx)

### store results
if(i==1){
  model3 <- c(run=run, par=par, phylo.res, ED2=as.numeric(phylo.res["ED"]),
              DivRate2=as.numeric(phylo.res["DivRate"]),
              Sp.Ages2=as.numeric(phylo.res["Sp.Ages"]))
}
else{
  model3 <- rbind(model3,
                  c(run=run, par=par, phylo.res, ED2=as.numeric(phylo.res["ED"]),
                    DivRate2=as.numeric(phylo.res["DivRate"]),
                    Sp.Ages2=as.numeric(phylo.res["Sp.Ages"])))
}
}

model3 <- data.frame(model3)

```

## 5. Table of results:

Statistical relationship between examined phylometrics, species richness, time and diversification rate, used in the simulations. Relationships are reported as Spearman's rank correlations.

```
kable(results, format = "markdown")
```

	Richness - rho		Richness - P-value		Time - rho		Time - P-value		Div. Rate - rho		Div. Rate - P-value	
PD	0.99	<0.001			0.91	<0.001			-0.97	<0.001		
MRD	0.93	<0.001			-0.01	0.719			-0.02	0.254		
MRD.time	0.91	<0.001			1.00	<0.001			0.97	<0.001		
ED	-0.96	<0.001			-0.01	0.324			-0.97	<0.001		
RBL	-0.96	<0.001			-0.01	0.324			-0.97	<0.001		
DR	1.00	<0.001			-1.00	<0.001			NA	NA		
MPD	-0.04	0.003			0.95	<0.001			-0.89	<0.001		
MNTD	-0.91	<0.001			0.83	<0.001			-0.97	<0.001		
PSV	-0.04	0.003			0.95	<0.001			-0.89	<0.001		
GAM	-0.56	<0.001			0.00	0.749			0.95	<0.001		
IMY	-0.15	<0.001			0.00	0.806			-0.02	0.198		
IMP	-0.71	<0.001			0.00	0.806			-0.02	0.198		
DivRate	0.96	<0.001			0.01	0.324			0.97	<0.001		
maxage	NA	NA			1.00	<0.001			NA	NA		

	Richness - rho	Richness - P-value	Time - rho	Time - P-value	Div. Rate - rho	Div. Rate - P-value
ED2	-0.01	0.549	-0.01	0.324	-0.97	<0.001
DivRate2	0.01	0.379	0.01	0.324	0.97	<0.001
Sp.Ages2	-0.02	0.121	0.83	<0.001	-0.97	<0.001

Please cite the original publication.

Contact [brunno.oliveira@me.com](mailto:brunno.oliveira@me.com) for any further information.