

Laura Rodrigues Vieira de Alencar



Doutoranda em Ecologia pela Universidade de São Paulo Título da tese: “A diversificação de espécies e da morfologia em serpentes da família Viperidae: padrões e processos” Orientador: Marcio Martins Co-orientador: Tiago B. Quental

[exec](#)

[Laboratório de Ecologia, Evolução e Conservação de Vertebrados](#)

[Laboratório de Macroevolução e Macroecologia](#)

Proposta de trabalho final

PLANO A - Quanto mais melhor...

Estudar a evolução de aspectos ecológicos e morfológicos em um grupo de espécies aparentadas requer muitas vezes a montagem de um banco de dados. Tal banco de dados pode incluir tanto dados de morfometria (e.g. comprimento rostro-cloacal “CRC”, comprimento da cauda “CC”), quanto registros de itens de presa, tamanho de ninhada etc. É essencial que se tenha um controle do “n” amostral para cada espécie. Ou seja, é importante saber quantos indivíduos foram medidos para cada variável morfológica e/ou quantos registros de cada aspecto ecológico foram coletados para determinada espécie. Dessa forma, os esforços na coleta de dados podem ser otimizados e direcionados àquelas espécies com menos informações.

Para tal, a função que vou desenvolver irá mostrar quais as espécies do grupo de estudo possuem um n abaixo de um limiar previamente estabelecido. Assim, irei saber quais espécies precisam da compilação de mais dados. Além disso, a função fornecerá a representatividade (porcentagem de espécies que possuem amostras acima do limiar estabelecido) de cada nó da filogenia do grupo (veja como exemplo a figura abaixo).

Vale ressaltar que quando trabalhamos com um grupo de organismos composto por, por exemplo, dez espécies, talvez esta função não seja tão necessária. Entretanto, quando o grupo abrange mais de 300, esta função pode ser extremamente útil.



```
library(ape)

taxa.sampling<-function(plan,n,filo,na=0)
{
  if(class(plan)=="data.frame")
  {
    if(class(n)=="numeric")
    {
```

```
    filtro<-subset(plan,plan[,2]<n)
  }
  else
  {
    cat("\n\n\t argumento deve ser um valor numérico \n\t")
  }
}
else
{
  cat("\n\n\t objeto não pertence a classe dataframe \n\t")
}
plan<-data.frame(plan,"prop"=rep(NA,dim(plan)[1]))
plan[,"prop"]<-replace(plan[, "prop"],list=(plan[,2]>=n),1)
plan[,"prop"]<-replace(plan[, "prop"],list=(plan[,2]<n),0)
if(class(filo)== "phylo")
{
  plot(filo, show.tip.label=TRUE)
  filo$tip.label <- plan[,3]
  x<-c(rep(NA,filo$Nnode))
  for(i in (filo$Nnode[1]):1)
  {
    y <- which(filo$edge[,1]==length(filo$tip.label)+i)
    y
    if(filo$edge[y[1],2]<=length(filo$tip.label)&filo$edge[y[2],2]<=length(filo$tip.label))
    {
      x[i] <-
      (filo$tip.label[filo$edge[y[1],2]]+filo$tip.label[filo$edge[y[2],2]])/2
    }
    if(filo$edge[y[1],2]>length(filo$tip.label)&filo$edge[y[2],2]>length(filo$tip.label))
    {
      numerico1<-
      as.numeric(extract.clade(filo,filo$edge[y[1],1])$tip.label)
      soma<-sum(numerico1)
      x[i] <-
      (soma)/length((extract.clade(filo,filo$edge[y[1],1])$tip.label))
    }
    if(filo$edge[y[1],2]>length(filo$tip.label)&filo$edge[y[2],2]<=length(filo$tip.label))
    {
      numerico2<-
      as.numeric(extract.clade(filo,filo$edge[y[1],1])$tip.label)
      soma2<-sum(numerico2)
      x[i] <-
      (soma2)/length((extract.clade(filo,filo$edge[y[1],1])$tip.label))
    }
    if(filo$edge[y[2],2]>length(filo$tip.label)&filo$edge[y[1],2]<=length(filo$tip.label))
    {
```

```
    numerico3<-  
as.numeric(extract.clade(filo,filo$edge[y[1],1])$tip.label)  
    soma3<-sum(numerico3)  
    x[i] <-  
(soma3)/(length((extract.clade(filo,filo$edge[y[1],1])$tip.label)))  
  }  
}  
rx<-round(x,2)  
nodelabels(rx,frame="circle",col="black",bg="white")  
}  
  
return(list(filtro,filo))  
}
```

[taxa_sampling_code_final.r](#)

taxa.sampling

package:unknown

R Documentation

The function shows the representativity of a species sample in their phylogeny

Description:

The taxa.sampling function works with data from phylogenetic related species. It returns a data frame showing the species that don't have a sample considered adequate for a "n" previously stipulated. It also displays the phylogeny of the group with the representativeness of the species that already have an adequate sample in the node labels.

Usage:

```
taxa.sampling(plan,n,filo,na=0)
```

Arguments:

plan	A data frame containing the species and their respective sample size
n	Minimum sample size for species to be used in later analyzes
filo	A phylogeny of the species group being analyzed
na	Any NA values in the data frame will correspond to 0

Details:

The argument `n` can vary in different studies and/or for different samples of the same study (e.g. imagine a spreadsheet containing behavioral observations for a given group of species or containing the number of specimens analyzed for each species to be used in morphometric analysis).

Value:

The function returns a list in which the `comp1` is a data frame showing the species that don't have an adequate sample regarding the argument `n` and `comp2` presents information concerning the tree structure. The function also displays the tree with the representativeness of the species with an adequate sample in the node labels.

Warning:

Note that the package "Ape" must be loaded

Rows in the data frame must be in the same order of the phylogeny's tip labels. It is important to note that the first species in the Phylogeny corresponds to the first one from bottom to top.

In the argument "plan" species must be in the first column and the respective sample in the second

Author(s):

Laura R. V. Alencar

laura.alencar@globo.com

References:

Paradis, E. 2012. Analysis of phylogenetics and evolution with R. Springer, New York. 386pp.

Examples:

```
library(ape)
```

```
tree<-read.tree(text="((sp1,(sp2,sp3)),(sp4,sp5));")
```

```
data<-
```

```
data.frame("especies"=c("sp1","sp2","sp3","sp4","sp5"), "dados"=c(5,7,3,8,2))  
data
```

```
taxa.sampling(data,5,tree,na=0)
```

Acknowledgments

I would like to thank Hamanda Badona Cavalheri and Daniel Caetano "Musgo" for their essential help during the development of this function

[help_taxa_sampling.r](#)

PLANO B - Quão conservativo é o dimorfismo sexual?

O dimorfismo sexual em tamanho é frequentemente observado em animais e plantas, sendo um tópico que me interessa bastante, principalmente quando analisado sob um contexto evolutivo. Assim, meu plano B seria desenvolver uma função que calculasse o grau de dimorfismo sexual em tamanho (degree of sexual size dimorphism, DSSD, ver fórmula abaixo) para o tamanho corporal das espécies pertencentes a um mesmo grupo filogenético. Além disso, a função apontaria quais das espécies analisadas possuem um DSSD muito maior do que o esperado a partir dos valores das espécies restantes. Adicionalmente, estes valores seriam plotados em uma filogenia do grupo para analisar se as espécies com os valores mais similares correspondem àsquelas mais próximas filogeneticamente. Este último passo forneceria uma idéia do quão forte é o sinal filogenético para o DSSD no grupo.



Comentários

PI

O que deixa o plano A interessante e desafiador é mostrar a representatividade ou tamanho de amostras no cladograma. Isto depende do quão rápido vc se entenderá com os pacotes do R que lidam com cladogramas. Sugiro q vc coloque um prazo para este estudo dos pacotes.

Para o segundo plano, não entendi o que seria o esperado, ou como os valores esperados são calculados, para comparar com os observados. Explique também como testar se as espécies com os valores mais similares correspondem àsquelas mais próximas filogeneticamente.

De modo geral, defina claramente o que serão as entradas e saídas de sua função, antes de começar.

From:
<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2013:alunos:trabalho_final:lauraherpeto:start

Last update: **2020/08/12 06:04**