

Ejercicio Macroecología Evolutiva 2022

Axel Arango

2022-09-03

Contents

En este ejercicio se probará la regla de Rapoport en un set de más de 6000 registros para especies de la familia *Icteridae* (Aves: Passeriformes) y se compararán los procedimientos utilizando dos paqueterías: **letsR** y EcoPhylo Mapper **epm**

epm utiliza los paquetes **sf** y **terra** para análisis espaciales, lo cuál podría resultar útil cuándo **rgdal** y **raster** se hayan discontinuado.

Primero, se cargarán los paquetes a utilizar:

```
library(epm)
library(rgdal)
library(sf)
library(caper)
library(geiger)
library(letsR)
library(ggplot2)
```

Después cargaremos el set de datos

```
ict<-read.csv("ict.csv",header=T)
```

Posteriormente limpiaremos los registros, asegurándonos que no haya ningún NA ni registros repetidos:

```
ict<-na.omit(ict)
ict<-unique(ict)
names(ict)<-c("taxon","x","y")
```

Finalmente, cambiaremos los espacios por guiones bajos en los nombres de las especies. Esto con el fin de que puedan ser reconocidas al momento de utilizar alguna filogenía

```
ict$taxon<-gsub(" ","_",ict$taxon)
head(ict)
```

Una vez que hemos cargado y limpiado los registros, podemos empezar a obtener los atributos de interés para evaluar la regla de Rapoport en esta familia de aves.

Para esto necesitamos saber el área de distribución de las especies y su ubicación geográfica.

Usando **letsR**:

Primero, creamos una matriz de presencia-ausencia

```
presab<-lets.presab.points(cbind(ict$x,ict$y),ict$taxon)
```

Después, obtenemos el área de distribución de las especies

```
ranges<-lets.rangesize(presab)
head(ranges)
```

Finalmente, necesitamos la ubicación de las especies.

Afortunadamente, `letsR` posee una práctica función que permite obtener el punto medio de la ubicación geográfica de las especies

```
midp<-lets.midpoint(presab)
head(midp)
```

Ahora que tenemos los datos, podemos poner a prueba la relación

```
letsrap<-lm(log(ranges)~abs(midp$y))
summary(letsrap)
```

La relación entre el área de distribución y la latitud es significativamente positiva en *Icteridae*

ahora observemosla de manera más gráfica:

```
dfram<-data.frame(midp,ranges);names(dfram)<-c("sp","x","y","ranges")
p<-ggplot(dfram,aes(abs(y),log(ranges)))+
  geom_point(color="black")+
  xlab("Latitud absoluta")+
  ylab("Log (tamaño área)")+
  geom_smooth(method= lm , color="red", fill="#69b3a2", se=TRUE)+
  theme_classic()

plot(p)
```

Eso fue relativamente sencillo.

Ahora utilicemos el paquete `epm` para tratar de hacer lo mismo:

`epm` no puede utilizar sólo los puntos de longitud y latitud de nuestros registros, primero hay que transformar esos puntos en un “spatial feature”:

```
sp_ict<-st_as_sf(ict, coords = c("x","y"), crs= st_crs("+datum=WGS84 +proj=longlat"))
head(sp_ict)
```

Ahora que nuestros registros son un objeto espacial, ya podemos utilizar `epm` para crear un grid con nuestras especies en el espacio:

```
grix<-createEPMgrid(sp_ict,resolution = 1)
```

Con este grid, ya podemos obtener el área de distribución de las especies, donde `cell` es área de distribución

```
Cellcount<-data.frame(grix$cellCount)
names(Cellcount)<- "cell"
head(Cellcount)
```

¿Qué tanto se parece la distribución calculada con `letsR` a la calculada con `epm`?

```
par(mfrow=c(1,2))
hist(dfram$ranges, breaks=15, main = "LetsR", xlab="área de distribución")
hist(Cellcount$cell, breaks =15, main= "EPM", xlab = "área de distribución")
```

Bastante parecidos.

Para la ubicación geográfica de las especies, en el caso de `epm` no se cuenta con una función específica para obtener los puntos medios de la distribución de las especies.

Por lo tanto, tenemos que cargar una función personalizada:

```
load("st_midpoints.R")
```

Esta función utiliza la nueva paquetería **sf** para crear un polígono para cada especie con todas las coordenadas de nuestros registros, para posteriormente calcular el centroide de dicho polígono, obteniendo así los puntos medios de la ubicación geográfica de nuestras especies.

```
midst<-st_midpoints(ict)
head(midst)
```

Ahora ya tenemos los datos necesarios para evaluar la relación:

```
summary(lm(log(Cellcount$cell)~abs(midst$Y)))
```

El resultado es una relación significativamente positiva en *Icteridae* entre el área de distribución y la latitud.

Esta relación es muy parecida a la obtenida utilizando el procedimiento **letsR**. Las diferencias observadas pueden deberse a 1) el cálculo de los grids, pues **letsR** utiliza celdas y nosotros calculamos hexagonos con **epm** ó 2) el cálculo de los puntos medios, pues para **epm** dicho cálculo es más o menos artesanal.

Ahora que hemos evaluado la regla de Rapoport utilizando dos métodos y paqueterías distintas, podemos pensar en el efecto del componente evolutivo sobre dicha relación.

Para esto, primero debemos cargar una filogenia para nuestras especies. En este caso utilizaremos una filogenia para *Emberizidae* (la súper familia que contiene a *Icteridae*), usada en Arango et al., 2022 modificada y estandarizada de Barker et al., 2015.

```
tre<-read.tree("corrected_mcc_cap3.tre")
tre
```

Ahora crearemos un *data.frame* el cual contenga el área de distribución y los puntos medios para cada especie. Para motivos del ejercicio, haremos uso de los resultados obtenidos con **epm**.

¡Pero ustedes pueden hacerlo con los resultados obtenidos con **letsR** y podemos comparar las relaciones finales!

```
rapp<-data.frame(midst,Cellcount$cell,row.names = NULL);names(rapp)<-c("sp","x","y","area")
head(rapp)
```

Una vez que tenemos nuestra tabla de datos, creamos un objeto *comparative.data*, el cual juntará nuestros datos geográficos con nuestros datos filogenéticos

```
compict<-comparative.data(tre,data =rapp,names.col = "sp",vcv = T)
```

Ahora ya podemos ajustar un modelo de regresión considerando las relaciones filogenéticas de nuestras especies utilizando un PGLS (*Phylogenetic Generalized Least Squares*):

```
raprule<-pgls(log(area)~abs(y),data=compict,lambda="ML")
summary(raprul)
```

¡La relación se mantiene!

Ahora, si bien es mucho más fácil obtener estos resultados utilizando **letsR**, **epm** hace uso de las nuevas paqueterías espaciales, además de servir para mapear otros componentes de diversidad y macroevolución.

Por ejemplo, podemos obtener la diversidad filogenética (PD), las tasas de diversificación (DR) y los endemismos filogenéticos de cada comunidad (o celda) y responder preguntas que involucren métricas de este tipo y la geografía.

Para esto, primero debemos agregarle una filogenia a nuestro grid creado con ‘epm’

```
grix<-epm::addPhylo(grix,tre)
```

Ya con la filogenía incorporada a nuestro grid de `epm`, calculamos para cada celda PD:

```
pd<-gridMetrics(grix,"pd")
```

DR:

```
dr<-gridMetrics(grix,"DR")
```

ó Filoendemismos:

```
phyendm<-gridMetrics(grix,"phyloWeightedEndemism")
```

Veamos estas métricas mapeadas en el espacio:

```
par(mfrow=c(1,3))
plot(pd,use_tmap=F, legend=F)
addLegend(pd,location="left",label = "PD")
plot(dr,use_tmap=F, legend=F)
addLegend(dr,location="left",label = "DR")
plot(phyendm,use_tmap=F, legend=F)
addLegend(phyendm,location="left",label = "Phyloendemisms")
```

Con estos datos, podemos hacernos preguntas cómo:

¿Las comunidades que poseen mayores tasas de diversificación tienen una mayor diversidad filogenética?

```
summary(lm(pd$grid$pd~dr$grid$DR))
```

En el caso de *Icteridae*, aparentemente no.

```
plot(dr$grid$DR,pd$grid$pd,xlab="Tasa de diversificación",ylab="Diversidad filogenética", pch=16)
```

... Al menos no lineal

¿Los endemismos filogenéticos siguen un patrón de gradiente latitudinal?

Obtengamos las coordenadas de cada celda y su valor de filoendemismos:

```
grdtemp<-grix$grid$gridTemplate
stcen<-st_centroid(grdtemp)
stcoor<-st_coordinates(stcen)
coordsx<-as.data.frame(stcoor)
phylem<-as.data.frame(phyendm$grid$phyloWeightedEndemism)
endmcoords<-cbind(coordsx,phylem); names(endmcoords)<-c("x","y","phyloendemism")
head(endmcoords)
```

Ahora, pongamos a prueba la pregunta:

```
latendm<-lm(log(phyloendemism)~abs(y),data= endmcoords)
summary(latendm)
```

Parece que los endemismos disminuyen al alejarse del ecuador, siguiendo un gradiente latitudinal.

Observemoslo:

```
p2<-ggplot(endmcoords,aes(abs(y),log(phyloendemism)))+
  geom_point(color="black")+
  xlab("Latitud absoluta")+
  ylab("log (Filoendemismos)")+
  geom_smooth(method= lm, , color="red", fill="#69b3a2", se=TRUE)+
```

```
theme_classic()
```

```
plot(p2)
```

¿Qué podemos concluir de esto?

¿Qué otras cosas podríamos hacer utilizando estos paquetes?