

Padrões de diversidade filogenéticos

Juan P. Quimbayo, Thiago Mendes, Murilo S. Dias, Mariana Bender, Augusto Flores

Contents

1	Obtendo dados e criando mapas	2
1.1	Instalando pacotes	2
1.2	Dados de peixes recifais	2
1.2.1	Onde encontrar os dados	2
1.2.2	Preparando para baixar o arquivo	3
1.2.3	Distribuição das espécies peixes borboletas	3
2	Riqueza de espécies na América	6
3	Padrões geográficos de diversidade filogenética	10
3.1	Base de dados filogenéticos	10
3.2	Carregando e limpando a árvores filogenética	12
3.3	Calculando as métricas filogenéticas	13
3.4	Plotando os índices no mapa	14
	Bibliography	18

1 Obtendo dados e criando mapas

1.1 Instalando pacotes

Para realizarmos esta prática, você precisará instalar alguns pacotes. Utilize o código abaixo para instalar os pacotes *devtools*, *maptools* e *letsR*. Pode ser que demore um pouco, mas se você tiver uma conexão internet o R vai baixar e instalar todo o necessário. Copie e cole os seguintes códigos em teu script ou terminal

```
# instalando pacotes install.packages('devtools')
# install.packages('rgdal') install.packages('maptools')
# install.packages('maps') library(devtools)
# install_github('macroecology/letsR',dependencies = T)
# install.packages('letsR')#o pacote do repositório Git é o
# mais atual
```

```
# carregando pacotes usados
library(letsR)
```

```
## Loading required package: raster
```

```
## Loading required package: sp
```

```
## Loading required package: maps
```

```
library(maptools)
```

```
## Checking rgeos availability: TRUE
```

```
library(maps)
```

```
library(rgdal)
```

```
## rgdal: version: 1.4-4, (SVN revision 833)
```

```
## Geospatial Data Abstraction Library extensions to R successfully loaded
```

```
## Loaded GDAL runtime: GDAL 2.2.2, released 2017/09/15
```

```
## Path to GDAL shared files: /usr/share/gdal/2.2
```

```
## GDAL binary built with GEOS: TRUE
```

```
## Loaded PROJ.4 runtime: Rel. 4.9.2, 08 September 2015, [PJ_VERSION: 492]
```

```
## Path to PROJ.4 shared files: (autodetected)
```

```
## Linking to sp version: 1.3-1
```

1.2 Dados de peixes recifais

1.2.1 Onde encontrar os dados

Muitos são os repositórios de dados espaciais. Dados geográficos para os organismos mais comumente acessados (e.g., *mamíferos*, *anuros*, etc) estão disponíveis no seguinte link: [*http://www.iucnredlist.org/technical-documents/spatial-data*](http://www.iucnredlist.org/technical-documents/spatial-data). Verifique as possibilidades disponíveis.

Nós utilizaremos aqui os dados de *peixes borboletas* (*Chaetodontidae*) do Novo Mundo pois são poucas as espécies e os dados são leves o suficiente para uma aula prática rápida. Se quiser testar algo mais

real, tente baixar os dados de *mamíferos* (cerca de 350Mb!) em casa e utilizá-lo no lugar dos peixes Mas antes, tenha certeza que teu computador é potente o suficiente para tratar tamanho volume de dados, pois você poderá perder **algumas horas (ou até dias)** para executar os exercícios descritos abaixo.

1.2.2 Preparando para baixar o arquivo

Inicialmente, precisamos criar uma pasta no diretório de trabalho que receberá o arquivo depois de terminado o download. Para criar a pasta, fazer o download do arquivo, descomprimir e finalmente carregá-lo no R será preciso executar os seguintes códigos.

```
# criar pasta
dir.create(file.path(getwd(), "/data"))
```

1.2.3 Distribuição das espécies peixes borboletas

Para importar os dados e ver o que tem dentro deles, faça

```
# Importante o shapefile
borb <- readOGR(dsn = "data/Chaetodontidae_NewWorld.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/murilo/Dropbox/Disciplina UnB_2sem2019/6_Macroecologia e Biogeografia Marinh
## with 29 features
## It has 27 fields
## Integer64 fields read as strings: id_no
```

```
# gymno <-readOGR(dsn = 'data/GYMNOPHIONA.shp')
```

```
# visualizando o conteúdo dos dados
head(borb@data)
```

```
##      id_no          binomial presence origin seasonal compiler
## 0 165627 Prognathodes aculeatus      1      1          1      IUCN
## 1   4364      Chaetodon robustus      1      1          1      IUCN
## 2 165688 Chaetodon flavirostris      1      1          1      IUCN
## 3   4363 Prognathodes obliquus      1      1          1      IUCN
## 4 165646 Prognathodes guyanensis      1      1          1      IUCN
## 5 165621      Prognathodes aya      1      1          1      IUCN
##      yrcompiled          citation source
## 0      2010 International Union for Conservation of Nature (IUCN) <NA>
## 1      2010 International Union for Conservation of Nature (IUCN) <NA>
## 2      2010 International Union for Conservation of Nature (IUCN) <NA>
## 3      2010 International Union for Conservation of Nature (IUCN) <NA>
## 4      2010 International Union for Conservation of Nature (IUCN) <NA>
## 5      2010 International Union for Conservation of Nature (IUCN) <NA>
##      dist_comm island subspecies subpop          legend tax_comm kingdom
## 0      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 1      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 2      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
```

```

## 3      <NA>  <NA>          <NA>  <NA> Extant (resident)  <NA> ANIMALIA
## 4      <NA>  <NA>          <NA>  <NA> Extant (resident)  <NA> ANIMALIA
## 5      <NA>  <NA>          <NA>  <NA> Extant (resident)  <NA> ANIMALIA
##   phylum      class      order_      family      genus category
## 0 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes LC
## 1 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Chaetodon LC
## 2 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Chaetodon LC
## 3 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes DD
## 4 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes LC
## 5 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes LC
##   marine terrestrial freshwater SHAPE_Leng SHAPE_Area
## 0   True          False          False      555.8      221.5
## 1   True          False          False      191.2       70.2
## 2   True          False          False      676.2      327.2
## 3   True          False          False        5.6        2.5
## 4   True          False          False      623.9      213.9
## 5   True          False          False      283.8       74.7

```

Este último comando te permitirá ver tudo todos os descritores (i.e., colunas) dentro do arquivo *.shp*.

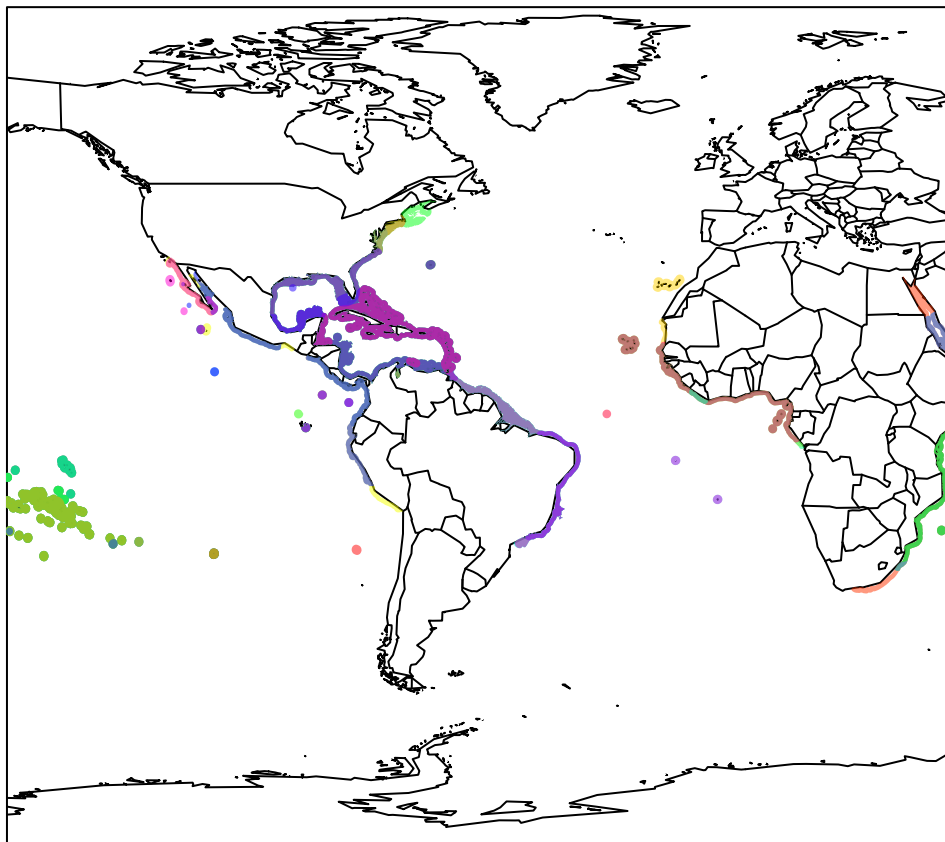
Agora precisaremos fazer um mapa para visualizar o que há dentro deste objeto *borb* que acabamos de criar. Utilizaremos algumas linhas de código a mais para fazer um mapa colorido para melhorar a compreensão dos dados. Utilizaremos a coluna referente ao nome científico de cada uma das espécies (i.e. *binomial*) para criar cores aleatórias e utilizá-las no nosso mapa. O parâmetro **alfa** dentro da função **rainbow** faz referência ao nível de transparência de cada cor; assim, poderemos ver polígonos mesmo que eles estejam cobertos por outros polígonos.

```

## usando cores aleatórias nos nossos mapas
colors <- rainbow(length(levels(borb@data$binomial)), alpha = 0.5)
position <- match(borb@data$binomial, levels(borb@data$binomial))
colors <- colors[position]

## Plot call
map("world", xlim = c(-150, 40))
box()
plot(borb, col = colors, lty = 0, add = T)

```



Cada polígono colorido se refere à área de distribuição de cada espécie.

2 Riqueza de espécies na América

A manipulação e tratamento de dados espaciais tipo *.shp* exigem um profundo conhecimento sobre os dados e a forma como eles são armazenados. Se quiserem saber mais sobre como utilizar o R para tratar este tipo de dados, recorram a livros como **Applied spatial data analysis with R** (Bivand et al. 2013) ou **An introduction to R for spatial analysis and mapping** (Brunsdon & Comber 2015). Felizmente, existe um pacote chamado **letsR** (Vilela & Villalobos 2015) que possui uma grande variedade de ferramentas úteis para macroecólogos. Nós já instalamos e carregamos este pacote no início da aula e podemos utilizar suas funções para fazer mais alguns mapas.

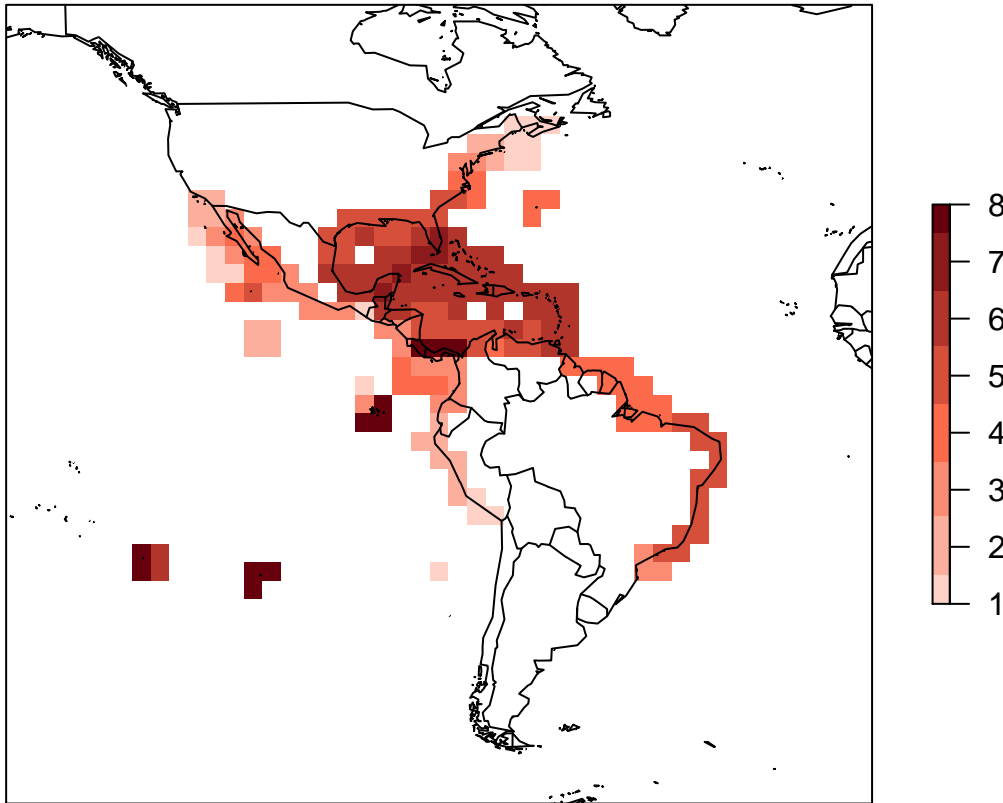
Vamos utilizar a função **lets.presab** para criar um raster (i.e., arquivo tipo imagem que armazena informações em cada um de seus pixels; digite ‘arquivo raster’ no Google para mais detalhes) resumindo a riqueza de espécie de peixes borboletas na América do Sul. Neste caso, a resolução do nosso arquivo de raster será de 3x3 graus (i.e., cada pixel ou célula tem tamanho 3 por 3 graus); mude a *resol* para valores 4 ou 5 e você verá a diferença na figura final. Já o argumento *cover* garante que somente pixels que tenham mais de 1% de sua área sobre do continente sejam mantidos (mude estes valores e veja a diferença no mapa resultante). Como nosso objetivo é um mapa da América, as coordenadas limites foram definidas para realizar o exercício somente ao longo da extensão desejada (argumentos *xmn*, *xmx*, *ymn* e *ymx* da função **lets.presab**).

```
borb_maps <- lets.presab(borb, resol = 3,
                        xmn = -130, xmx = -30, #América range
                        ymn = -65, ymx = 65, #América range
                        cover = 0.01,
                        remove.cells = F) #evita a remocao de celulas sem info
```

```
## Warning in `proj4string<-`(`*tmp*`, value = <S4 object of class structure("CRS", package =
## +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
## without reprojecting.
## For reprojecting, use function spTransform
```

```
plot(borb_maps, axes=F, main="Riqueza de Peixes Borboletas")
```

Riqueza de Peixes Borboletas



Aproveite e use o seguinte código para ver o que há no objeto criado pela função

```
# see what we have  
summary(borb_maps)
```

```
##  
## Class: PresenceAbsence  
## _ _  
## Number of species: 23  
## Number of cells: 1419  
## Cells with presence: 190  
## Cells without presence: 1229  
## Species without presence: 0  
## Species with the largest range: Chaetodon ocellatus  
## _ _  
## Grid parameters  
## Resolution: 3, 3 (x, y)  
## Extention: -130, -31, -64, 65 (xmin, xmax, ymin, ymax)  
## Coord. Ref.: +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
```

```
str(borb_maps)
```

```
## List of 3  
## $ Presence_and_Absence_Matrix: num [1:1419, 1:25] -128 -126 -122 -120 -116 ...  
## ..- attr(*, "dimnames")=List of 2  
## .. ..$ : NULL
```

```

## .. ..$ : chr [1:25] "Longitude(x)" "Latitude(y)" "Amphichaetodon melbae" "Chaetodon auri
## $ Richness_Raster :Formal class 'RasterLayer' [package "raster"] with 12 slots
## .. ..@ file :Formal class '.RasterFile' [package "raster"] with 13 slots
## .. .. . . .@ name : chr ""
## .. .. . . .@ datanotation: chr "FLT4S"
## .. .. . . .@ byteorder : chr "little"
## .. .. . . .@ nodatavalue : num -Inf
## .. .. . . .@ NChanged : logi FALSE
## .. .. . . .@ nbands : int 1
## .. .. . . .@ bandorder : chr "BIL"
## .. .. . . .@ offset : int 0
## .. .. . . .@ toptobottom : logi TRUE
## .. .. . . .@ blockrows : int 0
## .. .. . . .@ blockcols : int 0
## .. .. . . .@ driver : chr ""
## .. .. . . .@ open : logi FALSE
## .. ..@ data :Formal class '.SingleLayerData' [package "raster"] with 13 slots
## .. .. . . .@ values : num [1:1419] 0 0 0 0 0 0 0 0 0 0 ...
## .. .. . . .@ offset : num 0
## .. .. . . .@ gain : num 1
## .. .. . . .@ inmemory : logi TRUE
## .. .. . . .@ fromdisk : logi FALSE
## .. .. . . .@ isfactor : logi FALSE
## .. .. . . .@ attributes: list()
## .. .. . . .@ haveminmax: logi TRUE
## .. .. . . .@ min : num 0
## .. .. . . .@ max : num 8
## .. .. . . .@ band : int 1
## .. .. . . .@ unit : chr ""
## .. .. . . .@ names : chr ""
## .. ..@ legend :Formal class '.RasterLegend' [package "raster"] with 5 slots
## .. .. . . .@ type : chr(0)
## .. .. . . .@ values : logi(0)
## .. .. . . .@ color : logi(0)
## .. .. . . .@ names : logi(0)
## .. .. . . .@ colortable: logi(0)
## .. ..@ title : chr(0)
## .. ..@ extent :Formal class 'Extent' [package "raster"] with 4 slots
## .. .. . . .@ xmin: num -130
## .. .. . . .@ xmax: num -31
## .. .. . . .@ ymin: num -64
## .. .. . . .@ ymax: num 65
## .. ..@ rotated : logi FALSE
## .. ..@ rotation:Formal class '.Rotation' [package "raster"] with 2 slots
## .. .. . . .@ geotrans: num(0)
## .. .. . . .@ transfun:function ()
## .. ..@ ncols : int 33
## .. ..@ nrows : int 43
## .. ..@ crs :Formal class 'CRS' [package "sp"] with 1 slot

```



```
## .. ..@ projargs: chr "+proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0"
## .. ..@ history : list()
## .. ..@ z : list()
## $ Species_name : chr [1:23] "Amphichaetodon melbae" "Chaetodon auriga" "Cha
## - attr(*, "class")= chr "PresenceAbsence"
```

E se precisarmos criar somente uma matriz para saber qual espécie ocorre em cada uma das células do raster, você pode acessar a matriz direto do objeto **borb_maps** ou refazer a mesma função pedindo como resultado somente a matriz de presença/ausência (*show.matrix=TRUE* dentro da função *lets.presab*; digite *?lets.presab* no console do R e leia o help desta função e seus argumentos).

```
# Presence/absence matrix
borb_pa <- borb_maps$Presence_and_Absence_Matrix[, -c(1:2)]
```

3 Padrões geográficos de diversidade filogenética

Para entendermos melhor como podemos utilizar diferentes informações em larga escala, vamos proceder com um exercício sobre **diversidade filogenética**. Muito da macroecologia tem sido feito no sentido de incorporar informações filogenéticas às análises espaciais/geográficas. Desta forma, os padrões que analisamos em larga escala contém informações não somente da distribuição dos organismos mas também da evolução/história evolutiva das linhagens (*vide o nome desta disciplina*). Eu não entrarei em detalhes sobre a teoria por trás destes tipos de métricas, mas para maiores informações vejam papers recentes na area (e.g., (Ives & Helmus 2010; Mouquet *et al.* 2012; Miller *et al.* 2016; Tucker *et al.* 2016)).

3.1 Base de dados filogenéticos

Para iniciar, nós precisamos encontrar uma filogenia do grupo com o qual nós estamos trabalhando. Para *aves*, *mamíferos*, possivelmente *répteis* e alguns *anfíbios* este tipo de informação pode ser encontrada em algum trabalho já publicado¹. Como estamos tratando de peixes, boa parte das espécies já está descrita e a filogenia da maior parte delas também encontra-se disponível. Esta filogenia usada aqui foi encontrada a partir de uma publicação recente (ver (Leprieur *et al.* 2015)).

Primeiro tenham certeza de que os arquivos estão na pasta onde vocês estão rodando este script e que vocês conseguem acessá-los a partir do R. Depois, é só carregar a árvore filogenética e prosseguir com algumas manipulações simples para visualizá-la. No meio do código abaixo, nós precisaremos criar um outro nome para as espécies de peixes para que tanto o shape **borb_maps** quanto o objeto **filo** contenham exatamente o mesmo nome. Se houver pequenas diferenças (e.g., “*Tachyglossus aculeatus*” em vez de “*Tachyglossus_aculeatus*”) as funções vão travar e você não conseguirá computar os índices e fazer os procedimentos.

```
# Caso não tenha install.packages('ape')
# install.packages('picante')
```

```
# lendo arvore filogenetica
library(ape)
```

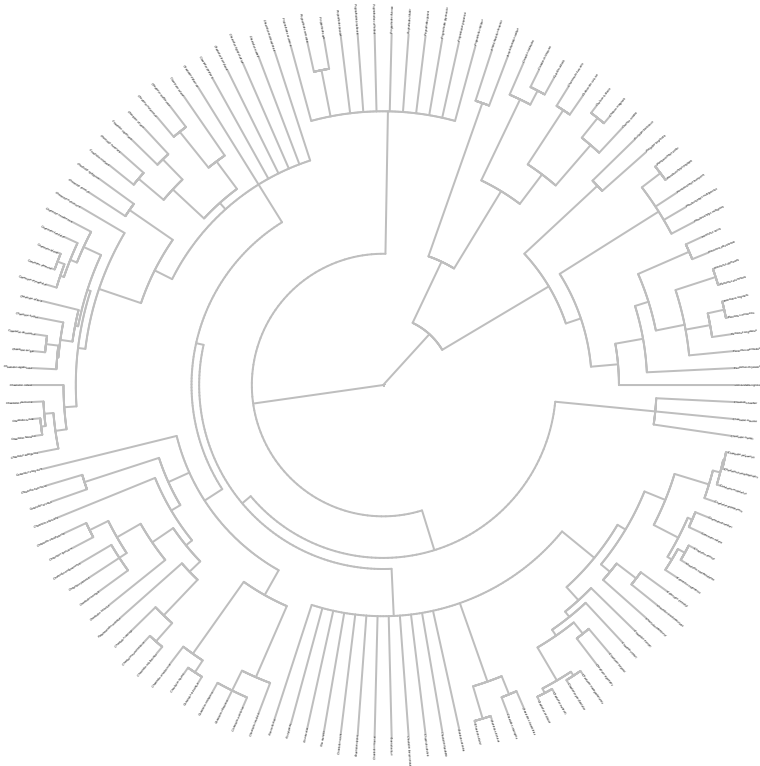
```
##
## Attaching package: 'ape'
## The following objects are masked from 'package:raster':
##
##   rotate, zoom
```

```
filo <- read.tree("newChaetNode.txt")
filo
```

```
##
## Phylogenetic tree with 130 tips and 95 internal nodes.
##
## Tip labels:
## Johnrandallia_nigrirostris, Heniochus_chrysostomus, Heniochus_intermedius, Heniochus_sing
## Node labels:
## N1, N2, N3, N4, N5, N6, ...
```

¹Busquem o termo *supertree* no portal Periodicos Capes e encontrarão papers com filogenias.

```
##
## Rooted; includes branch lengths.
plot.phylo(filo, type = "fan", edge.color = "gray", cex = 0.1)
```



```
# algumas informacoes uteis is.rooted(filo)
# is.ultrametric(filo) filo$tip.label

# mudando nome de mami de acordo com a filogenia funcao para
# mudar nome
muda_nome = function(x) {
  unlist(lapply(strsplit(x = as.character(x), " "), function(x) {
    paste(x[1], x[2], sep = "_")
  })))
}

# mudando nome da matrix local vs especie
colnames(borb_pa) = muda_nome(colnames(borb_pa))

# especies presentes tanto na filogenia quanto no arquivo
# mami separando somente as especies em ambos os objetos
sp <- intersect(filo$tip.label, colnames(borb_pa))

# excluindo as especies do mapa para reter somente as
# especies presentes na filogenia
borb <- borb[borb$binomialMod %in% sp, ]
dim(borb)
```

```
## [1] 0 27
# excluindo as especies da matrix para reter somente as
# especies presentes na filogenia
borb_pa <- borb_pa[, colnames(borb_pa) %in% sp]
dim(borb_pa)
```

```
## [1] 1419 23
rm(sp)
```

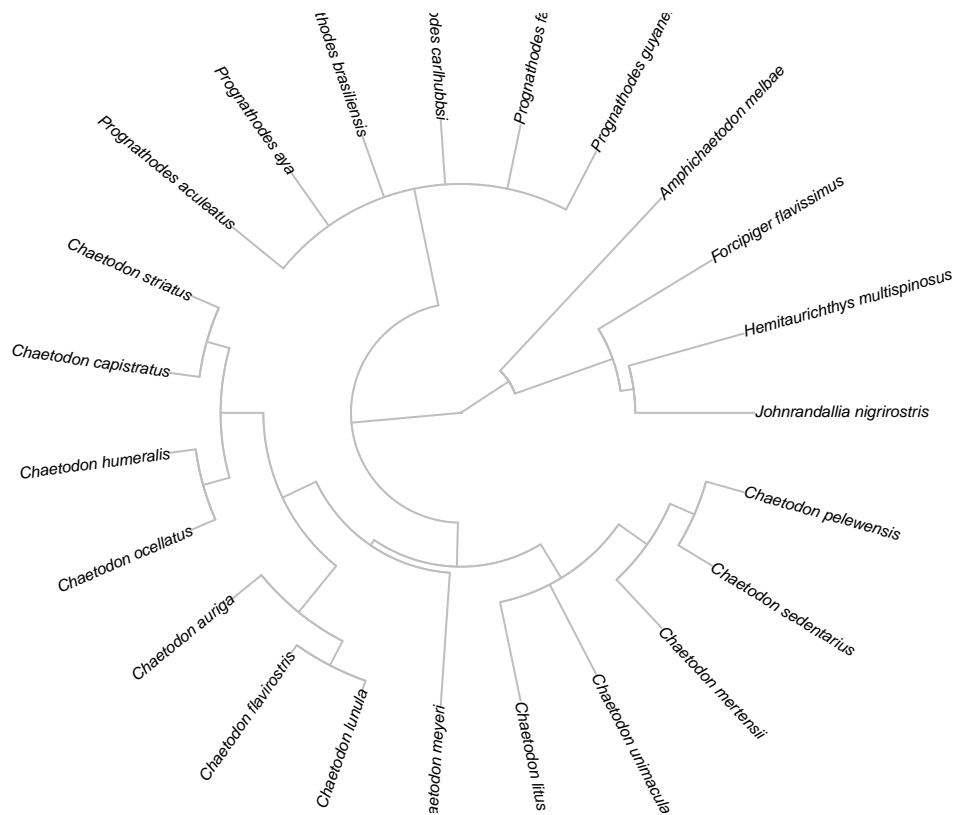
3.2 Carregando e limpando a árvores filogenética

Outro ponto importante é retirar da filogenia todas as espécies que não serão utilizadas. Isto facilitará o cálculo e manterá somente as espécies presentes em **mat**.

```
library(picante)

## Loading required package: vegan
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.5-4
## Loading required package: nlme
##
## Attaching package: 'nlme'
## The following object is masked from 'package:raster':
##
##   getData
# deixar na filo as sp da matrix
filo_gamb <- prune.sample(samp = borb_pa, phylo = filo)

plot.phylo(filo_gamb, type = "fan", edge.color = "gray", cex = 0.5)
```



3.3 Calculando as métricas filogenéticas

Existem diversas métricas filogenéticas que trazem informações distintas sobre a comunidade. A mais comum e mais antiga é **Diversidade Filogenética**, que nada mais é do que a soma dos ramos da árvore filogenética correspondentes às espécies presentes na comunidade.

```
# Faith's Phylogenetic diversity
PD <- pd(samp = borb_pa, tree = filo_gamb, include.root = T)
```

Uma outra informação complementar é fornecida pela distância média dos ramos filogenéticos correspondentes às espécies presentes em uma assembleia. Esse índice dá uma ideia de quão filogeneticamente próximas são as espécies presentes em uma comunidade. Se os valores são pequenos, mais relacionadas filogeneticamente são as espécies; se mais distantes, mais diferentes elas são em termos de história evolutiva. Vamos calcular este novo índice e juntá-lo àqueles anteriormente calculados.

```
# more indices: Mean pairwise distance separating taxa in a
# community
phydist <- cophenetic(filo_gamb) #distancia filogenetica entre especies
mpd.result <- mpd(borb_pa, phydist)
rm(phydist)
```

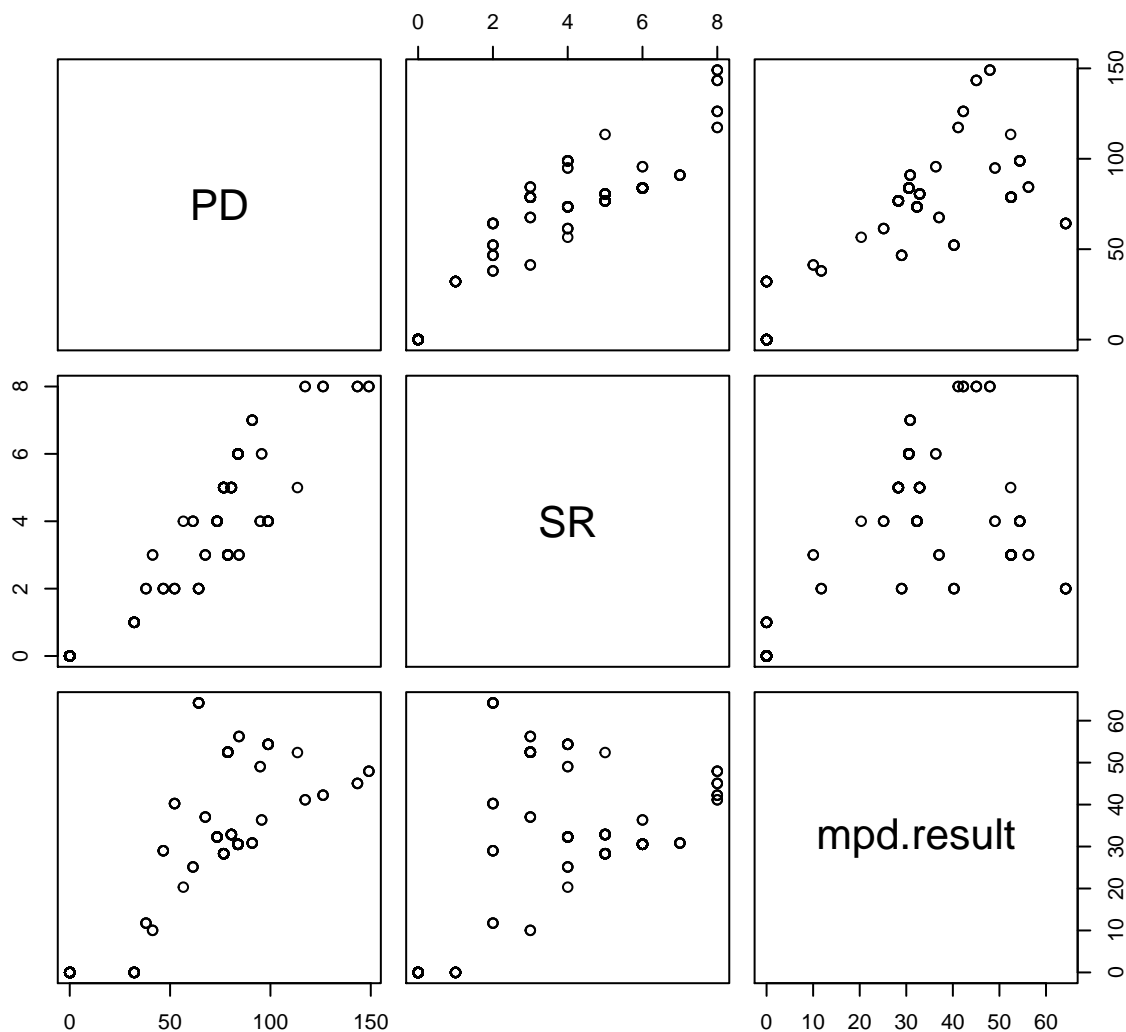
```
PD <- data.frame(PD, mpd.result)
PD[is.na(PD)] <- 0

head(PD[order(PD$SR, decreasing = T), ], 10)
```

```
##      PD SR mpd.result
```

```
## 610 126 8      42
## 611 126 8      42
## 612 126 8      42
## 707 143 8      45
## 739 143 8      45
## 740 143 8      45
## 958 117 8      41
## 991 117 8      41
## 997 149 8      48
## 998 149 8      48
```

```
# relação entre todos os índices calculados
plot(PD)
```



3.4 Plotando os índices no mapa

Agora que já temos uma planilha contendo cada um destes índices, nos podemos utilizar o mapa de riqueza para mapear os índices no espaço.

```
# criar uma copia para evitar perder info
borb2 <- borb_maps
```

```
# matriz, raster e variaveis calculadas estao na mesma ordem
```

```
all.equal(rowSums(borb_pa), values(borb2$Richness_Raster))
```

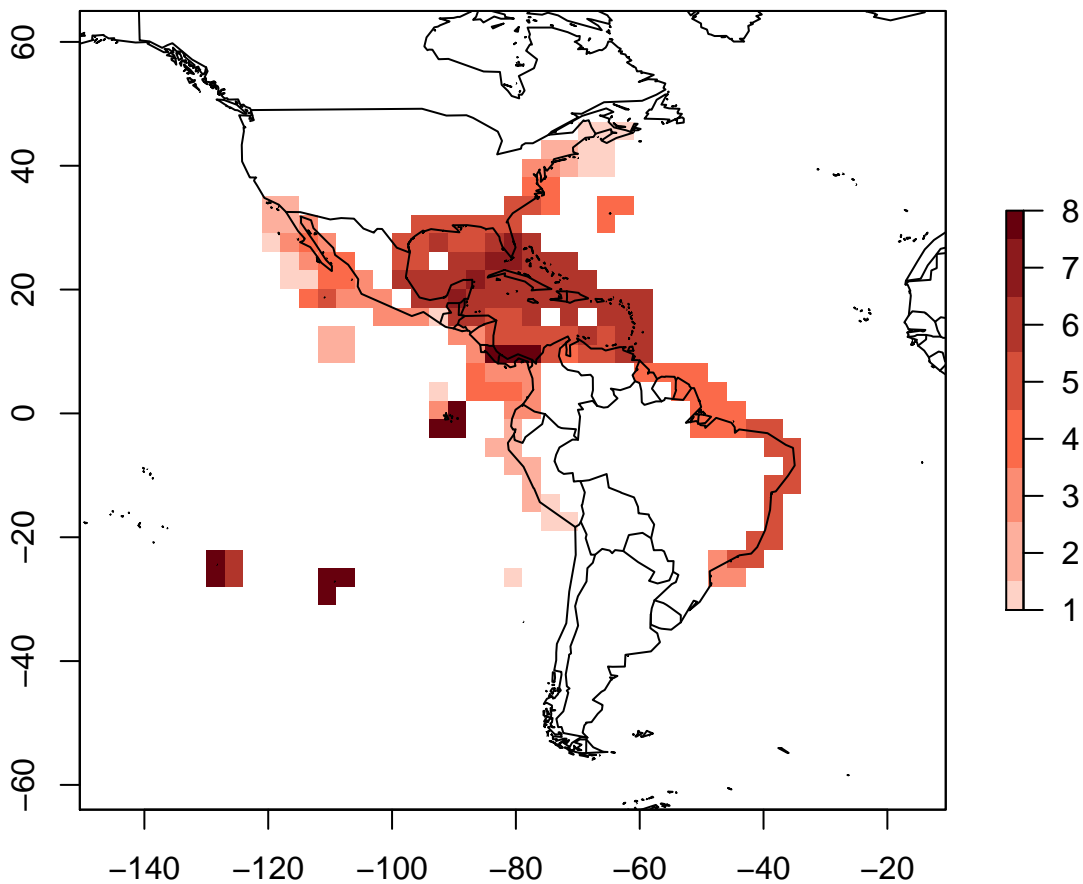
```
## [1] TRUE
```

```
all.equal(rowSums(borb_pa), PD$SR)
```

```
## [1] TRUE
```

```
plot(borb2, main = "Riqueza de especies")
```

Riqueza de especies



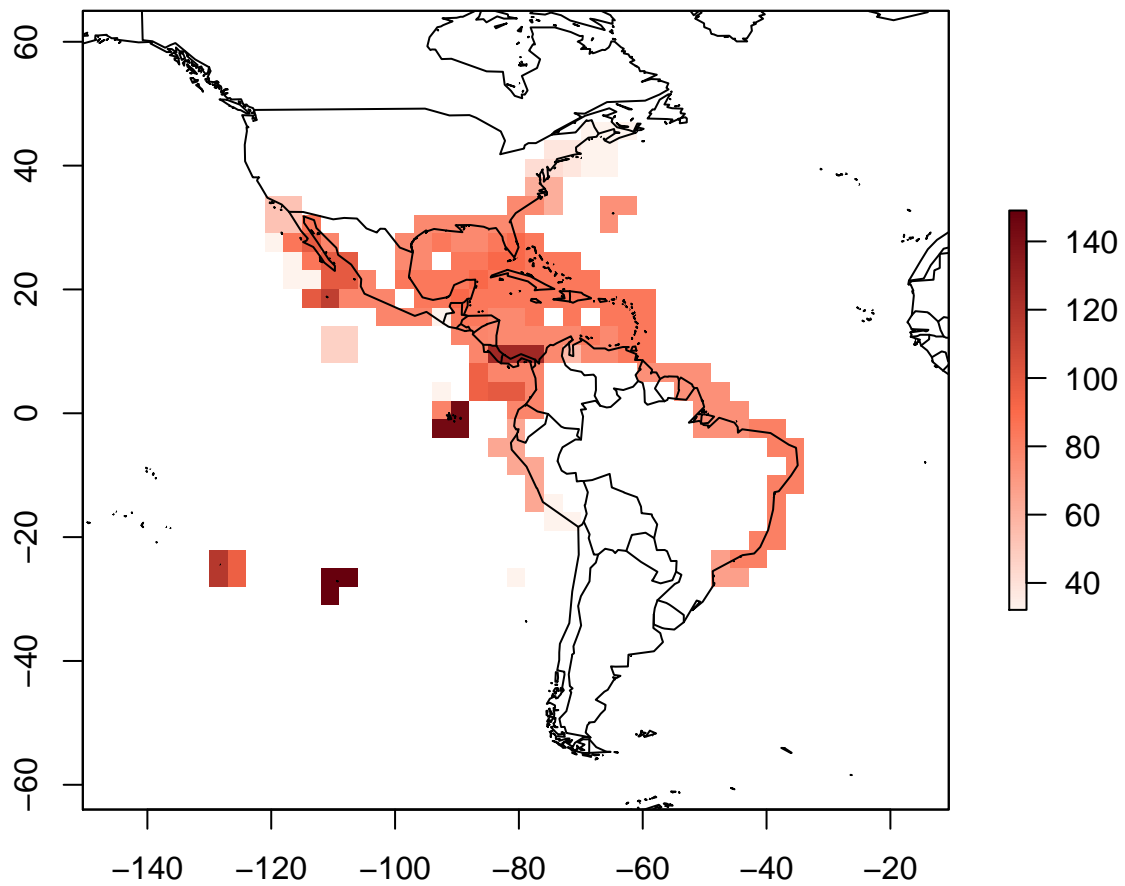
```
# basta mudar os valores do raster original pela metrica
```

```
# desejada PDiv
```

```
values(borb2$Richness_Raster) <- PD$PD
```

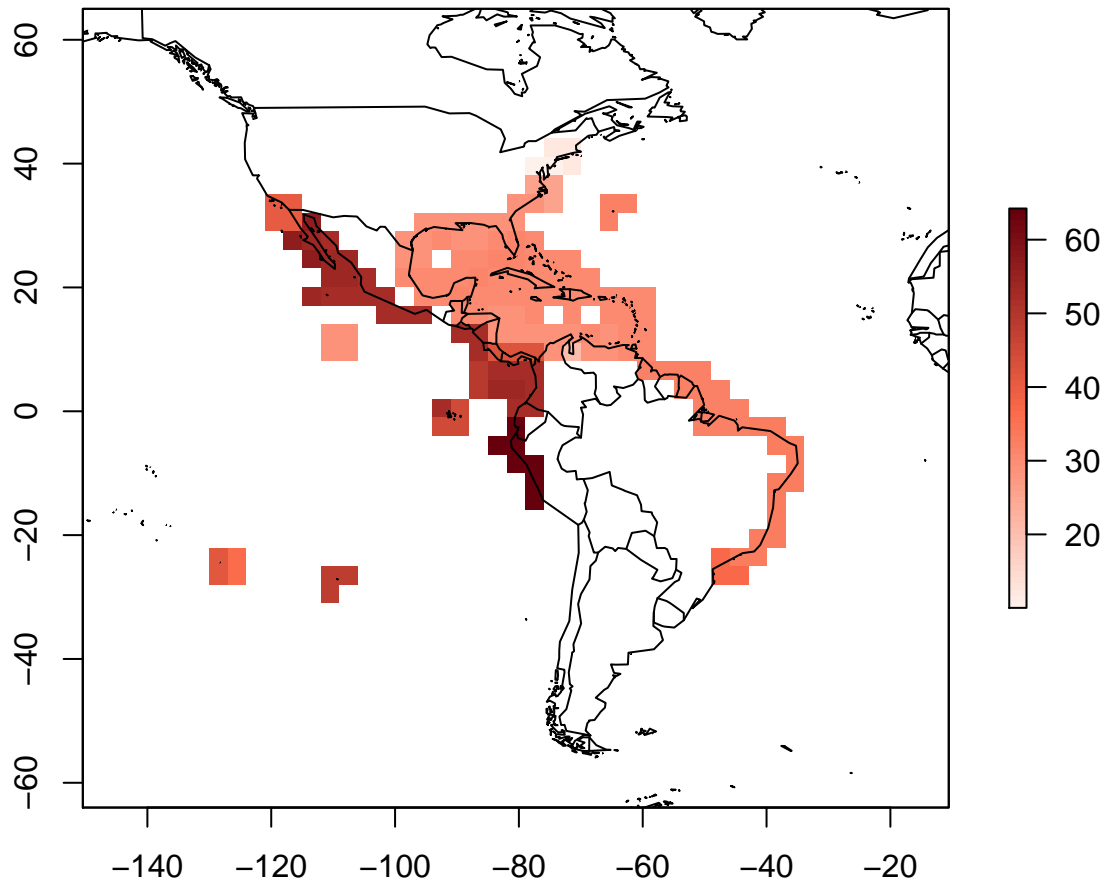
```
plot(borb2, main = "Diversidade filogenetica")
```

Diversidade filogenética



```
# MPDist  
values(borb2$Richness_Raster) <- PD$mpd.result  
plot(borb2, main = "Mean Pairwise Distance")
```


Mean Pairwise Distance



Que padrões vocês enxergam? Eles são os mesmos entre os índices? Quais são os mecanismos que explicam estes padrões?

Bibliography

1.
Ives, A.R. & Helmus, M.R. (2010). Phylogenetic metrics of community similarity. *The American Naturalist*, 176, E128–E142.
2.
Leprieur, F., Colosio, S., Descombes, P., Parravicini, V., Kulbicki, M. & Cowman, P.F. *et al.* (2015). Historical and contemporary determinants of global phylogenetic structure in tropical reef fish faunas. *Ecography*, 39, 825–835.
3.
Miller, E.T., Farine, D.R. & Trisos, C.H. (2016). Phylogenetic community structure metrics and null models: A review with new methods and software. *Ecography*, 40, 461–477.
4.
Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F. & Chave, J. *et al.* (2012). Ecophylogenetics: Advances and perspectives. *Biological Reviews*, 87, 769–785.
5.
Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S. & Fritz, S.A. *et al.* (2016). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, 92, 698–715.
6.
Vilela, B. & Villalobos, F. (2015). letsR: A new r package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*, 6, 1229–1234.