

# Distribución espacial de las especies de la familia *Malvaceae* en una parcela de 50 ,ha. Caso: Isla Barro Colorado, Panamá.

**Ana Hilda Valera Arias**    *Estudiante, Universidad Autónoma de Santo Domingo (UASD)*

---

La familia *Malvaceae* son plantas generalmente de tipo herbáceas, leñosas o de arbustos donde su distribución en el espacio estará condicionado principalmente por el factor clima. Este estudio fue realizado por la aplicación de la ecología numérica y tuvo como objetivo conocer la asociación de las diferentes especies y las variables ambientales que influyen en ello. Además, del análisis de la forma en que se agrupan y el patrón de distribución que presentan, al igual, la identificación de los indicadores ambientales que interfieren. Asimismo, examinar en qué volumen se encuentran representadas cada una y distinguir los sitios con especies alpha y beta. Los datos obtenidos fueron por medio de los censos realizados en varios años en la BCI y sometidos a distintas métricas de funciones ecológicas para la satisfacción de los resultados esperados. Existen 16 grupos de dicha familia y se agrupan de manera discontinua dentro de la parcela permanente de 50 hectárea en la isla, mayormente se concentran en sitios húmedos en presencia de los elementos zinc (Zn), Boro (B), potasio (K), donde el nivel del pH sea idóneo y en la zonas de la llanura, la vaguada, la vertiente y el espolón que pueden ser aptas para su crecimiento y desarrollo. Las especies reconocidas como indicadoras fueron la *Quararibea asterolepis* y la *Sterculia apetala* por tener los índices más altos 0.98% y 0.91% respectivamente. Asimismo, existe una distribución equitativa entre las especies con un rango de riqueza de 5 a 13 y de abundancia de 31 a 127 grupos por sitios, siendo la *Quararibea asterolepis* la principal contribuidora a la diversidad beta según el correlograma con 0.18%.

*Keywords:* ecología numérica, zinc, boro, *Quararibea asterolepis*

---

## 1 Introducción

La vegetación terrestre está constituida por un conjunto de plantas conocidas que alcanza cientos de miles de especies conocidas, cada una perteneciente a una familia en específico. Las familias se subdividen a su vez en géneros y especies para nombrarlas apropiadamente. En el caso de *Malvaceae* hay registrados 243 géneros y más de 4,300 especies; son plantas con flores son hermafroditas, pocas veces unisexuales, solitarias o fasciculadas en las axilas de las hojas o agrupadas en inflorescencia (Bayer, 2003; Marín, Hilario, & Andino, n.d.).

Dentro de los géneros más comunes en la familia *Malvaceae* están el *Abutilon* constituido por arbustos, subarbustos y hierbas bienales con pelo estrellados y tallos velloso, son carente del epicáliz conjunto de apéndices que por lo regular tienen otros grupos de dicha familia, así como de tener alrededor de 150 especie nativa en los trópicos y subtrópicos de América, África, Asia y Australia (Lorenzo-Cáceres, 2007). También, está el *Hibiscus* donde los segmentos del epicáliz están libres o unidos en la base, con estigmas alargados, semillas reniformes y numerosas (ORTIZ, 2010). Del mismo modo, se encuentra la *Althaea*, *Lavatera* y la *Malva* cada una contienen sus respectivas especies, las cuales pueden encontrarse en mayor o menor proporción en un espacio determinad,

en función de factores abioticos incidentes entre ellos, lo que implicaría la necesidad de utilizar técnicas y análisis cuantitativos para conocer su asociación y distribución.

La implementación de análisis numéricos en las investigaciones ecológicas permiten conocer, en términos cuantificables, la forma en que se encuentran asociadas y el tipo de patrón que presenta algunas especies, es por ello la importancia de la estratificación y zonificación del objeto de estudio en cuestión. De acuerdo con González (2006), esto permite conjugar en un mismo grupo información de aquellos organismos que pueden ser calculable junto con otros que son reproductivos y de manera general con toda la vegetación.

En tal sentido, este estudio por medio de la ecología numérica busca conocer cómo están asociadas las diferentes especies de la familia *Malvaceae* y si las variables ambientales existentes en la zona influyen en dicha asociación. También, analizar como están organizados los grupos y qué patrón presenta en su distribución, asimismo, establecer los indicadores ambientales que interfieren. De igual manera, examinar en qué volumen se encuentran representadas cada una y distinguir los sitios con especies alpha y beta. Por consiguiente, este estudio contribuirá al conocimiento de la dinámica ecológica espacial que envuelven las plantas pertenecientes a esta familia en la isla de Barro Colorado que en lo adelante será llamado BCI y con la misma gestionar estrategias para el cuidado y conservación de ésta.

## 2 Metodología

### 2.1 Área de estudio y datos fuente

La isla de Barro Colorado se encuentra ubicada en el canal de Panamá en las proximidades del lago Gatún. Según Pérez et al. (2005) esta se formó cuando se construyó dicho canal embalsando las aguas del río Chagres. La isla se localiza entre las coordenadas geográficas (latitud 9° 9'N, longitud 79° 50') y cubre una extensión de tierra de 1,500 hectáreas (ver figura 1). El clima se caracteriza por ser de bosque tropical, la temperatura promedio es de 27 grados centígrados, con temporadas lluviosas durante los meses mayo a diciembre y secas desde mediados de diciembre hasta abril, las tormentas convectivas son prevenidas por los vientos alisios dictando así las estaciones del año (Sugasti, Eng, & Pinzón, 2018).

Esta isla, por sus características físicas sirve de hábitat para muchos animales e insectos y por consiguiente para una variedad de especies vegetales, y es a la vez un espacio con una tradición de investigación de mucha importancia. Dentro de ella se encuentra la parcela permanente de 50 hectáreas de BCI (en lo adelante, BCI), dentro de la cual es posible caracterizar la distribución de la familia *Malvaceae* a través de los numerosos censos realizados con anterioridad.

Las informaciones obtenidas datan de los censos registrados durante los años (1981-1983 y 2010-2015, entre otros) donde se marcaron y cartografiaron todos los tallos leñosos independientes de al menos 10 mm de diámetro de altura (Hubbell, Condit, & Foster, 2021). En los mismos, se observó que en el año 1981 y para el 2015 de las 253 especie identificadas y medidas la *Quararibea asterolpis*, *Ceiba pentandra*, *Apeiba membranacea* y *Luehea seemannii* estuvieron entre las plantas de tallos reducidos aunque para el último año esta dimensión cambió relativamente, al contrario de la *Cavanillesia platanifolia* que en el primer censo fueron poco visibles en comparación con los posteriores que tuvieron mayor visibilidad.

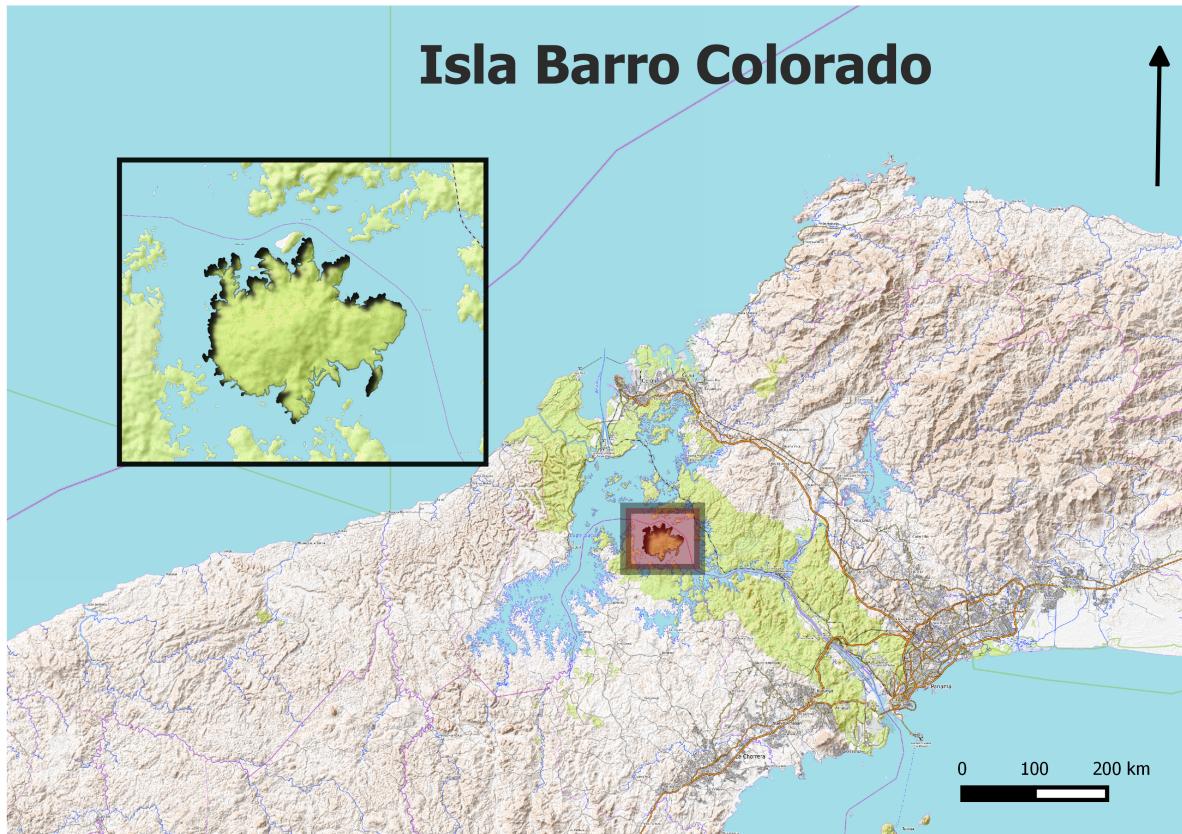


Figure 1: Ubicación de la isla Barro Colorado

### 2.1.1 Materiales y Técnicas de investigación

Para alcanzar los objetivos planteados, se efectuaron análisis aplicados a la base de datos censales de BCI. Específicamente, se realizaron análisis exploratorios sobre medición de asociación, de diversidad y de ecología espacial.

Los análisis exploratorios incluyeron la transformación de la matriz ambiental en columna de hábitat y el agrupamiento de las plantas por quadrat de 1,ha, para obtener resúmenes sobre la riqueza global y abundancia de especie por cada sitio. Asimismo, exploró la relación de cada una de las variables con las unidades geomorfológicas y los elementos químicos del suelo, para generar mapas de cuadros para representar patrones espaciales, así como para evaluar relación con grupos de plantas. Por otra parte, los de medición de asociación contribuyeron a conocer la relación existente entre cada especie utilizando una matriz de transpuesta para la comparación por medio de la transformación Chi. Se aplicaron distintas métricas en modo Q y R, tales como: el índice de correlación de Pearson, el *rho* de Spearman o *tau* de Kendall con las que se exploró, de manera preliminar, el grado de asociación entre descriptores.

El análisis de agrupamiento de los sitios de BCI basados en su composición usando una matriz de comunidad transformada por el método de “cuerdas” (o “matriz normalizada”), se realizó aplicando cuatro métodos distintos entre jerárquicos y aglomerativos. Específicamente, se incluyeron los métodos por enlace simple, completo y enlace promedio (mejor conocido como UPGMA) (Borcard, 2018). Estos métodos se basan en el agrupamiento secuencial y jerárquico (cada elemento

forma parte de un grupo), siendo el rasgo distintivo de cada uno de ellos el criterio de conexión que se realiza, ya sea por medio de valores mínimos (enlace simple), máximos (completo) o promedios. Adicionalmente, se utilizó el orden de la varianza mínima de Ward (que es también jerárquico), el cual aplica un criterio similar al de la regresión lineal por mínimos cuadrados. Cada método produjo una agrupación, que fue representado por medio de árboles. Para elegir un sistema idóneo, se utilizó el criterio de la correlación cofenética, que consiste en calcular, elemento a elemento, la correlación entre las matrices de distancias originales y cofenéticas. En tal sentido, el que obtuvo mayor valor de correlación cofenética, fue considerado como candidato oportuno, pero no fue elegido de manera automática. Para la clasificación de sitio se aplicó un segundo criterio, que fue el de determinación de grupos aptos por medio de las anchuras de siluetas. Se comenzó prefiriendo aquel método que alcanzara mayor conveniencia cofenética; si tras cortar su árbol en el número de grupos sugerido por las anchuras de siluetas, se obtenía al menos un grupo representado por tres o menos sitios sería descartado y se prefería el siguiente con mayor idoneidad, hasta obtener el de máxima adecuación y que no produjese grupos compuestos por tres o menos sitios.

Adicionalmente, como criterio complementario, se evaluaron los grupos obtenidos mediante remuestreo por bootstrap multiescalar (Borcard, 2018). Esta técnica determina, por medio de probabilidades bootstrap aplicada a submuestras de tamaño variable, cuáles grupos son realmente consistentes en un determinado agrupamiento, tanto por medio de la probabilidad bootstrap ("BP") como por medio de la probabilidad aproximadamente insesgada ("AU"). De modo, que los grupos que superan esta prueba pueden considerarse como "consistentes", por lo que la técnica ayuda a validar o rechazar clasificaciones realizadas por otros procedimientos. Con los resultados del agrupamiento elegido, se realizaron dos análisis adicionales: 1) Evaluación de la homogeneidad de promedios entre los grupos establecidos a partir del agrupamiento elegido, tanto por medio de diagramas de cajas como por pruebas estadísticas t-Student y de la suma de rangos de Wilcoxon; 2) Realización de análisis de asociación de especies a grupos de sitios, por medio de las métricas "valor indicador" (IndVal) y el coeficiente de correlación biserial puntual (Borcard, 2018).

Los análisis de diversidad contribuyeron al conocimiento de la abundancia de especies por cada sitio a través de dos componentes: riqueza y equidad. Se generó una matriz de comunidad combinada para la identificación de la diversidad alpha y las plantas que contribuyen a la diversidad beta. En tal sentido, para la medición se utilizó el índice de **equidad de Pielou** donde se midió la homogeneidad de abundancia de especie, el **índice de concentración de Simpson** para calcular el nivel de concentración de cada especie en el área. También se utilizaron las ratios de *Hill* para realizar comparaciones, dado que dichos índices tienen baja probabilidad de quedar influenciados en los casos de una matriz de comunidad con fuerte dominancia (Borcard, 2018). Además, se aplicaron dos enfoques de estimación de riqueza: asintóticos y no asintóticos, que permitieron comparar la estimación de la riqueza de cada espacio. Con el asintótico se realizó un estimado de abundancia y con el no asintótico se hizo una rarefacción y extrapolación basado en el tamaño y cobertura de la muestra.

Finalmente, los análisis de ecología espacial abordaron la autocorrelación de las especies en el espacio por medio de una matriz de comunidad transformada desde cuatro métricas distintas: mediante correlograma, por prueba Mantel (matrices de distancia), pruebas de permutación para el I de Moran local y global. Con la primera se conoció el grado de autocorrelación de la abundancia transformada de especie como variables de respuesta, con la segunda la abundancia por posición geográfica, la tercera la riqueza por zonas en comparación con las circundantes, y con la cuarta métrica se identificaron grupos de celdas que presentan parecido (autocorrelación) de su abundancia de especies. Es decir, en general, se identificó la riqueza de cada especie según su

ubicación y el nivel de dependencia espacial de las mismas, ya fuese localmente como para toda la parcela.

En la parte programática, para la realización de los análisis descritos anteriormente, se utilizó el software de R (R Core Team, 2019), y se emplearon varios paquetes de ecología numérica y de análisis de datos. De estos últimos, destacan *dplyr* y *tidyR*, que ayudaron a reorganizar los datos para formar las matrices de comunidad y ambiental, y que permitió identificar las diferentes especies que abundan, en qué cantidad y orden de acuerdo a su pH. También, se utilizó el paquete *simple Features* (*sf*) para crear área de hábitat por cuadros y obtener la densidad de cada especie por metro cuadrado y así conocer la abundancia y riqueza global. De los paquetes de ecología numérica, *vegan* y *BiodiversityR* fueron los más empleados, y se utilizaron para caracterizar y analizar el orden y disimilaridades entre cada especie, diversidad alpha y beta, entre otras métricas (Kindt & Coe, 2005). De manera general, también se usaron los paquetes *ez* para examinar correlación entre variables. Asimismo, los paquetes *ggplot2*, *graphics*, *psych* y *mapview* para la representación gráfica de cada datos. Por último, cada *script* utilizado fue adaptado a partir del repositorio de Batlle (2020), el cual sirvió como fuente. Adicionalmente, se empleó el programa de información geográfica Qgis (QGIS Development Team, 2009) para actualizar el mapa de localización de la BCI.

### 3 Resultados

Los árboles generados por los métodos de enlaces simple y completo, UPGMA y Ward, aplicados a la matriz de comunidad normalizada, mostraron las típicas organizaciones de los sitios en función de su altura (ver figura 2). A excepción del árbol producido por el método Ward, los generados por otros sistemas, mostraron singularidades que los hacen difíciles de interpretar y de reutilizar en posteriores análisis. En general, el principal problema de los árboles generados por enlaces simple, completo, y por UPGMA, fue que produjeron agrupamientos integrados por un único sitio o por un puñado de éstos. Por ejemplo, varios árboles aislaron a los sitios 31 y 37 en una rama singular, o igualmente al sitio 13 como integrante único de un “grupo”. Si bien, este último lugar es singular por la abundancia de *Luehea seemannii* y la rareza de *Quararibea asterolepis*, no menos cierto es que dicha zona guarda similaridad con el resto si se analiza la representación de otras especies (ver figura 3).

Asimismo, el remuestreo por bootstrap multiescalar utilizando como base el agrupamiento Ward, no destacó singularidad alguna para el sitio 13. Por estas razones, se prefirió utilizar el agrupamiento realizado con Ward, cortado en dos grupos, siguiendo los resultados obtenidos por el criterio de anchura de siluetas. El grupo más numeroso, denominado “1” en lo adelante, está representado por 42 sitios, y el grupo denominado 2 lo integran 8 sitios.

Se evaluaron homogeneidades de promedios de las variables ambientales entre los dos grupos establecidos por el método Ward. Las variables que presentaron inhomogeneidades significativas fueron N, pH, Cu, B, Mn, y el porcentaje de relieve de piedemonte en el cuadro (ver figura 4 y tabla ). De manera particular, los sitios agrupados en el grupo 1, presentan valores significativamente elevados de las variables N, B, pH (entre los ácidos, los menos ácidos del conjunto), Cu y Mn. Asimismo, mediante el IndVal y el coeficiente de correlación biserial puntual, se detectaron varias especies asociadas con los hábitats representados por cada uno de los grupos. Concretamente, se detectó a *Quararibea asterolepis* como especie asociada al grupo 1, mientras que asociadas al grupo 2, se detectaron *Luehea seemannii*, *Sterculia apetala* y *Apeiba tibourbou*.

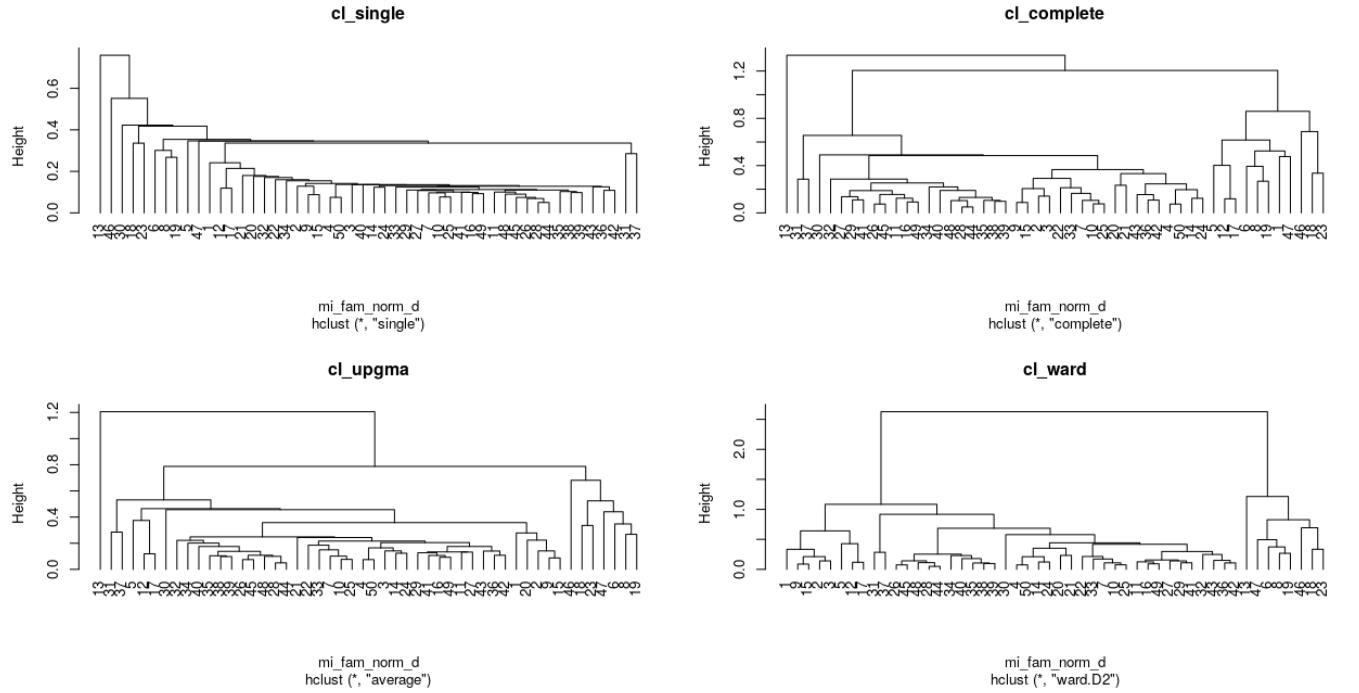


Figure 2: Matriz de Comunidad Normalizada

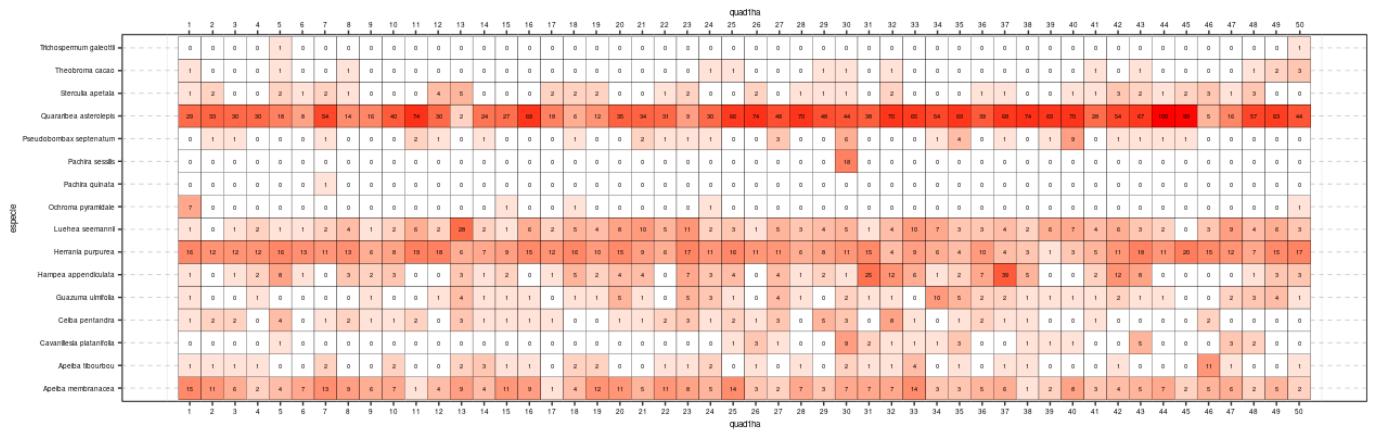


Figure 3: Representación de especies por abundancia de sitio

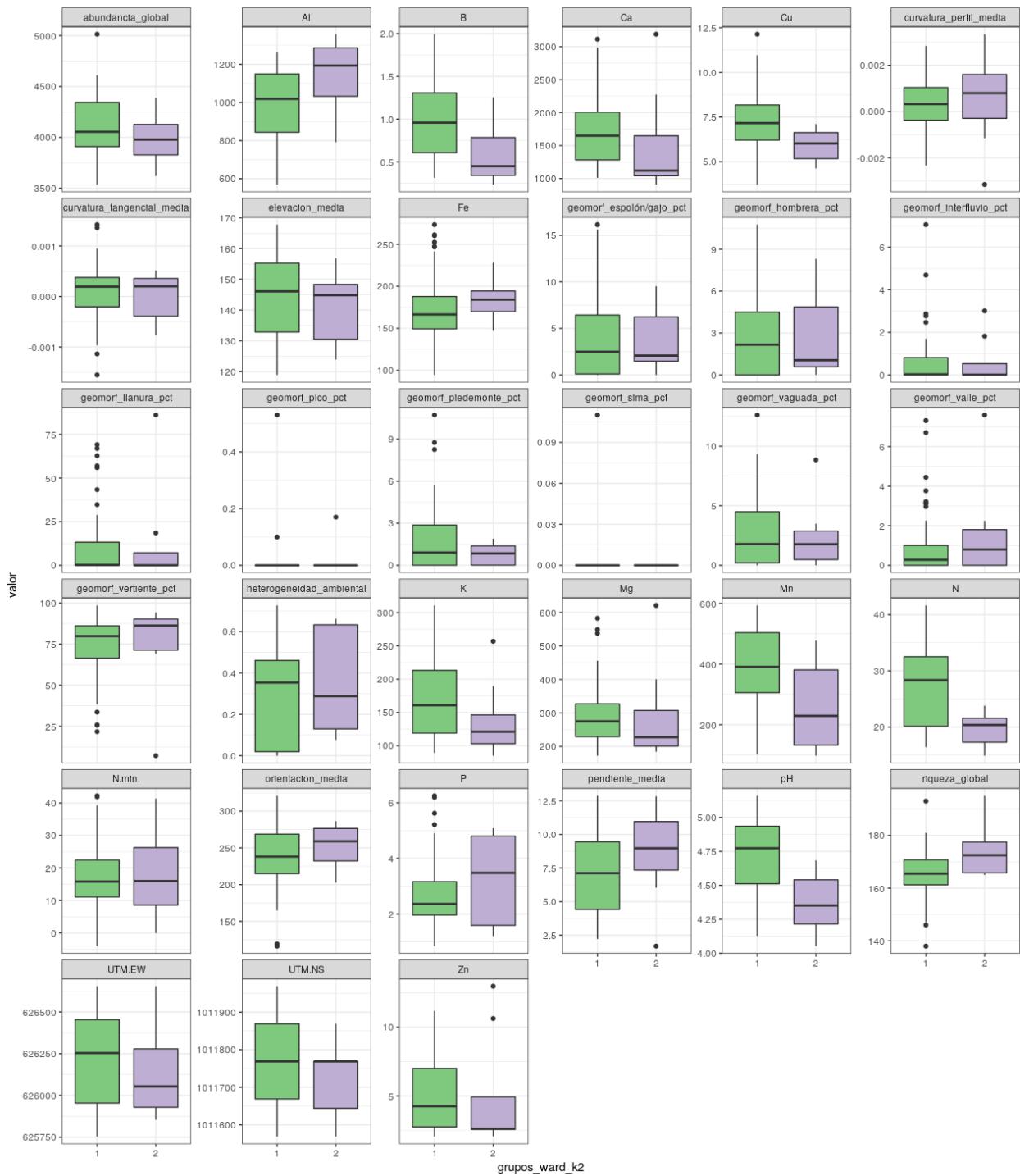


Figure 4: Representación de especies por abundancia de sitio

Table 1: Homogeneidad de promedios de variables ambientales

variable	Valor de p, prueba t	Valor de p, prueba Wilcoxon
N	0.00	0.02
pH	0.00	0.00
Cu	0.01	0.03
B	0.02	0.02
geomorf_piedemonte_pct	0.02	0.43
Mn	0.04	0.03
riqueza_global	0.05	0.04
Al	0.11	0.05

Al aplicar las diferentes métricas como la equidad de *Shannon*, la antropía *simpson* y la ratio de *Hill* para conocer la abundancia y riqueza de especies por cada sitio, se observó una alta correlación entre cada una a través del método de *Pearson* a excepción de la ratio de *Hill* donde fue menor. También, en conformidad con las variables ambientales la abundancia de la familia fue mayor en presencia del zinc (Zn), nitrogeno (N) y el pH con valores de 0.45%, 0.47% y 0.66% respectivamente. En tanto, que la riqueza global predominó en la zona donde el pH fue más idóneo con estimado de -0.48% (ver figura 7). Por otro lado, en los modelos de abundancia de especies los sitios alpha con mayor y menor riqueza son el 30 y 45 con 13 y 5 especies sucesivamente, los de máxima y mínima abundancia fue el 6 y 37 con un rango entre 31 y 127. Siendo el lugar más rico el 30 donde abundan 110 grupos y el más pobre el 45 con 123. Mientras, que los sitios que contribuyen a la diversidad beta al implementar el modelo de un único número con la matriz de comunidad con el método de *Helliger* fueron el 13 con 0.12% y 46 con 0.82%. En tanto, que las especies contribuidoras son la *Quararibea asterolepis*, *Hampea appendiculata* y *Herrania purpurea* con 0.18%, 0.14% y 0.11%, respectivamente (ver tabla 2).

En tanto, que la rarefacción a la abundancia más pequeña encontrada fue en la zona 45 con 5 y la estimación de individuo por abundancia de grupo de 3,393. Asimismo, al combinar los grupos y ubicarlo en un solo lugar por medio del enfoque asintótico, se observó que exceden un total de 3,789 individuos y 14 especie. Dentro de las mismas se encuentran 4 sujeto y 2 clases raras con un estimado de cobertura muestral de 0.67% para plantas extrañas, con grado de 16% de homogeneidad, con un nivel bajo de 16% y superior 17.72%. En cuanto, a la matriz con el enfoque no asintótico se registró un espacio con 3,792 individuo y 16 grupos, al relacionarla con la antropía de *Shannon* se estimó un 4.68% y con la equidad de *Simpson* un 2.77% de diversidad. Por otro lado, al agrupar la comunidad según el método de *Ward* se visualizó una cantidad de 3,393 sujeto, 13 clase, con 13 individuo y 3 grupos inusuales, con una estimación de 0.92% de especies.

De acuerdo los datos obtenidos en el análisis de la ecología espacial por medio de las pruebas de permutación para el *I de Moran* local se observó una alta correlación de algunas especies en varias zonas geomorfológicas de la BCI en el orden uno. Donde se estima un 0.42% en la llanura, 0.31% en el espolón, 0.29% en la vertiente y 0.38% en la vaguada, siendo el área del espolón el de más conveniencia. Mientras, que ante la presencia de ciertos elementos químicos algunas plantas presentaron una adecuación como fue con el zinc (Zn) con 0.85%, el potasio (K) 0.74%, el calcio (Ca) 0.69 y el pH con 0.72%. Mientras, que con el *I de Moran* global aplicado a la abundancia de las especies de la familia *Malvaceae* las de mayor grado de autocorrelación fue la *Pachira quinata* con valor de 1%, *Pachira sessilis* con 0.71%, *Guazuma ulmifolia* 0.62% y la *Ochroma pyramidale* 0.53%.

En cuanto a la matriz de distancia con la prueba Mantel se observó leve dependencia espacial de algunas especies inducida por variables ambientales, a medida en que aumenta ésta la autocorrelación es de forma negativa, siendo el orden 1 y 2 los de más conexión con 4.62% y -4.95% respectivamente. También, mediante el correlograma con una sola variable los grupos tuvieron una fuerte relación en los siete ordenes primero con estimaciones entre 0.21% y 0.72% en los tres iniciales y desde -0.36% a -0.66% en los cuatro siguientes. Por último, la clase que más abunda es *Herrania purpurea* con un valor estimado de 0.41%, seguido por la *Apeiba membranacea* con 0.36% (ver figura 5 & 6).

Table 2: Especie contribuidora a la diversidad beta

Especie	Valor
<i>Apeiba membranacea</i>	0.08
<i>Apeiba tibourbou</i>	0.07
<i>Hampea appendiculata</i>	0.14
<i>Herrania purpurea</i>	0.11
<i>Luehea seemannii</i>	0.09
<i>Quararibea asterolepis</i>	0.18

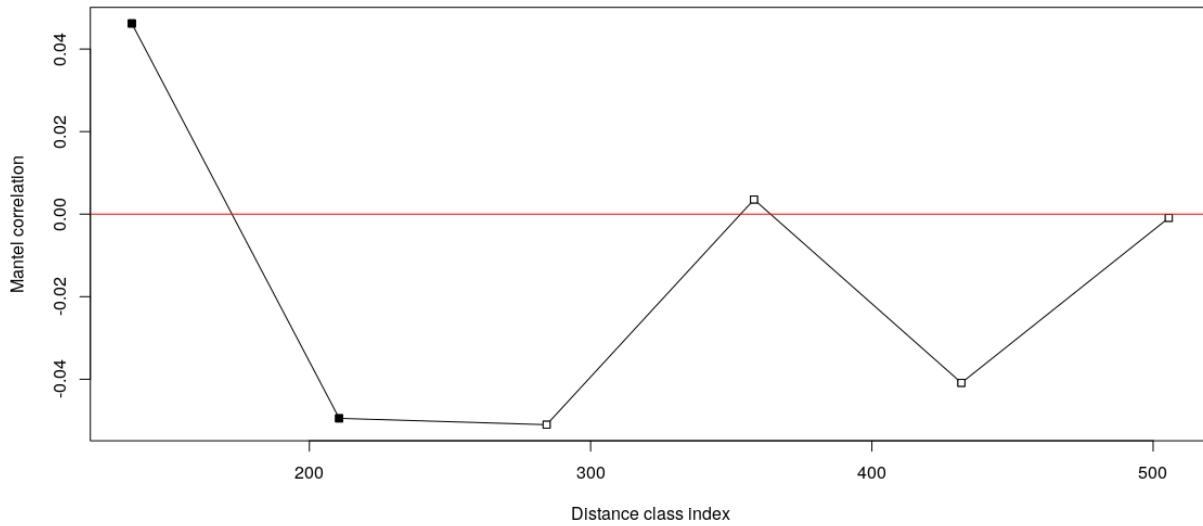


Figure 5: Distancia con prueba Mantel

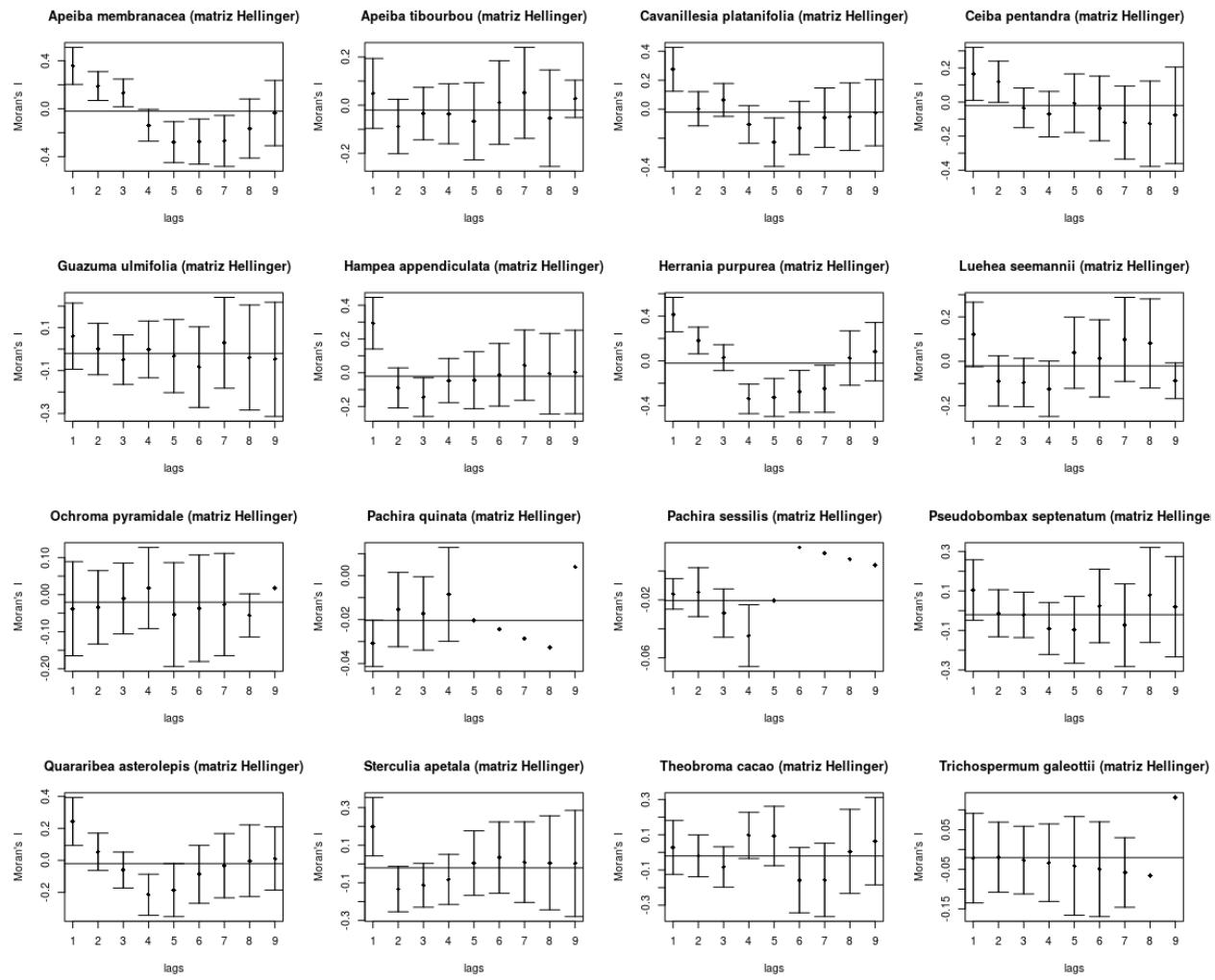


Figure 6: Correlación de las diferentes especies



Figure 7: Distribución de especies por ambiente

#### 4 Discusión

Los agrupamientos integrados producidos por varias especies en los dos grupos generados por los diferentes métodos de enlace, donde se identificaron las plantas *Quararibea asterolepis* y *Sterculia apetala* como las que más se asocian de manera combinada dentro de la familia, sugieren que alguna de estas especies sirve de refugio para otros grupos, provocando concentraciones de individuos que favorecen el crecimiento y desarrollo de otras especies (Alvarado-Hernández, 2011).

El nivel de representación de la diversidad de especies dado por hectárea, fue como resultado de las condiciones climáticas y a la disponibilidad del agua existente en la zona (Stevenson & Rodríguez, 2008), razón de que en los sitios más húmedo el nivel de abundancia y riqueza fue mayor. Además, de la presencia de otras clases que contribuyeron a la variación espacial de la diversidad entre los lugares como fue la *Apeiba membranacea*, *Hampea appendiculata*.

En cuanto a la ecología espacial, la formación de patrones discontinuo por aglomeración en presencia de los elementos químico como el zinc (Zn), el boro (B) y el potasio (K) especialmente en la zona de la Vaguada, la llanura, el espolón y la vertiente indica que ciertas especies de la familia de la Malvaceae son propensas a desarrollarse en sitios con alto volumen de materia orgánica y gran disponibilidad de agua. Aunque, el exceso de ciertos elementos pueden afectar la distribución y crecimiento de ciertas plantas (Clark, 2002), así, como el origen de formación del suelo los cuales ofrecen ciertas propiedades que definirían el tipo de colocación (Flores, Suvires, & Dalmasso, 2015). Esto posiblemente sucedió en los lugares del piedemonte, el valle, la sima y ante el conjunto de compuesto como el fósforo (P), cobre (Cu) y el aluminio (Al) donde la correlación fue menor. También, es posible que el modelo presentado sea el resultado por los períodos de sequía

que pudo padecer la parcela en el momento de algunos de los censos, razón de que este fenómeno en caso de prolongarse podría causar la extinción local de especies arbóreas (Springer, 1998).

El patrón presentado por las especies de la familia *Malvaceae* dentro de la parcela permanente de la BCI puede variar al pasar el tiempo por causa del cambio en el clima. Razón, de que algunos grupos tienen alta dependencia a desarrollarse en zonas húmedas las cuales pueden ser afectadas por una temporada de sequía que enfrente ese espacio, dejando como posibles consecuencias la alteración de las características físicas y la extinción de algunas especies. Posiblemente fue el caso de la *Apeiba hybrid* quién para los censos de 1981 y 1985 mostró un área basal de 0.005 y 0.003 respectivamente, mientras que para los restantes no presentó valor de medición (Hubbell et al., 2021). En tal sentido, se espera la desaparición de algunas de la clase que componen dicha familia si el ser humano no diseña estrategias oportunas para contrarestar las variaciones climáticas y la conservación de las plantas originarias del bosque.

## 5 Agradecimientos

Este estudio fue realizado gracias al Dr. José Ramón Martínez Batlle maestro de la asignatura de Biogeografía en la Universidad Autónoma de Santo Domingo (UASD) quién por medio de la escuela de geografía incentiva al conocimiento a través de la investigación de carácter científico. También, al Instituto Smithsonian de Investigaciones Tropicales por suministrar los datos, resultado de años de observaciones y puesto a la disposición mediante censos.

## 6 Información de soporte

```
\beginsupplement
```

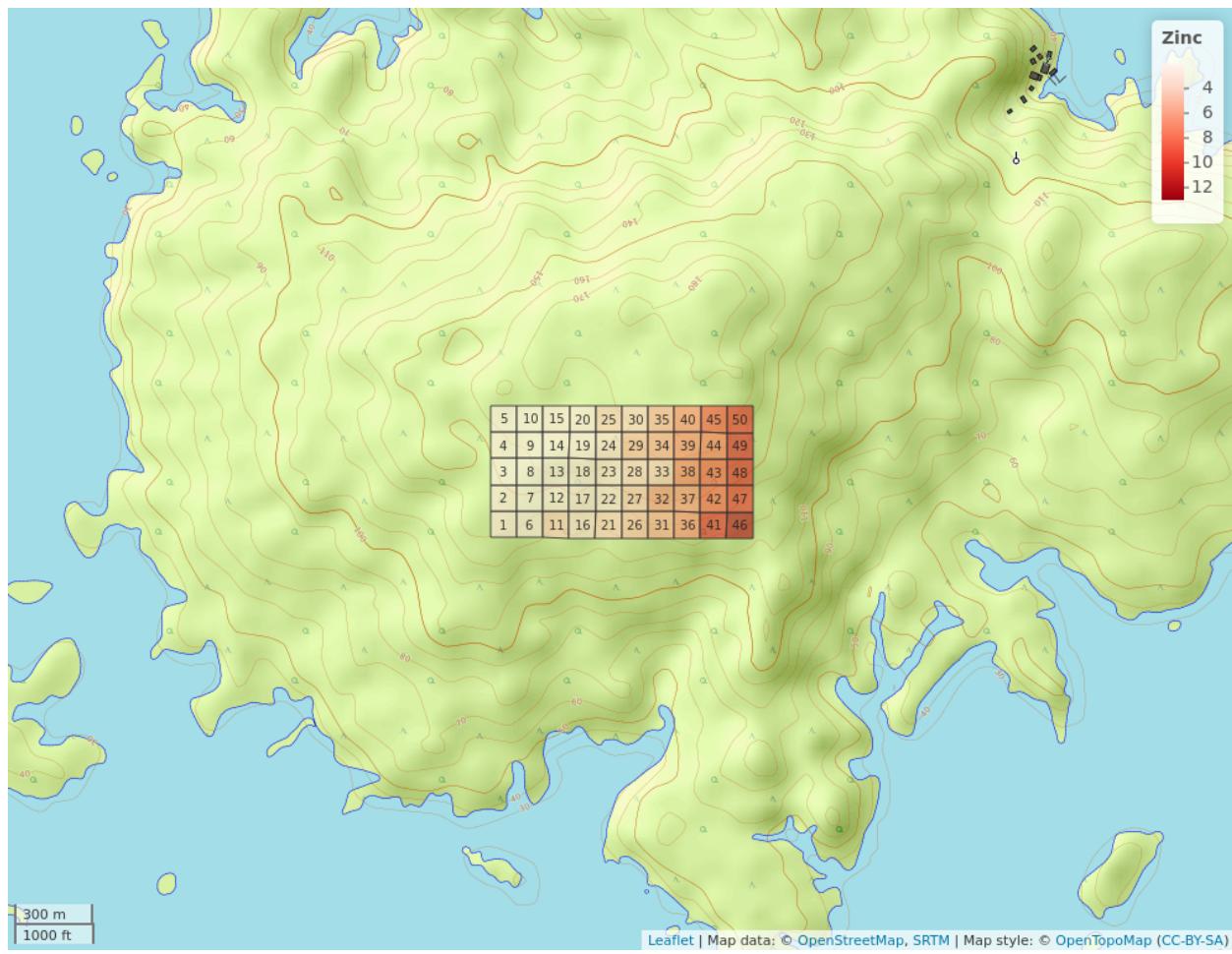


Figure 8: Mapa de asociación de especie con el elemento Zinc

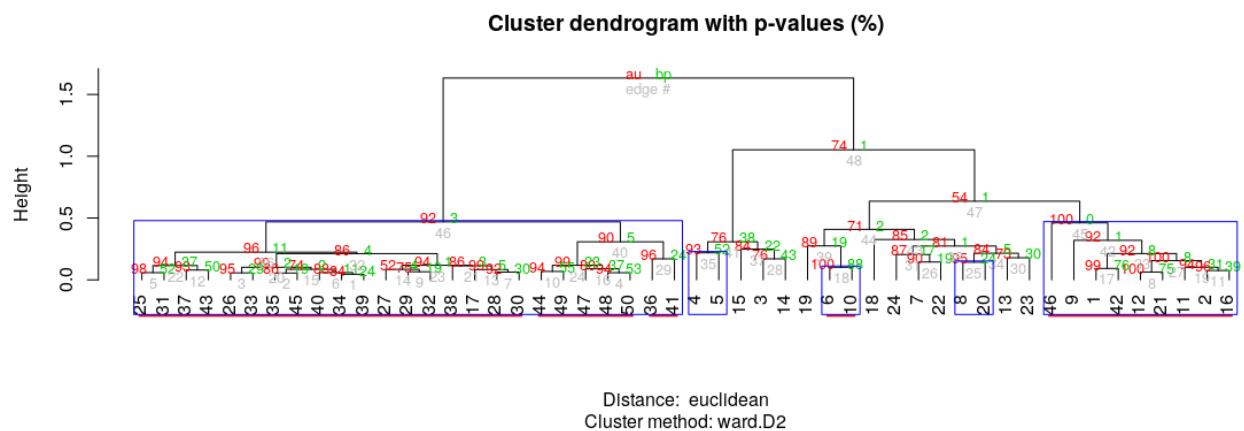


Figure 9: Dendograma de la distancia por sitios con el método Ward\_2

Cuadros de 1 Ha de BCI. Variables ambientales numéricas escaladas de 0 a 1

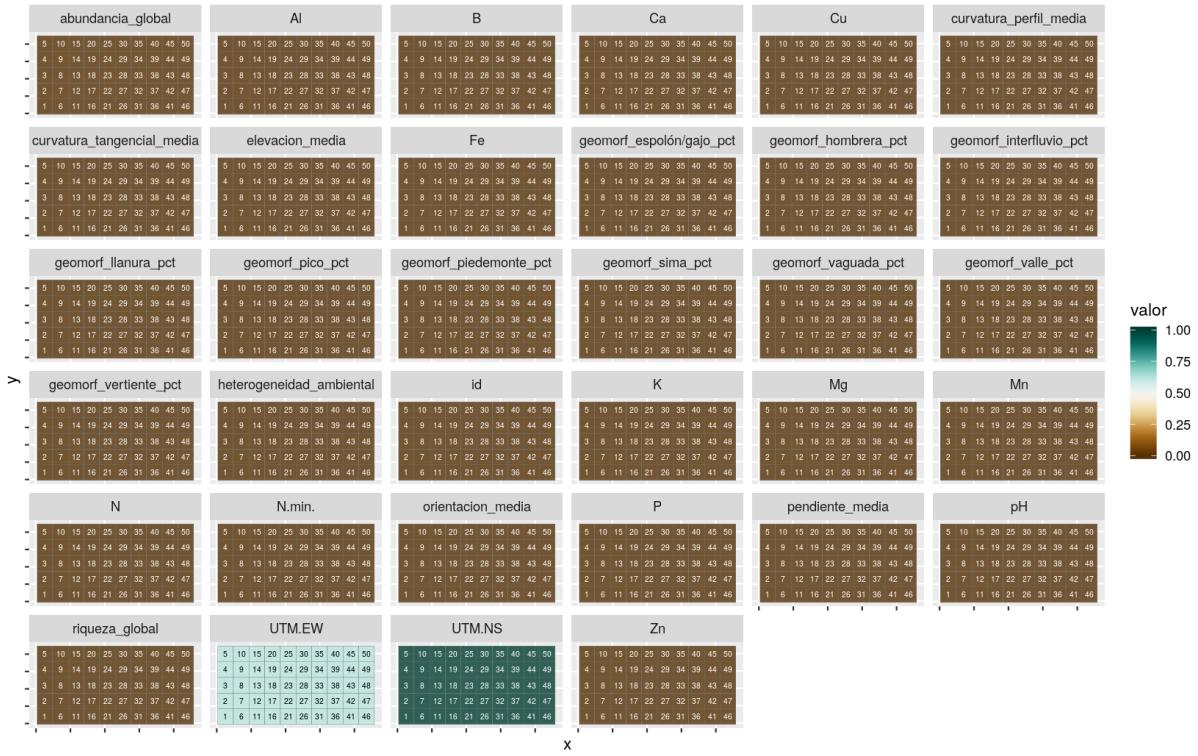


Figure 10: Variables ambientales numéricas

## Cuadros de 1 Ha de BCI. Variables ambientales nominales

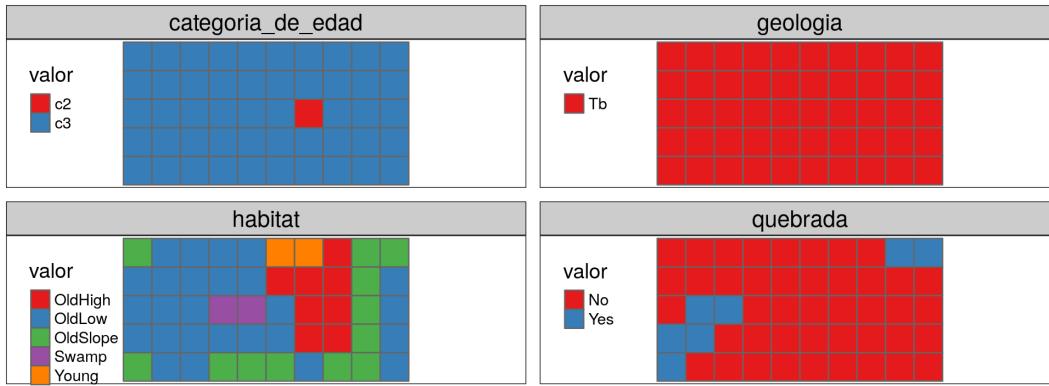


Figure 11: Variables ambientales nominales

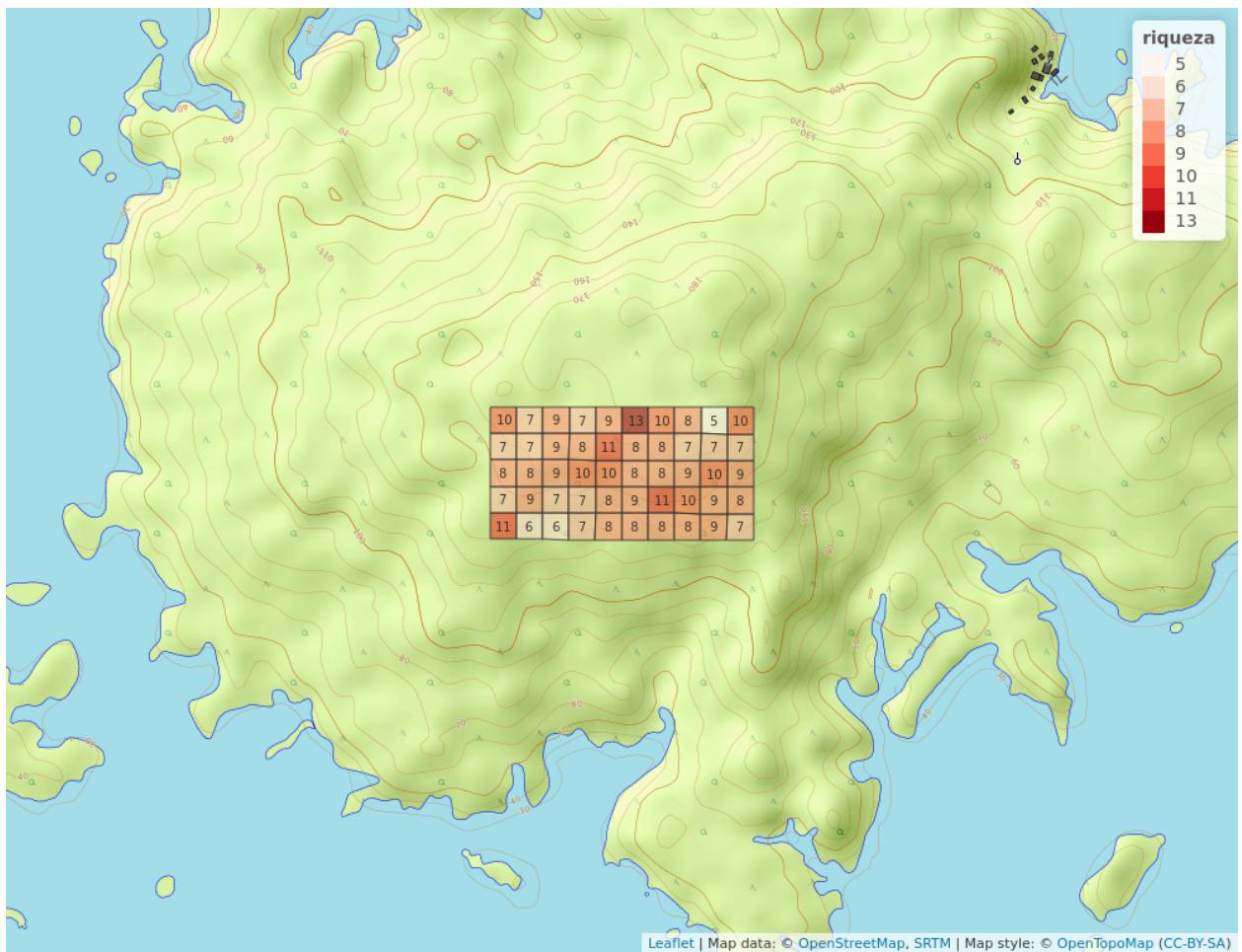


Figure 12: Mapa de riqueza por sitios de la familia *Malvaceae*

## 7 Script reproducible

```

library(vegan)
library(adespatial)
library(broom)
library(tidyverse)
library(sf)
library(gclus)
source('biodata/funciones.R')

load('biodata/matriz_ambiental.Rdata')
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")

mi_fam_t <- mc_malvc %>%
  rename_all(gsub, pattern = ' ', replacement = '\n') %>%

```

```

t()
mi_fam_t %>% tibble

mi_fam_t_chi <- decostand(mi_fam_t, "chi.square")
mi_fam_t_chi %>% tibble

mi_fam_t_chi_d <- dist(mi_fam_t_chi)
mi_fam_t_chi_d %>% tidy

coldiss(mi_fam_t_chi_d, diag = TRUE)

mi_fam_t_jac <- vegdist(mi_fam_t, "jaccard", binary = TRUE)
mi_fam_t_jac %>% tidy
coldiss(mi_fam_t_jac, diag = TRUE)

env_num <- bci_env_grid %>%
  dplyr::select_if(is.numeric) %>%
  dplyr::select(-id, -matches('^U.*')) %>%
  st_drop_geometry %>%
  mutate(
    riqueza_mifam = specnumber(mc_malvc),
    abundancia_mifam = rowSums(mc_malvc)) %>%
    rename_all(gsub, pattern = '_pct$', replacement = ',') %>%
    rename_all(gsub, pattern = '_| ', replacement = '\n')
env_num %>% tibble

```

```

#Medición de asociación con el modo Q: matriz de disimilaridad.
library(vegan)
library(adespatial)
library(broom)
library(tidyverse)
library(sf)
library(cluster)
library(gclus)
source('biodata/funciones.R')

load('biodata/matriz_ambiental.Rdata')
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")

#' matriz de distancia euclídea, utilizando la transformación *Hellinger*:
#
mi_fam_d_hel <- dist.ldc(mc_malvc, "hellinger", silent = T)
mi_fam_d_hel %>% tidy # Para evitar desbordar la consola

coldiss(mi_fam_d_hel, diag = T)

coldissgg(mi_fam_d_hel, ordered = T, nc = 4, fsz = 0)

```

```

coldissgg(mi_fam_d_hel, ordered = T, nc = 4, fsz = 1.5)

png(
  filename = 'matriz_disimilaridad_hellinger.png',
  width = 2400, height = 1200, pointsize = 32
)
coldiss(mi_fam_d_hel, diag = T)
dev.off()

**distancia de Jaccard** (**D<sub>J</sub>**) en un único paso usando la función `vegdist`.
#
mi_fam_jac <- vegdist(mc_malvc, method = 'jac', binary = T)
mi_fam_jac %>% tidy # Mostrando sólo las primeras 10 combinaciones en modo data.frame

coldiss(mi_fam_jac, diag = T)

#
(1 - mi_fam_jac) %>% tidy %>% rename(similaridad=distance) #Similaridad

#' La fórmula de la similaridad de Jaccard es **S<sub>J</sub>=a/(a+b+c)**, donde **a** es el númer
#
mi_fam_abc <- betadiver(mc_malvc)
mi_fam_abc %>%
  map(tidy) %>%
  map(slice, 1) %>%
  map_df(I, .id = 'tipo') %>%
  dplyr::select(tipo, n_especies=distance)

round(11/12*100,2) #Porcentaje de especies compartidas = similaridad

#' Índices de similaridad. Por ejemplo, el Jaccard se calcula así:
#
betadiver(mc_malvc, method = 'j') %>% tidy

#' Además de la distancia de Jaccard, otra distancia muy utilizada es la de Sorensen o Bray-Curt
#
mi_fam_sor <- vegdist(mc_malvc, method = 'bray', binary = T)
mi_fam_sor %>% tidy
coldiss(mi_fam_sor, diag = T)
#
#' ### Modo Q para datos cuantitativos, NO de abundancia de especies (variables ambientales)

env_suelo_punt_z <- bci_env_grid %>%
  st_drop_geometry() %>%
  dplyr::select(matches('^[A-T,Z] | ^pH$', ignore.case = F)) %>%
  scale()
env_suelo_punt_z_d <- dist(env_suelo_punt_z)

```

```

env_suelo_punt_z_d %>% tidy
coldiss(env_suelo_punt_z_d, diag = T)
#
#' ### Modo Q para datos cualitativos y cuantitativos (mixtos), NO de abundancia de especies (v)
#
env_mix <- bci_env_grid %>%
  st_drop_geometry() %>%
  dplyr::select(heterogeneidad_ambiental, habitat, quebrada)
env_mix_d <- daisy(x = env_mix, metric = 'gower')
env_mix_d %>% as.dist %>% tidy
env_mix_d %>% coldiss(diag = T)
#
#' Medición de asociación de especie modo R

library(vegan)
library(adespatial)
library(broom)
library(tidyverse)
library(sf)
library(gclus)
source('biodata/funciones.R')

load('biodata/matriz_ambiental.Rdata')
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
#
#' ## Modo R: matrices de dependencia entre variables (índice de correlación)
#
#' ### Modo R para datos cuantitativos de especies (abundancia)

mi_fam_t <- mc_malvc %>%
  rename_all(gsub, pattern = ' ', replacement = '\n') %>%
  t()
mi_fam_t %>% tibble
#
#' Segundo, transformo la matriz transpuesta usando estandarización *Chi*.
#
mi_fam_t_chi <- decostand(mi_fam_t, "chi.square")
mi_fam_t_chi %>% tibble
#
#' Tercero, calculo la distancia euclídea.
#
mi_fam_t_chi_d <- dist(mi_fam_t_chi)
mi_fam_t_chi_d %>% tidy
#
#' Finalmente, creo el "mapa de calor".

```

```

#'
coldiss(mi_fam_t_chi_d, diag = TRUE)

#' ### Modo R para datos binarios (presencia/ausencia)

mi_fam_t_jac <- vegdist(mi_fam_t, "jaccard", binary = TRUE)
mi_fam_t_jac %>% tidy
coldiss(mi_fam_t_jac, diag = TRUE)
#'
#' ### Modo R para datos cuantitativos, NO de abundancia de especies (variables ambientales)

env_num <- bci_env_grid %>%
  dplyr::select_if(is.numeric) %>%
  dplyr::select(-id, -matches('^U.*')) %>%
  st_drop_geometry %>%
  mutate(
    riqueza_mifam = specnumber(mc_malvc),
    abundancia_mifam = rowSums(mc_malvc)) %>%
  rename_all(gsub, pattern = '_pct$', replacement = ' ') %>%
  rename_all(gsub, pattern = '_| ', replacement = '\n')
env_num %>% tibble

```

```

#' Técnica de ordenación

library(vegan)
library(tidyverse)
library(sf)
library(mapview)
source('biodata/funciones.R')
#'
#' ### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_malvc
(colnames(mi_fam) <- make.cepnames(colnames(mi_fam)))
(df_equivelencias <- data.frame(
  nombre_original = colnames(mc_malvc),
  colnames(mi_fam)))
bci_env_grid %>% tibble
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k2 <- readRDS('grupos_ward_k2.RDS')
table(grupos_ward_k2)
#'Matriz de correlaciones y obtener vectores propios para el PCA.
#'
env_suelo <- bci_env_grid %>%

```

```

st_drop_geometry %>%
  dplyr::select(matches('^[A-T,Z] | ^pH$', ignore.case = F))
env_suelo %>% tibble
env_suelo_pca <- rda(env_suelo, scale = TRUE)
env_suelo_pca
summary(env_suelo_pca)

#'
screeplot(env_suelo_pca, bstick = TRUE)
#'
#' Usando función 'cleanplot.pca'
#'
par(mfrow = c(1, 2))
cleanplot.pca(env_suelo_pca, scaling = 1, mar.percent = 0.08, cex.char1 = 1.5)
cleanplot.pca(env_suelo_pca, scaling = 2, mar.percent = 0.04, cex.char1 = 1.5)
par(mfrow = c(1, 1))
#'
#' Comparar distribución de los sitios en biplots con distribución real en el mapa:
#'
#' #### Generar mapa de cuadros sin simbología
#'
mapa_cuadros <- mapView(
  bci_env_grid,
  col.regions = 'grey80',
  alpha.regions = 0.3,
  map.types = 'OpenTopoMap',
  legend = F, zoom = 14,
  zcol = 'id') %>% addStaticLabels() %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)
mapa_cuadros
#'
#' Comparar con resultados de un análisis de agrupamiento del mismo conjunto de datos. Primero a
#'
(env_agrupamiento <- hclust(dist(scale(env_suelo)), 'ward.D'))
(env_grupos <- cutree(env_agrupamiento, k = 3))
(mi_cluster <- factor(env_grupos))
(mi_cluster_l <- levels(mi_cluster))
(mi_cluster_l_seq <- 1:length(mi_cluster_l))

#'
(puntuaciones <- scores(env_suelo_pca, display = 'wa', scaling = 1))
#'
#' Luego creo el gráfico base, coloco los puntos sobre el gráfico usando las puntuaciones, les co
#'

```

```

grafico_base <- plot(
  env_suelo_pca,
  display = "wa",
  scaling = 1,
  type = "n",
  main = "PCA y grupos"
)
abline(v = 0, lty = "dotted")
abline(h = 0, lty = "dotted")
for (i in mi_cluster_l_seq) {
  points(puntuaciones[mi_cluster == i, ],
         pch = (14 + i),
         cex = 2,
         col = i + 1)
}
text(puntuaciones, row.names(env_suelo), cex = 1, pos = 3)
legend(
  "topright", # Otras alternativas: "bottomleft", "bottomright" y "topleft"
  paste("Grupo", c(mi_cluster_l_seq)),
  pch = 14 + c(mi_cluster_l_seq),
  col = 1 + c(mi_cluster_l_seq),
  pt.cex = 2
)

# (mi_cluster_anterior <- grupos_upgma_k2)
(mi_cluster_anterior <- grupos_ward_k2)
(mi_cluster_anterior_l <- levels(mi_cluster_anterior))
(mi_cluster_anterior_l_seq <- 1:length(mi_cluster_anterior_l))
grafico_base <- plot(
  env_suelo_pca,
  display = "wa",
  scaling = 1,
  type = "n",
  main = "PCA y grupos"
)
abline(v = 0, lty = "dotted")
abline(h = 0, lty = "dotted")
for (i in mi_cluster_anterior_l_seq) {
  points(puntuaciones[mi_cluster_anterior == i, ],
         pch = (14 + i),
         cex = 2,
         col = i + 1)
}
text(puntuaciones, row.names(env_suelo), cex = 1, pos = 3)
legend(
  "topright", # Otras alternativas: "bottomleft", "bottomright" y "topleft"
  paste("Grupo", c(mi_cluster_anterior_l_seq)),

```

```

    pch = 14 + c(mi_cluster_anterior_l_seq),
    col = 1 + c(mi_cluster_anterior_l_seq),
    pt.cex = 2
)

#' ##### PCA aplicado a datos de comunidad transformados
#'
mi_fam_hel <- decostand(mi_fam, method = 'hellinger')
mi_fam_hel %>% tibble
mi_fam_hel_pca <- rda(mi_fam_hel)
summary(mi_fam_hel_pca)
screeplot(
  mi_fam_hel_pca,
  bstick = TRUE,
  n pcs = length(mi_fam_hel_pca$CA$eig)
)
mi_fam_hel_pca_sc1 <- scores(mi_fam_hel_pca,
                               display = "species", scaling = 1)
mi_fam_hel_pca_sc2 <- scores(mi_fam_hel_pca,
                               display = "species", scaling = 2)
par(mfrow = c(1, 2))
cleanplot.pca(mi_fam_hel_pca, scaling = 1, mar.percent = 0.06, cex.char1 = 0.7)
cleanplot.pca(mi_fam_hel_pca, scaling = 2, mar.percent = 0.06, cex.char1 = 0.7)
par(mfrow = c(1, 1))

biplot(
  mi_fam_hel_pca,
  main = "PCA, escalamiento 2, ajuste a variables ambientales")
(mi_fam_hel_pca_envfit <- envfit(mi_fam_hel_pca, env_suelo, scaling = 2))
plot(mi_fam_hel_pca_envfit, p.max = 0.05 , col = 3)

env_num <- bci_env_grid %>%
  select_if(is.numeric) %>%
  select(-id) %>%
  st_drop_geometry
(mi_fam_hel_pca_envfit_num <- envfit(mi_fam_hel_pca, env_num, scaling = 2))
biplot(
  mi_fam_hel_pca,
  main = "PCA, escalamiento 2, ajuste a variables ambientales")
plot(mi_fam_hel_pca_envfit_num, p.max = 0.05 , col = 3)
biplot(
  mi_fam_hel_pca,
  main = "PCA, escalamiento 2, ajuste a variables ambientales")
plot(mi_fam_hel_pca_envfit_num, p.max = 0.1 , col = 3)

#' ### Análisis de correspondencia (CA)
#'

```

```

mi_fam_ca <- cca(mi_fam)
summary(mi_fam_ca)
summary(mi_fam_ca, scaling = 1)
#
#'
#' Screeplot
#'
screeplot(mi_fam_ca, bstick = TRUE, n pcs = length(mi_fam_ca$CA$eig))
#
#'
#' Biplots
#'
par(mfrow = c(1, 2))
plot(mi_fam_ca,
      scaling = 1,
      main = "Análisis de correspondencia, escalamiento 1"
)
plot(mi_fam_ca,
      scaling = 2, # Por defecto scaling=2, lo escribo sólo para fines didácticos
      main = "Análisis de correspondencia, escalamiento 2")
par(mfrow = c(1, 1))
#
#'
#' Excluyendo especie *Thevetia ahouai*, abreviada como *Thevahou*.
#'
mi_fam_ca <- cca(mi_fam[, -grep('Thevahou', colnames(mi_fam))])
summary(mi_fam_ca)
summary(mi_fam_ca, scaling = 1)
screeplot(mi_fam_ca, bstick = TRUE, n pcs = length(mi_fam_ca$CA$eig))
par(mfrow = c(1, 2))
plot(mi_fam_ca,
      scaling = 1,
      main = "CA, escalamiento 1, sin Thevetia ahouai"
)
plot(mi_fam_ca,
      scaling = 2,
      main = "CA, escalamiento 2, sin Thevetia ahouai")
par(mfrow = c(1, 1))
#
#'
#' Análisis de coordenadas principales (PCoA)

mi_fam_d_bray <- vegdist(mi_fam, method = 'bray') # En realidad, 'bray' es la opción por defecto
mi_fam_d_bray_pcoa <- cmdscale(
  mi_fam_d_bray,
  k = (nrow(mi_fam) - 1),
  add = T,
  eig = TRUE)
round(mi_fam_d_bray_pcoa$eig, 2)
round(sum(mi_fam_d_bray_pcoa$eig[mi_fam_d_bray_pcoa$eig<0]),2)
round(sum(mi_fam_d_bray_pcoa$eig[mi_fam_d_bray_pcoa$eig>=0]),2)

```

```

ordiplot(scores(mi_fam_d_bray_pcoa, choices = c(1, 2)),
         type = "t",
         main = "PCoA con promedios ponderados de especies")
abline(h = 0, lty = 3)
abline(v = 0, lty = 3)
mi_fam_d_bray_pcoa_wa <- wascores(mi_fam_d_bray_pcoa$points[, 1:2], mi_fam)
text(
  mi_fam_d_bray_pcoa_wa,
  rownames(mi_fam_d_bray_pcoa_wa),
  cex = 0.7, col = "red")
(mi_fam_d_bray_pcoa_env <- envfit(mi_fam_d_bray_pcoa, env_num))
plot(mi_fam_d_bray_pcoa_env, p.max = 0.05, col = 3)

```

```

# Técnica de ordenación 2

library(vegan)
library(tidyverse)
library(sf)
source('biodata/funciones.R')
#
# #### Cargar datos
#
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_malvc
(colnames(mi_fam) <- make.cepnames(colnames(mi_fam)))
(df_equivalencias <- data.frame(
  nombre_original = colnames(mc_malvc),
  colnames(mi_fam)))
bci_env_grid %>% tibble
#
# ## Ordenación restringida

# #### Análisis de redundancia (RDA)

# Matriz ambiental de variables suelo:
#
mi_fam_hel <- decostand(mi_fam, method = 'hellinger')
mi_fam_hel %>% tibble
env_suelo <- bci_env_grid %>%
  st_drop_geometry %>%
  dplyr::select(matches('^[A-T,Z] | ^pH$', ignore.case = F))
env_suelo %>% tibble
mi_fam_hel_rda_suelo <- rda(mi_fam_hel ~ ., env_suelo)
summary(mi_fam_hel_rda_suelo)

RsquareAdj(mi_fam_hel_rda_suelo)$adj.r.squared

```

```

vif.cca(mi_fam_hel_rda_suelo)

# Representación del modelo se realiza en un *tripplot*, que es un gráfico enriquecido, puesto qu
#
# Escalamiento 1:
#
plot(mi_fam_hel_rda_suelo,
      scaling = 1,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. suelo, escalamiento 1"
)
mi_fam_hel_rda_suelo_sc1 <-
  scores(mi_fam_hel_rda_suelo,
         choices = 1:2,
         scaling = 1,
         display = "sp"
  )
arrows(0, 0,
       mi_fam_hel_rda_suelo_sc1[, 1] * 0.9,
       mi_fam_hel_rda_suelo_sc1[, 2] * 0.9,
       length = 0,
       lty = 1,
       col = "red"
)
# Escalamiento 2
#
plot(mi_fam_hel_rda_suelo,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. suelo, escalamiento 2"
)
mi_fam_hel_rda_suelo_sc2 <-
  scores(mi_fam_hel_rda_suelo,
         scaling = 2,
         choices = 1:2,
         display = "sp"
  )
arrows(0, 0,
       mi_fam_hel_rda_suelo_sc2[, 1] * 0.9,
       mi_fam_hel_rda_suelo_sc2[, 2] * 0.9,
       length = 0,
       lty = 1,
       col = "red"
)
# Matriz ambiental con las variables que resultaron significativas en el ajuste *post-hoc* (pas
env_selec <- bci_env_grid %>%

```

```

select(
  heterogeneidad_ambiental,
  riqueza_global,
  UTM.EW,
  Al, B, Ca, Cu, Fe, K, Mg, Mn, P, Zn, N, N.min., pH) %>%
  st_drop_geometry
mi_fam_hel_rda_selec <- rda(mi_fam_hel ~ ., env_selec)
#'
vif.cca(mi_fam_hel_rda_selec)
# Comprobación por gráfico de asociación entre variables sin las flechas de especies:
#'
plot(mi_fam_hel_rda_selec,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec, escalamiento 2"
)

env_selec2 <- bci_env_grid %>%
  select(
    heterogeneidad_ambiental,
    riqueza_global,
    UTM.EW,
    Al, B, Cu, Fe, Mg, Mn, P, Zn, N, N.min., pH) %>%
  st_drop_geometry
mi_fam_hel_rda_selec2 <- rda(mi_fam_hel ~ ., env_selec2)
vif.cca(mi_fam_hel_rda_selec2)
plot(mi_fam_hel_rda_selec2,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec2, escalamiento 2"
)

env_selec3 <- bci_env_grid %>%
  select(
    heterogeneidad_ambiental,
    riqueza_global,
    UTM.EW,
    Al, Cu, Fe, Mn, P, Zn, N, N.min., pH) %>%
  st_drop_geometry
mi_fam_hel_rda_selec3 <- rda(mi_fam_hel ~ ., env_selec3)
vif.cca(mi_fam_hel_rda_selec3)
plot(mi_fam_hel_rda_selec3,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec3, escalamiento 2"
)
#'

```

```

#' Exclusión de la coordenada 'UTM.EW' para mejorar los VIF de las demás variables, como por ejem
#
env_selec4 <- bci_env_grid %>%
  select(
    heterogeneidad_ambiental,
    riqueza_global,
    Al, Cu, Fe, Mn, P, Zn, N, N.min., pH) %>%
  st_drop_geometry
mi_fam_hel_rda_selec4 <- rda(mi_fam_hel ~ ., env_selec4)
vif.cca(mi_fam_hel_rda_selec4)
plot(mi_fam_hel_rda_selec4,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec4, escalamiento 2"
)

summary(mi_fam_hel_rda_selec4)
RsquareAdj(mi_fam_hel_rda_selec4)$adj.r.squared
#
#'
#' Triplot, pero ahora con las flechas para las especies.
#'
#'
#' Escalamiento 1:
#'
plot(mi_fam_hel_rda_selec4,
      scaling = 1,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec4, escalamiento 1"
)
mi_fam_hel_rda_selec4_sc1 <-
  scores(mi_fam_hel_rda_selec4,
         choices = 1:2,
         scaling = 1,
         display = "sp"
  )
arrows(0, 0,
       mi_fam_hel_rda_selec4_sc1[, 1] * 0.9,
       mi_fam_hel_rda_selec4_sc1[, 2] * 0.9,
       length = 0,
       lty = 1,
       col = "red"
)
#
#'
#' Escalamiento 2
#'
plot(mi_fam_hel_rda_selec4,
      scaling = 2,
      display = c("sp", "lc", "cn"),

```

```

    main = "Triplot de RDA especies ~ var. selec4, escalamiento 2"
)
mi_fam_hel_rda_selec4_sc2 <-
  scores(mi_fam_hel_rda_selec4,
         scaling = 2,
         choices = 1:2,
         display = "sp"
  )
arrows(0, 0,
       mi_fam_hel_rda_selec4_sc2[, 1] * 0.9,
       mi_fam_hel_rda_selec4_sc2[, 2] * 0.9,
       length = 0,
       lty = 1,
       col = "red"
)
#'
#' #### Análisis de correspondencia canónica (CCA)
#'
#' ##### Ejemplo usando las matriz ambiental con variables seleccionadas:
#'
mi_fam_cca_selec4 <- cca(mi_fam ~ ., env_selec4)
summary(mi_fam_cca_selec4)
RsquareAdj(mi_fam_cca_selec4)
#'
#' Escalamiento 1
#'
plot(mi_fam_cca_selec4,
      scaling = 1,
      display = c("sp", "lc", "cn"),
      main = "Triplot de CCA especies ~ var. selec4, escalamiento 1"
)
#'
#' Escalamiento 2
#'
plot(mi_fam_cca_selec4,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de CCA especies ~ var. selec4, escalamiento 2"
)
#'
#' Excluyendo especies con abundancia menor a 100 individuos
#'
colSums(mi_fam) %>% sort
mi_fam_no_raras <- mi_fam %>% select_if(colSums(.) > 100)
intersect(colnames(mi_fam), colnames(mi_fam_no_raras))
setdiff(colnames(mi_fam), colnames(mi_fam_no_raras))
mi_fam_no_raras_cca_selec4 <- cca(mi_fam_no_raras ~ ., env_selec4)

```

```

summary(mi_fam_no_raras_cca_selec4)
RsquareAdj(mi_fam_no_raras_cca_selec4)
#'

#' Variables ambientales
library(mapview)
library(tidyverse)
library(sf)
library(RColorBrewer)

#' ### Cargar datos
load("~/unidad-0- asignacion-99-mi-manuscrito-anavalera29/biodata/Meliaceae.Rdata")

#' ### Paletas
azul <- colorRampPalette(brewer.pal(8, "Blues"))
rojo <- colorRampPalette(brewer.pal(8, "Reds"))
rojo_inv <- colorRampPalette(rev(brewer.pal(8, "Reds")))

#' ### Mapa de cuadros, simbología por pendiente
mapa_cuadros_pendiente <- mapView(
  bci_env_grid,
  layer.name = 'pendiente',
  alpha.regions = 0.4,
  map.types = 'OpenTopoMap',
  legend = T, zoom = 14,
  col.regions = rojo,
  zcol = 'pendiente_media') %>%
  addStaticLabels(label = round(bci_env_grid$pendiente_media, 1)) %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)
mapa_cuadros_pendiente
mapa_cuadros_pendiente %>% mapshot(file = 'mapa_cuadros_pendiente.png') #Genera archivo

#' ### Mapa de cuadros, simbología por Nitrógeno
mapa_cuadros_nit <- mapView(
  bci_env_grid,
  layer.name = 'N (mg/kg)',
  alpha.regions = 0.4,
  map.types = 'OpenTopoMap',
  legend = T, zoom = 14,
  col.regions = rojo,
  zcol = 'N') %>%
  addStaticLabels(label = round(bci_env_grid$N, 1)) %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)
mapa_cuadros_nit

```

```

    lat = 9.15097,
    zoom = 15)
mapa_cuadros_nit
mapa_cuadros_nit %>% mapshot(file = 'mapa_cuadros_nit.png')

#' #### Mapa de cuadros, simbología por pH
mapa_cuadros_ph <- mapView(
  bci_env_grid,
  layer.name = 'pH',
  alpha.regions = 0.4,
  map.types = 'OpenTopoMap',
  legend = T, zoom = 14,
  col.regions = rojo_inv,
  zcol = 'pH') %>%
  addStaticLabels(label = round(bci_env_grid$pH, 1)) %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)
mapa_cuadros_ph
mapa_cuadros_ph %>% mapshot(file = 'mapa_cuadros_ph.png')

```

```

#' Variables ambientales por Lote

library(tmap)
library(sf)
library(tidyverse)
library(RColorBrewer)
#'
#' #### Cargar datos
#'
load('biodata/matriz_ambiental.Rdata')
#'
#' ## Convertir a KML
#'
st_write(
  bci_env_grid %>% rename(Name = id),
  driver = 'KML',
  dsn = 'matriz_ambiental.kml')
st_write(
  bci_env_grid %>% rename(Name = id) %>% st_centroid(),
  driver = 'KML',
  dsn = 'matriz_ambiental_puntos.kml')

#'
#' ## Generar mapas por lotes
#'

```

```

#' ### Variables ambientales numéricas con 'ggplot2'
#
mapas_var_amb_num_gg <- bci_env_grid %>%
  select_if(is.numeric) %>%
  gather(variable, valor, -geometry) %>%
  group_by(variable) %>%
  mutate(
    valor = scales::rescale(valor, to = c(0, 1)),
    id = rep(1:50)) %>%
  ggplot +
  aes(geometry = geometry, fill = valor) +
  theme(axis.text = element_blank()) +
  geom_sf(lwd = 0.1, color = 'grey50', alpha = 0.8) + coord_sf() +
  scale_fill_gradientn(colours = brewer.pal(11, 'BrBG')) +
  geom_sf_text(aes(label = id, color = between(valor, 0.3, 0.7)), size = 1.75) +
  scale_color_manual(guide = FALSE, values = c("white", "black")) +
  facet_wrap(~ variable, ncol = 6) +
  ggtitle('Cuadros de 1 Ha de BCI. Variables ambientales numéricas escaladas de 0 a 1')
mapas_var_amb_num_gg
#
#' PNG
#
png(
  filename = 'mapas_variables_ambientales_numericas.png',
  width = 1700, height = 1080, res = 150)
mapas_var_amb_num_gg
dev.off()
#
#' ### Variables ambientales numéricas con 'tmap'
#
mapas_var_amb_num_tmap <- bci_env_grid %>%
  select_if(is.numeric) %>%
  gather(variable, valor, -geometry) %>%
  group_by(variable) %>%
  mutate(
    valor = scales::rescale(valor, to = c(0, 1)),
    id = rep(1:50)) %>%
  tm_shape() +
  tm_polygons(col = 'valor',
              palette = brewer.pal(11, 'BrBG'),
              style ='cont',
              legend.is.portrait = FALSE) +
  tm_facets(by = 'variable', ncol = 6, nrow = 6) +
  tm_layout(main.title="Cuadros de 1 Ha de BCI. Variables ambientales numéricas escaladas de 0 a 1",
            main.title.size = 0.7,
            legend.outside.position="bottom",
            legend.outside=TRUE,

```

```

    legend.width = 0.2,
    legend.text.size = 0.5,
    legend.stack="horizontal",
    outer.margins=0)
mapas_var_amb_num_tmap
#'
#' PNG
#'
png(
  filename = 'mapas_variables_ambientales_numericas_tmap.png',
  width = 1800, height = 1400, res = 350, pointsize = 12)
mapas_var_amb_num_tmap
dev.off()
#'
#' #### Variables ambientales nominales con 'tmap'
#'
mapas_var_amb_nom_tmap <- bci_env_grid %>%
  select_if(negate(is.numeric)) %>%
  gather(variable, valor, -geometry) %>%
  tm_shape() +
  tm_polygons(col = 'valor',
              palette = brewer.pal(8, 'Set1'),
              legend.show = T) +
  tm_facets(by = 'variable', ncol = 2, free.scales = T, free.coords = T) +
  tm_layout(main.title="Cuadros de 1 Ha de BCI. Variables ambientales nominales",
            main.title.size = 0.7,
            asp = 3.5,
            legend.text.size = 0.7)
mapas_var_amb_nom_tmap
#'
#' PNG
#'
png(
  filename = 'mapas_variables_ambientales_nominales_tmap.png',
  width = 2000, height = 1200, res = 350, pointsize = 12)
mapas_var_amb_nom_tmap
dev.off()

```

```

#' #### Área de cargar paquetes
library(vegan)
library(tidyverse)
library(sf)
source('biodata/funciones.R')

#' #### Área de cargar datos
#' Censo (el objeto se carga con prefijo "censo") y matriz de comunidad (prefijo "mc")
load('~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata')

```

```

load('biodata/matriz_ambiental.Rdata') #Matriz ambiental, se carga como "bci_env_grid"

#' #### Imprimir datos en pantalla (impresiones parciales con head)
head(censo_malvc)
head(mc_malvc)
bci_env_grid # No necesita imprimirse parcialmente

# Requiere que se haya cargado ya la colección tidyverse
censo_malvc %>% tibble
mc_malvc %>% tibble

# #### Lista de especies
sort(colnames(mc_malvc))

# #### Número de sitios, tanto en matriz de comunidad como en ambiental
# Verifica que coinciden
nrow(mc_malvc) #En la matriz de comunidad
nrow(bci_env_grid) #En la matriz ambiental

# #### Riqueza numérica de especies (usando matriz de comunidad) por quadrat
# Nota: cargar paquete vegan arriba, en el área de paquetes
specnumber(mc_malvc)
sort(specnumber(mc_malvc)) # Ordenados ascendentemente
summary(specnumber(mc_malvc)) # Resumen estadístico

# #### Abundancia de especies por quadrat
sort(rowSums(mc_malvc))
summary(rowSums(mc_malvc)) # Resumen estadístico

# #### Abundancia por especie
sort(colSums(mc_malvc))
summary(colSums(mc_malvc)) # Resumen estadístico

# #### Riqueza numérica de toda la "comunidad"
specnumber(colSums(mc_malvc))

# #### Abundancia de toda la comunidad
sum(colSums(mc_malvc))

# #### Una tabla
# Para esto, usaré la colección "tidyverse"
abun_sp <- censo_malvc %>%
  group_by(Latin) %>%
  count() %>%
  arrange(desc(n))
abun_sp

```

```

#' Gráfico de mosaicos de la abundancia por especie por cuadros
abun_sp_q <- crear_grafico_mosaico_de_mc(mc_malvc, tam_rotulo = 6)
abun_sp_q

#' Correlación ambiental

library(tidyverse)
library(sf)
library(ez)
library(psych)
library(vegan)

#' ### Cargar datos
load('biodata/matriz_ambiental.Rdata')
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")

#' ### Una correlación simple
cor(bci_env_grid$pendiente_media, bci_env_grid$geomorf_vertiente_pct)
plot(bci_env_grid$pendiente_media, bci_env_grid$geomorf_vertiente_pct)
cor.test(bci_env_grid$pendiente_media, bci_env_grid$geomorf_vertiente_pct)

#' ### Generar objeto de columnas numéricas

env_num <- bci_env_grid %>%
  dplyr::select_if(is.numeric) %>%
  dplyr::select(-id, -matches('^U.*')) %>%
  st_drop_geometry %>%
  mutate(
    riqueza_mifam = specnumber(mc_malvc),
    abundancia_mifam = rowSums(mc_malvc)) %>%
  rename_all(gsub, pattern = '_pct$', replacement = ',') %>%
  rename_all(gsub, pattern = '_| ', replacement = '\n')
env_num %>% tibble

#' ### Panel de correlaciones con herramientas del paquete 'graphics' y 'psych'
cor(env_num)
ncol(env_num)
pairs(env_num[,sample(1:33, 15)]) # paquete graphics
env_num[,sample(1:33, 15)] %>% pairs.panels #paquete psych

#' ### Panel de correlaciones con 'ez'

#' ##### Sólo suelo (elementos y pH), abundancia/riqueza
p_cor_suelo_ar <- env_num %>%
  dplyr::select(matches('^[A-T,Z] | abundancia|riqueza|~pH$', ignore.case = F)) %>%
  ezCor(r_size_lims = c(4,8), label_size = 3)

```

```

p_cor_suelo_ar

#' ##### Sólo heterogeneidad, geomorfología, abundancia/riqueza
p_cor_geomorf_ar <- env_num %>%
  dplyr::select(-matches('^[A-T,Z]|pH', ignore.case = F)) %>%
  ezCor(r_size_lims = c(4,8), label_size = 3)
p_cor_geomorf_ar

#' ##### Matriz de comunidad
p_cor_mc <- mc_malvc%>%
  rename_all(gsub, pattern = '_| ', replacement = '\n') %>%
  ezCor(r_size_lims = c(4,8), label_size = 3)
p_cor_mc

```

```

#'Análisis de agrupamiento 1

library(vegan)
library(magrittr)
library(broom)
source('biodata/funciones.R')

#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
mi_fam <- mc_malvc
#'
#' ## Características de las técnicas de agrupamiento

#' ## Agrupamiento jerárquico

#'Método de normalización y distancia euclídea (distancia de cuerdas o *chord*).
#'
mi_fam_norm <- decostand(mi_fam, "normalize")
mi_fam_norm_d <- vegdist(mi_fam_norm, "euc")
mi_fam_norm_d %>% tidy
#'
#' Es importante, para garantizar consistencia a lo largo del agrupamiento, asignar los nombres
#'
attr(mi_fam_norm_d, "labels") <- rownames(mi_fam)
#'
#' Posteriormente, el agrupamiento jerárquico lo realizaré con la función 'hclust' del paquete '
#'
(cl_single <- hclust(mi_fam_norm_d, method = 'single'))
#'
#' Dendrograma:
plot(cl_single, labels = rownames(mi_fam), hang = -1,
      main = "Sitios de BCI según composición de especies de Malvaceae\nEnlace simple a partir de"

```

```

    xlab = 'Sitios', ylab = 'Altura')
#
#' #### Agrupamiento "aglomerativo" por enlace completo

(cl_complete <- hclust(mi_fam_norm_d, method = 'complete'))
plot(cl_complete, labels = rownames(mi_fam), hang = -1,
      main = "Sitios de BCI según composición de especies de Malvaceae\nEnlace completo a partir
      xlab = 'Sitios', ylab = 'Altura')
#
#' #### Agrupamiento "aglomerativo" por enlace promedio

#'Dendrograma del método UPGMA.
#
(cl_upgma <- hclust(mi_fam_norm_d, method = 'average'))
plot(cl_upgma, labels = rownames(mi_fam), hang = -1,
      main = "Sitios de BCI según composición de especies de Malvaceae\nUPGMA a partir de matriz
      xlab = 'Sitios', ylab = 'Altura')
#
#' #### Agrupamiento por el método de Ward de varianza mínima

(cl_ward <- hclust(mi_fam_norm_d, method = 'ward.D2'))
plot(cl_ward, labels = rownames(mi_fam), hang = -1,
      main = "Sitios de BCI según composición de especies de Malvaceae\nMétodo de Ward a partir de
      xlab = 'Sitios', ylab = 'Altura')

```

*#'Análisis de agrupamiento 2*

```

library(vegan)
library(tidyverse)
library(broom)
library(cluster)
library(gclus)
library(pvclust)
library(sf)
source('biodata/funciones.R')
#
#' #### Cargar datos
#
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
mi_fam <- mc_malvc
load('biodata/matriz_ambiental.Rdata')
mi_fam %>% tibble
bci_env_grid %>% tibble
#
#' #### Generar matriz de distancias de cuerdas
#
mi_fam_norm <- decostand(mi_fam, "normalize")

```

```

mi_fam_norm_d <- vegdist(mi_fam_norm, "euc")
mi_fam_norm_d %>% tidy
#'
#' ## Interpretación visual de dendrogramas

#' Para la exploración visual, objetos de cluster dentro de una lista:
#'
lista_cl <- list(
  cl_single = hclust(mi_fam_norm_d, method = 'single'),
  cl_complete = hclust(mi_fam_norm_d, method = 'complete'),
  cl_upgma = hclust(mi_fam_norm_d, method = 'average'),
  cl_ward = hclust(mi_fam_norm_d, method = 'ward.D2')
)
#'
#' Un plot en panel 2x2 ayuda a visualizarlos todos de manera conjunta. En tu caso, observa y co
#'
par(mfrow = c(2,2))
invisible(map(names(lista_cl), function(x) plot(lista_cl[[x]], main = x, hang = -1)))
par(mfrow = c(1,1))

#' Usando la lista de objetos de clústers, cálculo de correlación cofenética dentro de un 'map'
#'
map_df(lista_cl, function(x) {
  coph_d <- cophenetic(x)
  corr <- cor(mi_fam_norm_d, coph_d)
  return(corr)
})

#'cálculos para UPGMA y Ward.
#'
#' Para UPGMA:
#'
anch_sil_upgma <- calcular_anchuras_siluetas(
  mc_orig = mi_fam,
  distancias = mi_fam_norm_d,
  cluster = lista_cl$cl_upgma)
anch_sil_upgma

u_dend_reord <- reorder.hclust(lista_cl$cl_upgma, mi_fam_norm_d)
plot(u_dend_reord, hang = -1)
rect.hclust(
  tree = u_dend_reord,
  k = anch_sil_upgma$n_grupos_optimo)
#'
#' Ahora compararé el dendrograma con el mapa de calor en un mismo gráfico, colocando los dendro
#'
heatmap(

```

```

    as.matrix(mi_fam_norm_d),
    Rowv = as.dendrogram(u_dend_reord),
    symm = TRUE,
    margin = c(3, 3),
    col = rev(cm.colors(4))
)

# Mostraré el resultado para Ward:
#
anch_sil_ward <- calcular_anchuras_siluetas(
  mc_orig = mi_fam,
  distancias = mi_fam_norm_d,
  cluster = lista_cl$cl_ward)
anch_sil_ward
#
w_dend_reord <- reorder.hclust(lista_cl$cl_ward, mi_fam_norm_d)
plot(w_dend_reord, hang = -1)
rect.hclust(
  tree = w_dend_reord,
  k = anch_sil_ward$n_grupos_optimo)
plot(w_dend_reord, hang = -1)
rect.hclust(
  tree = w_dend_reord,
  k = anch_sil_ward$n_grupos_optimo + 1)
#
# Comparando el dendrograma con el mapa de calor. Verificar si el número de grupos hace sentido
#
heatmap(
  as.matrix(mi_fam_norm_d),
  Rowv = as.dendrogram(w_dend_reord),
  symm = TRUE,
  margin = c(3, 3),
  col = rev(cm.colors(4)))
)

# ##### UPGMA
#
cl_pvclust_upgma <-
  pvclust(t(mi_fam_norm),
    method.hclust = "average",
    method.dist = "euc",
    iseed = 91, # Resultado reproducible
    parallel = TRUE)
# Añadir los valores de p
plot(cl_pvclust_upgma, hang = -1)
# Añadir rectángulos a los grupos significativos
lines(cl_pvclust_upgma)

```

```

pvrect(cl_pvclust_upgma, alpha = 0.91, border = 4)
#'
#' ##### Ward
#'
cl_pvclust_ward <-
  pvclust(t(mi_fam_norm),
           method.hclust = "ward.D2",
           method.dist = "euc",
           iseed = 191, # Resultado reproducible
           parallel = TRUE)
# Añadir los valores de p
plot(cl_pvclust_ward, hang = -1)
# Añadir rectángulos a los grupos significativos
lines(cl_pvclust_ward)
pvrect(cl_pvclust_ward, alpha = 0.91, border = 4)
#'
#' #### Recapitulando los grupos de sitios.
#'
#' ##### Patrones comunes y dispares

#' UPGMA:
(grupos_upgma_k2 <- as.factor(cutree(lista_cl$cl_upgma, k = 2)))

table(grupos_upgma_k2)

#' Ward:
#'
(grupos_ward_k2 <- as.factor(cutree(lista_cl$cl_ward, k = 2)))
table(grupos_ward_k2)
#'
#' Guardaré estos vectores en archivos para reutilizarlos en *scripts* posteriores:
#'
saveRDS(grupos_upgma_k2, 'grupos_upgma_k2.RDS')
saveRDS(grupos_ward_k2, 'grupos_ward_k2.RDS')

```

```

#'Análisis de agrupamiento 3

library(mapview)
library(tidyverse)
library(sf)
library(RColorBrewer)
source('biodata/funciones.R')
#'
#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')

```

```

grupos_ward_k2 <- readRDS('grupos_ward_k2.RDS')
table(grupos_ward_k2)
#'
#' #### Paletas
#'
rojo <- colorRampPalette(brewer.pal(8, "Reds"))
rojo_inv <- colorRampPalette(rev(brewer.pal(8, "Reds")))
colores_grupos <- brewer.pal(8, "Accent")
#'
#' ## Explorar efectos

#' Creación de un objeto que permita realizar tanto las pruebas como los diagramas de cajas.
#'
(m_amb_ward_k2 <- bci_env_grid %>%
  select_if(is.numeric) %>% select(-id) %>%
  mutate(grupos_ward_k2) %>%
  st_drop_geometry() %>%
  pivot_longer(-grupos_ward_k2, names_to = "variable", values_to = "valor"))
#'
#' A continuación, las pruebas:
#'
m_amb_ward_k2 %>%
  group_by(variable) %>%
  summarise(
    p_valor_t = t.test(valor ~ grupos_ward_k2)$p.value,
    p_valor_w = wilcox.test(valor ~ grupos_ward_k2, exact = F)$p.value) %>%
  arrange(p_valor_t) %>%
  rename(`Valor de p, prueba t` = p_valor_t, `Valor de p, prueba Wilcoxon` = p_valor_w) %>%
  ungroup() %>% slice(1:8) %>% knitr::kable(digits = 2)

#' Los gráficos:
#'
m_amb_ward_k2 %>%
  group_by(variable) %>%
  ggplot() + aes(x = grupos_ward_k2, y = valor, fill = grupos_ward_k2) +
  geom_boxplot() +
  scale_fill_brewer(palette = 'Accent') +
  theme_bw() +
  theme(legend.position="none") +
  facet_wrap(~ variable, scales = 'free_y')
#'
#' Mapas:
#'
mapa_ward_k2 <- mapView(
  bci_env_grid %>% mutate(grupos_ward_k2),
  layer.name = 'Grupos (2) ward_k2',
  alpha.regions = 0.6,

```

```

map.types = 'OpenTopoMap',
legend = T,
col.regions = colores_grupos[1:2],
zcol = 'grupos_ward_k2') %>%
addStaticLabels(label = bci_env_grid$id) %>%
leaflet::setView(
  lng = -79.85136,
  lat = 9.15097,
  zoom = 15)
mapa_ward_k2
mapa_ward_k2 %>% mapshot(
  file = 'mapa_ward_k2.png',
  remove_controls = c("zoomControl", "layersControl", "homeButton")
)
#
#'
#' Mapa de una de las variables donde se presentó efecto de su promedio ( $p<0.01$ ), en este caso,
#'
mapa_zn <- mapView(
  bci_env_grid,
  layer.name = 'Zinc',
  alpha.regions = 0.6,
  map.types = 'OpenTopoMap',
  legend = T,
  col.regions = rojo,
  zcol = 'Zn') %>%
addStaticLabels(label = bci_env_grid$id) %>%
leaflet::setView(
  lng = -79.85136,
  lat = 9.15097,
  zoom = 15)
mapa_zn
mapa_zn %>% mapshot(
  file = 'mapa_zinc.png',
  remove_controls = c("zoomControl", "layersControl", "homeButton")
)
#
#'
#' Pruebas de igualdad de promedios de las variables entre 3 grupos o más
#'
#' Objeto común:
#'
(m_amb_ward_k2 <- bci_env_grid %>%
  select_if(is.numeric) %>% select(-id) %>%
  mutate(grupos_ward_k2) %>%
  st_drop_geometry() %>%
  pivot_longer(-grupos_ward_k2, names_to = "variable", values_to = "valor"))

```

```

#'Análisis de agrupamiento 4

library(indicspecies)
source('biodata/funciones.R')

#'
#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
mi_fam <- mc_malvc
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k2 <- readRDS('grupos_ward_k2.RDS')
table(grupos_ward_k2)
#'
#' ## Análisis de especies indicadoras mediante IndVal
#'
#' #### UPGMA
#'
iva_upgma_k2 <- multipatt(
  x = mi_fam,
  cluster = grupos_upgma_k2,
  func = 'IndVal.g',
  max.order = 1,
  control = how(nperm = 999))
summary(iva_upgma_k2, indvalcomp = TRUE)
colSums(mi_fam)
(p_upgma_adj <- p.adjust(iva_upgma_k2$sign$p.value))
(iva_upgma_boot <- strassoc(
  X = mi_fam,
  cluster = grupos_upgma_k2,
  func = "IndVal.g",
  nboot = 1000))
#'
#' Ward
#'
iva_ward_k2 <- multipatt(
  x = mi_fam,
  cluster = grupos_ward_k2,
  func = 'IndVal.g',
  max.order = 2,
  control = how(nperm = 999))
summary(iva_ward_k2, indvalcomp = TRUE)
colSums(mi_fam)
(p_ward_adj <- p.adjust(iva_ward_k2$sign$p.value))
(iva_ward_boot <- strassoc(
  X = mi_fam,

```

```

cluster = grupos_ward_k2,
func = "IndVal.g",
nboot = 1000))
#'
## ## Análisis de especies con preferencia por hábitat mediante el coeficiente de correlación bi
#'
### UPGMA
#'
phi_upgma_k2 <- multipatt(
  mi_fam,
  grupos_upgma_k2,
  func = "r.g",
  max.order = 1,
  control = how(nperm = 999))
summary(phi_upgma_k2)
colSums(mi_fam)
(phi_upgma_boot <- strassoc(
  X = mi_fam,
  cluster = grupos_upgma_k2,
  func = "r.g",
  nboot = 1000))
#'
#' Ward
#'
phi_ward_k2 <- multipatt(
  mi_fam,
  grupos_ward_k2,
  func = "r.g",
  max.order = 2,
  control = how(nperm = 999))
summary(phi_ward_k2)
colSums(mi_fam)
(phi_ward_boot <- strassoc(
  X = mi_fam,
  cluster = grupos_ward_k2,
  func = "r.g",
  nboot = 1000))

```

```

#Análisis de la diversidad de especie alpha

#'
library(vegan)
library(adespatial)
library(plyr)
library(RColorBrewer)
library(tidyverse)
library(sf)

```

```

library(SpadeR)
library(iNEXT)
source('biodata/funciones.R')
#'
#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_malvc
bci_env_grid %>% tibble
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k2 <- readRDS('grupos_ward_k2.RDS')
table(grupos_ward_k2)
#'
#' ## Diversidad alpha

#' - Cinco especies, 55 individuos, abundancias desiguales:
#'
foo1 <- c(25, 16, 9, 4, 1)
diversity(foo1)
#'
#' - Cinco especies, 55 individuos, abundancias homogéneas
#'
foo2 <- c(11, 11, 11, 11, 11)
diversity(foo2)
log(5)

#' **Índices, entropías, equidades, ratios**
#'
(indices <- alpha_div(mi_fam))
pairs(indices,
      lower.panel = panel.smooth,
      upper.panel = panel.cor,
      diag.panel = panel.hist,
      main = "Pearson Correlation Matrix")
indices_env <- bind_cols(
  indices,
  bci_env_grid %>%
    select_if(is.numeric) %>%
    st_drop_geometry %>%
    select(-id) %>%
    select(-matches('geom.*pct$'))))
indices_env %>% tibble
ezCorM(indices_env, r_size_lims = c(3,5), label_size = 4)
#'
#' **Modelos de abundancia de especies**

```

```

#'
mi_fam_mae <- radfit(mi_fam)
plot(mi_fam_mae)
#'
#' **Rarefacción**
#'
#' Riqueza por sitios
#'
riqueza <- specnumber(mi_fam)
riqueza %>% sort
#'
#' Sitios con riqueza mínima y máxima
#'
riqueza[riqueza == min(riqueza)]
riqueza[riqueza == max(riqueza)]
range(riqueza)
#'
#' Abundancia por sitios
#'
abundancia <- rowSums(mi_fam)
abundancia %>% sort
#'
#' Sitios con abundancias mínima y máxima
#'
abundance[abundance == min(abundance)]
abundance[abundance == max(abundance)]
(rango_abun <- range(abundance))
#'
#' Abundancia en el sitio más pobre
#'
abundance[riqueza == min(riqueza)]
#'
#' Abundancia en el sitio más rico
#'
abundance[riqueza == max(riqueza)]
#'
#' Riqueza en el sitio con menor abundancia
#'
riqueza[abundance == min(abundance)]
#'
#' Riqueza en el sitio con mayor abundancia
#'
riqueza[abundance == max(abundance)]
#'
#' Rarefacción a la abundancia más pequeña encontrada
#'
riqueza_menor_abun <- rarefy(mi_fam, sample = rango_abun[1])

```

```

sort(riqueza)
sort(round(riqueza_menor_abun))
rarecurve(
  mi_fam,
  step = 1,
  sample = rango_abun[1],
  xlab = "Número de individuos (tamaño de muestra)",
  ylab = "Especies",
  label = TRUE,
  col = "blue"
)
#
#'
#' #### Riqueza de especies, estimación y comparación, "completitud de muestra" (existe en el diccionario)
#'
# Aproximación básica:
#'
specpool(mi_fam)
specpool(mi_fam)[[1]]/specpool(mi_fam)*100

# Generar la matriz de comunidad combinada, en la que todos los sitios forman uno.
#'
mi_fam_combinada <- colSums(mi_fam)
mi_fam_combinada %>% sort
mi_fam_combinada_chao <- estimacion_riqueza_chao(
  mc = mi_fam_combinada,
  n_raras = 10)
mi_fam_combinada_chao$asintoticos_estimacion
mi_fam_combinada_chao$no_asintoticos_rarefaccion_extrapolacion
mi_fam_combinada_chao$no_asintoticos_rarefaccion_extrapolacion_grafico

# #' Generar matriz de comunidad agrupada según el método de Ward (tres grupos), procedente de la matriz combinada
#'
mi_fam_k3 <- mi_fam %>%
  mutate(g=grupos_ward_k2) %>%
  group_by(g) %>%
  summarise_all(sum) %>%
  select(-g) %>%
  data.frame
mi_fam_k3 %>% rowSums %>% sort
mi_fam_k3_chao <- estimacion_riqueza_chao(
  mc = mi_fam_k3,
  n_raras = 10)
mi_fam_k3_chao$asintoticos_estimacion
mi_fam_k3_chao$no_asintoticos_rarefaccion_extrapolacion
mi_fam_k3_chao$no_asintoticos_rarefaccion_extrapolacion_grafico

```

```

#'Análisis de diversidad beta

library(vegan)
library(adespatial)
library(plyr)
library(tidyverse)
library(sf)
library(vegetarian)
library(mapview)
source('biodata/funciones.R')
#'
#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_malvc
bci_env_grid %>% tibble

#' #### Diversidad beta con un único número
#'
beta_multiplicativa <- calcular_beta_multiplicativa(
  mc = mi_fam,
  orden = 0:20)
beta_multiplicativa

#'
#' #### Contribución de especies a la diversidad beta (SCBD, *species contribution to beta diversity*)
#'
determinar_contrib_local_y_especie(
  mc = mi_fam,
  alpha = 0.05,
  nperm = 9999,
  metodo = 'hellinger')
#'
mapa_cuadros <- mapView(
  bci_env_grid,
  col.regions = 'grey80',
  alpha.regions = 0.3,
  map.types = 'OpenTopoMap',
  legend = F, zoom = 14,
  zcol = 'id') %>% addStaticLabels() %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)

contrib_local_especie_guardado <- determinar_contrib_local_y_especie(

```

```

mc = mi_fam,
alpha = 0.05,
nperm = 9999,
metodo = 'hellinger')
contrib_local_especie_guardado$especies_contribuyen_betadiv %>%
  bind_rows %>% gather(especie, valor) %>%
  mutate(valor = round(valor, 2)) %>%
  knitr::kable()

#'
#' Utilizar el mapa de cuadros para identificar aquellos que contribuyen más a la diversidad beta
#'
mapa_cuadros

```

```

#'Análisis de ecología espacial

library(ape)
library(spdep)
library(ade4)
library adegraphics
library adespatial
library vegan
library tidyverse
library sf
library gridExtra
library grid
library gtable
source('biodata/funciones.R')
source('https://raw.githubusercontent.com/maestria-geotel-master/unidad-3-asignacion-1-vecindad-')
#'
#' #### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-anavalera29/biodata/Malvaceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_malvc
mi_fam %>% tibble
bci_env_grid %>% tibble

#'
#' #### Generar matriz Hellinger
#'
mi_fam_hel <- decostand (mi_fam, "hellinger")
#'
#' #### Transformar matriz ambiental en objeto 'sp', generar vecindad
#'
bci_env_grid_sp <- bci_env_grid %>% as_Spatial
centroides <- bci_env_grid %>% st_centroid

```

```

bci_xy <- centroides %>% st_coordinates %>% as.data.frame
(vecindad <- bci_env_grid_sp %>% poly2nb)
(pesos_b <- nb2listw(vecindad, style = 'B'))
#+ fig.width=12, fig.height=6
plot(bci_env_grid_sp)
plot(vecindad, coords = bci_xy, add=T, col = 'red')
#'
#' ## Autocorrelación espacial mediante correlograma
#'
#' #### Una y solo una variable ambiental
#'
var_ph <- bci_env_grid %>% st_drop_geometry %>% pull(pH)
ph_correl <- sp.correlogram(vecindad,
                             var_ph,
                             order = 9,
                             method = "I",
                             zero.policy = TRUE)
print(ph_correl, digits = 2, p.adj.method = 'holm')
#+ fig.width=12, fig.height=6
plot(ph_correl)
#'
#' #### Múltiples variables
#'
#' ##### Abundancias de especies (matriz de comunidad transformada)
#'
suppressWarnings(auto_spp_hel <- calcular_autocorrelacion(
  df_fuente = mi_fam_hel,
  orden = 9,
  obj_vecindad = vecindad,
  pos_var = '(matriz Hellinger)'))
print(auto_spp_hel, digits = 2, p.adj.method = 'holm')
dim_panel <- rev(n2mfrow(ncol(mi_fam_hel)))
#+ fig.width=12, fig.height=10
par(mfrow = dim_panel)
suppressWarnings(invisible(lapply(auto_spp_hel, function(x) plot(x, main = x$var))))
#'
#' #### Variables ambientales (matriz ambiental)
#'
bci_env_grid_num <- bci_env_grid %>%
  st_drop_geometry %>%
  select_if(is.numeric) %>%
  select(-id, -UTM.EW, -UTM.NS)
suppressWarnings(auto_amb <- calcular_autocorrelacion(
  df_fuente = bci_env_grid_num,
  orden = 9,
  obj_vecindad = vecindad))
print(auto_amb, digits = 2, p.adj.method = 'holm')

```

```

dim_panel <- rev(n2mfrow(ncol(bci_env_grid_num)))
#+ fig.width=12, fig.height=14
par(mfrow = dim_panel)
suppressWarnings(invisible(lapply(auto_amb, function(x) plot(x, main = x$var))))
#'
#' ## Autocorrelación espacial mediante prueba Mantel (matrices de distancia)

mi_fam_sin_tendencia <- resid(
  lm(as.matrix(mi_fam_hel) ~ .,
     data = bci_xy))
mi_fam_sin_tendencia_d <- dist(mi_fam_sin_tendencia)
(mi_fam_correlograma <- mantel.correlog(
  mi_fam_sin_tendencia_d,
  XY = bci_xy,
  nperm = 999))
#+ fig.width=12, fig.height=6
plot(mi_fam_correlograma)
#'
#' ## Autocorrelación espacial por medio de pruebas de permutación para el I de Moran
#'
#' #### I de Moran global aplicado a abundancia de especies transformadas sin tendencia
#'
(autocor_global_residuos <- sapply(
  dimnames(mi_fam_sin_tendencia)[[2]],
  function(x)
    moran.mc(
      x = mi_fam_sin_tendencia[,x],
      listw = pesos_b,
      zero.policy = T,
      nsim = 9999),
    simplify = F))
#'
#' #### I de Moran local
#'
#' ##### Aplicado a variables ambientales
#'
bci_env_grid_num_sf <- bci_env_grid %>%
  select_if(is.numeric) %>%
  select(-id, -UTM.EW, -UTM.NS)
bci_env_grid_num_sf %>% tibble
lisamaps_amb <- sapply(grep('geometry', names(bci_env_grid_num_sf), invert = T, value = T),
                        function(x) {
                          m <- lisamap(objesp = bci_env_grid_num_sf[x],
                                       var = x,
                                       pesos = pesos_b,
                                       tituloleyenda = 'Significancia ("x-y", léase como "x" rodea',
                                       leyenda = F,

```

```

        anchuratitulo = 50,
        tamanotitulo = 10,
        fuentedatos = '\nhttp://ctfs.si.edu/webatlas/datasets/bci/'
        titulomap = paste0('Clusters LISA de "', x, '')})
    return(m$grafico)
}, simplify = F
)
lisamaps_amb$leyenda <- gtable_filter(ggplot_gtable(ggplot_build(lisamaps_amb[[1]] + theme(legend.position = 'none'))
grid.arrange(do.call('arrangeGrob', c(lisamaps_amb[1:12], nrow = 3)), lisamaps_amb$leyenda, height = 100),
grid.arrange(do.call('arrangeGrob', c(lisamaps_amb[13:22], nrow = 3)), lisamaps_amb$leyenda, height = 100),
grid.arrange(do.call('arrangeGrob', c(lisamaps_amb[23:31], nrow = 3)), lisamaps_amb$leyenda, height = 100))
#'
## ##### Aplicado a abundancias de especies transformadas
#
mi_fam_hel_sf <- bci_env_grid %>% select %>% bind_cols(mi_fam_hel)
lisamaps_mifam <- sapply(
  grep('geometry', names(mi_fam_hel_sf), invert = T, value = T),
  function(x) {
    m <- lisamap(objesp = mi_fam_hel_sf[x],
      var = x,
      pesos = pesos_b,
      tituloleyenda = 'Significancia ("x-y", léase como "x" rodeado de "y")',
      leyenda = F,
      anchuratitulo = 50,
      tamanotitulo = 10,
      fuentedatos = '\nhttp://ctfs.si.edu/webatlas/datasets/bci/',
      titulomap = paste0('Clusters LISA de "', x, ''))

    # dev.new();print(m$grafico)
    return(m$grafico)
  }, simplify = F
)
lisamaps_mifam$leyenda <- gtable_filter(ggplot_gtable(ggplot_build(lisamaps_mifam[[1]] + theme(legend.position = 'none'))
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam[1:8], nrow = 3)), lisamaps_mifam$leyenda, height = 100),
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam[9:16], nrow = 3)), lisamaps_mifam$leyenda, height = 100))
#'
## ##### Aplicado a abundancias de especies transformadas sin tendencia
#
mi_fam_sintendencia_sf <- bci_env_grid %>% select %>% bind_cols(mi_fam_sin_tendencia %>% as.data.frame())
lisamaps_mifam_sintendencia <- sapply(
  grep('geometry', names(mi_fam_sintendencia_sf), invert = T, value = T),
  function(x) {
    m <- lisamap(objesp = mi_fam_sintendencia_sf[x],
      var = x,
      pesos = pesos_b,
      tituloleyenda = 'Significancia ("x-y", léase como "x" rodeado de "y")',
      leyenda = F,
      anchuratitulo = 50,
      tamanotitulo = 10,
      fuentedatos = '\nhttp://ctfs.si.edu/webatlas/datasets/bci/',
      titulomap = paste0('Clusters LISA de "', x, ''))

    # dev.new();print(m$grafico)
    return(m$grafico)
  }, simplify = F
)

```

```
    tamanotitulo = 10,
    fuenteDatos = '\nhttp://ctfs.si.edu/webatlas/datasets/bci/',
    tituloMapa = paste0('Clusters LISA de ', x, ''))
# dev.new(); print(m$grafico)
return(m$grafico)
}, simplify = F
)
lisamaps_mifam_sintendencia$leyenda <- gtable_filter(ggplot_gtable(ggplot_build(lisamaps_mifam_s
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam_sintendencia[1:8], nrow = 3)), lisamaps_mif
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam_sintendencia[9:16], nrow = 3)), lisamaps_mi
\ldots
```

## Referencias

- Alvarado-Hernández, A. M. (2011). *Caracterización florística de los hábitats utilizado por el tepzcuintle cuniculus paca, (LINNEO, 1766; rodentia: Cuniculidae), en el piso basal del parque nacional carara, costa rica.*
- Batlle, J. R. M. (2020). Biogeografia-master/scripts-de-analisis-BCI;coding sessions (Version v0.0.9000). <https://doi.org/10.5281/zenodo.4402362>
- Bayer, K., Clemens y Kubitzki. (2003). Malvaceae. In Springer (Ed.), *Plantas con flores textperiodcentered dicotiledóneas*.
- Borcard, F. . ois y L., Daniel y Gillet. (2018). *Ecología numérica con r* (Springer, Ed.).
- Clark, D. B. (2002). Los factores edáficos y la distribución de las plantas. *Ecología Y Conservación de Bosques Neotropicales. LUR, Cartago, Costa Rica*, 193–221.
- Flores, D. G., Suvires, G., & Dalmasso, A. (2015). Distribución de la vegetación nativa en ambientes geomorfológicos cuaternarios del monte Árido central de argentina. *Revista Mexicana de Biodiversidad*, 86(1), 72–79.
- González, A. R. (2006). *Ecología: Métodos de muestreo y análisis de poblaciones y comunidades*. Pontificia Universidad Javeriana.
- Hubbell, S., Condit, R., & Foster, R. (2021). *Parcela del censo forestal en la isla de barro colorado*.
- Kindt, R., & Coe, R. (2005). *Tree diversity analysis. a manual and software for common statistical methods for ecological and biodiversity studies*. Retrieved from <http://www.worldagroforestry.org/output/tree-diversity-analysis>
- Lorenzo-Cáceres, J. M. S. de. (2007). Las especies del género abutilon mill.(Malvaceae) cultivadas en españa. *PARJAP: Boletín de La Asociación Española de Parques Y Jardines*, (45), 45–49.
- Marín, J. Z., Hilario, R. F., & Andino, O. O. (n.d.). *Análisis filogenético de la familia malvaceae*.
- ORTIZ, D. G. (2010). *Claves para los taxones y cultones del género hibiscus l.(Malvaceae) cultivados y comercializados en la comunidad valenciana (e españa)*.
- Pérez, R., Aguilar, S., Condit, R., Foster, R., Hubbell, S., & Lao, S. (2005). Metodología empleada en los censos de la parcela de 50 hectareas de la isla de barro colorado, panamá. *Centro de Ciencias Forestales Del Tropico (CTFS) Y Instituto Smithsonian de Investigaciones Tropicales (STRI)*, 1–24.
- QGIS Development Team. (2009). *QGIS geographic information system*. Retrieved from <http://qgis.osgeo.org>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from <https://www.R-project.org/>
- Springer (Ed.). (1998). Implicaciones ecológicas de los cambios en los patrones de sequía: Cambios en la composición de los bosques en panamá. In *Impactos potenciales del cambio climático en los ecosistemas de bosques tropicales* (pp. 273–287).
- Stevenson, P. R., & Rodríguez, M. E. (2008). Determinantes de la composición florística y efecto de borde en un fragmento de bosque en el guaviare, amazonía colombiana. *Colombia Forestal*, 11, 5–17.
- Sugasti, L., Eng, B., & Pinzón, R. (2018). *Medición continua de flujo de co2 ensuelo en una parcela de bosque tropical en isla barro colorado, canal de panamá*.