

# Análisis de la Familia de Moraceae en una Parcela de 50 Hectáreas, Dentro de la Isla Barro Colorado

**José Abreu Díaz** *Estudiante de geografía, Universidad Autónoma de Santo Domingo (UASD)*

El presente artículo, está basado en el análisis de Moraceae existentes en la Isla de Barro Colorado, para el mismo se emplean mapas y tablas con muestras en un parcela de 50 hectáreas, dividida en celdas de una hectárea cada una. Empleando un estudio inferencial, se destacan las variables propuestas como son medidas de pH, abundancia global, abundancia específica (familia), riqueza global, riqueza específica, mapa de pendientes, los cuadros de variables ambientales numéricas y nominales, medición de asociación, análisis de agrupamiento jerárquico, identificación de especies indicadoras, análisis de diversidad, y el análisis ecológico espacial. En todas ellas se aprecia un crecimiento exponencial de la familia bajo estudio. Con el objetivo de entender la razón de existencia de las condiciones que han dado origen a la amplia riqueza y abundancia de BCI, en el caso de la familia en cuestión, resalto el contexto político e histórico en que ha surgido la isla, y su situación geográfica. Dejando claro que pudieran existir otros grandes laboratorios de alcance global, aunque claro, no con la misma posición geográfica estratégica que tiene esta isla en el istmo de Panamá

*Keywords:* coeficiente de Spearman, distancia de Jaccard, matriz de Hellinger, pH, variable ambiental numérica, variable ambiental nominal, Moraceae, pendiente, abundancia global, riqueza global, Canal de Panamá.

## 1 Introducción

Toda familia taxonómica convencional, presenta unos criterios de agrupación que hacen de cada grupo botánico, verdaderas unidades identificables (Van Devender et al., 2010). Los senderos a seguir nos podrían encaminar por un sinfín de rutas alternativas, según, Van Devender et al. (2010), otras zonas próximas y muy parecidas a la isla Barro Colorado, dentro del istmo, presentan grandes concentraciones en diversidad de plantas y animales. La organización de las plantas en familias constituye una respuesta específica al orden en que la ecología, como ciencia que estudia las categorías agrupacionales que presentan los organismos en sus interacciones, ordena las unidades de grupos de plantas.

La ecología es la forma sistemática de estudiar las diversas dimensiones que resultan de la interacción entre organismos diversos, me apoyo en la biogeografía, con la intención de conocer los emplazamientos que ocupan esos grupos de organismos, especialmente de plantas. En este manuscrito se pretenden entender los aspectos fundamentales de la familia de plantas conocida como Moraceae, partiendo del estudio de gráficos, tablas y diagramas procedentes de investigaciones realizadas en la isla Barro Colorado, en Panamá.

Es cierto que el condicionamiento de diversas especies de plantas en distintos espacios termina creando casos de escasa homogeneidad, que de una u otra forma se rompe el patrón de similitud. Es por ello que en este trabajo se maneja más un enfoque particular de la realidad existente en BCI. Algunos estudios demuestran que las Moraceae no solo sirven de alimento a los frugívoros, o como soporte en la industria maderera, también sirven de huésped a algunas especies como es el caso de las avispas, las cuales se desarrollan en las flores femeninas de dichas plantas según

Cardona, De Ulloa, & Kattan (2007). Otros estudios abordan las propiedades químicas de dichas plantas, este trabajo busca definir las características más notables y básicas de esta familia de plantas apoyándome en el escenario de BCI.

Según Clement & Weiblen (2009), la familia Moracea comprende 37 géneros y aproximadamente 1100 especies en regiones tropicales y templadas en todo el mundo. En América se aprecia una amplia distribución de las especies de esta familia, como lo es en los territorios de Panamá y otras regiones próximas de la América Central, siendo de gran importancia para la comunidad de frugívoros que se alimentan de las mismas. Entre las especies reconocidas, se destacan: *Brosimum alicastrum*, *Maclura tinctoria*, *Ficus mexicana*, *Ficus petiolaris*, *Ficus cotinifolia*, *Ficus spp.* De esta misma distribución existen importantes trabajos, tales como el de Magallanes, Rocha, & Terán (n.d.), y Piedra-Malagón, Ramírez Rodríguez, & Ibarra -Manríquez (2006). En ese orden pretendo identificar otros territorios del mundo donde podrían existir ejemplares de estas plantas correspondientes a los tipos de especies, y cuáles regiones presentan condiciones favorables para el predominio de las mismas, partiendo del análisis de los trabajos realizados en isla Barro Colorado.

Aunque en un principio he hecho mención de regiones próximas al istmo de Panamá, todo este análisis ecológico está volcado en la medición de abundancia y riqueza de especies correspondientes a la familia de Moraceae en la isla mencionada. De hecho todo el material empleado procede de estudios precedentes, extraídos del universo biogeográfico que constituye la Isla Barro Colorado.

La ciencia en un sentido epistemológico es como el principal puerto en el cual se puede anclar el raciocinio, por ende debe ser depositaria dicha profesión de respeto y realización con cordura, pero sobretodo de la mano del talento y el compromiso. A sambientes de que la isla Barro Colorado no es un espacio en donde se venga realizando ciencia de antaño, es decir, no existe allí partiendo del territorio, una cultura científica milenaria que conserve los grandes hitos del quehacer intelectual, el pensamiento viaja de unas sociedades a otras y en cada latitud quiere matices que le hacen ser.

Por encima de estas limitaciones, que en un sentido estricto no representan ningún obstáculo, las investigaciones científicas en lugares como isla Barro Colorado crecen de una forma acelerada, en una escala de tiempo sumamente corta si se toman en cuenta los criterios de periodización de la historia humana. En contraste con otros lugares, que son objeto de investigaciones desde antes de la era actual, y que tal vez para muchos permanezcan en el anonimato.

Y que bueno que la investigación florece en sitios relativamente jóvenes, es un indicio de que no siempre estaremos atados al patrón clásico de hacer las cosas, seguir el mismo camino que tantas veces se ha recorrido en virtud de obtener resultados diferentes empleando las mismas técnicas y buscando los mismos objetivos que se estudian y persiguen desde hace siglos. En este contexto se trata de estudios recientes sobre espacios relativamente jóvenes, de modo que si las circunstancias nos llevan por caminos equivocados, pues no tendríamos reparos en volver a comenzar. Porque, riqueza y abundancia inmensas, conjugadas en un ambiente ecológico reducido es de inferir que hacen previsibles los resultados de cualquier análisis de esta categoría.

## 2 Metodología

Apoyándome en los censos realizados en la isla de Barro Colorado, mediante un análisis induutivo, pretendo dilucidar los datos estadísticos más pertinentes sobre la familia de plantas conocidas como Moraceae. Dejando claro que en otras zonas del continente existen importantes trabajos sobre esta familia de plantas, susceptibles de estudios debido en gran parte al soporte que constituyen en la alimentación de una amplia comunidad de frugívoros. Tomando como escenario esa

pequeña isla del istmo de Panamá, empleando un serie de recursos que han derivado de profundas investigaciones, intento crear una nueva fisonomía a esta familia de plantas, o por lo menos generar algo nuevo para quienes consumen este tipo de artículos. El objetivo de este análisis es conocer la dinámica estadística de las principales especies de Moraceae de la isla Barro Colorado, empleando los censos que se describen en los materiales empleados.

La isla se formó tras la creación del lago Gatún en 1913, durante la construcción del Canal de Panamá (1904-1914), su superficie de 15 kilómetros cuadrados, alberga una de las estaciones más antiguas de investigación tropical del mundo donde se han llevado a cabo estudios por más de 100 años.

La Isla Barro Colorado o BCI, por sus siglas en inglés, es uno de los puntos geográficos más estudiados en el mundo, es como una especie de Meca para los investigadores que quieren desarrollar allí sus estudios, muchas veces acompañados de sus estudiantes, estudiantes que en lo adelante también realizarán sus propias investigaciones por separado. En dicha isla todo está registrado, cada comunidad de insectos, plantas y otros animales, están debidamente mapeados, con inscripciones por doquier. Tal vez por la diversidad de organismos que allí predominan, se hace esta clase de inferencia en esta isla que por su situación geográfica reúne las condiciones que hacen posible estos estudios de interés para el conocimiento de una comunidad que hace ciencia y para otros individuos que sacan provecho de estos conocimientos para el fortalecimiento de la enseñanza y para el crecimiento del campo de estudio que nos compete.

Algunos estudios como los de Fredericksen, Justiniano, Rumiz, McDonald, & Aguape (n.d.), resaltan la importancia de las Moraceae en el sentido de que sirven de alimento a una amplia comunidad de frugívoros, pero lo cierto es que también, tienen gran importancia en el sector maderero. Como dije en un principio, todo el material de apoyo, proviene de las investigaciones de la isla de Barro Colorado, el cual consiste en el empleo de tablas, cuadros, mapas, diagramas, dendrogramas y una serie de recursos estadísticos que proceden de sendas investigaciones. Para el análisis exploratorio de datos se han empleado los cuadros de 1 hectárea de BCI, que corresponden a las variables ambientales.

La isla de Barro Colorado se presenta en un espacio insular fragmentado, para muestra de ello, ver a continuación el mapa adjunto (ver figura ##1).

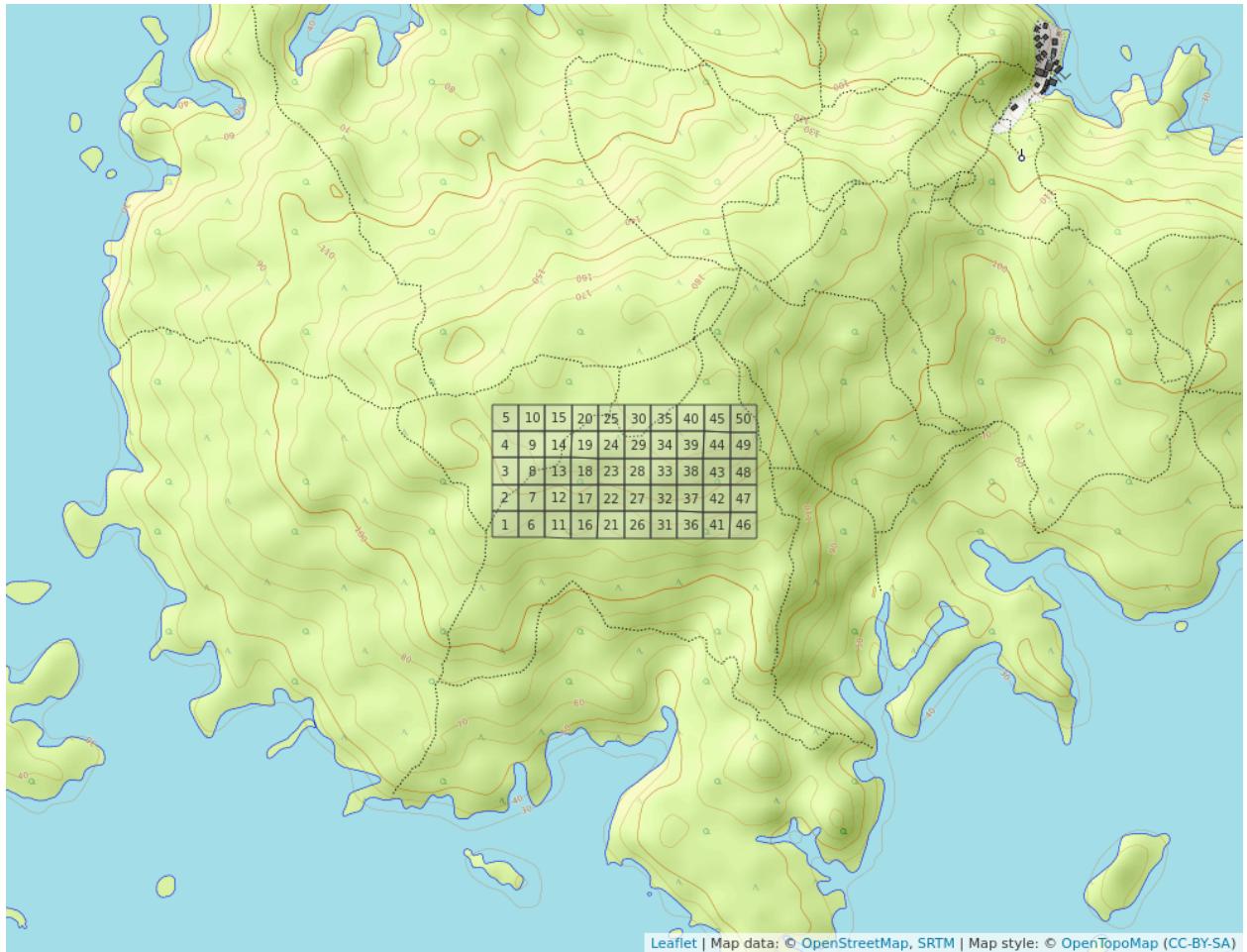


Figure 1: mapa de isla Barro Colorado cuadros

Por igual empleo los cuadros de variables ambientales nominales y numéricas de 1 hectárea de BCI, constituidas por categorías de edad, habitat, geología y quebrada. Otros recursos de vital importancia son las variables ambientales nominales de una hectárea de BCI, las matrices, en especial de Spearman, de riqueza y abundancia de suelo, y las de Hellinger de medición de asociación. La interacción con estas fuentes de información ayudan a tener una aproximación más concreta sobre la diversidad con relación a la familia tratada.

Los métodos que hasta el momento he señalado corresponden a los mapas en los cuales aparece la parcela de 50 hectáreas en la isla Barro Colorado, en donde se describen las variables ambientales ilustradas en matrices que describen el comportamiento de dichas variables, y los patrones de movilidad en el espacio que presentan las mismas, lo cual en lo adelante permitirá comprender el mecanismo que siguen la fuerte asociación, abundancia y riqueza de la familia en cuestión.

Todo lo anterior obedece a la parte de datos exploratorios, con excepción de las matrices de Spearman y Hellinger que ya fueron mencionadas y que corresponden a la sección de medición de asociación. Está presente el segmento correspondiente a la medición de asociación, en el que se presentan los modos Q y R, el primero referido a la medición de asociación de sitios, en este caso por medio de la distancia, y la similaridad de Jaccard. El modo R se refiere a la medición de la asociación entre pares descriptores como son variables o especies, también hago uso de la matriz

de Hellinger. En el método de análisis jerárquico, también se emplean los diagramas del árbol o dendrogramas (método de cuerda y de Ward).

El análisis de agrupamiento, consiste en una agrupación sucesiva basada en la repetición de un procedimiento dado de grupos de objetos, hasta que estos encuentren su lugar, se basa en la construcción de dendrogramas. El análisis jerárquico se caracteriza por tener un enfoque aglomerativo, lo que implica un algoritmo ascendente, y la ordenación de los subgrupos de objetos en un único grupo. Los algoritmos del análisis jerárquico son los criterios de enlace, y estos a su vez se integran por los siguientes tipos: Enlace simple, enlace completo y enlace promedio. Los métodos de representación utilizados son: Agrupamiento aglomerativo por enlace simple, agrupamiento aglomerativo por enlace completo, y agrupamiento aglomerativo por enlace promedio.

También es notable la sección correspondiente a las técnicas de ordenación y a las especies indicadoras.

La ordenación es una técnica de medición que busca determinar el grado de asociación y de agrupamiento. Un objeto se caracteriza por sus propiedades en un espacio n-dimensional donde cada dimensión es una variable, un descriptor. A diferencia de la técnica de agrupamiento, o como complemento de este, el análisis de ordenación abarca un conjunto de técnicas que busca abarcar el tamaño (dimensionalidad) de los datos. El análisis de ordenación puede ser no restringido y restringido, en el primer caso, las tendencias detectadas en el conjunto de datos suelen estar no restringidas por otro conjunto. En el segundo caso las tendencias detectadas en un conjunto suelen estar asociadas a otro conjunto.

Se realiza el análisis de las especies indicadoras por medio del estadístico de IndVal. Para el análisis de diversidad se emplean los modos alpha y beta. Dentro de las técnicas de ordenación (restringida y no restringida) las pruebas PCA, RDA, y CCA. Estas pruebas buscan medir la similaridad entre los sitios con variables ambientales determinadas. Y por último, para el análisis de ecología espacial se empleó el correlograma con estadístico de Moran's I, que determina la existencia de una correlación positiva o negativa.

En un correlograma, se aprecia un rectángulo cortado a la mitad de forma horizontal, sobre este suelen alternarse barras verticales surcadas a la mitad por un punto que representa el valor estimado (Moran's I) el cual se lee asumiendo los valores por debajo o por encima de cero (correlación negativa o positiva), que coincidan con dicho punto, estos son los valores de la barra lateral izquierda. Estos puntos al franquear la barra que representa los valores positivos y negativos pues recrean la autocorrelación positiva y negativa, que se expresa en sitios ubicados a cierta cantidad de metros entre sí.

A menudo se define la biodiversidad como la variabilidad de organismos presente en un lugar determinado, y muchos estudiosos han dejado su impronta sobre este concepto a lo largo del tiempo. Por ejemplo, Harper y Hawsworth, defienden que es el estudio de tres aspectos fundamentales: Intraespecífica, interespecífica, y de ecosistemas, con los adjetivos de genética, de organismos, y ecológica. Hubbell ofrece una definición más restringida y ajustada a la realidad actual: Biodiversidad es sinónimo de riqueza de especies y de abundancia relativa de especies en el espacio en el tiempo. Magurran utiliza el término diversidad biológica y biodiversidad como sinónimas y la define como la variedad y la abundancia de especies en un estudio.

### 3 Resultados

#### 3.1 Análisis Exploratorio de Datos

A continuación muestro las mediciones de pH, mapa de pendientes, mapas de abundancia, tanto global como de la familia, mapa de riqueza global y de la familia, variables ambientales

numéricas, y nominales en una parcela de 50 hectáreas, integrada por subdivisiones de una hectárea, todo ello correspondiente a la familia de las Moraceae. En el caso del pH tenemos que es ácido, según los mapas y diagramas consultados, estas condiciones para las Moraceae, y cualquier otra especie en condiciones extremas, sería perjudicial. El pH del suelo es importante porque los vegetales solo pueden absorver a los minerales disueltos y la variación del pH modifica el grado de solubilidad de los minerales y nutrientes. En el siguiente mapa, se pueden apreciar las medidas de pH presentes en la isla de Barro Colorado en una parcela de 50 hectáreas, con un patrón de distribución al oeste de la parcela en su variedad ácida (ver figura ##2).

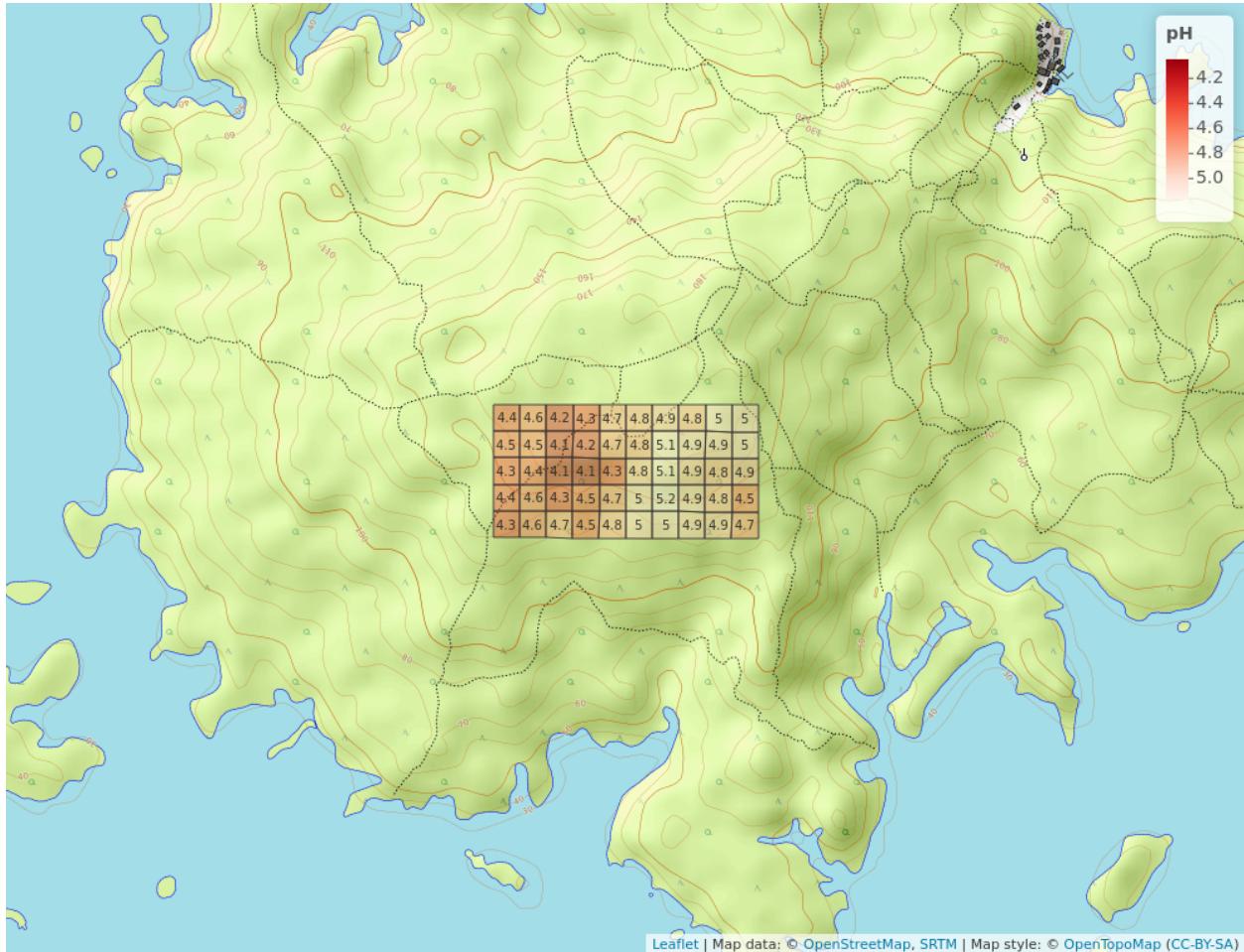


Figure 2: mapa de la isla Barro Colorado cuadros ph

Como se puede ver, el pH es eminentemente ácido, en una parcela de 50 hectáreas, puesto que oscila en una escala de 0 a 6, con una distribución que indica mayor concentración al oeste de dicha parcela.

En el mapa de abundancia global se advierten valores no lejanos de lo ideal cuando se emplean parámetros de la muestra comprendidos entre 3600 y 5000, que en este caso sería una escala, en una parcela de 50 hectáreas. El Valor promedio que representaría cada hectárea en una media hipotética sería de aproximadamente casi 4000 individuos de esta familia, sin embargo tiende a cambiar cuando hablamos de riqueza. Si bien existen dentro de esta familia, especies representati-

vas de casi todas las formas de vida leñosas, las formas más comunes entre las especies de bibosi, son por un lado plantas hemiepífitas estranguladoras o matapalos y por otro, plantas no epífitas con sistema propio de sustento o higuerones (Fredericksen et al., n.d.), (ver figura ##3).

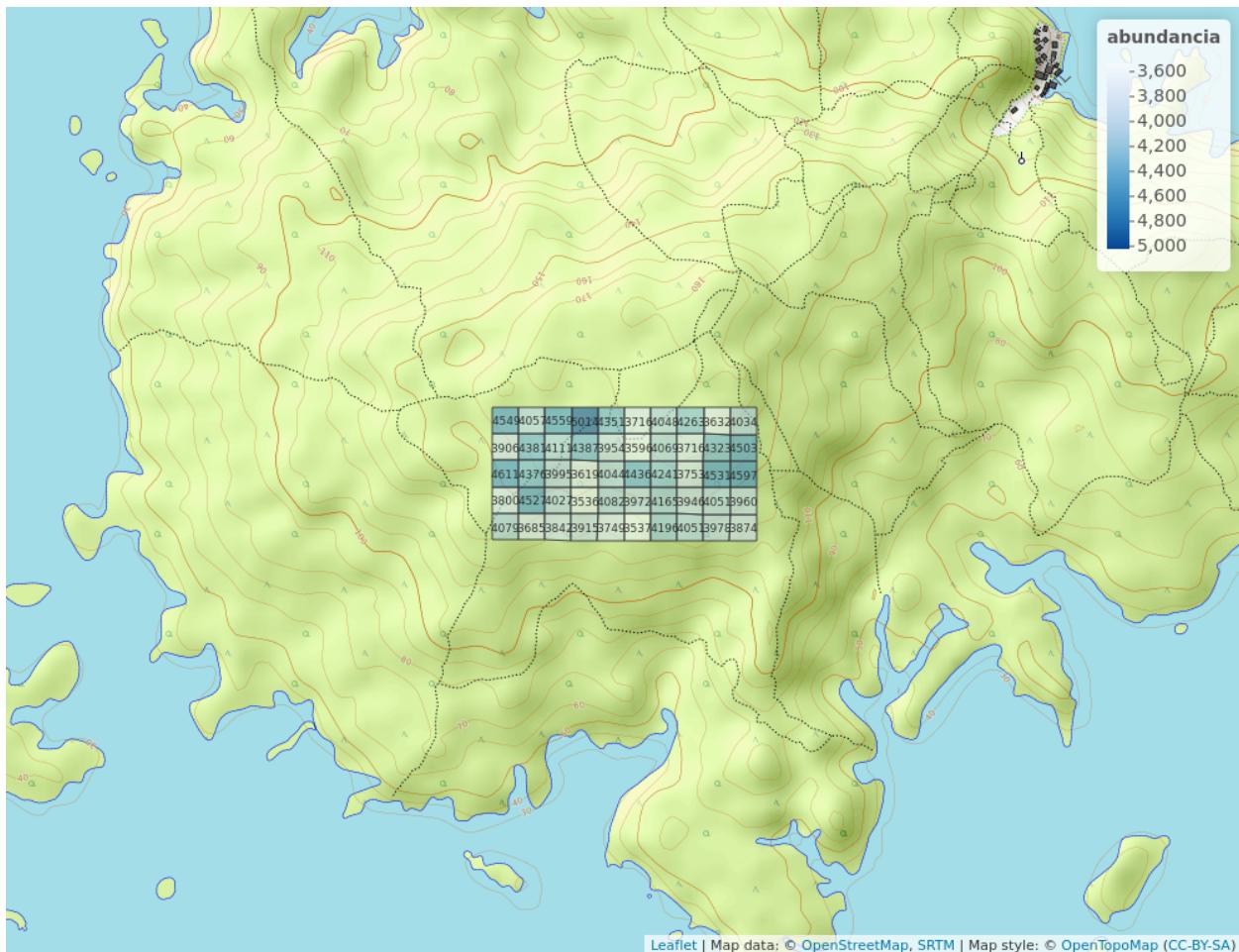


Figure 3: mapa de la isla Barro Colorado abundancia global

En el mapa de abundancia de mi familia no se presenta tampoco que la muestra por hectárea se aleja de forma marcada del parámetro empleado para la parcela de 50 hectáreas, en este caso en un parámetro de entre 200 y 700 individuos, podría verificarse una especie de relación directa entre el primer y segundo mapa.

En ambos casos cada hectárea adquiere valores próximos a la dimensión numérica de individuos presentes en la parcela, para ello adjunto el siguiente mapa sobre abundancia de mi familia (ver figura ##4).

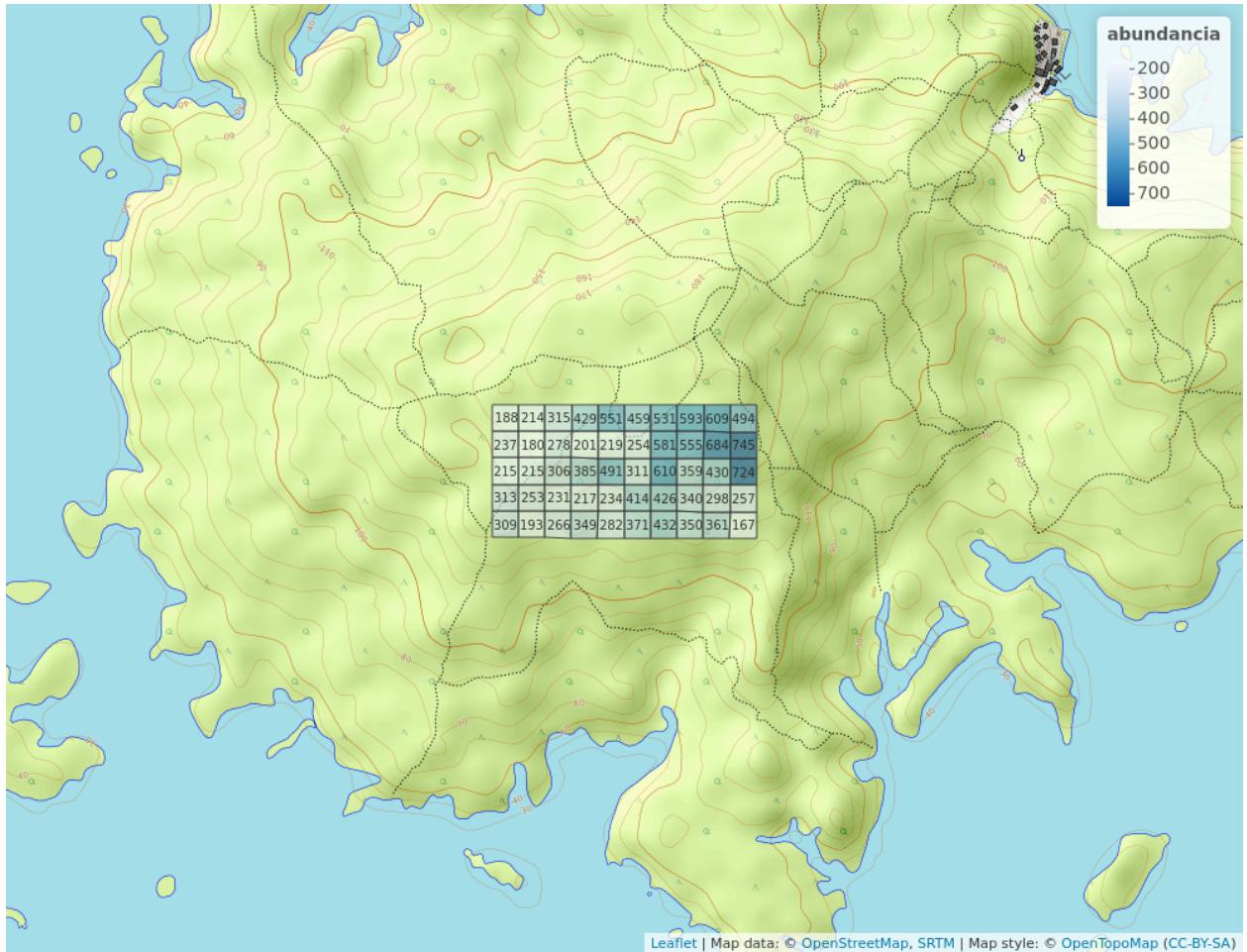


Figure 4: mapa de la isla Barro Colorado abundancia de mi familia

Como puede comprobarse existen muchos individuos, en este caso citando la abundancia, de una determinada especie de la familia de Moraceae en una hectárea de la isla Barro Colorado, con un patrón de concentración al noreste de la parcela, lo que muestra una amplia densidad para una parcela de 50 hectáreas. Siendo estas plantas muchas veces ejemplares de gran dimensión en muchos de los casos, se infiere que estas forman un dosel impenetrable, típico de muchos bosques de la selva tropical centroamericana, en donde BCI no es la excepción. No solo estaríamos hablando de Moraceae, sino de otras especies que forman la cubierta de esta región intertropical.

La isla de Barro Colorado es realmente joven dentro de lo que es la escala de tiempo geológico, solo tiene algo más de 100 años de existencia lo que nos advierte de un relieve poco accidentado, debido en gran parte a su tamaño, pero sobre todo a su edad.

En este sentido sería bueno apreciar las pendientes, y ver el valor de la altitud en esta especie de enclave marítimo, lo que también favorece la existencia abundante de ejemplares de Moraceae (ver figura ##5).

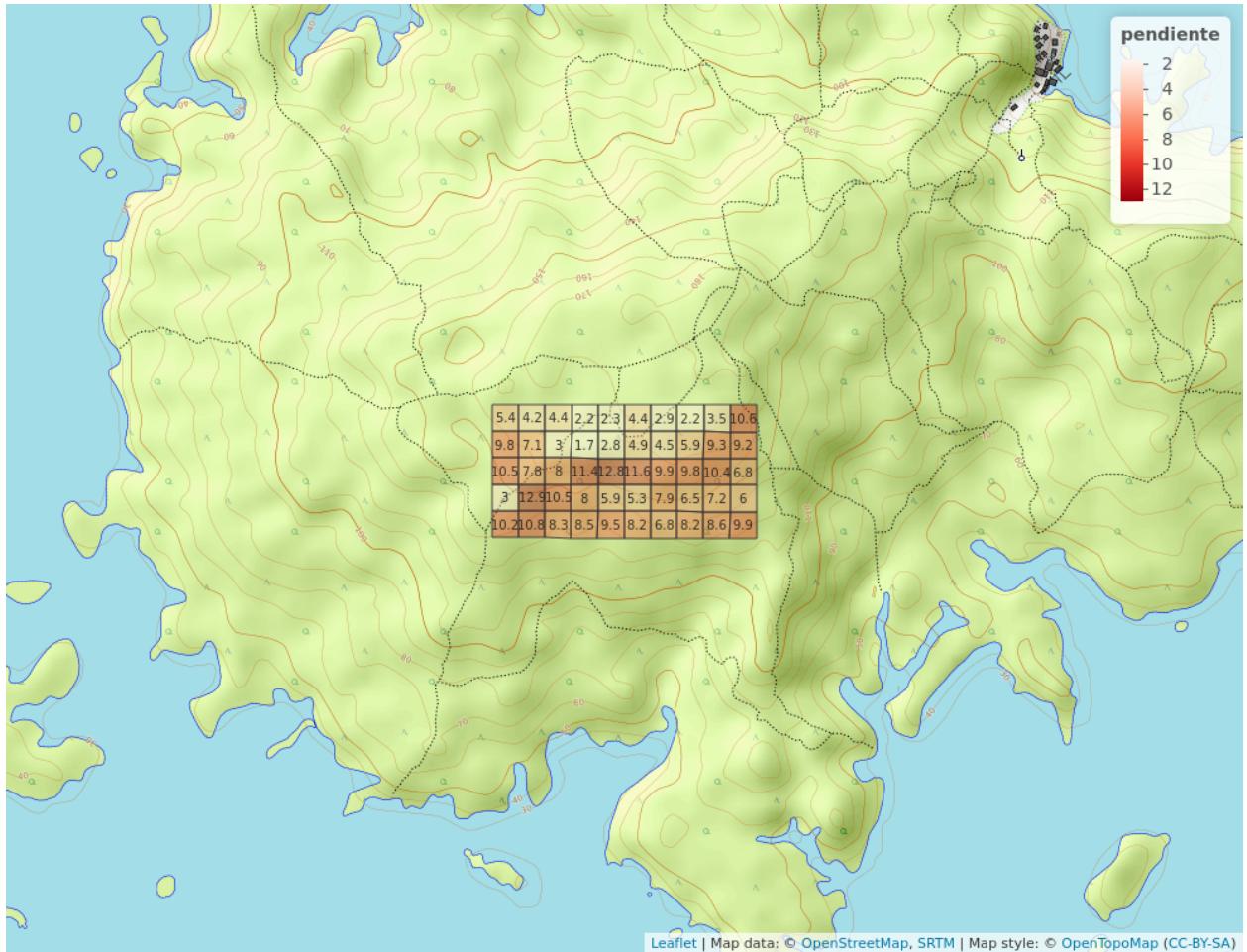


Figure 5: mapa de la isla Barro Colorado cuadros y pendientes

Para una consideración general de la parcela de 50 hectáreas se destaca gran pendiente, con un patrón de distribución al sur de la misma.

En el caso de las variables ambientales numéricas se obtiene los siguientes valores (ver figura ##6)

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

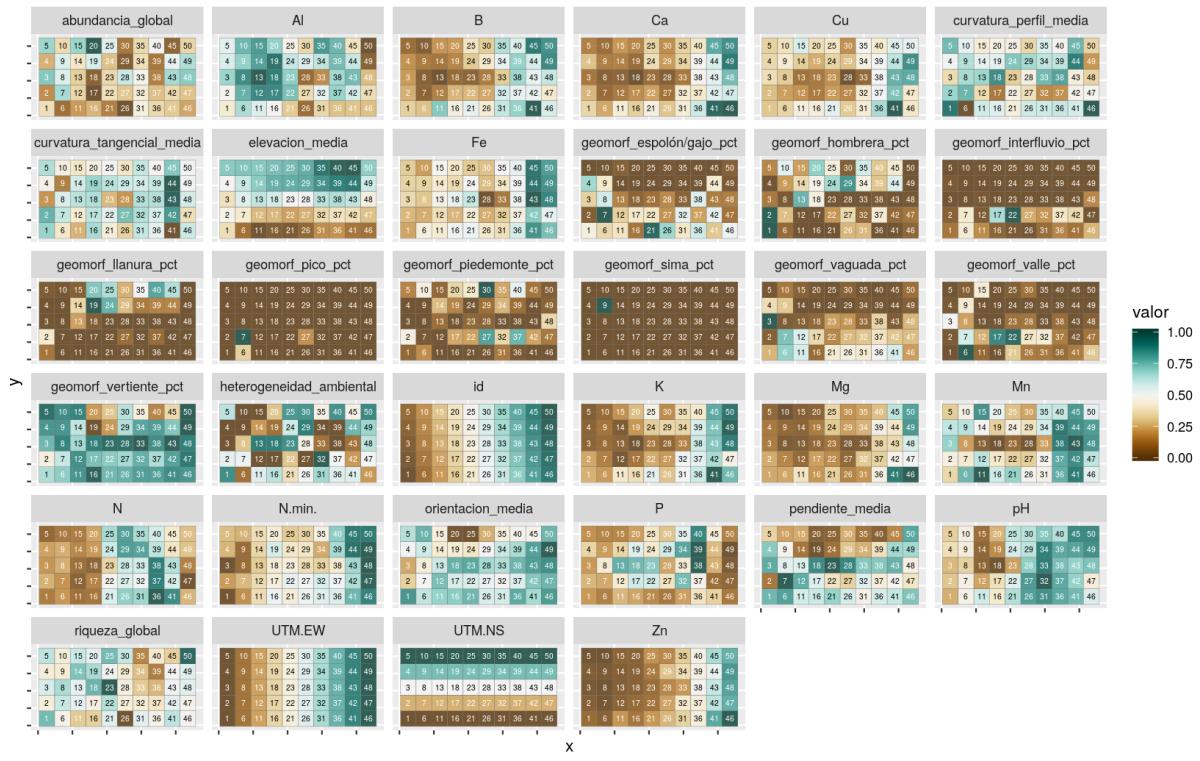


Figure 6: mapa de la isla Barro Colorado variables ambientales numéricas

Aquí podemos ver una especie de síntesis de las informaciones que ofrecen las variables ambientales de una hectárea, escaladas de 0 a 1. En el caso de las geoformas, correspondientes a llanuras, se obtienen valores grandes en un segmento de la escala que va de 0.00 a 0.25, lo que refleja un terreno eminentemente llano. Con relación a las vertientes ocurre todo lo contrario, se obtienen valores grandes, en el segmento de la escala que va de 0.75 a 1.00, lo que se refiere a mayor inclinación en contraste con las tierras llanas. Otra geoforma como es el caso de los valles, presenta valores muy pronunciados en los segmentos bajos de la escala citada. Los valores de la escala muestran amplias diferencias entre unas variables y otras, pero dentro de la misma escala, suele verse la superioridad de una sobre otra.

En el mapa de variables ambientales nominales, se aprecia la dinámica de algunas variables ambientales (ver figura ##7)

Cuadros de 1 Ha de BCI. Variables ambientales nominales

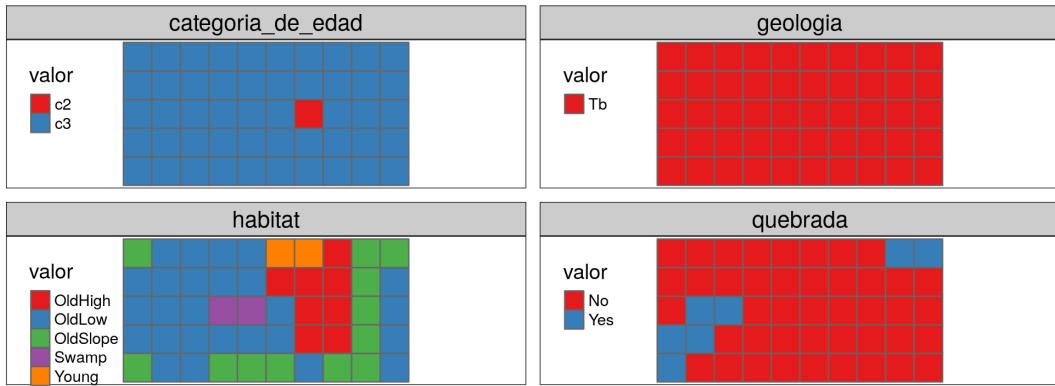


Figure 7: mapa de la isla Barro Colorado variables ambientales nominales

En esta matriz se distinguen variables como la categoría de edad, geología de la región, hábitat, y la quebrada. El hábitat de bosque viejo en terreno alto (OldHigh), se destaca por su relativa minoría. En tanto que el bosque viejo en relieve bajo (OldLow) es el más abundante. El bosque viejo con pendiente (OldSlope), es relativamente abundante. El pantano es escaso (Swamp), así también el bosque joven (Young). En cuanto a la geología tenemos que el tipo de roca predominante es de tipo intrusivo o plutónica. Las quebradas o barrancos son estrechos. Existen dos grandes categorías dentro de la variable de edad, el bosque viejo en terreno bajo, y el bosque viejo en terreno alto. De lo anterior se infiere que existe más terreno bajo con bosque viejo en BCI en una base gelógica joven, que con las demás variables.

El mapa de riqueza presenta un formato similar a los de abundancia, con comportamientos diferenciados por hectárea, comparados con la parcela de 50 hectáreas. En el siguiente mapa puede apreciarse la riqueza de la familia botánica bajo estudio (Moraceae) por hectárea (ver figura ##8).

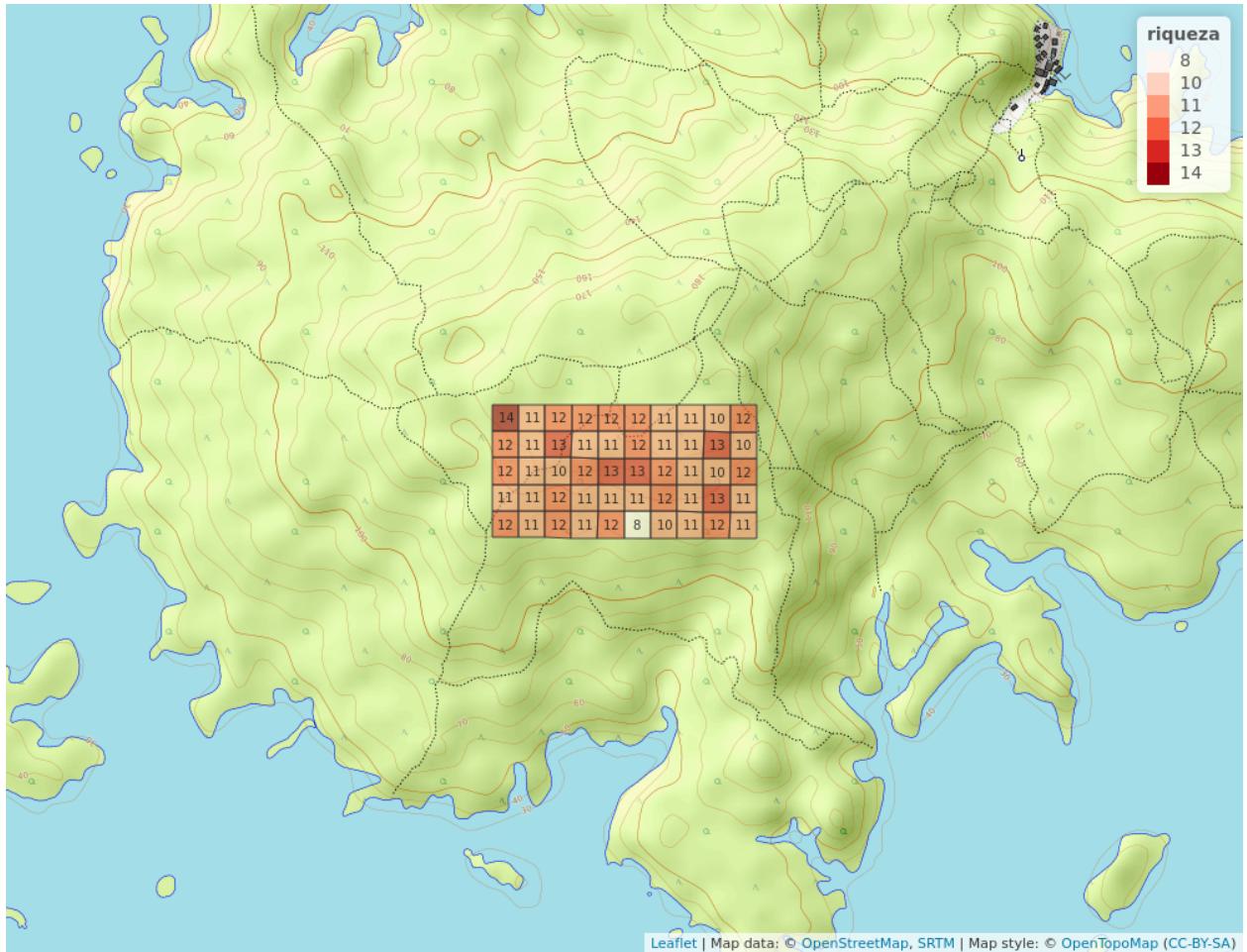


Figure 8: mapa de la isla Barro Colorado cuadro de riqueza mi familia

En esta parte puede apreciarse que la riqueza es casi la misma por hectárea, acercándose al nivel superior de la escala, así se puede asumir que el grado de dispersión de algunos individuos de la familia Moraceae es casi nulo. Existe una gran riqueza de estas plantas que ayudan en gran parte a mantener este grandioso ecosistema.

Ya para el mapa de riqueza global se advierten mayores contrastes por hectárea, pero igual se impone la diversidad, la multiplicidad de individuos es asombrosa (ver figura ##9). La riqueza global presenta un patrón de distribución noroeste-sureste.

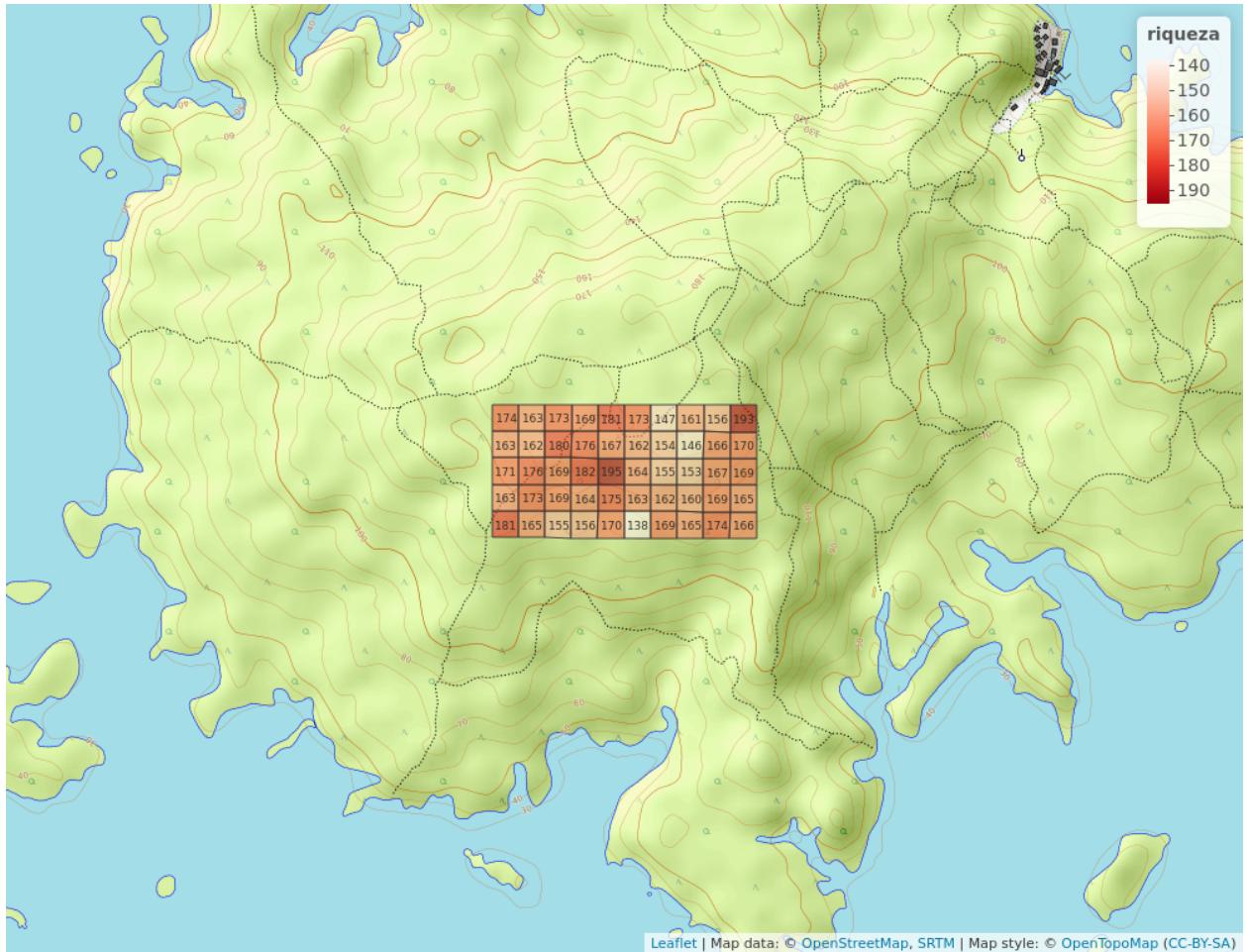


Figure 9: mapa de la isla Barro Colorado riqueza global

En los cuadros de variables ambientales, se registran las diferencias entre unas y otras, se distinguen tierras bajas de otras no muy altas. Pero no solo se analizan los rasgos geomorfológicos, otras variables ofrecen también allí su concurso. Aparecen valores de pH, riqueza global por igual, abundancia global, entre otras variables.

### 3.2 Medición de Asociación (Modos Q y R)

En esta parte se aborda lo que es la medición de asociación, tomando en cuenta lo que son los valores de disimilitud o distancia, así como las matrices de correlación. Dentro de los modelos de medición de asociación existen los modos Q y R, en este caso de la distancia o disimilitud, tenemos el modo Q que describe la distancia entre objetos cuantitativos. La paradoja de Orlóci plantea que la distancia euclídea es más pequeña entre sitios que no comparten especies que entre sitios que sí las comparten. A continuación se presenta una matriz de disimilitud (ver figura ##10).

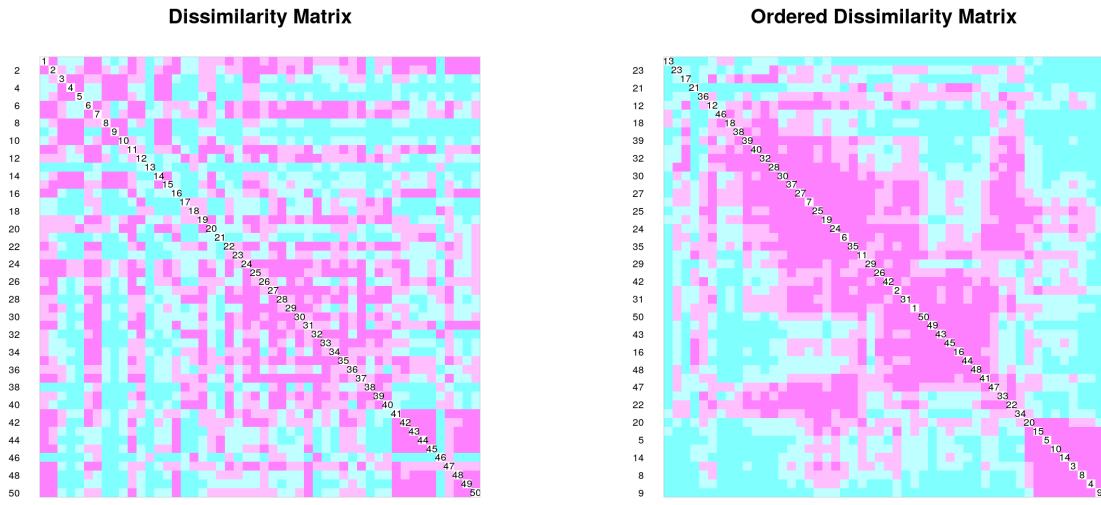


Figure 10: mapa de la isla Barro Colorado matriz disimilaridad hellinger

La matriz de disimilaridad es una representación de la medición de asociación entre objetos en el modo Q, como serían los sitios de muestreo. En esta matriz se puede apreciar una fuerte asociación entre objetos y esto lo indica la continuidad del color rosado (cuadro de la derecha). En este caso nos referimos a la matriz de la derecha, que es la matriz ordenada, a partir de la primera. Por ejemplo entre los sitios 40 y 35 existe una distancia muy pequeña entre sitios parecidos para la matriz ordenada. La transformación Hellinger permite hacer la interpretación en base a dos colores representativos: Color fucsia (magenta, rosa) que equivale a corta distancia, muy similar, y el cian o celeste que significa gran distancia o poco similar, todo esto con relación a los elementos de la familia de las moraceae. Algunos de los valores de disimilaridad o de distancia entre objetos que podemos apreciar dentro de la transformación Hellinger son los siguientes: Entre los sitios 2 y 1, tenemos que la distancia (dbl) es de 0.187, entre 3 y 1 es de 0.439, entre 4 y 1 es de 0.499, entre 5 y 1 es de 0.430, entre 6 y 1 de 0.297, entre 7 y 1 es de 0.280, entre 8 y 1 es de 0.492, entre 9 y 1 es de 0.584, entre 10 y 1 es de 0.425, y entre 11 y 1 es de 0.269 lo que revela una escasa distancia o disimilaridad entre los objetos.

A continuación presento una versión mejorada de la matriz ordenada de datos cuantitativos de especies (abundancia) con fuerte asociación en la diagonal (ver figura ##11).

## Ordered dissimilarity matrix

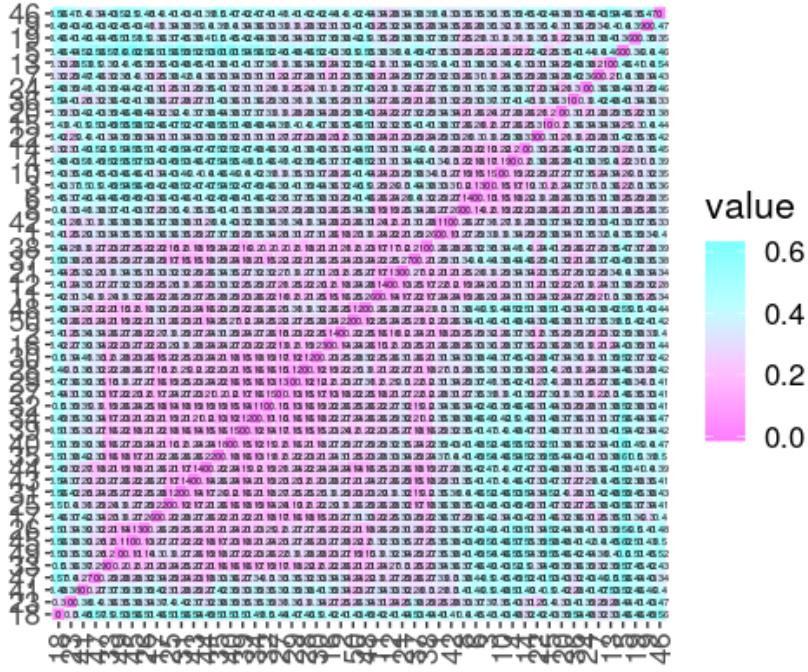


Figure 11: mapa de la isla Barro Colorado matriz disimilaridad

En esta matriz de disimilaridad o de distancia vienen a recrearse los anteriores casos de abundancia de los mapas de la Isla de Barro Colorado. En la diagonal se pueden apreciar los sitios que poseen características similares con la menor distancia(color rosado), mientras que en los demás lugares se aprecian lugares con características similares a mayor distancia y con una menor asociación (color cian o celeste). Hasta el momento me he centrado en el modo Q que mide el nivel de asociación por medio de la disimilaridad o similitud. Algo sumamente importante dentro del modo Q, en este caso basándome en datos binarios (presencia/ausencia) es la famosa distancia de Jaccard, la misma se puede expresar como la proporción de especies no compartidas. Para el caso correspondiente de los datos de moraceae, la distancia (dbl) muestra los siguientes resultados para algunos sitios: Entre 2 y 1, la distancia es de 0.22, para 3 y 1 es de 0.33, entre 4 y 1 es de 0.222, para 5 y 1 es de 0.3, para 6 y 1 es de 0.333, para 7 y 1 es de 0.222, para 8 y 1 es de 0.333, para 9 y 1 es de 0.333, para 10 y 1 es de 0.4, y para 11 y 1 es de 0.333, evidenciando en este caso una fuerte asociación.

La similaridad es la proporción de objetos compartidos entre sitios, en la siguiente gráfica de la matriz pueden apreciarse las especies compartidas entre sitios (ver figura ##12).

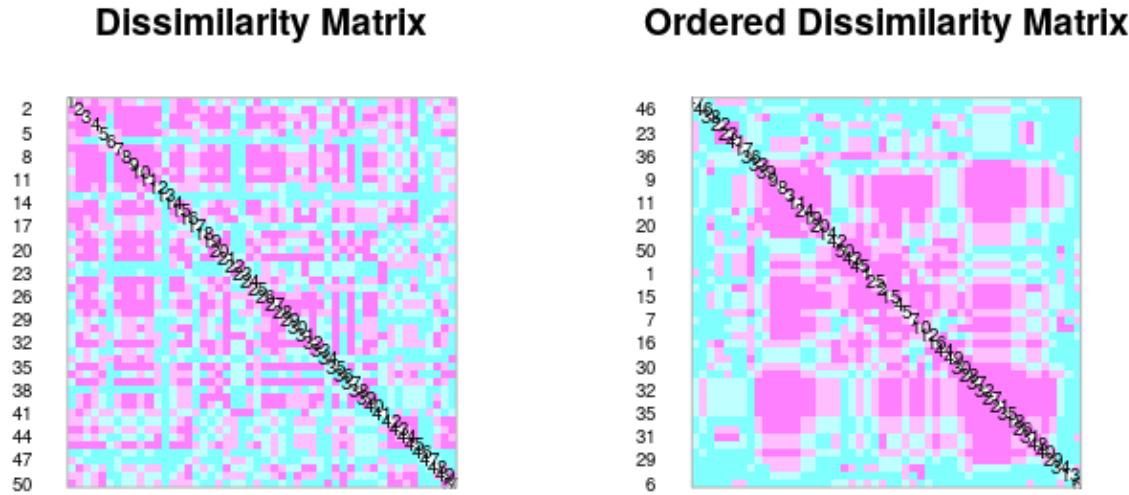


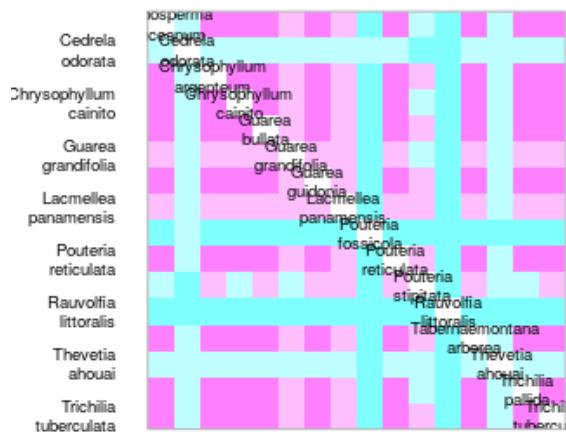
Figure 12: mapa de la isla Barro Colorado matriz disimilaridad

Pueden notarse en los bloques de colores las especies compartidas entre sitios, y que son exclusivas de dichos sitios. Por ejemplo para las variables  $a$ ,  $b$  y  $c$ , que se desprenden de la fórmula, incluida en los scripts reproducibles que sirven de soporte a este trabajo, de la similaridad, donde  $a$  representa el número de especies compartidas,  $b$  es el número de especies exclusivas del sitio 2 y  $c$  el numero de especies exclusivas del sitio 1. La cantidad de especies compartidas entre sitios con relación a las moraceae es grande, esto debido a la ausencia de ceros, hecho que en la paradoja de Orlóci refleja cercanía según la distancia euclídea que hace ver próximos los lugares o sitios que no comparten especies o individuos.

En el modo R se mide la asociación entre pares descriptores, es decir variables o especies, mediante la covarianza y la correlación. Por ejemplo para las variables  $a$ ,  $b$  y  $c$ , que se desprenden de la fórmula de la similaridad, donde  $a$  representa el número de especies compartidas,  $b$  es el número de especies exclusivas del sitio 2 y  $c$  el numero de especies exclusivas del sitio 1.

Un ejemplo ilustrativo sería el mapa de calor (ver figura ##13)

### Dissimilarity Matrix



### Ordered Dissimilarity Matrix

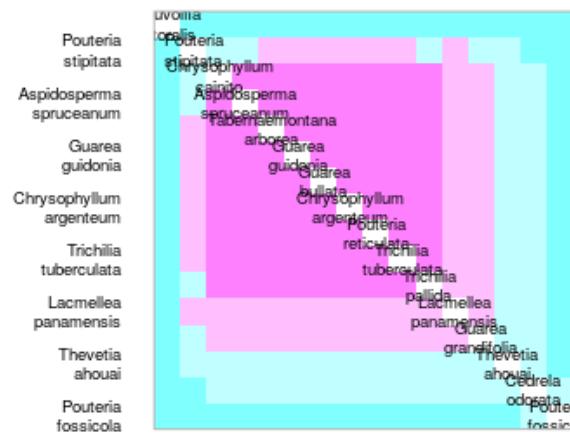


Figure 13: mapa de la isla Barro Colorado mapa de calor

En el mapa de calor mostrado se percibe una relación de dependencia en la diagonal de la derecha, esto es lo que se mide en el modo R, la asociación entre elementos descriptores, en este caso la abundancia de especies como variables. Entre la *Cainito aspidosperma* y la *trichinia tuberculata* se nota una fuerte dependencia, evidenciado por la continuidad del color fucsia o rosa. Lo mismo no ocurre con los colores de los bordes que reflejan franjas más pequeñas, y a su vez degradadas, lo que evidencia ausencia de dependencia. Esto así cuando se trata del modo R aplicado a datos cuantitativos de especies, que reflejan la abundancia, al igual que el modo Q, el Modo R se aplica para datos binarios, así tambien para la abundancia de variables (coeficientes de spearman y pearson).

A continuación muestro una matriz de abundancia y riqueza de suelo según el coeficiente de Spearman, con valores que aumentan y disminuyen de este a oeste, y de norte a sur, los valores en rojo presentan los valores más altos, y los azules los más bajos (ver figura ##14).



Figure 14: mapa de la isla Barro Colorado abundancia y riqueza de suelo

En la medición de asociación en los modos antes mencionados, se obtiene una apreciación estadística de la distribución de sitios, especies y variables ambientales presentes en la isla Barro Colorado que reflejan la cuatificación de las comunidades de moraceas presentes en dicha isla, así como la mayor o menor concentración de las mismas con relación al grado de asociación presente en los sitios.

### 3.3 Agrupamiento Jerárquico

A continuación se muestran los resultados de esta sección correspondiente a la familia de Moraceae.

### 3.3.1 Agrupamiento Aglomerativo por Enlace Simple

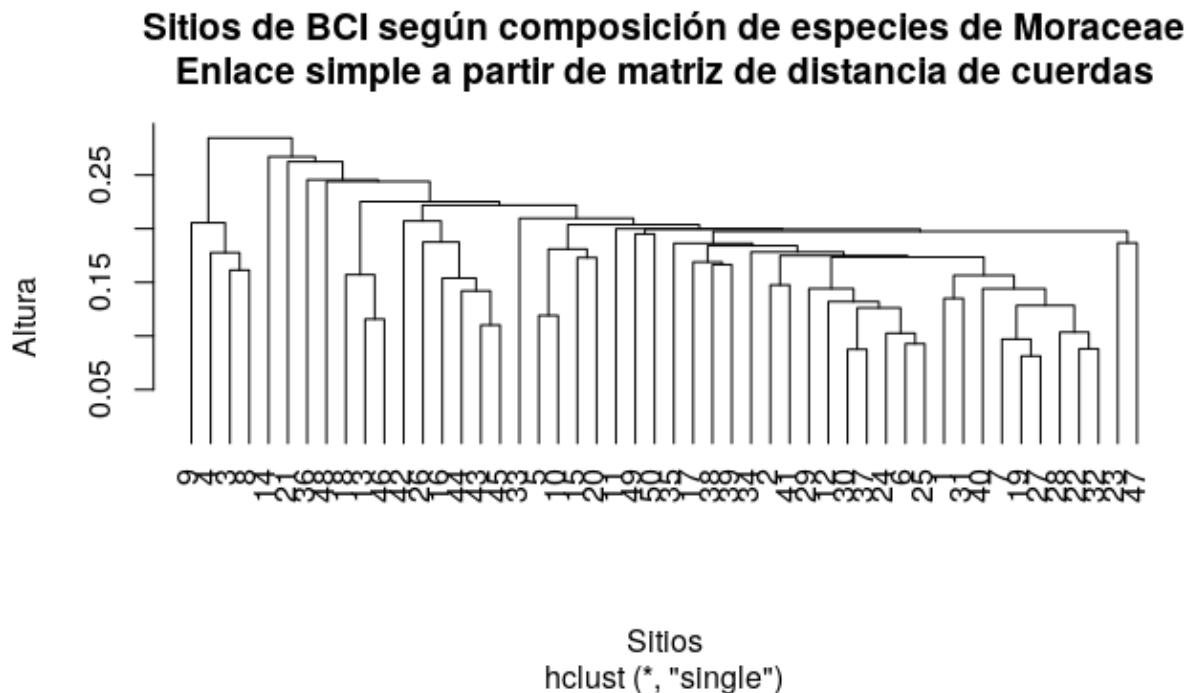


Figure 15: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

Cuando se construye este dendrograma, empleando la menor distancia entre sitios, se obtiene una matriz de cuerdas o matriz normalizada (ver figura ##15). En este dendrograma se puede apreciar que en la parcela de 50 hectáreas aún cuando se construyeran subgrupos cada vez más pequeños para la familia de Moraceae, se formaría una innumerable cantidad de los mismos, lo que pone sobre relieve la gran similaridad entre sitios que albergan a especies de dicha familia, ya sea en función de la abundancia o de la riqueza. Este dendrograma recuerda la forma de una escalera (ver figura ##15).

### 3.3.2 Agrupamiento Aglomerativo por Enlace Completo

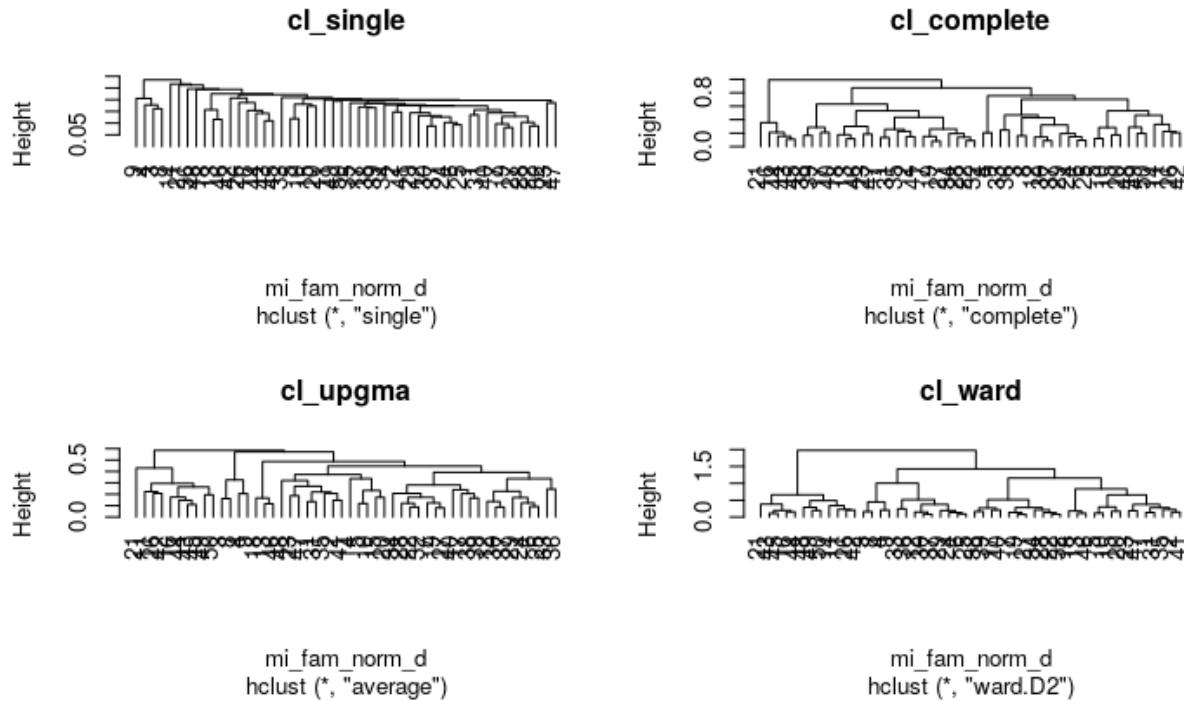


Figure 16: mapa de la isla Barro Colorado análisis interpretación dendrogramas

En esta imagen, aunque se ilustran los tipos considerados de dendrogramas (ver figura ##16), nos interesa el completo. En el agrupamiento aglomerativo por enlace completo, el criterio empleado es la mayor distancia o menor similaridad, el vecino más lejano. En este dendrograma se muestra una escalera pero con menor sucesión, se distinguen grupos de dendrogramas como en el caso anterior, solo que en esta ocasión los grupos con mayor distancia son poco numerosos en contraste con los los grupos que presentan la menor distancia, y la altura dominante para los cuadros mayores es de 0.8 en su generalidad.

### 3.3.3 Agrupamiento Aglomerativo por Enlace Promedio

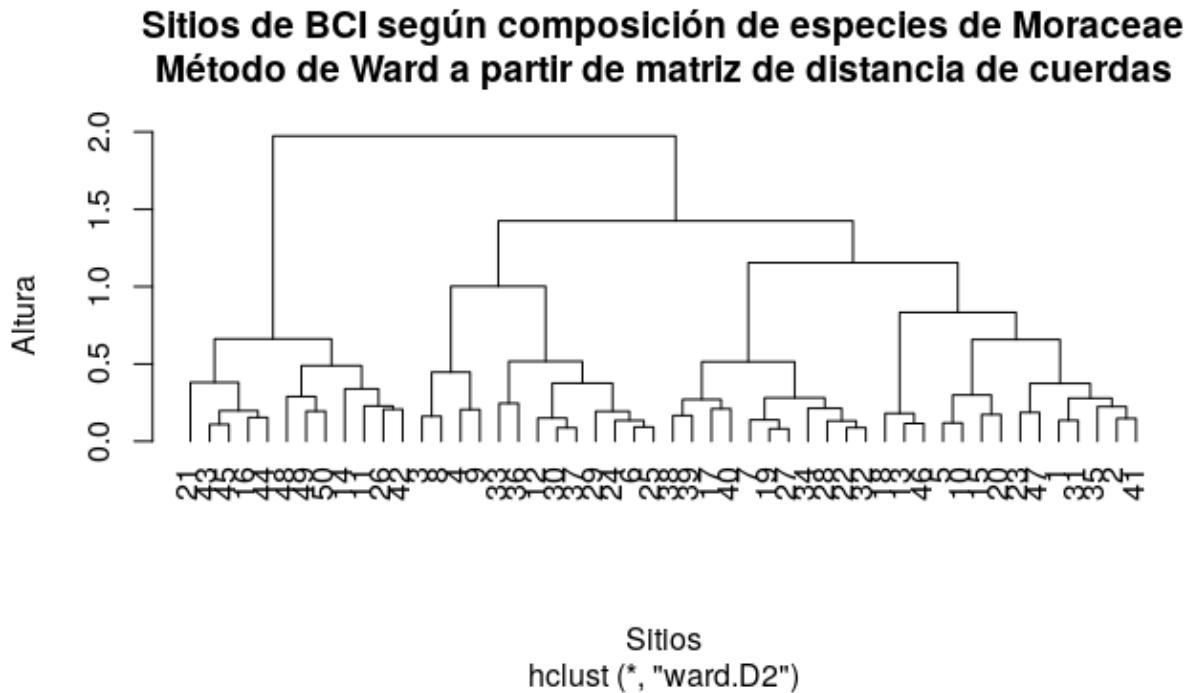


Figure 17: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

El método de Ward, aunque no es un método por enlace (ver figura ##17), lo elijo por su fácil comprensión (en este caso sería el UPGMA, para el caso de la media o promedio), recuerda la silueta de una pirámide toscamente dibujada, con una interpretación visual más factible que en los casos anteriores. Existe una regularidad en los subgrupos, lo que refleja, al igual que en las secciones anteriores una fuerte asociación, mayor similaridad que disimilaridad en las partes dominantes o más visibles. Un ejemplo de visualización lo ofrece el modelo de agrupación aglomerativa por enlace promedio, UPGMA, en el cual aparecen siluetas aglomerativas y repetitivas que reflejan una continuidad en la similaridad de los subgrupos.

### 3.3.4 Interpretación de Agrupamiento Jerárquico

En esta parte estamos a la altura de elegir métodos que nos permitan minimizar la mayor cantidad posible de unidades de grupos en los dendrogramas. Se establece la correlación cofenética y para la misma los métodos convenientes son los de Ward y el UPGMA. En esencia consiste en obtener la menor cantidad posible de siluetas que se distinguen por la existencia de pequeños grupos de figuras. En la siguiente imagen pueden apreciarse los cortes de las siluetas en tamaños menos complicados de analizar y no muy distantes en su dimensión, por lo menos en los dos primeros cuadros (ver figura ##18).

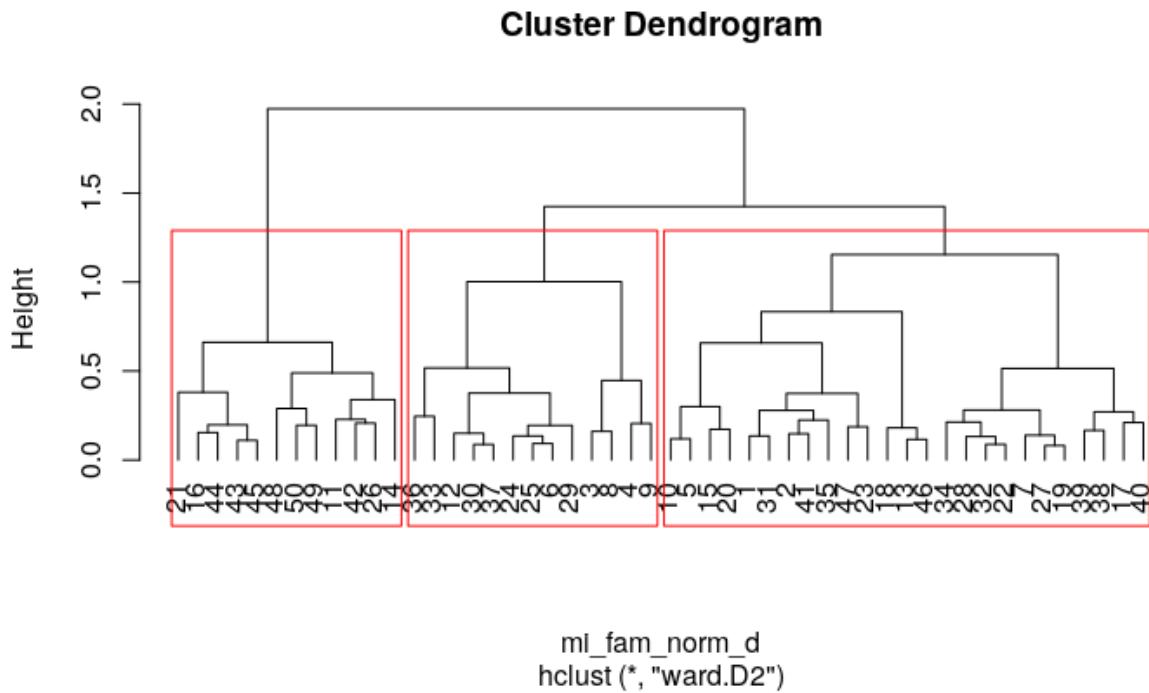


Figure 18: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

El método Ward es uno de los más ideales para elegir un dendrograma único que permita apreciar por simple inspección el tamaño de las siluetas, y el número de sitios que contienen, en este caso serían tres cuadros de 12 y 13 sitios para el primer y segundo cuadro respectivamente, y de 25 sitios para el tercero, todos a una altura de 1.3 aproximadamente dentro del dendrograma. En este dendrograma pueden apreciarse las siluetas más fácilmente que en los ejemplos anteriores.

Una forma de ilustrar mejor la disimilaridad y la similaridad entre los agrupamientos o dendrogramas es por medio del mapa de calor de la matriz de comunidad fusionado con el dendrograma. En esta ilustración se distinguen los sitios que presentan mayor grado de asociación. Puede verse que el dendrograma está cortado a la mitad, y los sitios que presentan mayor similaridad o menor distancia están al centro y en los extremos, por ejemplo, los grupos 13 y 18 presentan una fuerte asociación, y en general los grupos dispuestos de forma vertical presentan mayor similaridad que los dispuestos de forma horizontal (ver figura ##19).

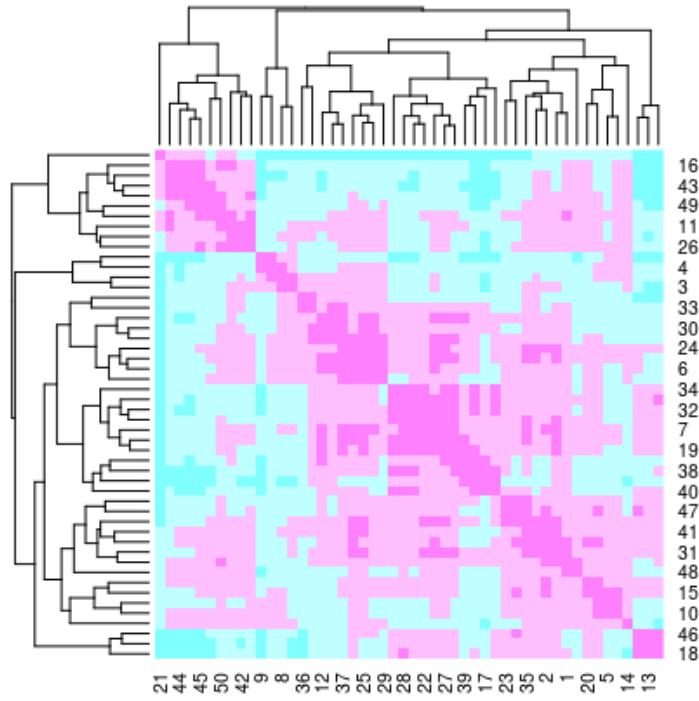


Figure 19: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

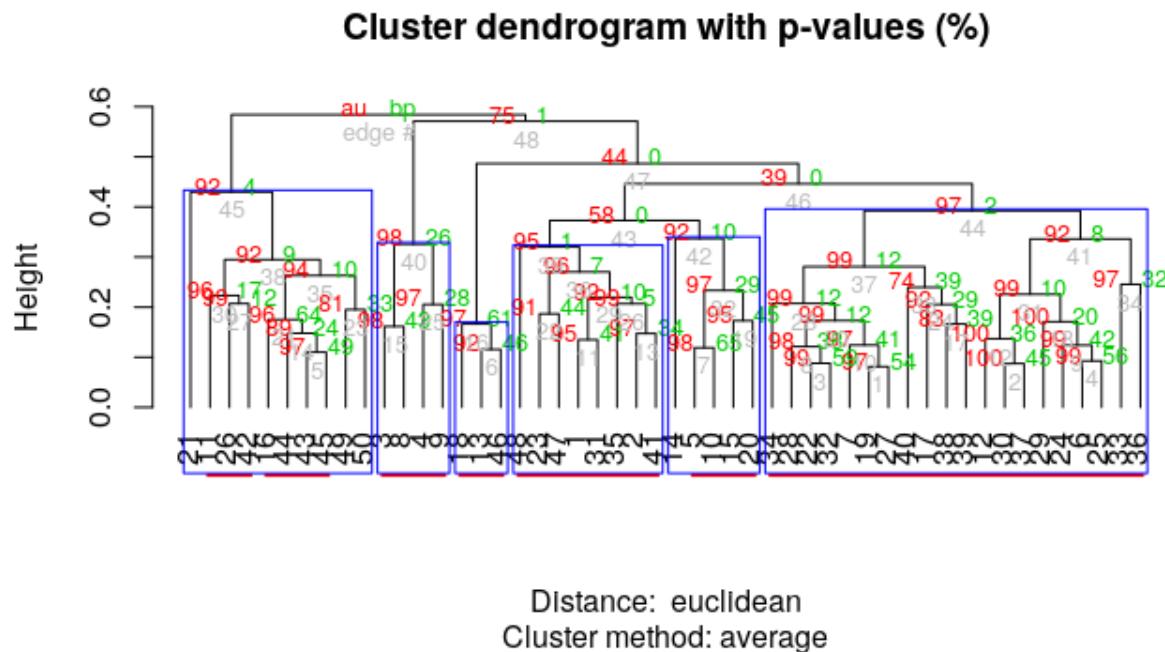


Figure 20: mapa de la isla Barro Colorado remuestreo multiescalar por Bootstrap

En el mapa anterior pueden verse las probabilidades, y la anchura de las siluetas de los cuadros que contienen los subgrupos con mayor similaridad, lo que refleja los posibles cortes que se podrían hacer sobre el dendrograma (ver figura ##20).

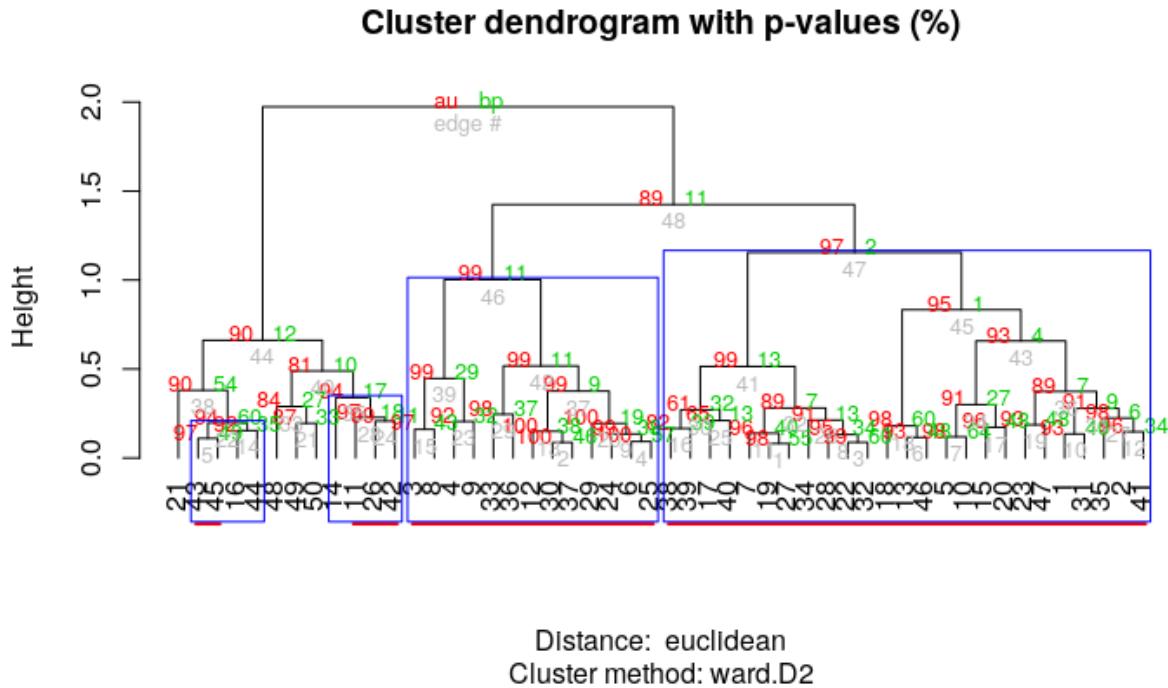


Figure 21: mapa de la isla Barro Colorado dendrogramas y porcentajes

En el dendrograma anterior, también pueden notarse las probabilidades por el método Ward (ver figura ##21)

Los dendrogramas, al igual que en los métodos de medición de asociación permiten ver gráficamente la dinámica de la distribución de las especies de moraceae, sin embargo, estos nos ayudan a segmentar en grupos los sitios con mayor cantidad de individuos o especies, dividiéndolos en cuadros a una mayor o menor altura, sería una forma de fraccionar en grupos representativos lo que en principio parece desordenado y monótono.

### 3.4 Análisis de Variables Ambientales y Mapas

Dentro del análisis de agrupamiento, se haya incrustado el estudio de las variables ambientales y los mapas. Desde un principio, en este artículo, tanto la riqueza global como la de mi familia, han presentado unos valores altos, y unos patrones de concentración y desplazamiento por igual grandes.

Así se puede visualizar en el siguiente mapa (ver figura ##22)

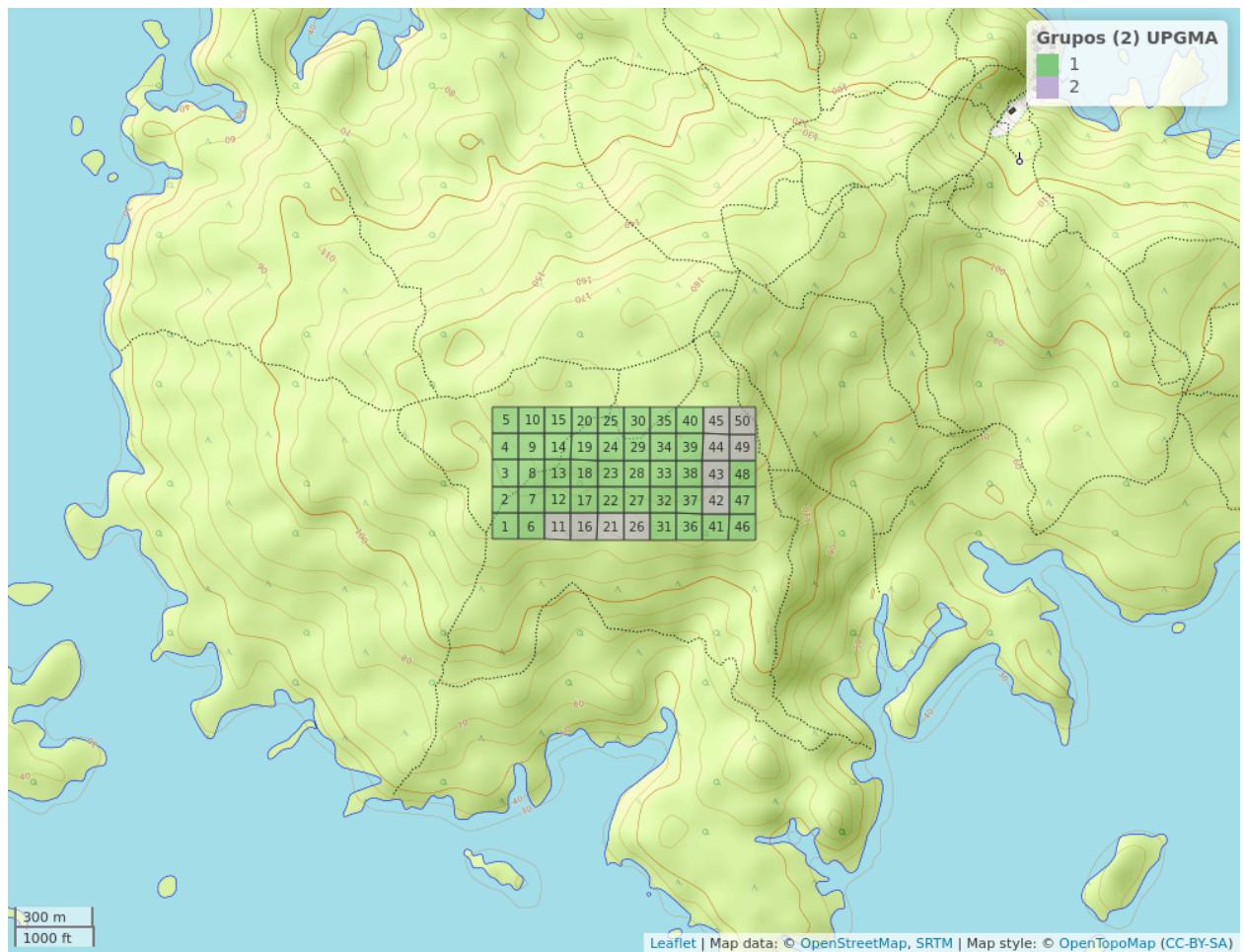


Figure 22: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

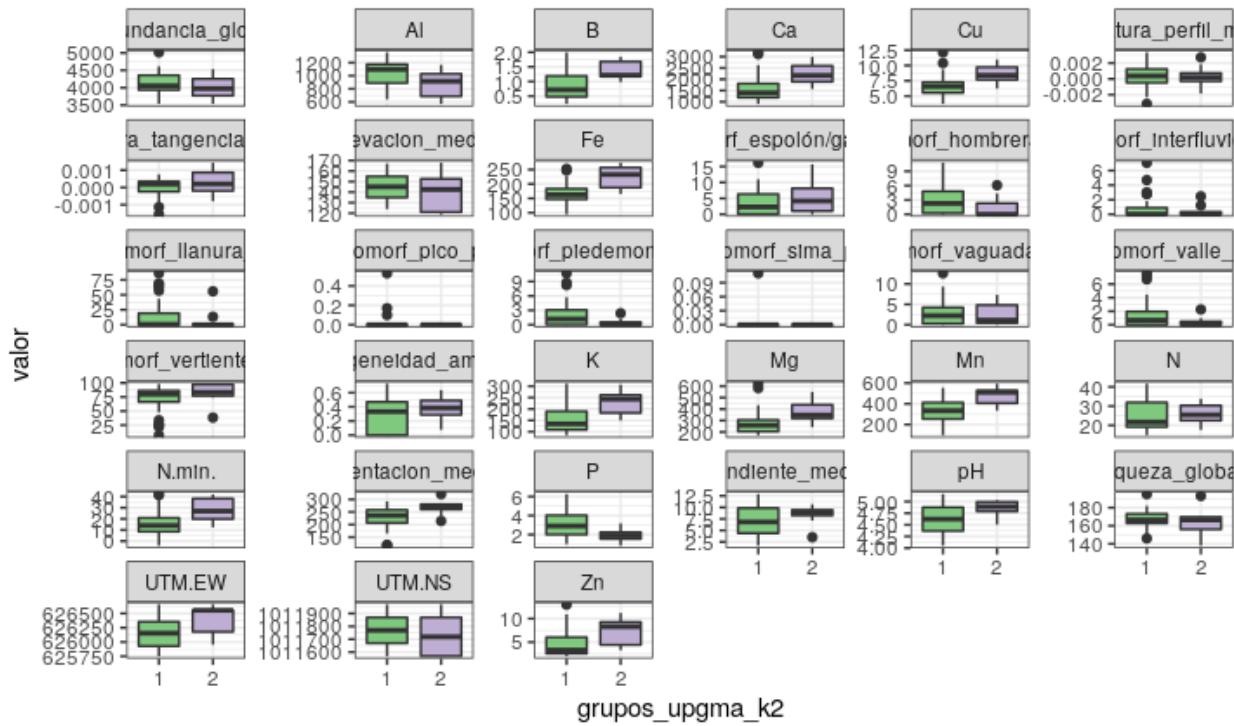


Figure 23: mapa de la isla Barro Colorado mapa de variables ambientales

En el mapa adjunto, así como en el diagrama de cajas (ver figura ##23), puede visualizarse el comportamiento de la mediana de las variables ambientales, de la riqueza, así como también de abundancia. El método UPGMA, presenta un patrón de distribución al este y al sur para el grupo 2 de las variables ambientales dentro de la parcela de 50 hectáreas. Las variables del grupo 1 de la parcela presentan un patrón en todos los rumbos. Dentro de la riqueza, y la abundancia, los grupos 1 y 2, presentan valores no muy distantes, lo que evidencia el condicionamiento ambiental de la isla, y que nuevamente refleja la fuerte riqueza y asociación de las especies de Moraceae en dicho espacio insular. Dentro de las variables ambientales de tipo geomorfológico, la mediana para el grupo 1, naturalmente presenta mayores valores que aquellos sitios de la misma parcela que se presentan en el grupo 2, pero en cuanto a la concentración de elementos químicos en un número de hectáreas más reducido (en este caso el grupo 2), representan mayor valor para la mediana.

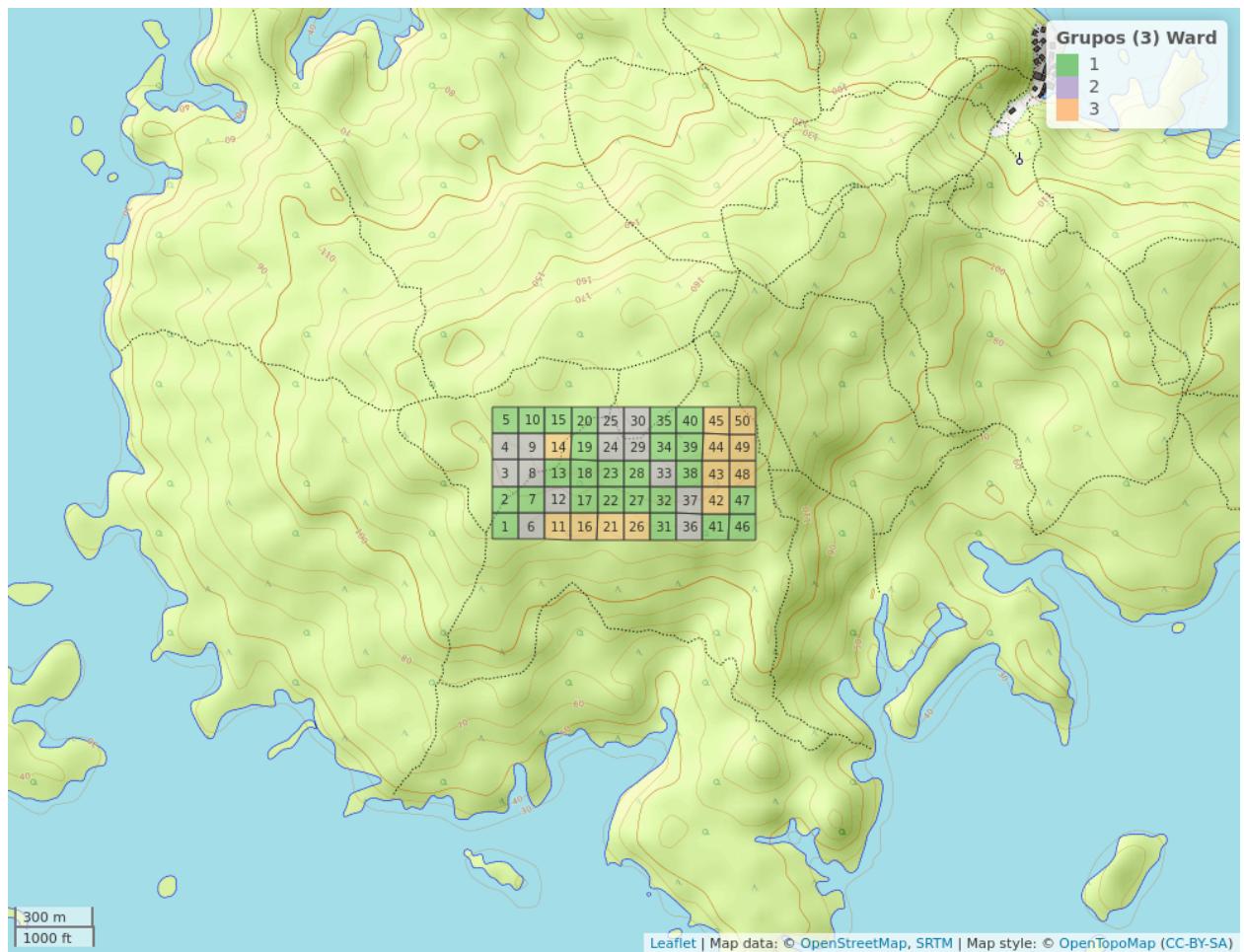


Figure 24: mapa de la isla Barro Colorado grupos Ward k3

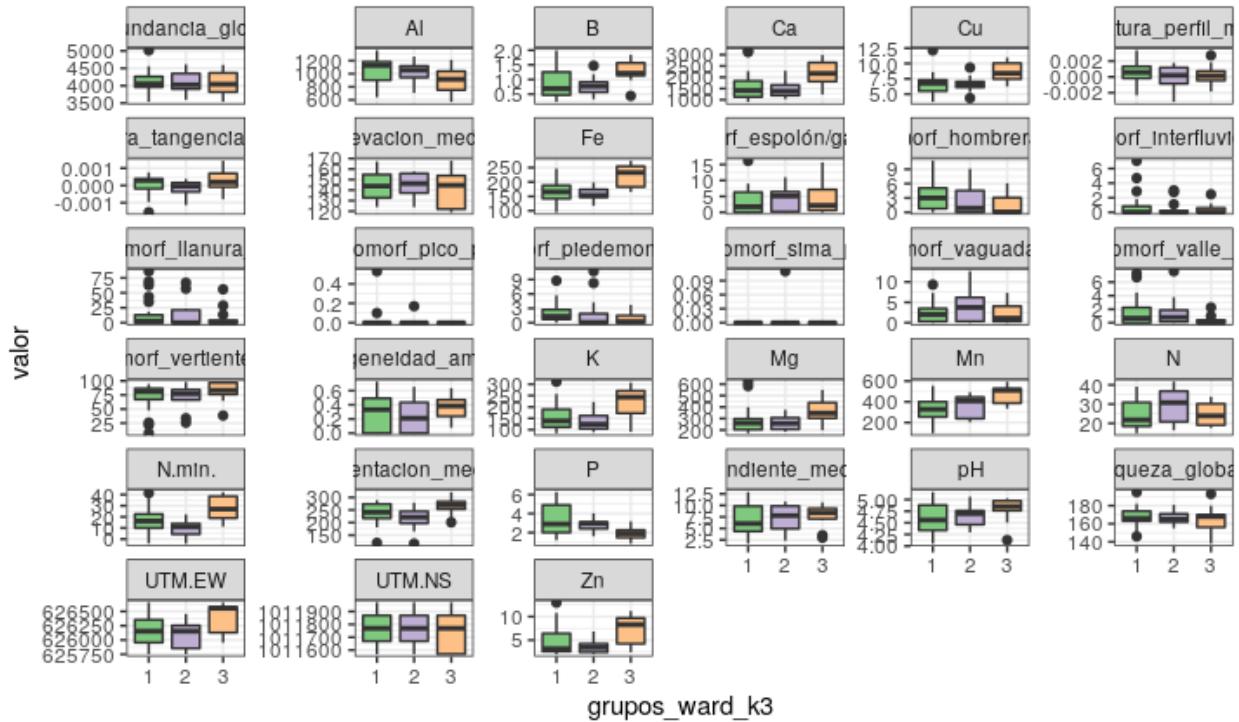


Figure 25: mapa de la isla Barro Colorado cuadro de cajas Ward

Para el método Ward en el comportamiento de las variables se presentan tres sitios que muestran la dinámica de las variables ambientales dentro de la parcela de 50 hectáreas con un patrón de distribución distinto al caso anterior. Cuando relacionamos las variables del diagrama de cajas (ver figura ##25) que también se aprecian en el mapa (ver figura ##24), puede corroborarse que estas se concentran más al centro de la parcela, algunas al sur, y otras al este. Variables como el pH, presentan altos valores de mediana en casi todas las hectáreas de la parcela. También ocurre lo mismo con el nitrógeno. Para el caso de la abundancia y la riqueza el valor de la mediana presenta valores elevados, aunque en el caso de la primera las cajas o sitios ocupan mayor espacio.

### 3.5 Especies Indicadoras

Para las especies indicadoras, tenemos que por especie indicadora se entiende el grado de predicción que esta tenga dentro de un espacio bajo estudio. Pueden darse casos de abundancia, y esos ejemplos de abundancia, nos pueden llevar a la concepción de grupos dentro de una familia. Dichos grupos tienden a identificarse por el predominio de una variable ambiental, ya sea un tipo de bosque, una concentración de pH, o de nitrógeno. La aparición de estos grupos bajo el condicionamiento de dichas variables es lo que hace posible la existencia ordenada de los mismos.

Existe el análisis de especies indicadoras por medio del método de Inval, este método se aplica tanto para UPGMA, así como para Ward, se verifica cuál es la especie que presenta el mayor indicador o estadístico, la abundancia y la permutación. Para UPGMA, en el caso de las Moraceae se registran dos grupos de una y cuatro especies respectivamente. En el caso del primer grupo sería *Perebea xanthochyma* con una abundancia de 226, y un indicador (Inval) de 0.32, tenemos en el segundo grupo: *Pulsenia armata*, cuya abundancia es de 993, y su indicador 0.774, *Ficus tonduzzi*

23 y 0.459, *Tropis caucana* 136 y 0.454, y *Tropis racemosa* con 218 y 0.420. En este primer grupo la especie indicadora es *Perebea xanthochyma* con los valores indicados y una permutación de 0.037, en el segundo grupo: *Pusenia armata* es la especie predictora de este grupo por la abundancia que presenta, con un valor de permutación igual 0.001. El intervalo de confianza es pequeño, por ejemplo para *Brosimum alicastrum* va de -0.3597668 (-0.36) a 0.2896983 (0.29).

En el caso Ward, se registran por igual dos grupos, el grupo 3 y el grupo 4 que resultaría de la suma del 2 y el 3 y que contiene una especie indicadora. Sus valores de de abundancia e indicador estadístico respectivamente son: *Poulsenia armata* 993 y 0.785, *Trophis caucana* 136 y 0.538, *Trophis racemosa* 228 y 0.435, y *Ficus tonduzii* 23 y 0.417.

Para el segundo grupo, tenemos la *Maquira guianensis* con 1315 y 0.575. Para el tercer y cuarto grupo, las especies indicadoras son: *Pulsenia armata*, y para el cuarto grupo *Maquiara guinensis* con valores de permutación de 0.001 para ambos casos. Para el caso Ward el intervalo de confianza es para la misma *Brosimum alicastrum*, desde su límite inferior con -0.1289961 (-0.129) hasta su límite superior de 0.41576092 (0.42).

### 3.6 Análisis de Diversidad Alpha y Beta

En este artículo se emplean dos modalidades de estudio: Alpha y beta.

Para la diversidad alpha sobre la familia en cuestión se muestran los resultados siguientes,

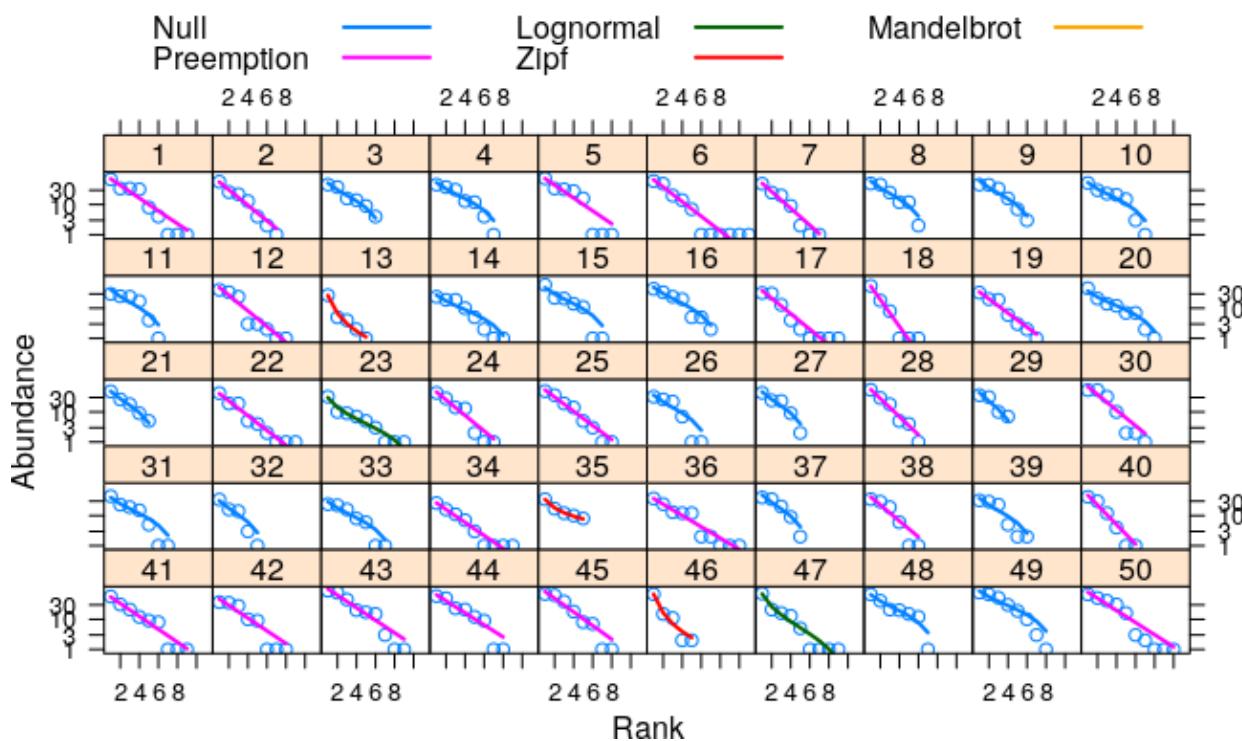


Figure 26: mapa de la isla Barro Colorado modelo abundancia de especies

La gráfica anterior (ver figura ##26) representa la abundancia por sitios, sitios de una hectárea, pero cada línea representa una métrica. El gráfico también presenta cuáles sitios podrían mostrar mayor equidad. Encontramos que las líneas de color verde y azul representan la mayor abundancia

de los lugares en los cuales se encuentran.

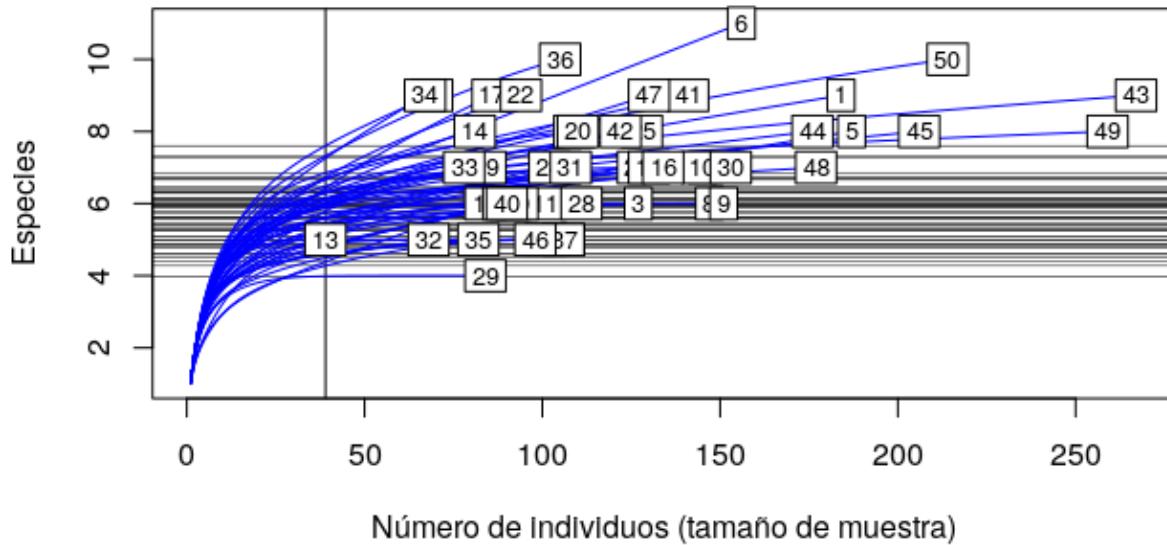


Figure 27: mapa de la isla Barro Colorado número de individuos

En el digrama anterior (ver figura ##27) de curvas de rarefacción puede apreciarse también el número de individuos por sitios. En este diagrama puede apreciarse que prácticamente la totalidad de de sitios poseen una cantidad de individuos superior a los 50 individuos, con excepción del sitio 13 que posee unos 45 aproximadamente, y un escaso número de sitios con una cantidad superior a los 200 individuos. La mayor cantidad de sitios comprende una riqueza en especies que va de los 50 a los 200 individuos.

### 3.7 Diversidad Beta

La diversidad beta se puede apreciar en siguiente mapa de cuadros (ver figura ##27), seguido de la gráfica (ver figura).

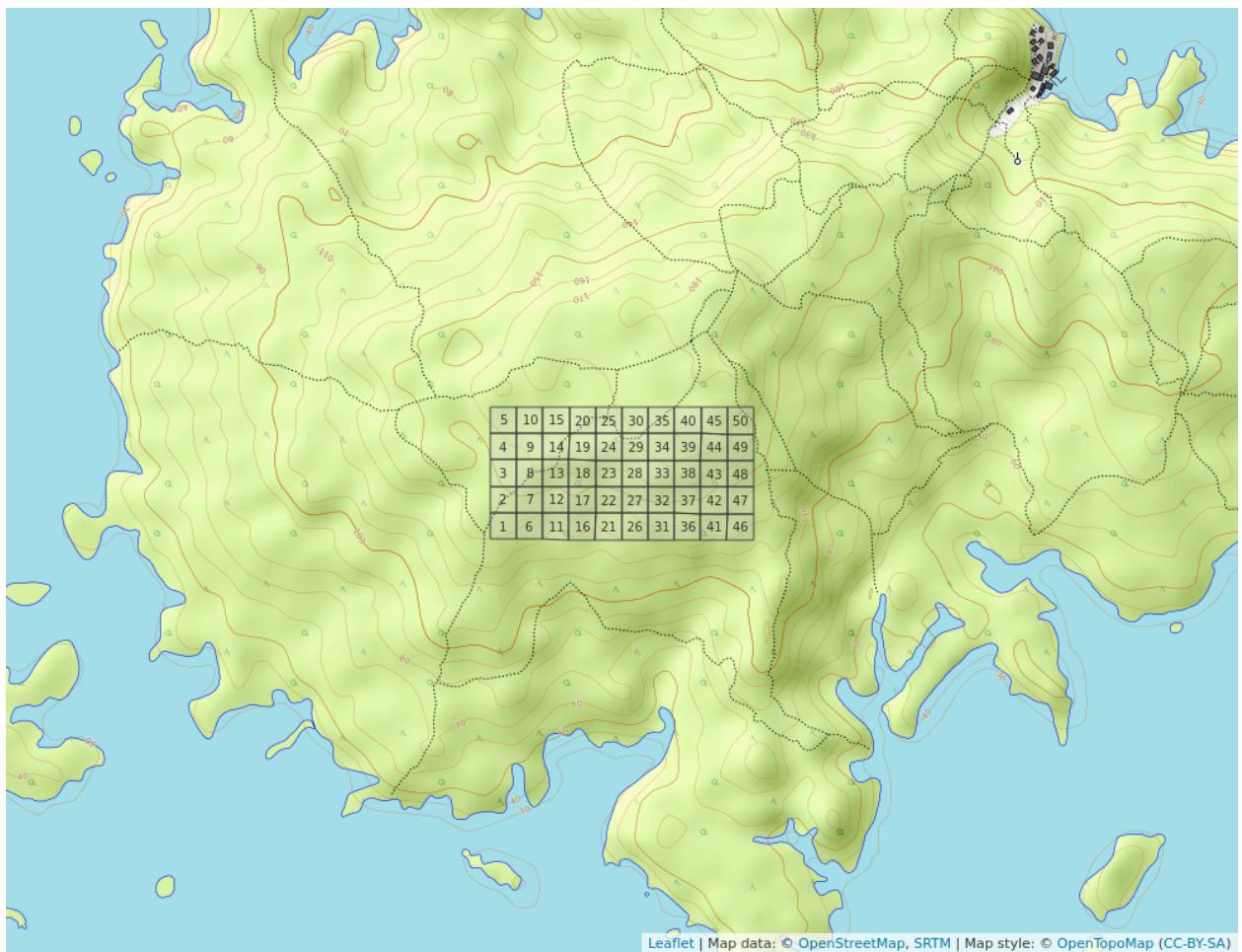


Figure 28: mapa de la isla Barro Colorado mapa cuadro diversidad beta

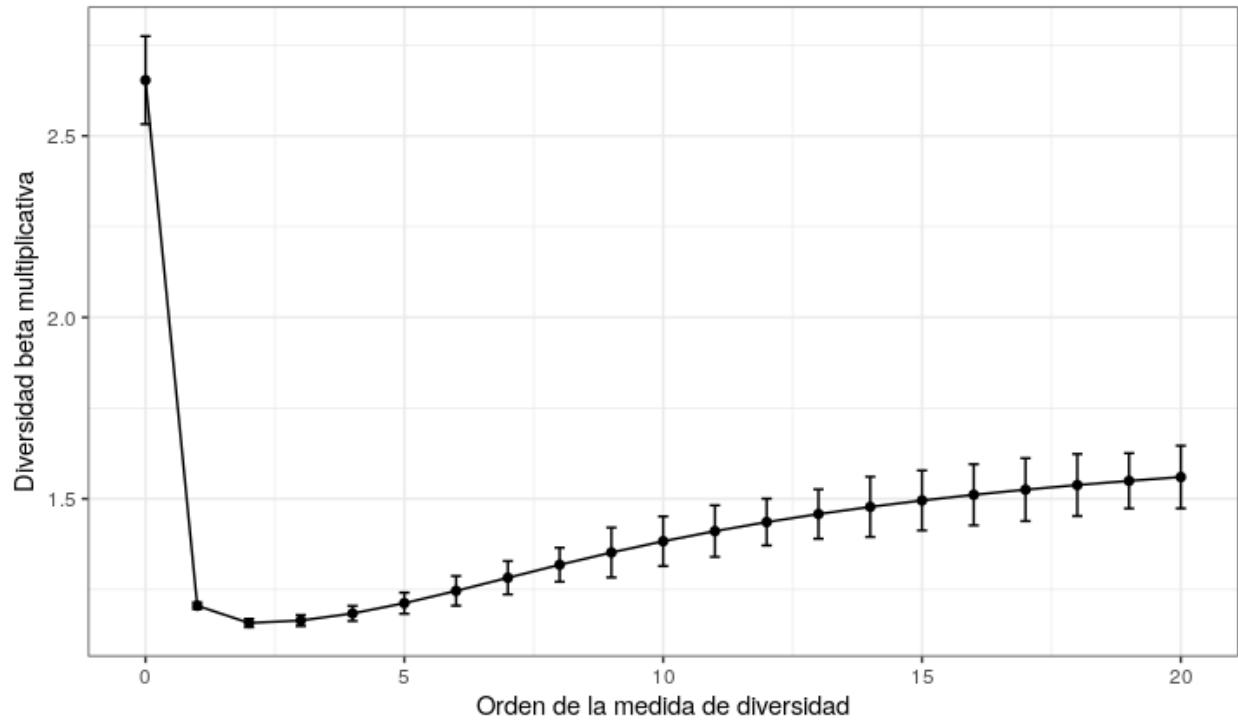


Figure 29: mapa de la isla Barro Colorado mapa de calor

Para este tipo de diversidad, se presenta un patrón de distribución al este de la parcela de 50 hectáreas y se aprecian las parcelas de una hectárea que poseen la mayor cantidad de individuos por especies, al oeste se aprecia una diversidad mucho más reducida. Este modo de visualizar la variabilidad es más comprensible que en los casos anteriores. En el gráfico (ver figura ##28) también puede apreciarse mayor diversidad al este de la parcela.

### 3.7.1 Análisis de Ordenación Simple (no restringida y restringida)

Las principales técnicas de análisis no restringido son: PCA (Análisis de Componentes Principales), CA (Análisis de Correspondencia), PCoA (Análisis de Coordenadas Principales), y NMDS (Escalamiento multidimesional No Métrico).

Así se muestran para la familia de moraceae, empleando la técnica de escalamiento PCA, la cual compara las componentes principales.

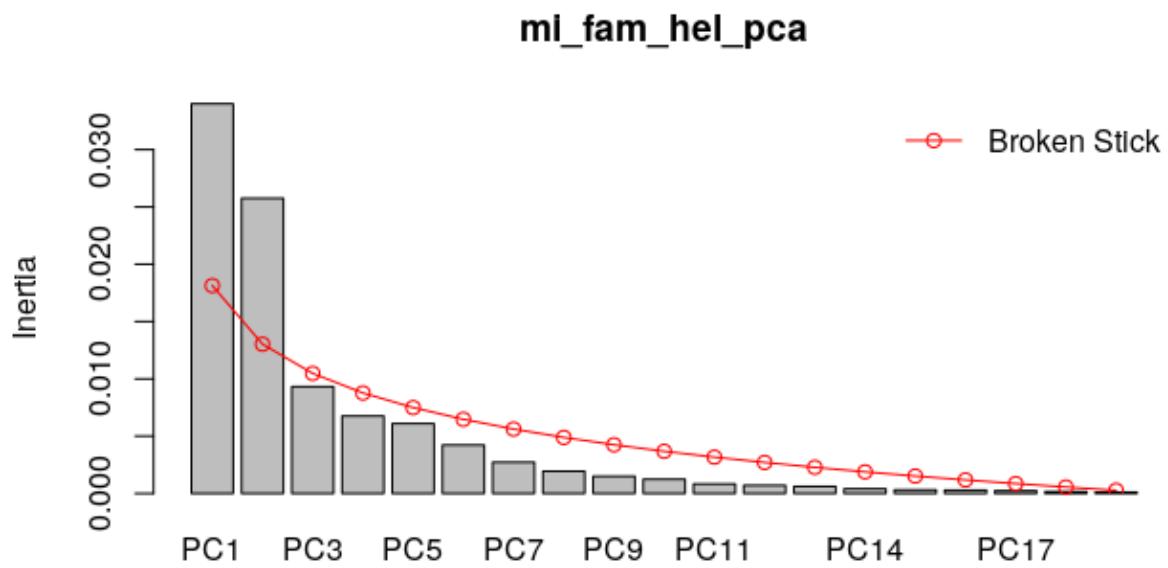
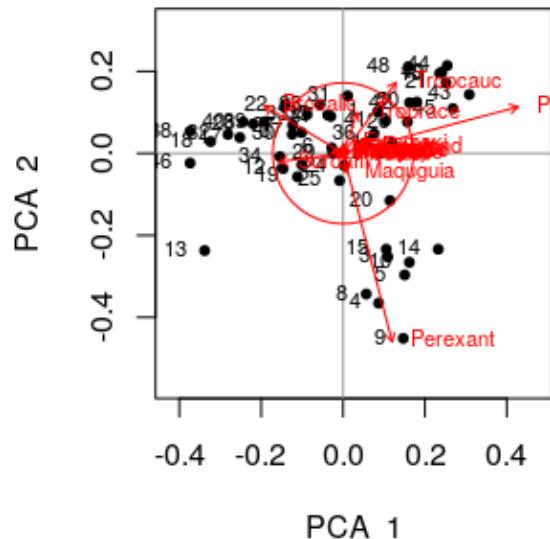


Figure 30: mapa de la isla Barro Colorado diagrama de componentes principales variables

En este diagrama de componentes principales se puede visualizar un decenso brusco de la medida de la inercia que pasa de 0.030 en la componente principal 1 a 0.00 en la componente principal 17 (ver figura ##29).

**Biplot PCA, escalamiento 1**



**Biplot PCA, escalamiento 2**

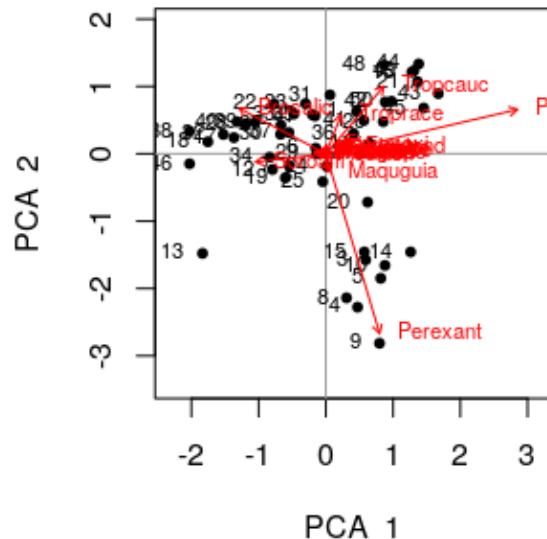


Figure 31: mapa de la isla Barro Colorado diagrama de escalamiento 1 y 2

En esta técnica de escalamiento pueden visualizarse los grados de asociación mediante un plano de coordenadas ortogonales. La cercanía que describen los ángulos de las variables ambientales representa el grado de asociación de las mismas. En el primer cuadrante (ver figura ##30) predomina la distancia euclídea, en el segundo, la distancia de Mahalanovich. Se forman nubes de puntos en los cuadrantes, estos puntos representan los sitios, y como se puede ver en los gráficos, estos se concentran más en el cuadrante negativo. Para ambos casos hay mayor asociación al norte del plano.

## PCA, escalamiento 2, ajuste a variables ambientales

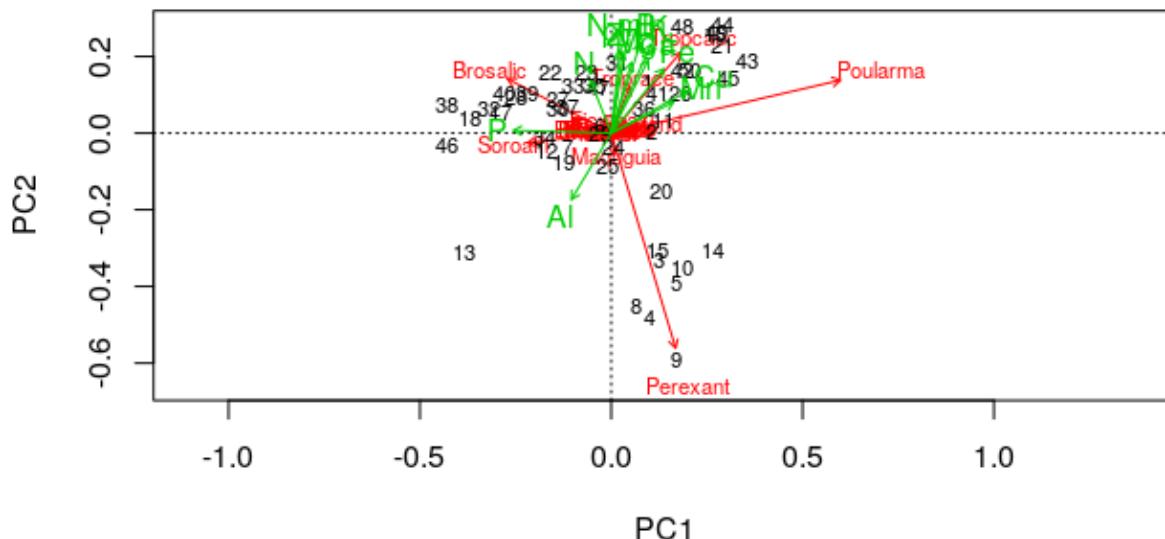


Figure 32: mapa de la isla Barro Colorado diagrama asociación de variables

En este diagrama (ver figura ##31) puede percibirse la fuerte asociación que existe en las variables ambientales (color verde), por la cercanía de los ángulos en la parte superior del plano. Hay mayor asociación en la componente principal 2 que en la componente principal 1.

### 3.7.2 Ordenación Simple Restringida:

Para la ordenación restringida, las tendencias asociadas a un grupo de ordenación se asocian a otro grupo. Las principales técnicas de ordenación restringida son: Análisis de redundancia (RDA), Análisis de redundancia basado en la distancia (db-RDA), Análisis de Correspondencia Canónica (CCA), Análisis Discriminante Lineal (LDA), Curvas de respuestas Principales (PRC), Análisis de Correspondencia conjunto (CoCA, Análisis de Correlación Canónica (CCorA) y Análisis de Inercia Conjunto (ColA).

En esta parte se mostrarán los resultados para RDA Y CCA. RDA combina la regresión y el análisis de componentes principales.

## ot de RDA especies ~ var. suelo, escala ot de RDA especies ~ var. suelo, escala

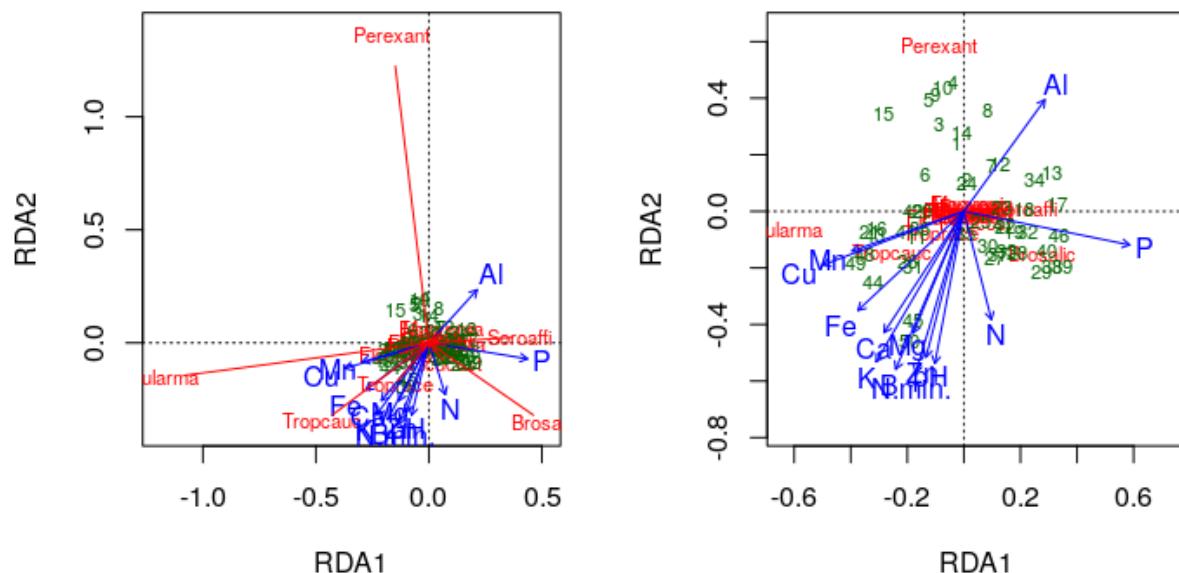


Figure 33: mapa de la isla Barro Colorado RDA diagrama de especies variables

CCA especies no raras ~ var. selec4, es CCA especies no raras ~ var. selec4, es

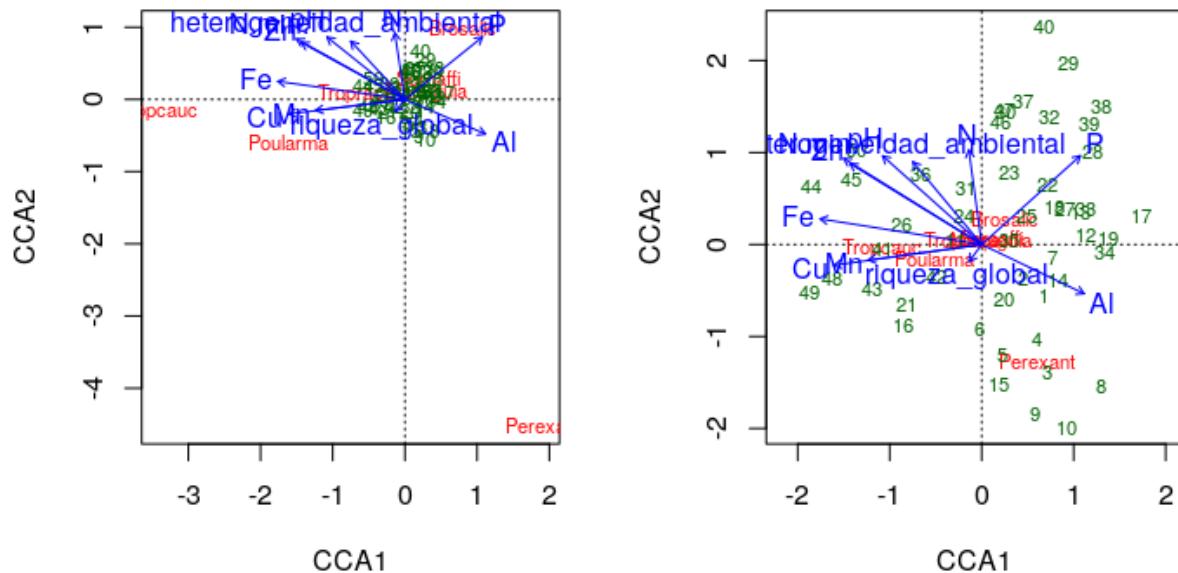


Figure 34: mapa de la isla Barro Colorado RDA especies no raras variables

## 3.7 Análisis de CCA especies ~ var. selec4, escala

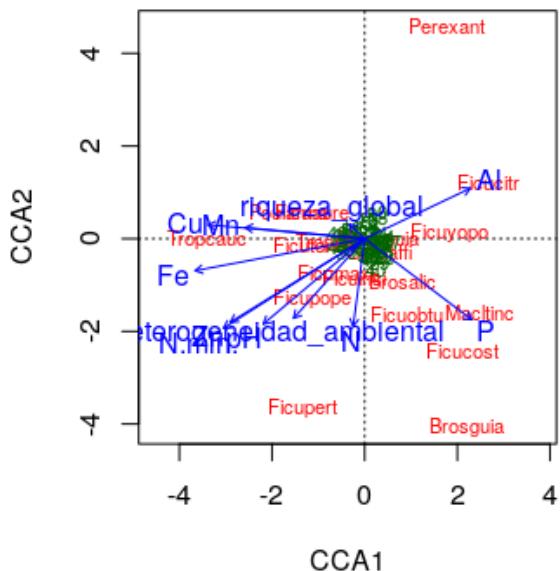


Figure 35: mapa de la isla Barro Colorado RDA digrama de especies variables

En estos planos bidimensionales puede apreciarse el comportamiento de las variables ambientales incluyendo el recurso suelo, de las especies, y de especies no raras. en el primer diagrama (RDA) (ver figura ##32) puede verse la fuerte asociación de los elementos químicos que integran el suelo, descrita por las líneas azules que representan los ángulos en el tercer cuadrante del plano. En el digrama de especies no raras (CCA) (ver figura ##33) se percibe una fuerte concentración de sitios en el tercer cuadrante, y es precisamente allí donde existe una menor asociación entre elementos químicos (variables ambientales). Por ejemplo el fósforo está presente en la mayor cantidad de sitios, y es un elemento que se encuentra a una distancia angular considerada con relación a los demás, lo mismo pasa con el aluminio. Para el cobre, el manganeso y el hierro ocurre todo lo contrario.

En el último diagrama (ver figura ##34) correspondiente a CCA, existe mayor asociación entre las especies en las zonas con predominio de fósforo y aluminio, correspondiente a los cuadrantes 1 y 2 del plano.

### 3.8 Análisis Espacial de Datos Ecológicos (Autocorrelación)

Esta es una forma de observar los valores de las variables Ambientales desde una perspectiva espacial, por medio de segmentos verticales, parecidos a bastones, encerrados en cajas, se pueden ver de una forma más comprensiva la dinámica del valor de correlación de las especies de Moraceae que se vienen tratando, y que se pueden apreciar en el gráfico o correlograma (ver figura ##35).

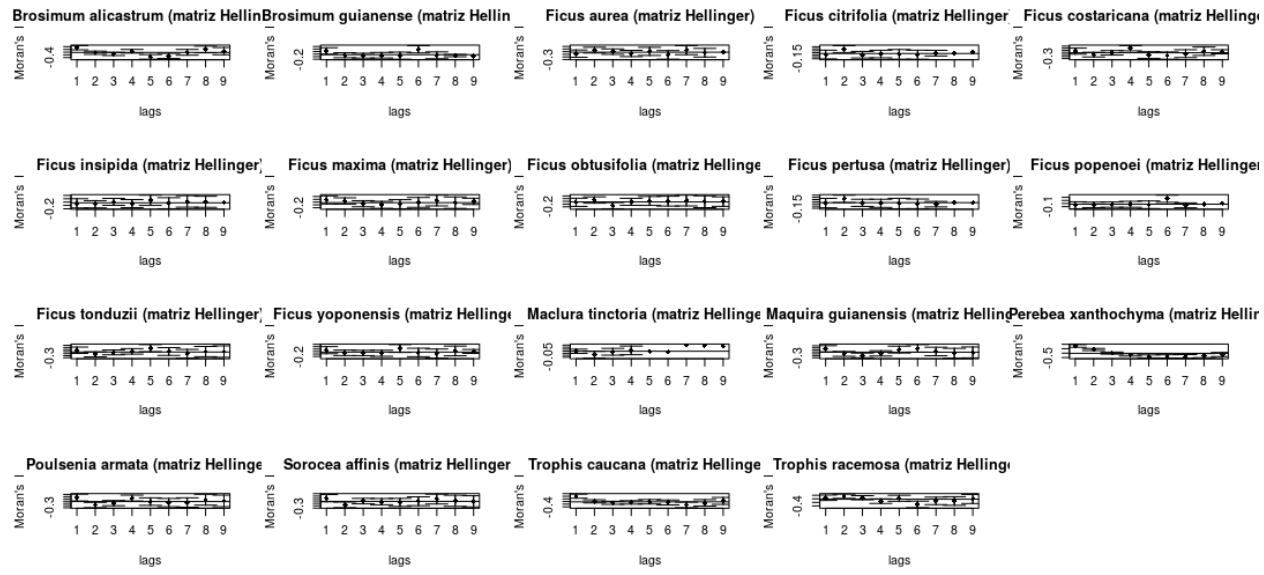


Figure 36: mapa de la isla Barro Colorado RDA diagrama de especies variables

A partir de esta pequeña descripción del correlograma se puede entender la dinámica espacial de algunas especies de Moraceae. Como puede verse existe una correlación más positiva que negativa entre las distintas especies de moraceae. Lo que nuevamente viene a recrear la gran riqueza predominante en las distintas especies de Moraceae, la similaridad qe se hace visible en cada modo de estudio empleado.

#### 4 Discusión

La abundancia y riqueza de la familia Moraceae, dentro de la isla Barro Colorado, es vieja y asentada en terreno bajo, en su generalidad, con una adaptación bastante ajustada a las condiciones de los factores físico químicos de dicha isla, por lo general son plantas perennes, de gran porte que sirven de alimentación a los individuos frugívoros que allí habitan, constituyendo con su dosel el típico paisaje de la selva tropical crentroamericana.

Como se pudo ver en los mapas, y los cuadros de variables, se presenta gran riqueza y abundancia en espacios muy reducidos. Todo esto nos hace pensar que se trata un gran bioma en miniatura en el que se reducen las distancias, y se incrementan los ejemplares que hacen confluir la biología y la geografía. Tal vez en el aspecto geomorfológico no se puedan hacer conjjeturas similares. Pero desde la biogeografía, el sendero es complejo. El hecho de que exista este conglomerado de especies de plantas y animales, requiere de unas condiciones óptimas para hacer posible la vida de este gigantesco ecosistema. En los mapas de abundacia, tanto global como de la familia se advierte un crecimiento casi perfecto, bajo condiciones físicas y químicas que le son favorables a la vida en esta parte del mundo.

No se ignora que a la misma latitud se encuentren condiciones similares en otros rincones de la tierra, de hecho las zonas intertropicales, las tropicales y las ecuatoriales se caracterizan más por la riqueza que por la abundancia. Pero en ese sentido tenemos millones de kilómetros bajo estudio si nos adentraramos en las grandes selvas que se conocen de antiguo, y tal vez nos toparíaamos

más con abundancia que con riqueza. Las grandes expediciones de la historia se han distinguido siempre por cubrir amplios itinerarios, que a su vez implican grandes distancias.

Dentro de la selva del istmo, BCI, es como un extracto sintético susceptible de múltiples investigaciones, por su amplia concentración de flora y fauna. Si bien la isla Barro Colorado es un enclave marítimo, y de investigación científica, contando por ende con las riquezas antes descritas, es motivo de cuestionamiento, el cómo un espacio, esta pequeña isla, haya acumulado estas riquezas en tan poco tiempo, a sabiendas de la zona geoastronómica en que se encuentra. Sería más aceptable la abundancia.

La construcción de riquezas ecológicas, ademas del emplazamiento geográfico, requieren de eventos geológicos, geográficos y tal vez antrópicos, en este caso tenemos el Canal de Panamá, el recuerdo mejor conservado de la injerencia yanqui en el Caribe. Una isla que se formó en algo más de un siglo, pues de la misma forma corre el riesgo de desaparecer. Pero esto solo compete a un asunto geológico, que no concierne a este trabajo, solo de manera parcial.

En el caso de la pendiente, esta es favorable para formaciones boscosas como la que he asumido, aunque inservible para la agricultura bajo cualquier consideración que se haga a favor de la vegetación que no sea referente a la anteriormente citada. Sin embargo la altitud guarda una relación acorde con las dimensiones de la isla. Las plantas como las Moraceae mantienen una relación favorable con estos grados de inclinación del terreno. Ya en espacios controlados y con plantaciones con las qe se persigue un determinado propósito se tiene un conocimiento sobre el manejo de la pendiente, a sabiendas de que en este escenario han crecido de forma silvestre sin intervención alguna de la mano del hombre.

La isla constituye uno de los ecosistemas más grandes del mundo, lo que deja al descubierto la gran adaptabilidad que presentan algunas plantas a este medio. Por tratarse de un lugar sumamente pequeño, los rasgos orográficos han de ser casi imperceptibles, pero sobre todo la edad de la isla no la hace poseedora de formas del relieve que estén consolidadas dentro de lo qes es la periodización geológica. Pues la isla se formó hace algo más de un siglo, tiempo en el qe no caben eventos geológicos trascendentales.

BCI constituye como una especie de interrupción del sistema orográfico de la región istmica panameña y centroamericana, pues la espina dorsal de toda la costa del Pacífico, aunque no se eleva de forma tan imponente como en los Andes o en las Montañas Rocosas, en Centro América, poseen continuidad, aunque con una escasa altitud. Es como si se tratara de un lugar construido para la investigación, en medio de sistemas tan extremos, es decir franja estrecha de tierra y masas imensas de agua que aguardan a ambos lados del Canal de Panamá (situación ístmica).

La familia de plantas que nos interesa sobreabunda en estas condiciones. Muy probablemente las familias de Moraceae presentes en otros lugares prosperen bajo condiciones muy diferentes. Sin embargo los lugares a una misma latitud en el planeta, se caracterizan por operar sobre los individuos bajo las mismas condiciones, por ende las condiciones de pH en otras regiones tal vez no se alejen tanto de las que por el momento se citan.

El pH es un factor que en lo adelante incidirá en la utilidad de la planta, especialmente a nivel comercial, la presencia de uno o varios elementos, o la ausencia de estos determinará el valor utilitario de estas. La quimiotaxonomía o qimiosistemática evalúa la presencia de elementos químicos en especies vegetales, el aspecto químico de la clasificación de las plantas, se basa en sus constituyentes es decir en sus características moleculares. Estas al igual que las características geomorfológicas son controladas químicamente.

Dentro los factores físico químicos las medidas de pH no suelen presetar cambios significativos, en las inmediaciones de las aguas del Canal de Panamá, pese a que se presentan ligeras fluctuaciones en las medidas de este indicador, que en esencia muestran la sensibilidad de las aguas circundantes, este por lo general es regular (Simmonds, Gómez, & Villalaz, 2002,). Esto sugiere un

grado de adaptabilidad bastante desarrollado por las especies de la familia en cuestión, no solo al medio de la isla, sino tambien de las áreas circundantes. La fenología reprodctiva de la Isla de Barro Colorado ha sido descrita extensamente por algunos autores. Los árboles más grandes presentan un pico de floración entre febrero y junio, y alcanzan un máximo en marzo y abril, justamente al final de la estación seca (Williams-Linera & Meave, 2002,). Tal vez este régimen de floración y de adaptación permitan a las Moraceae mantener una fenología por encima de los patrones físico químicos, en este caso especialmente el pH.

Como pudo verse en los demás bloques que corresponden a las mediciones de asociación, análisis de agrupamiento jerárquico, diversidad, ordenación, y análisis ecológico espacial. La fuerte asociación prevalece en las especies de Moraceae. Algunas variables ambientales como el pH mantienen sun primacía en su modalidad ácida. Pero en todo este trabajo, lo que generalmente prevalece es la similaridad entre sitios que contienen especies, denotado en los diferentes métodos empleados y en las matrices descritas, existe una fuerte cohesión entre individuos similares y abundantes en espacios sumamente reducidos.

Las Matrices de Hellinger, de Spearman, así como la distancia de Jaccard, vienen a hacer síntesis de los niveles de asociación predominantes en la parcela, determinando la disimilaridad, y la similaridad y los contrastes que existen cuando se emplean otros métodos.

Nuestro interés se centra, en esta parte, en las condiciones que han creado este espacio insular tan singular, en un tiempo tan breve, en el istmo de Panamá. Si bien sabemos que América es la masa continental que alcanza mayor extensión latitudinal, es probable que en todo el tiempo precedente a la construcción del canal, se dieran las migraciones de especies tanto del norte como del sur. Aunque, una vía interoceánica artificial como lo es el Canal de Panamá, probablemente no provocaría cambios notables en el equilibrio biológico de las comunidades de individuos que poblaban esta parte del continente.

La sostenibilidad de las especies de plantas y animales que allí existen no se puede poner en duda, por las condiciones climáticas que en dicho lugar se mantienen. Es como si tratara de un ambiente controlado y planificado. De modo que las investigaciones siempre tienen una página que arañar, dando motivos de colgarse una mochila y salir a hacer trabajo de campo. Es muy probable que para muchos, BCI siga siendo el secreto mejor guardado dentro de las regiones de investigación más llamativas del mundo,lo cual resulta ser asombroso por el peso que posee el istmo de Panamá a escala comercial.

El presente artículo ha sido un intento por abordar la información básica de la familia de las Moraceae, desde una perspectiva científica, teniendo como escenario la isla de Barro Colorado. En el desarrollo me he volcado más en lo que es la distribución de individuos por hectárea, en una parcela de 50 hectáreas, tanto desde una perspectiva global, como específica, en las condiciones del terreno, así como en la situación geográfica, pretendiendo dar un enfoque lo más abarcador posible, siguiendo estas variables, que han de ser de las más pertinentes para la consideración de un trabajo de campo.

Sabemos que esta familia de angiospermas posee gran riqueza y abundancia en esta parte del istmo, y que las condiciones típicas del suelo son idóneas para su sostenimiento. En un principio mencioné el caso de otras regiones próximas, por sus estadísticas en cuanto a predominio de plantas, especialmete de la familia que estoy tratando.

Panamá es uno de los países más diversos, ya que tiene el 10% de la fauna y flora global. México tiene el 11%, pero Panamá cabe en México unas 26 veces. Panamá tiene 1010 especies registradas, de esa cantidad un tercio le corresponde a la isla Barro Colorado. El bosque de Barro Colorado se cierra muy abajo y es difícil ver algunas especies de aves, solo se pueden escuchar. Las especies de Moraceae por lo general corresponden a árboles de hojas anchas (latifoliados), lo que hace de los suelos de estas especies, lugares prácticamente impenetrables por la radiación

solar, algo típico de la selva en cualquiera de sus denominaciones. En el bosque la temperatura promedio es de 27 grados centígrados (80° fahrenheit), y la humedad es bastante alta.

Existen muchos lugares en el Mar Caribe que como BCI, pueden ser grandes laboratorios, a partir de los cuales se pueda hacer ciencia, producir material científico de utilidad, pero la dependencia es peor que el analfabetismo. Por ello tal val tenemos una ciencia estacionaria, incapáz de dar el santo de la libertad.

Y se espera que el caudal de inforamción que se genera en las investigaciones alcance a estos lugares, porque la ciencia genera identidad a los lugares y los hace crecer. BCI, tiene algo más de 100 años, pero Panamá como nación es más vieja, y no es mi incumbencia el tema político, ni de este artículo. Pero hago la aclaración porque entiendo que se puede crecer en la ciencia, siempre dejando claro que cada pueblo tiene contextos distintos y que el error siempre será tratar de copiar. Este ha sido un trabajo en virtud del interés biológico, y geográfico. Esperando que sea del agrado de quienes puedan leerlo.

## 5 Agradecimientos

Quiero dedicar todo mi agradecimiento, primero a Dios porque es quien hace posible todas las cosas, luego al maestro José Ramón Martínez Batlle, maestro de la escuela de ciencias geográficas en la facultad ciencias de la Universidad Autónoma de Santo Domingo, en las áreas de biogeografía, geomorfología, y anteriormente de geología por enrutarlos en este mar de investigación que siempre termina arrojando cosas nuevas. Por último a mis compañeros de carrera, Carolain Pérez Ureña, Ana Hilda Valera, Darihanna Linares y Wilson Robles Rosario, sin ellos esto no habría sido posible, muchas gracias.

## 6 Información de soporte

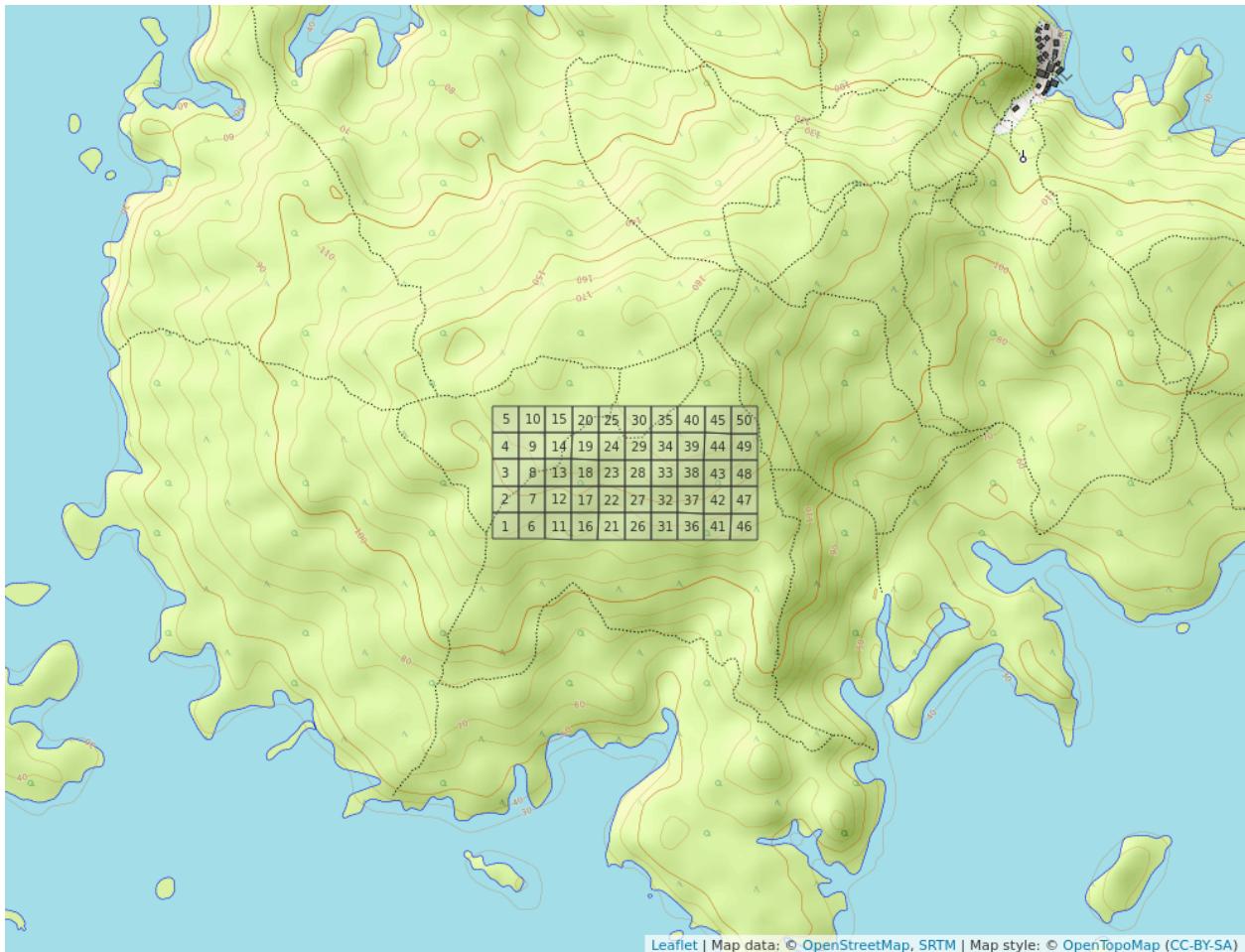


Figure 37: mapa de isla Barro Colorado cuadros

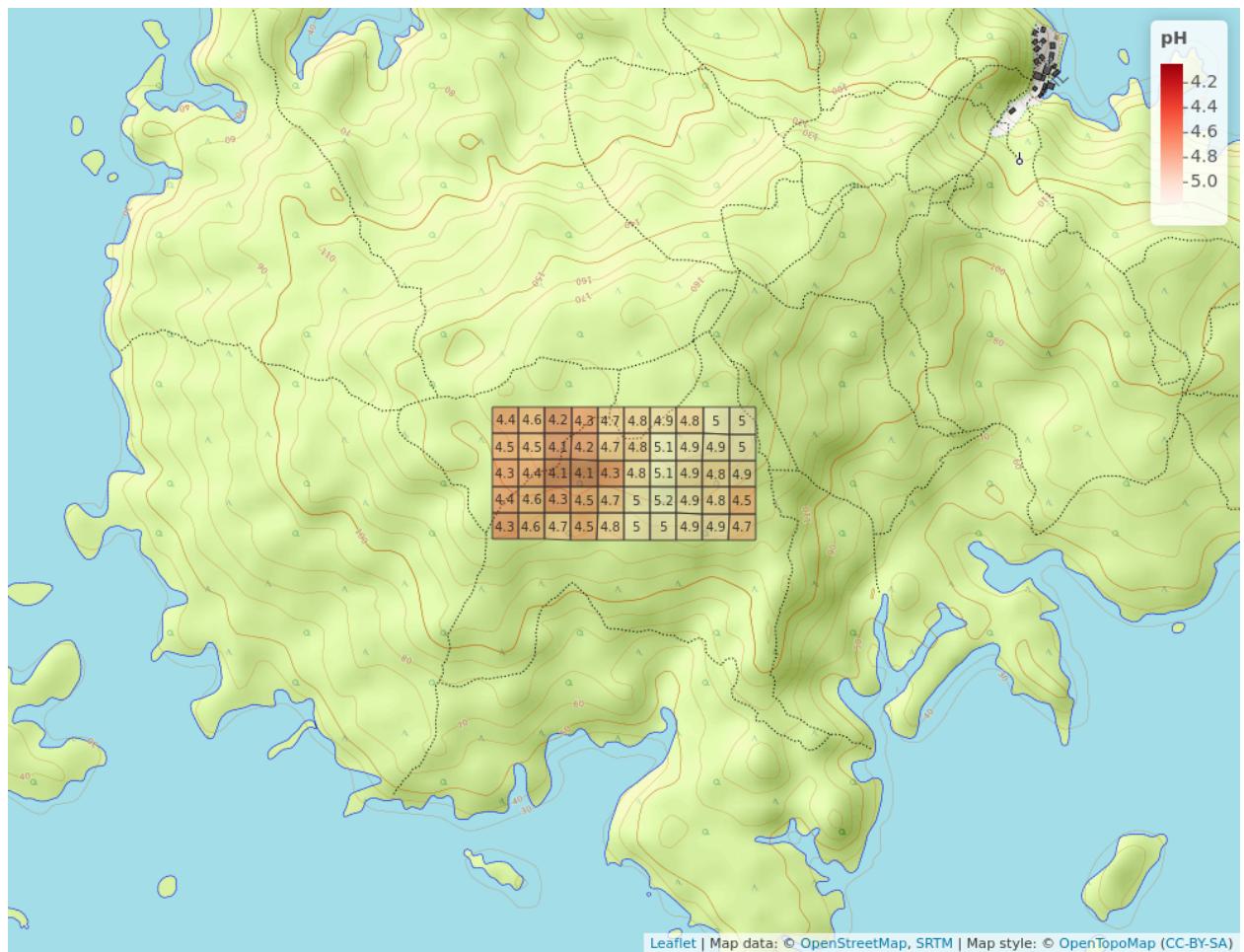


Figure 38: mapa de la isla Barro Colorado cuadros ph

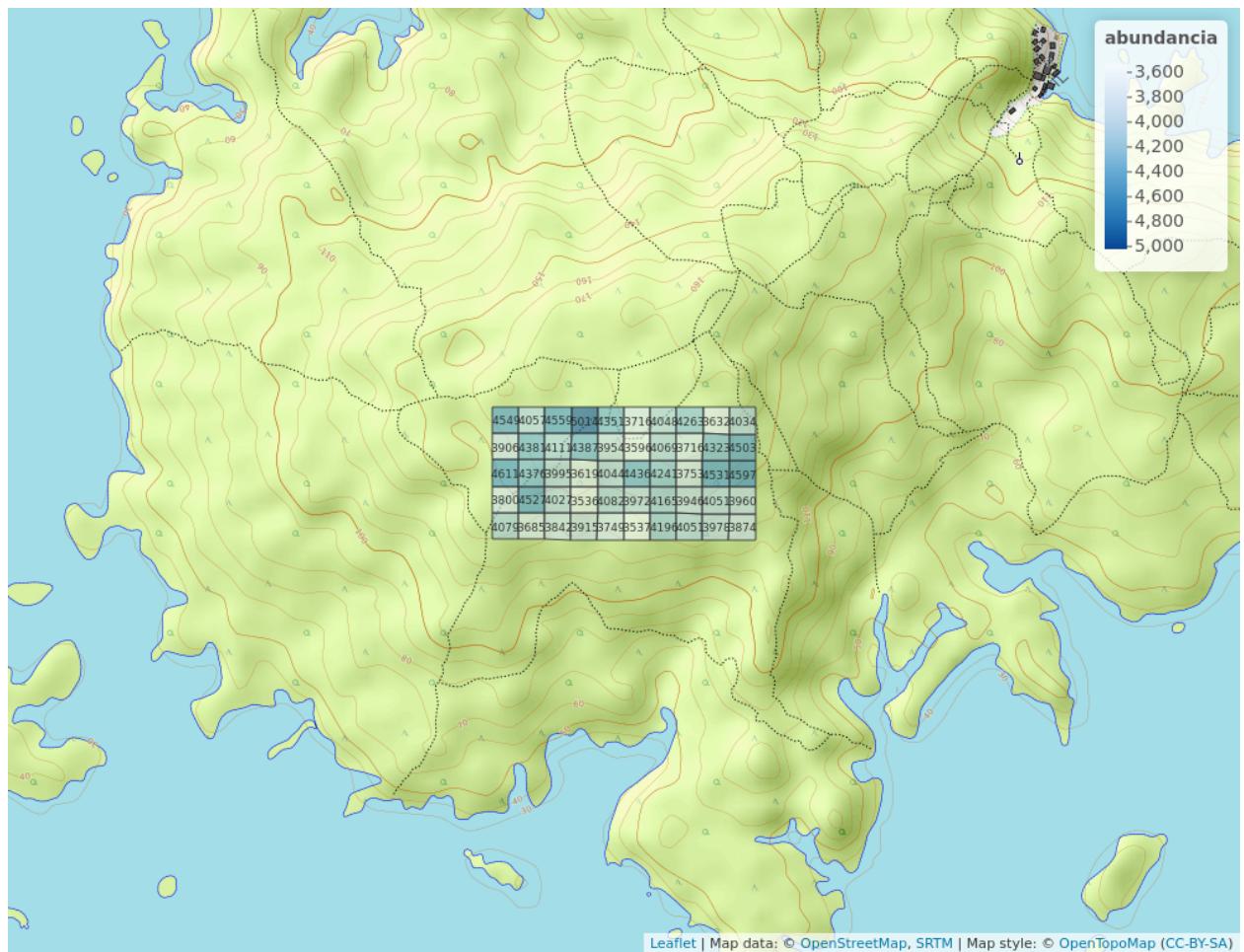


Figure 39: mapa de la isla Barro Colorado abundancia global

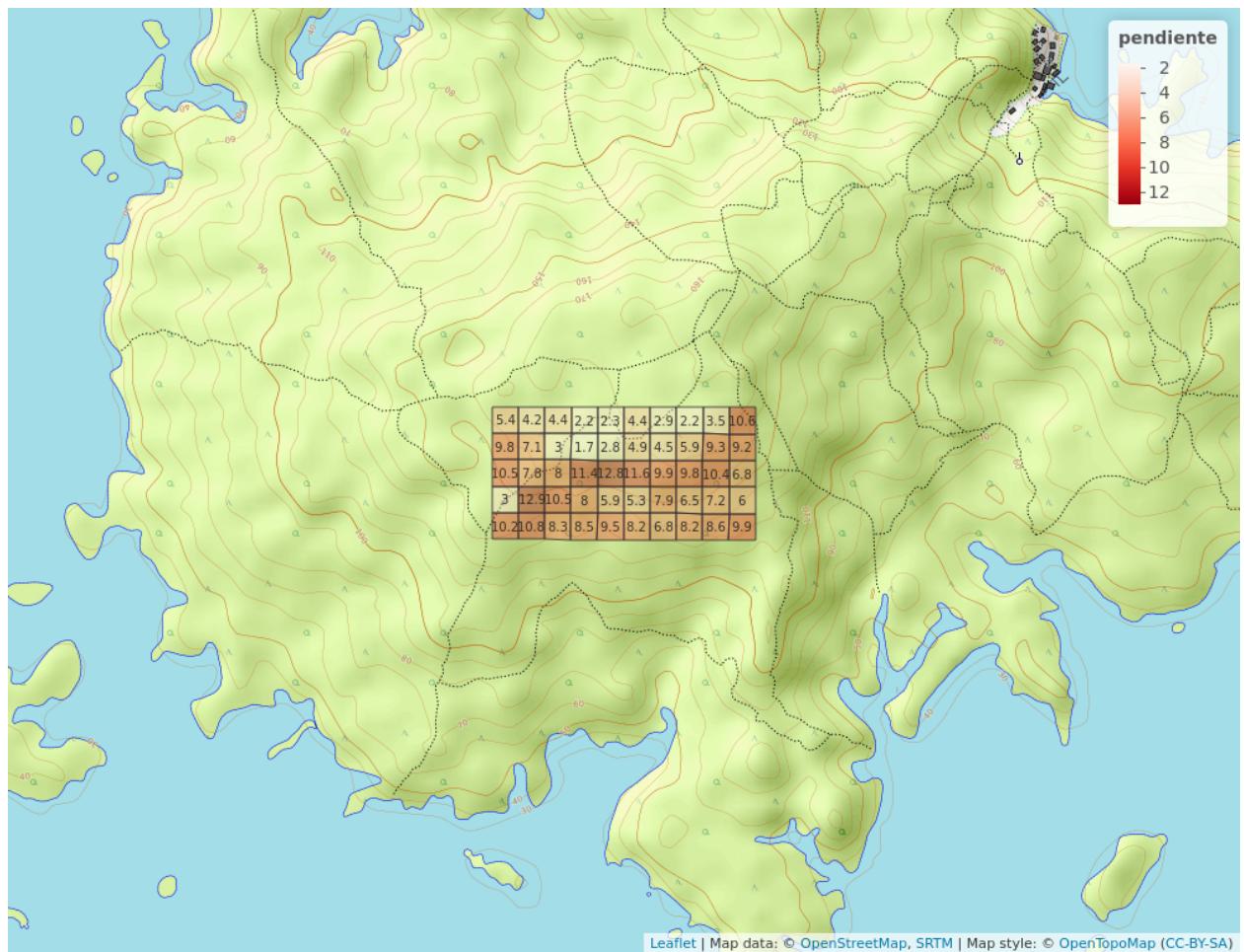


Figure 40: mapa de la isla Barro Colorado cuadros y pendientes

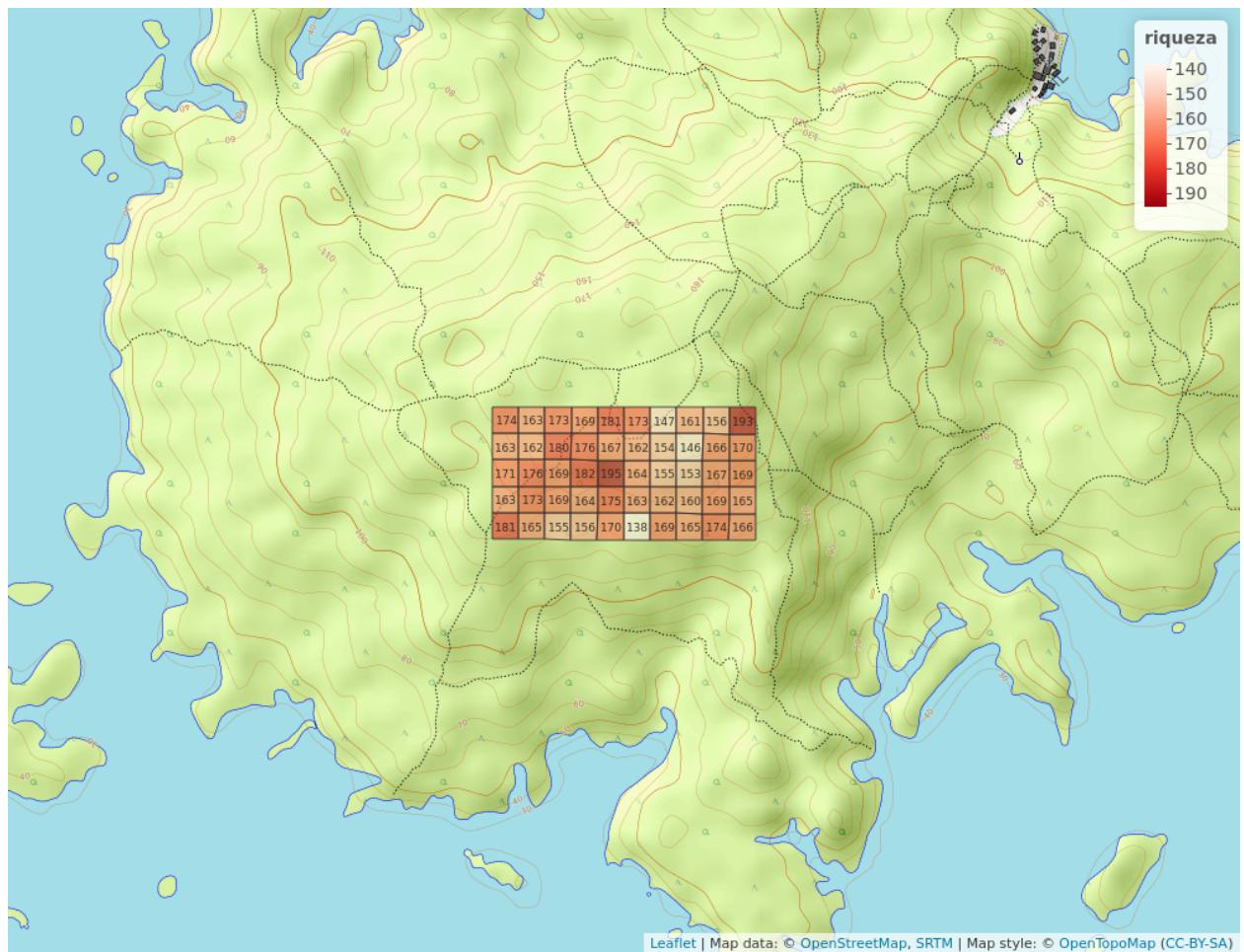


Figure 41: mapa de la isla Barro Colorado riqueza global

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

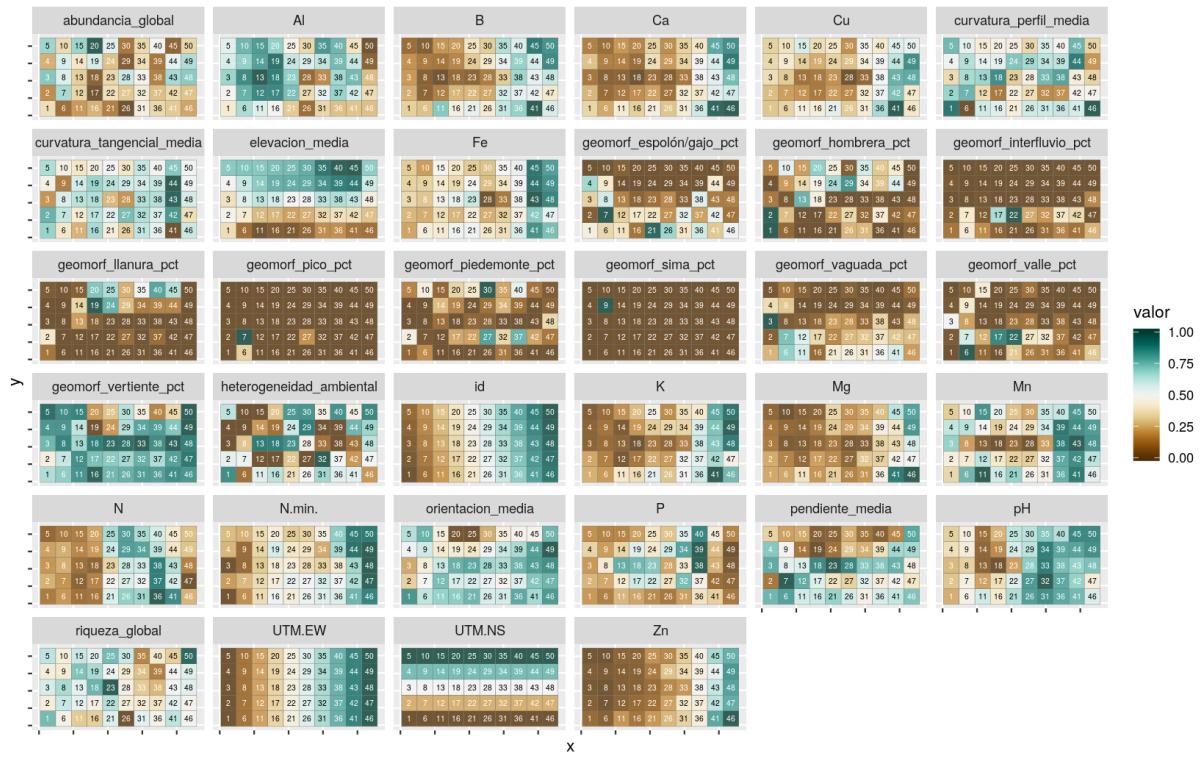


Figure 42: mapa de la isla Barro Colorado variables ambientales numéricicas

Cuadros de 1 Ha de BCI. Variables ambientales nominales

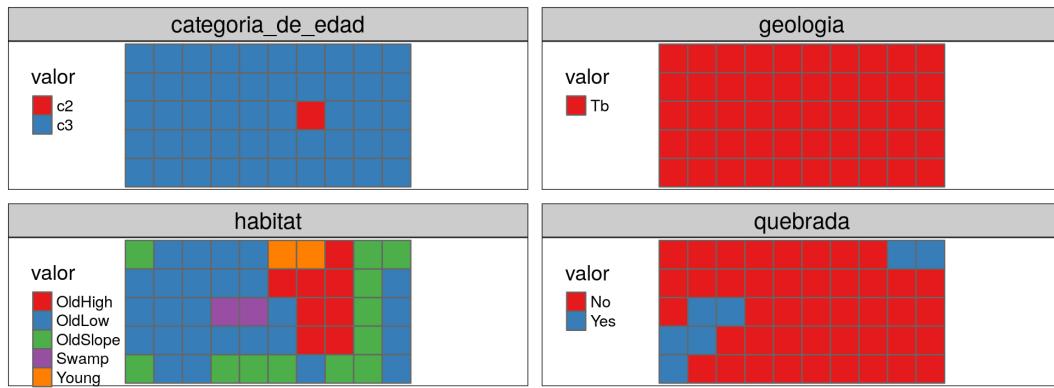


Figure 43: mapa de la isla Barro Colorado variables ambientales nominales

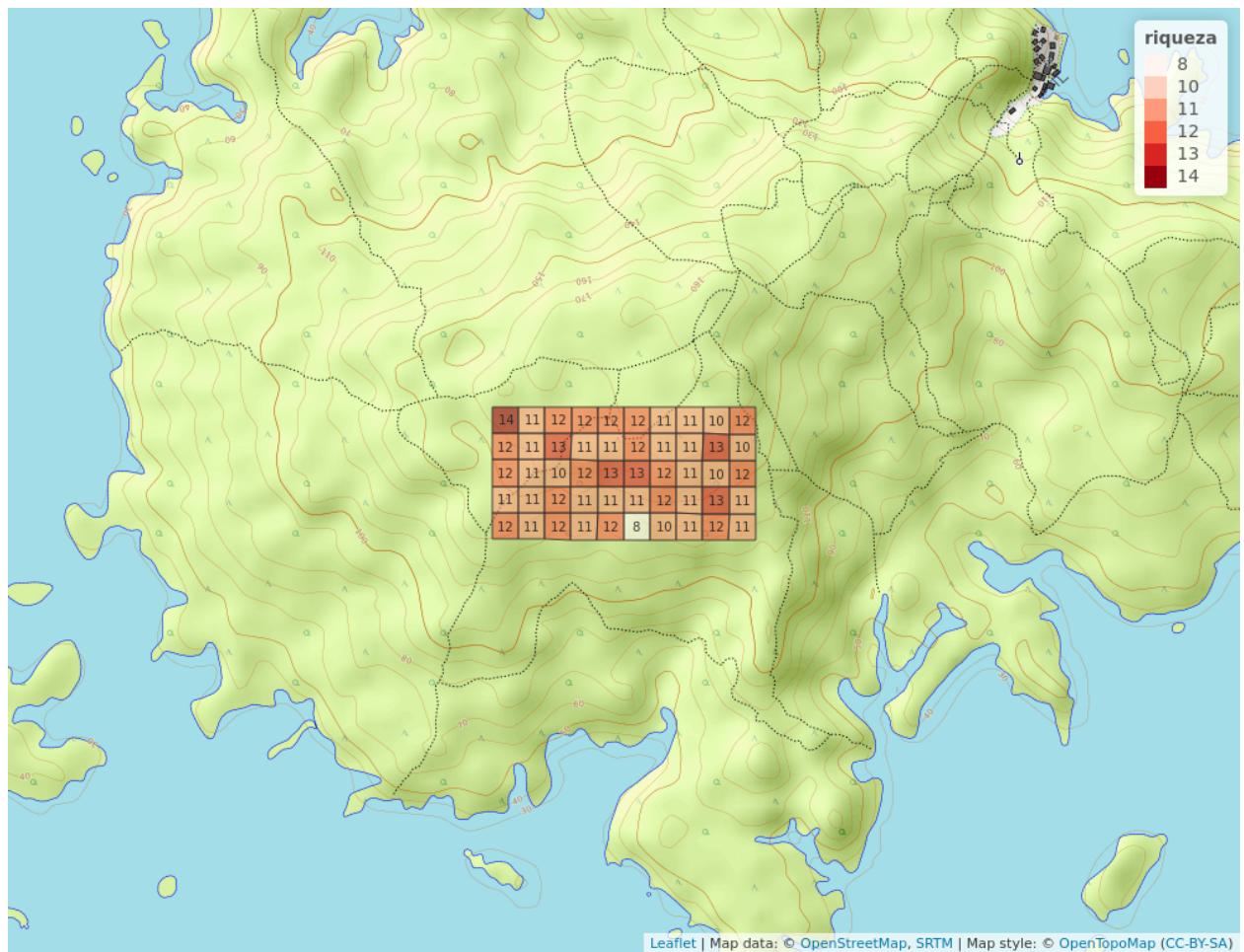


Figure 44: mapa de la isla Barro Colorado cuadro de riqueza mi familia

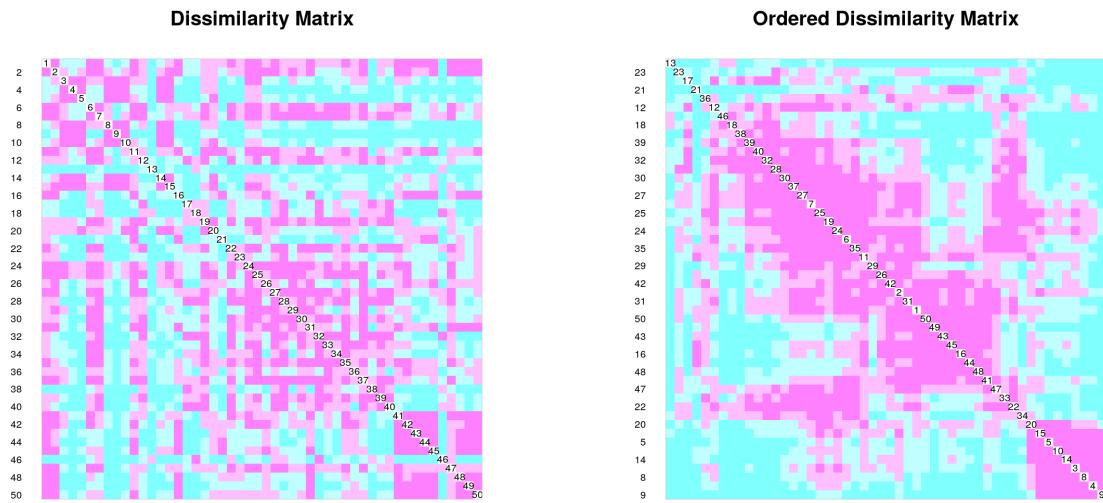


Figure 45: mapa de la isla Barro Colorado matriz disimilaridad hellinger

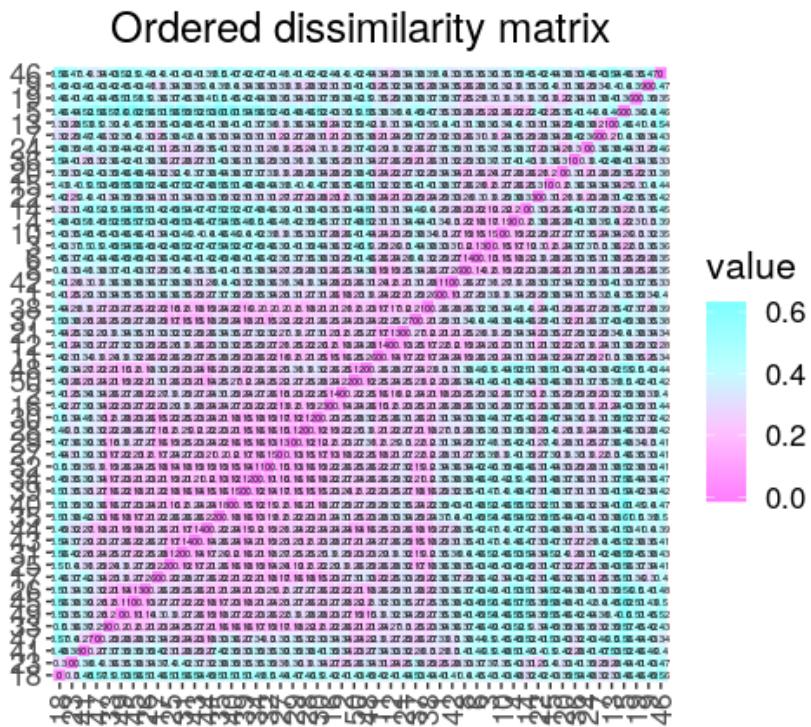
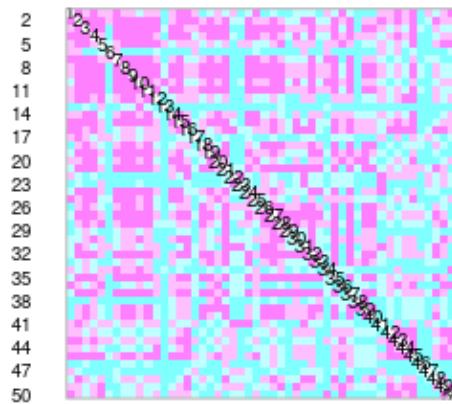


Figure 46: mapa de la isla Barro Colorado matriz disimilaridad

**Dissimilarity Matrix**



**Ordered Dissimilarity Matrix**

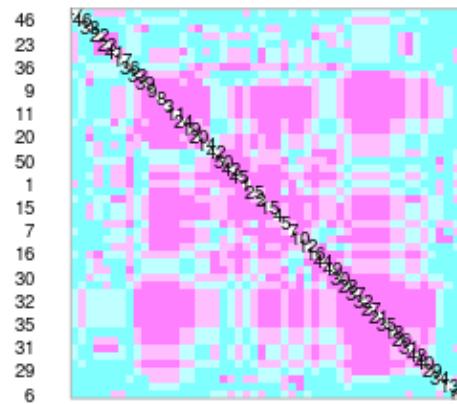
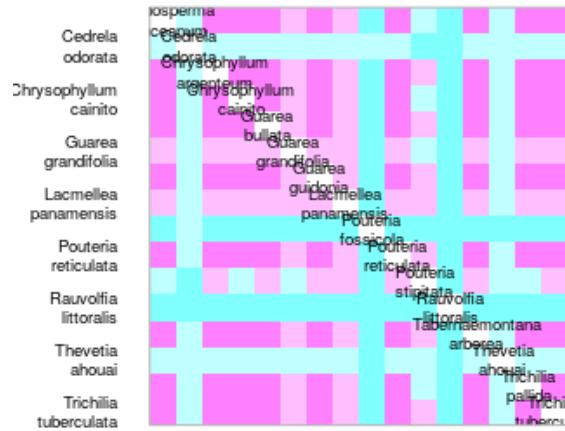


Figure 47: mapa de la isla Barro Colorado matriz disimilaridad

**Dissimilarity Matrix**



**Ordered Dissimilarity Matrix**

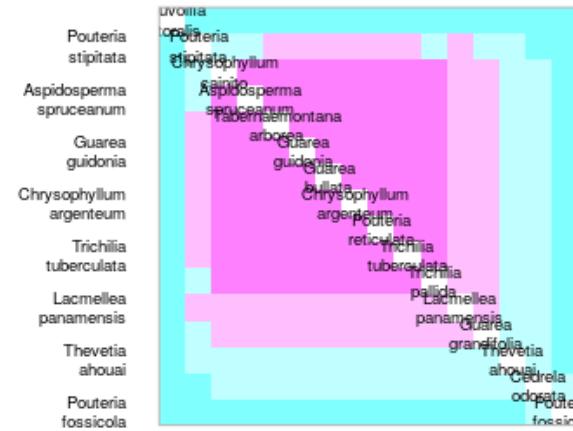


Figure 48: mapa de la isla Barro Colorado mapa de calor



Figure 49: mapa de la isla Barro Colorado abundancia y riqueza de suelo

## Sitios de BCI según composición de especies de Moraceae Enlace simple a partir de matriz de distancia de cuerdas

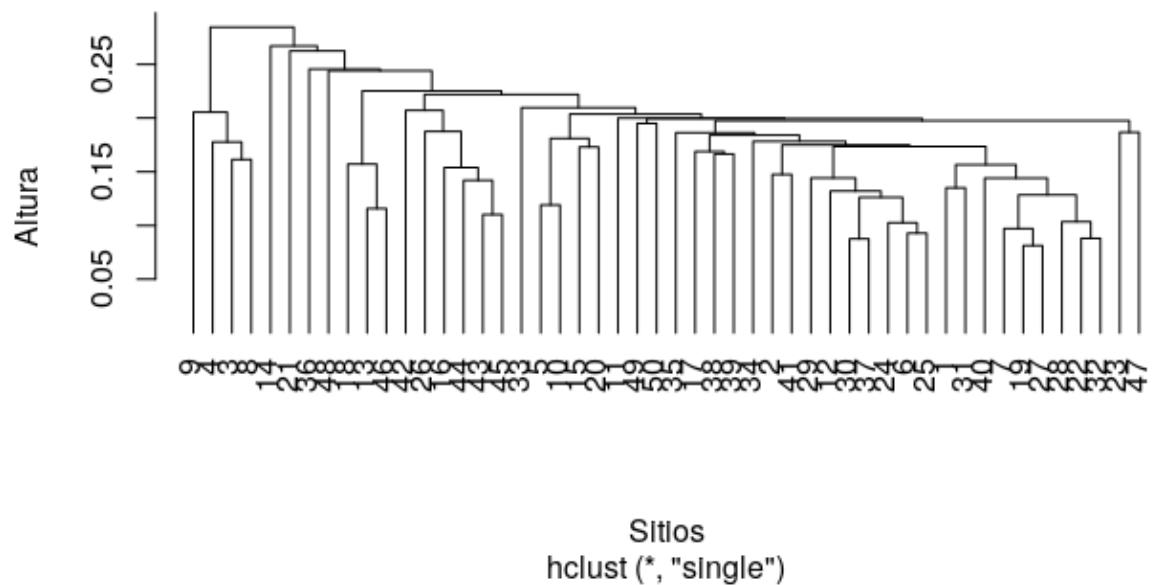


Figure 50: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

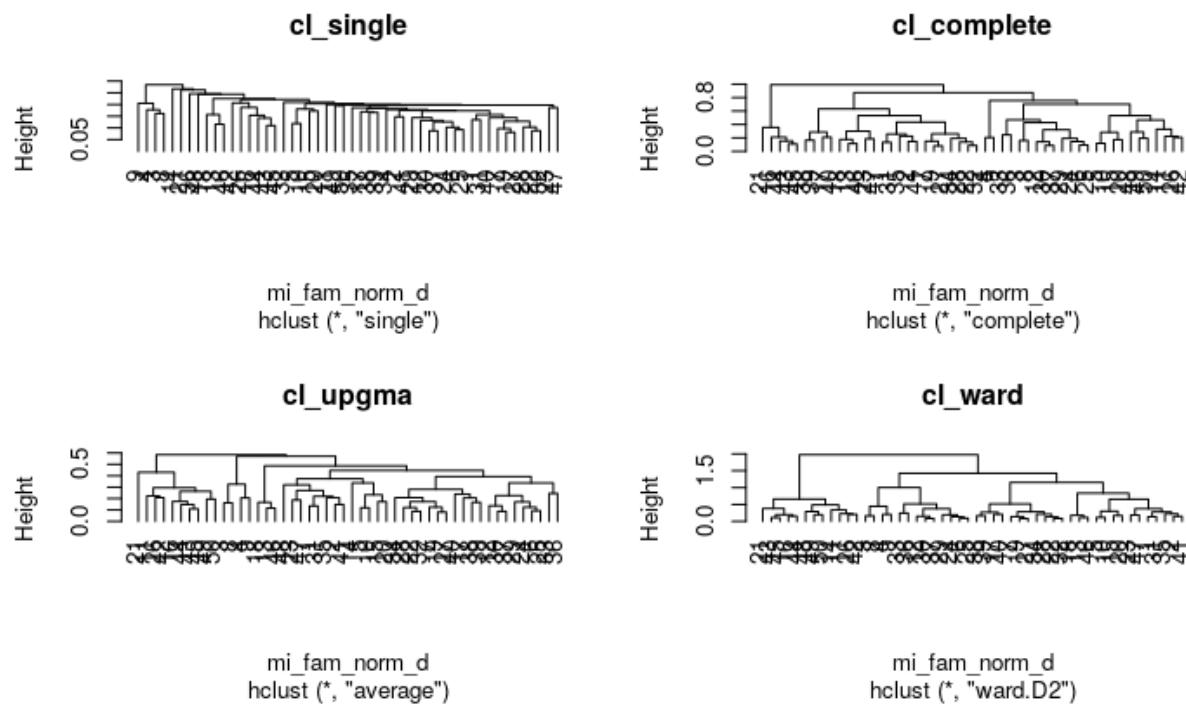


Figure 51: mapa de la isla Barro Colorado análisis interpretación dendrogramas

**Sitios de BCI según composición de especies de Moraceae**  
**Método de Ward a partir de matriz de distancia de cuerdas**

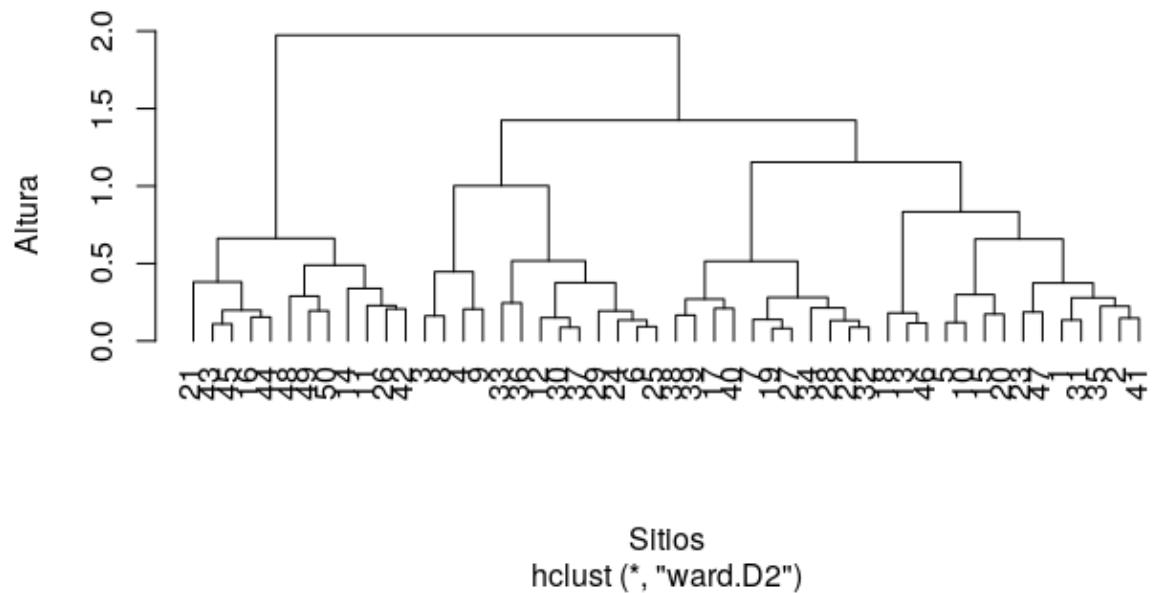


Figure 52: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

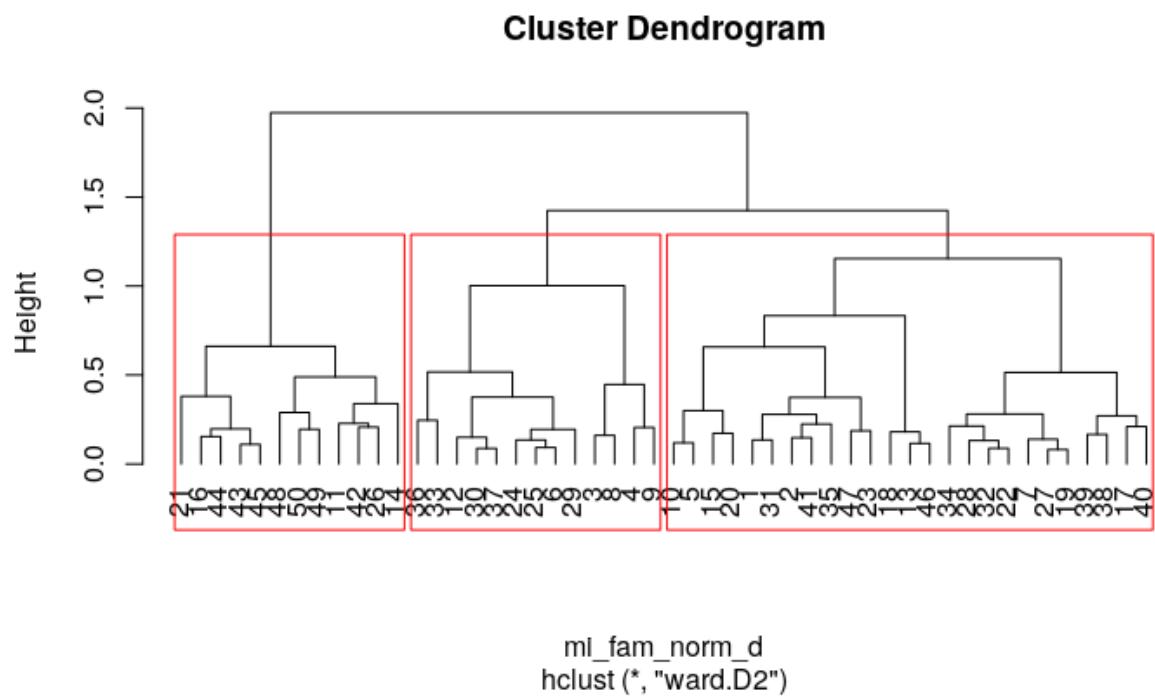


Figure 53: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

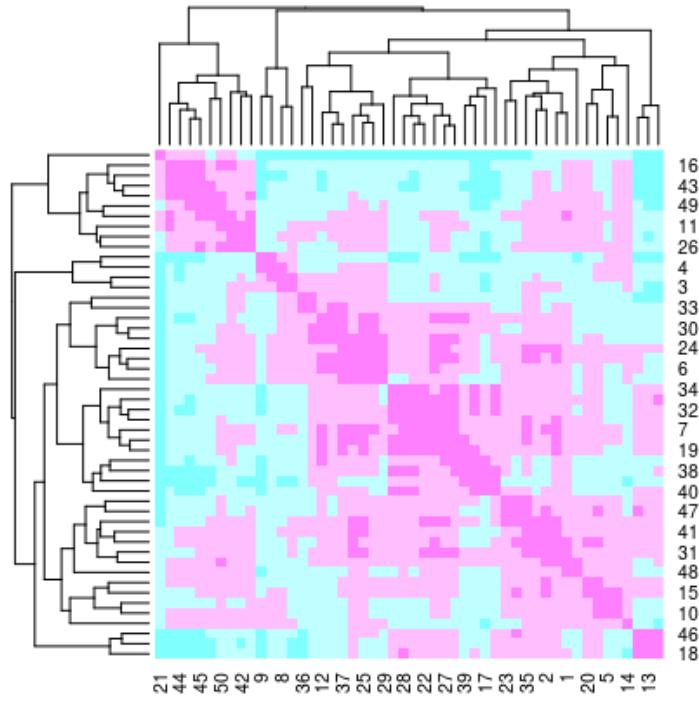


Figure 54: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

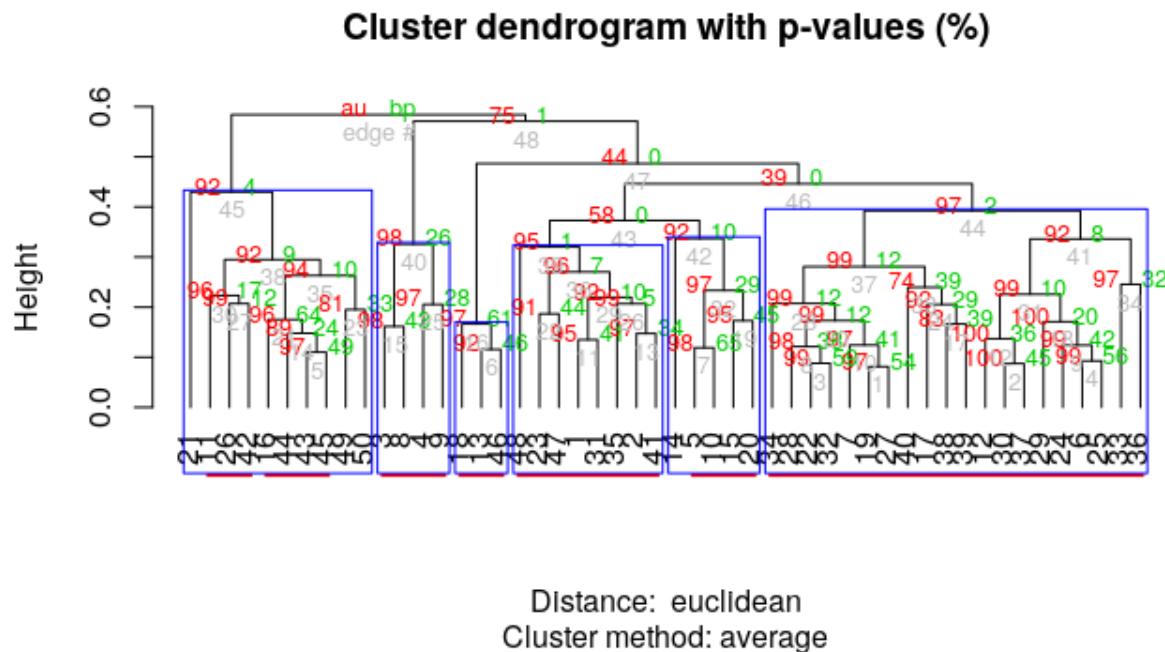


Figure 55: mapa de la isla Barro Colorado remuestreo multiescalar por Bootstrap

### Cluster dendrogram with p-values (%)

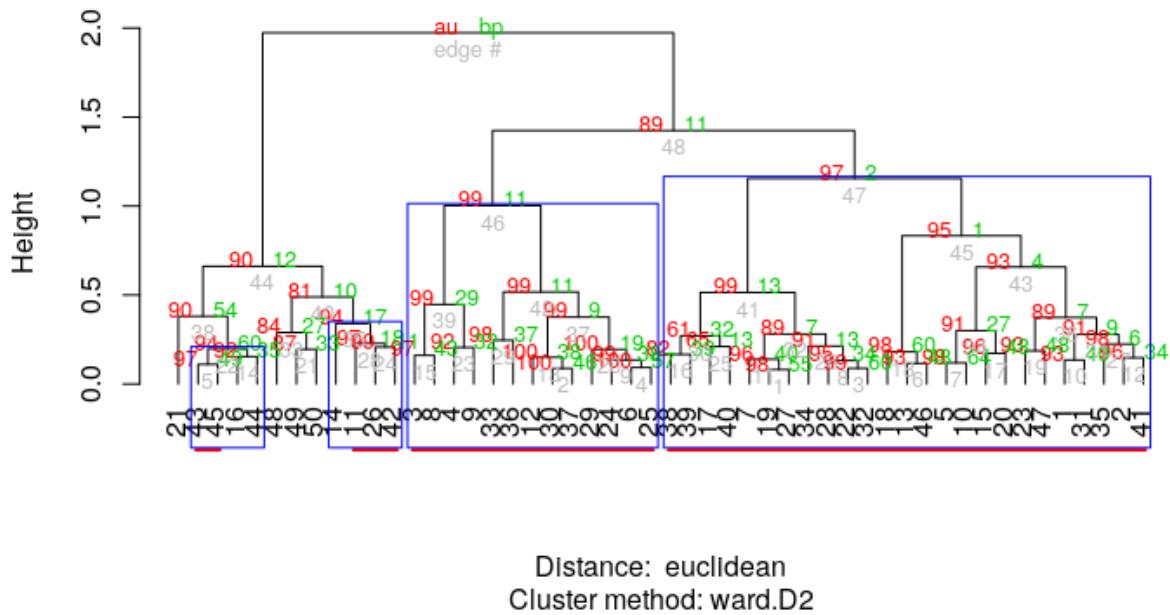


Figure 56: mapa de la isla Barro Colorado dendrogramas y porcentajes

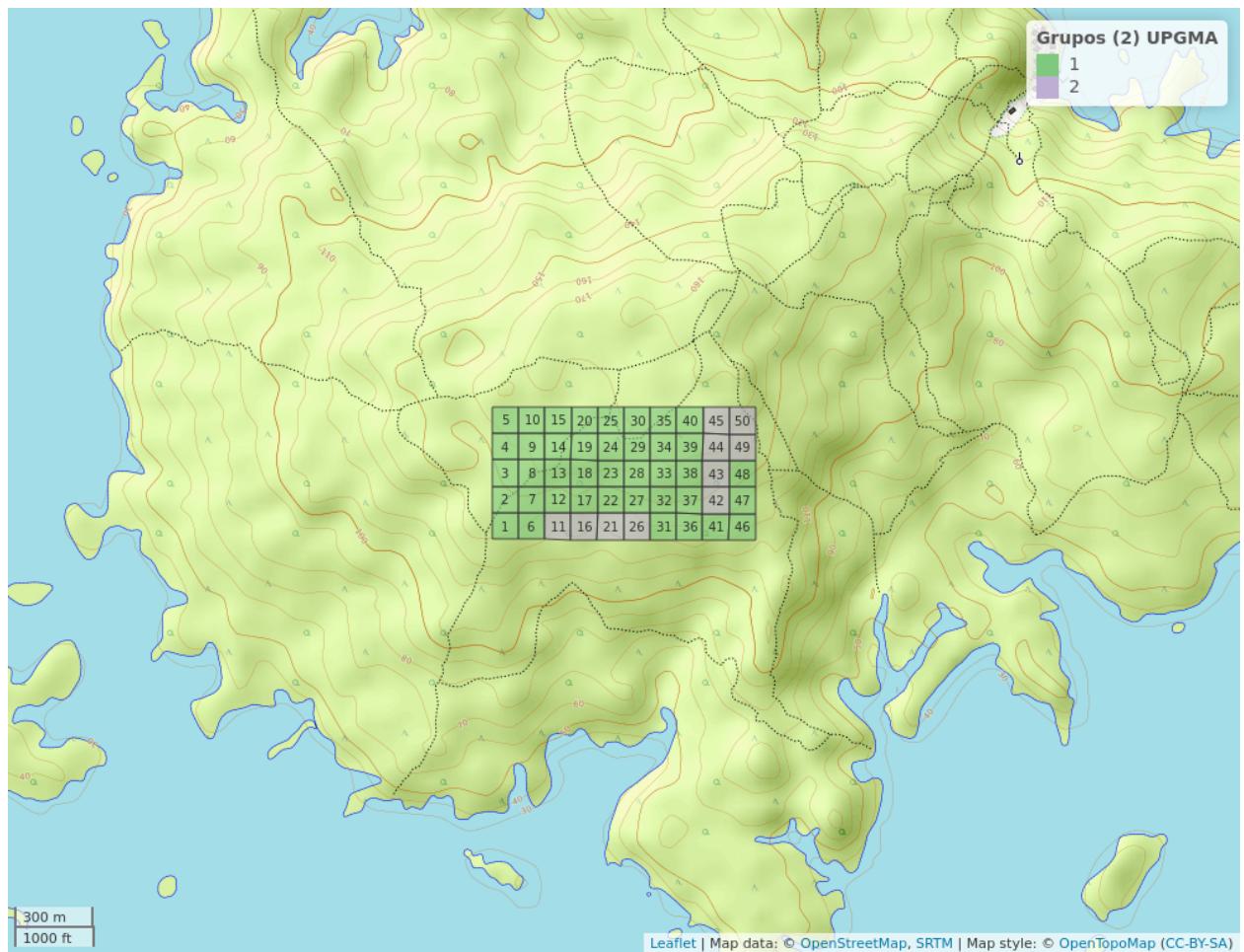


Figure 57: mapa de la isla Barro Colorado sitios de composición por distancia de cuerda

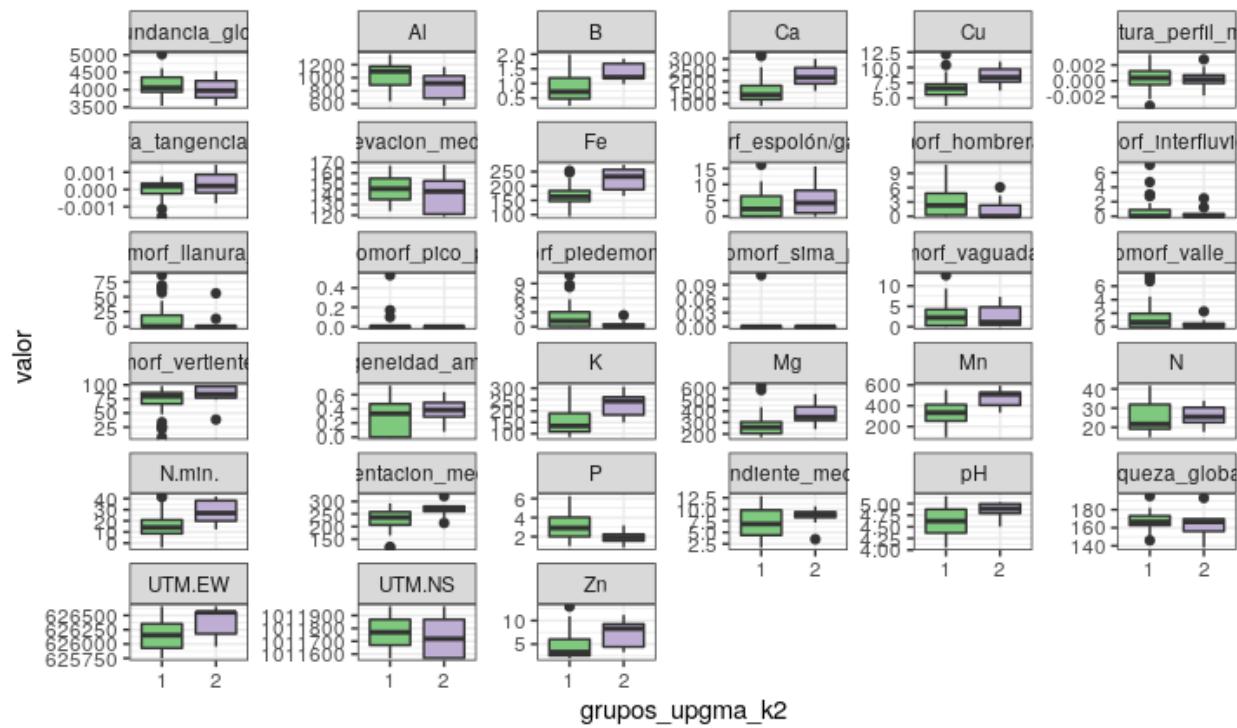


Figure 58: mapa de la isla Barro Colorado mapa de variables ambientales

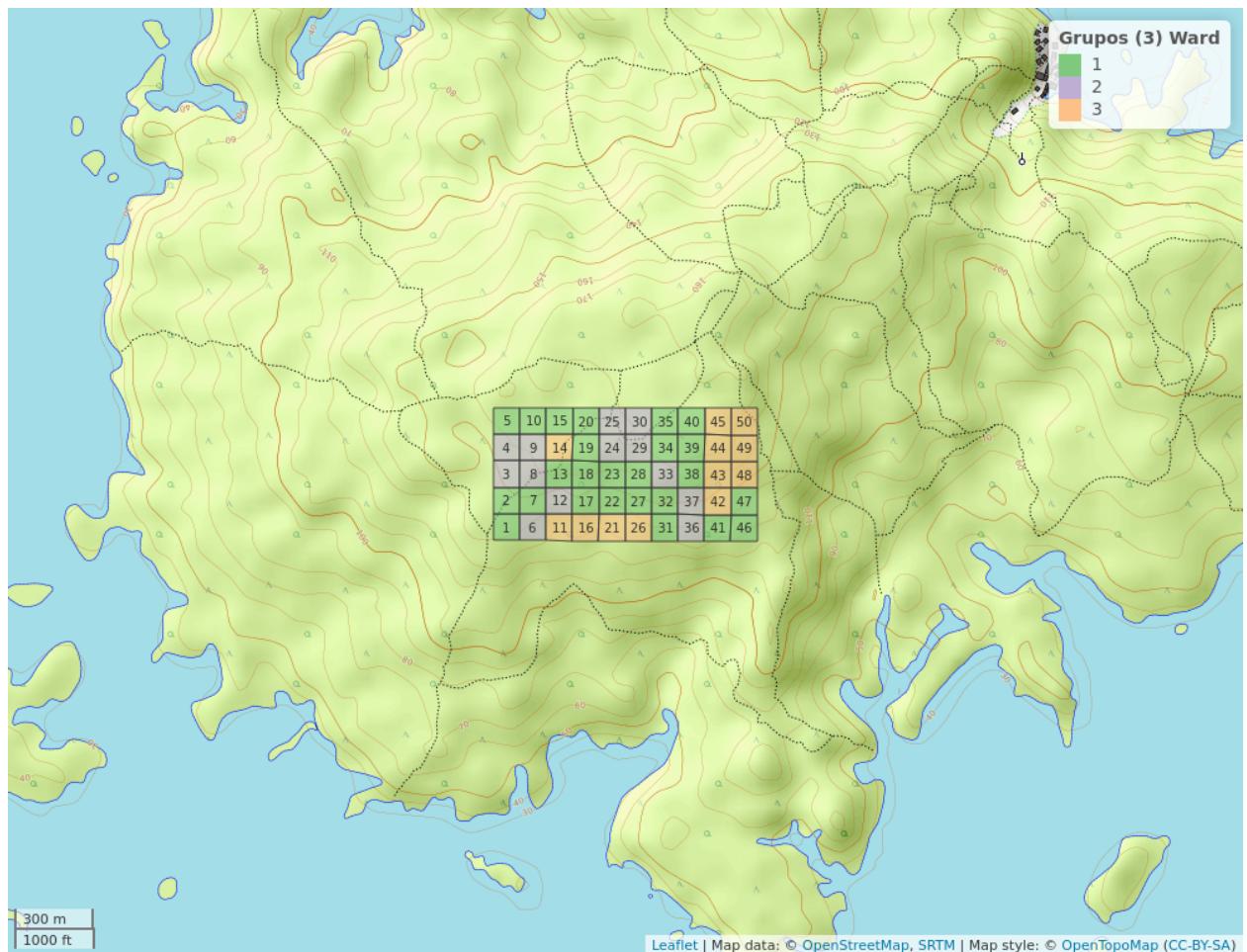


Figure 59: mapa de la isla Barro Colorado grupos Ward k3

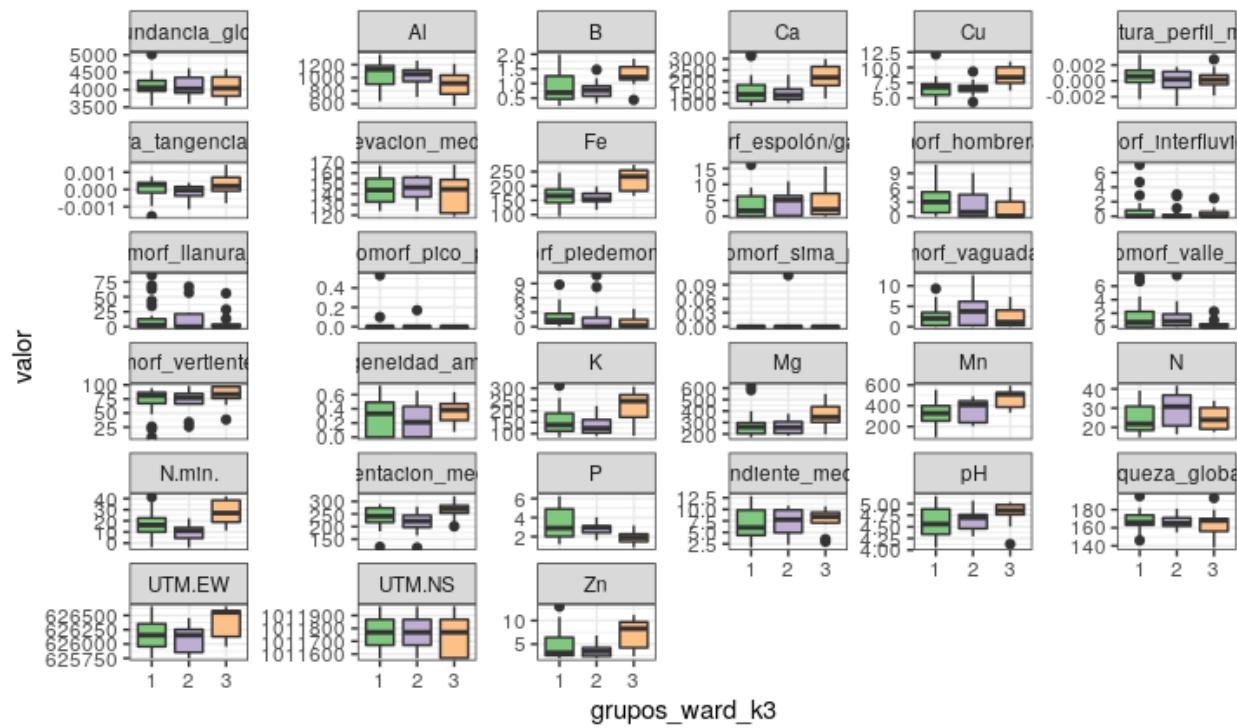


Figure 60: mapa de la isla Barro cuadro de cajas Ward

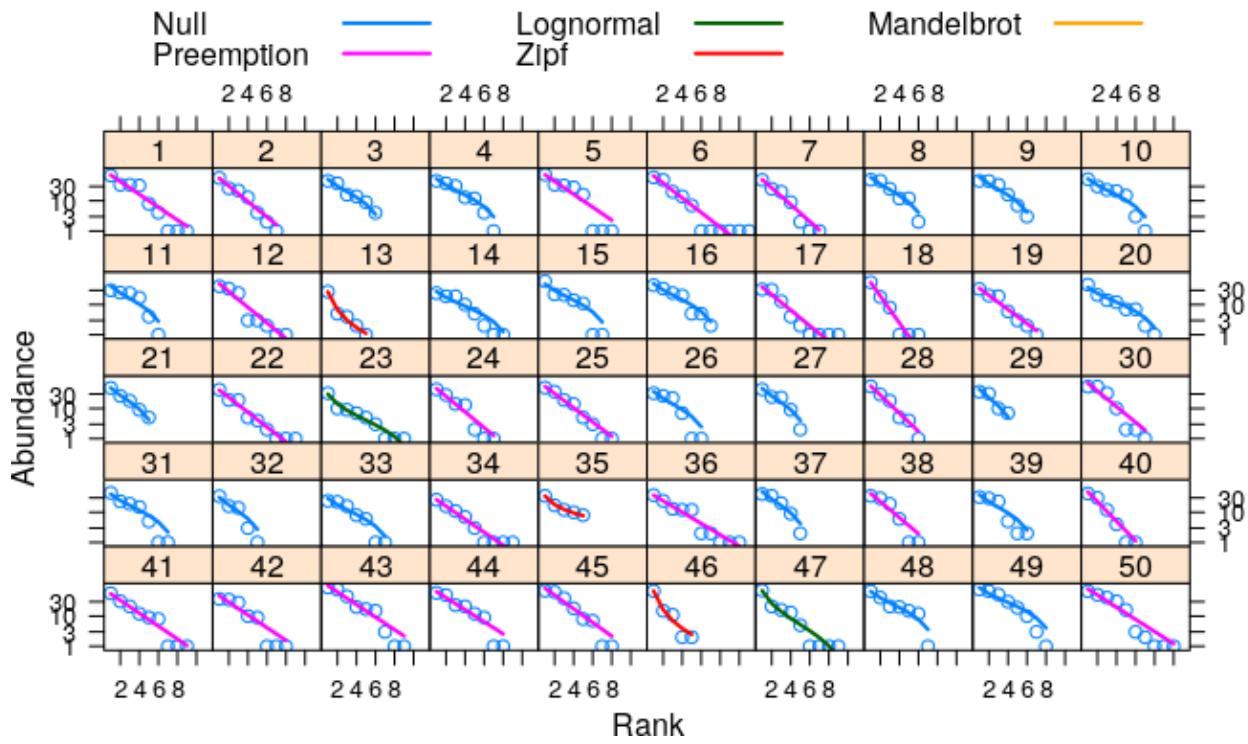


Figure 61: mapa de la isla Barro Colorado modelo abundancia de especies

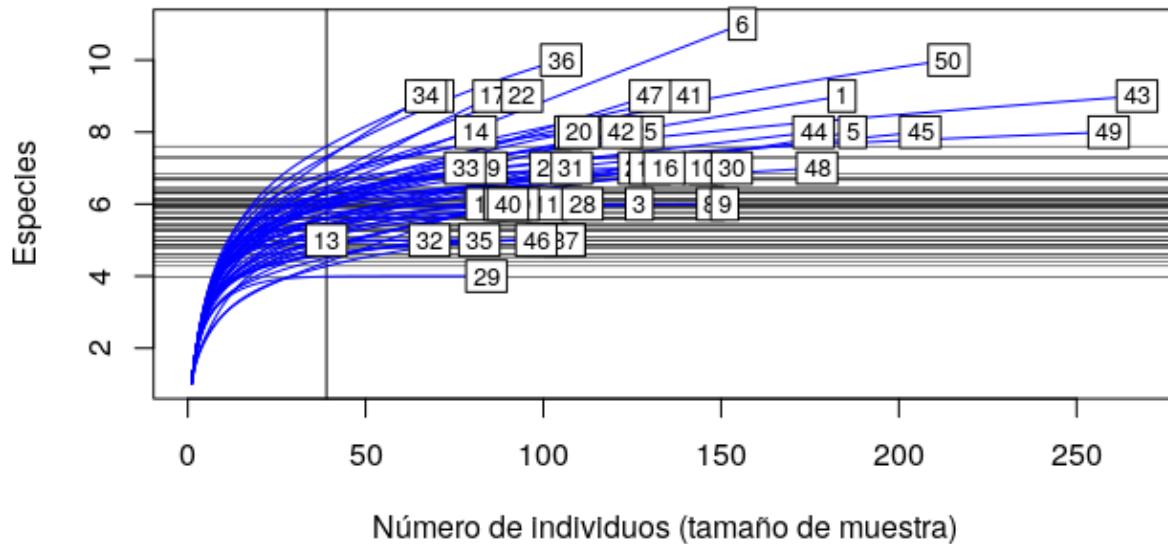


Figure 62: mapa de la isla Barro Colorado número de individuos

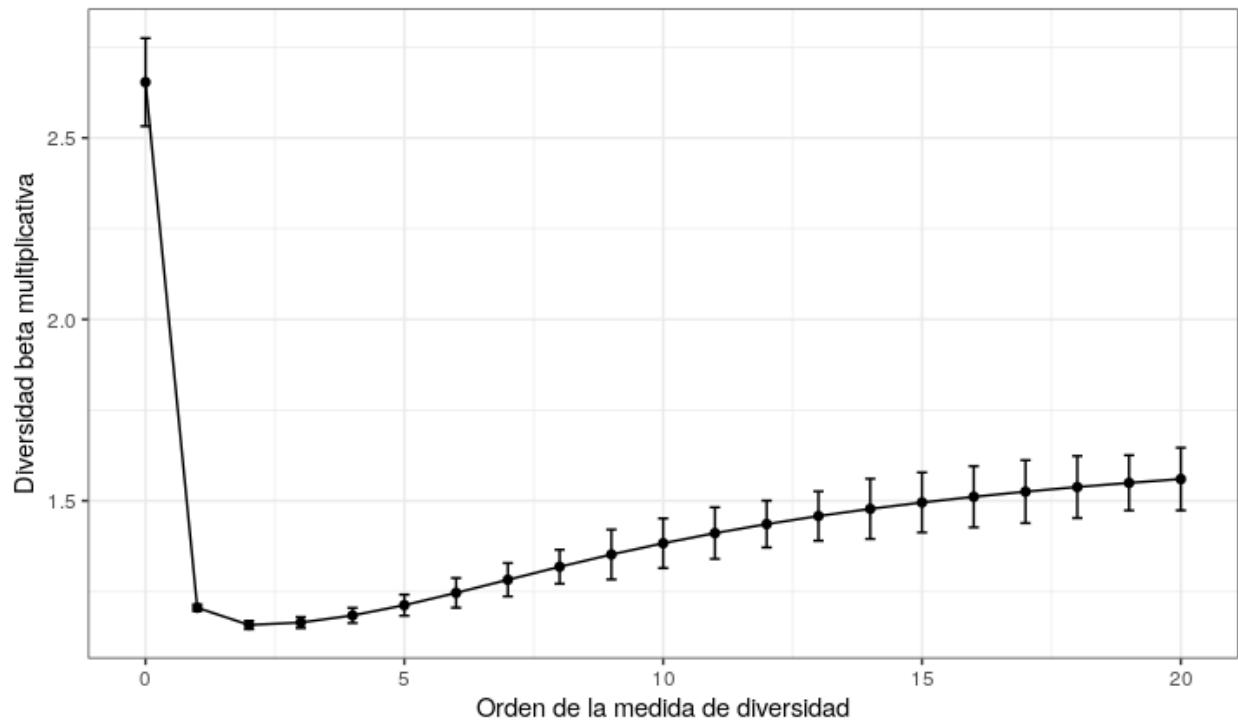


Figure 63: mapa de la isla Barro Colorado mapa de calor

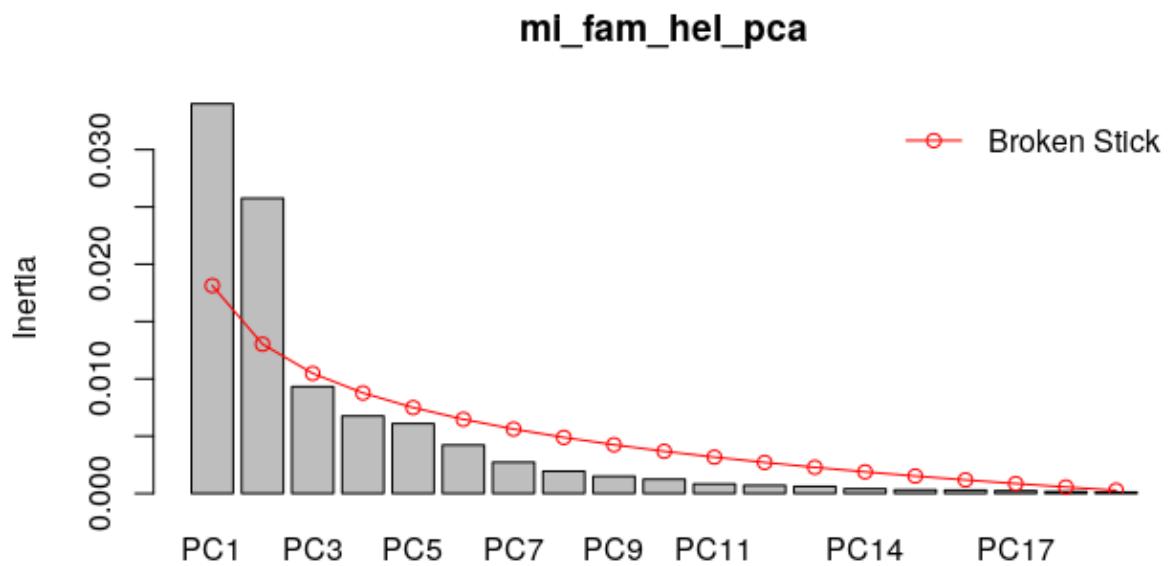


Figure 64: mapa de la isla Barro Colorado diagrama de componentes principales variables

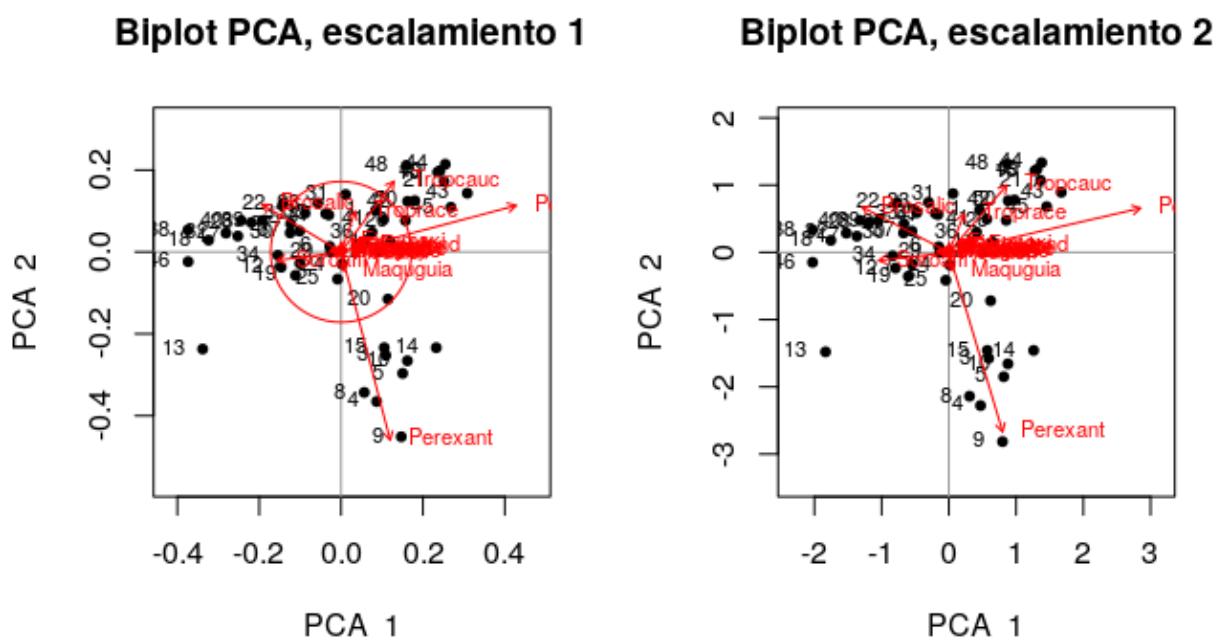


Figure 65: mapa de la isla Barro Colorado diagrama de escalamiento 1 y 2

## ot de RDA especies ~ var. suelo, escala ot de RDA especies ~ var. suelo, escala

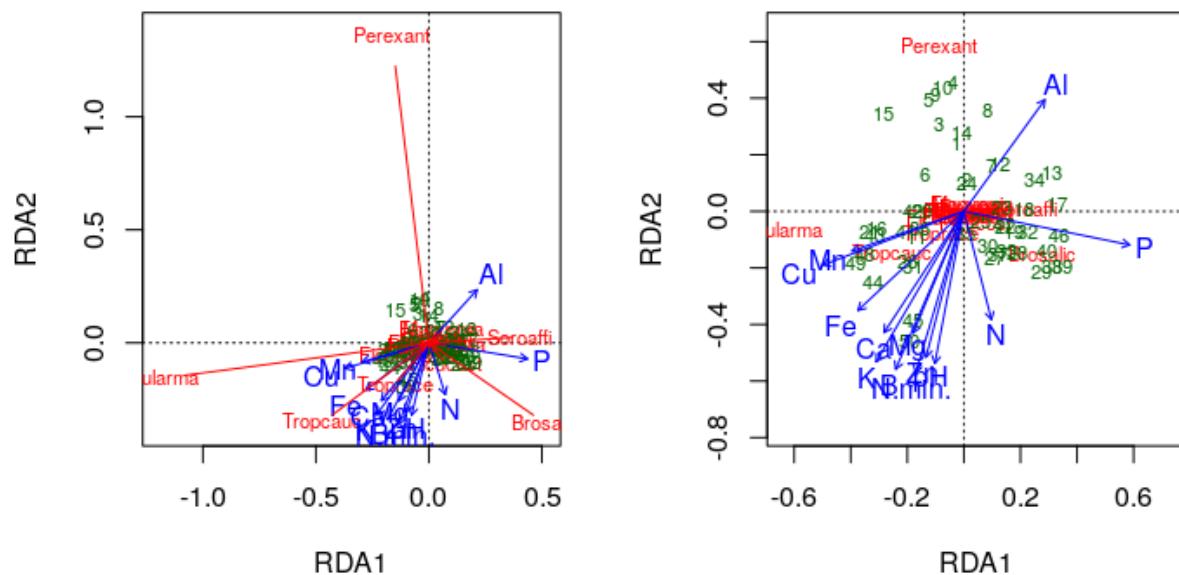


Figure 66: mapa de la isla Barro Colorado RDA diagrama de especies variables

## CCA especies no raras ~ var. selec4, e CCA especies no raras ~ var. selec4, e

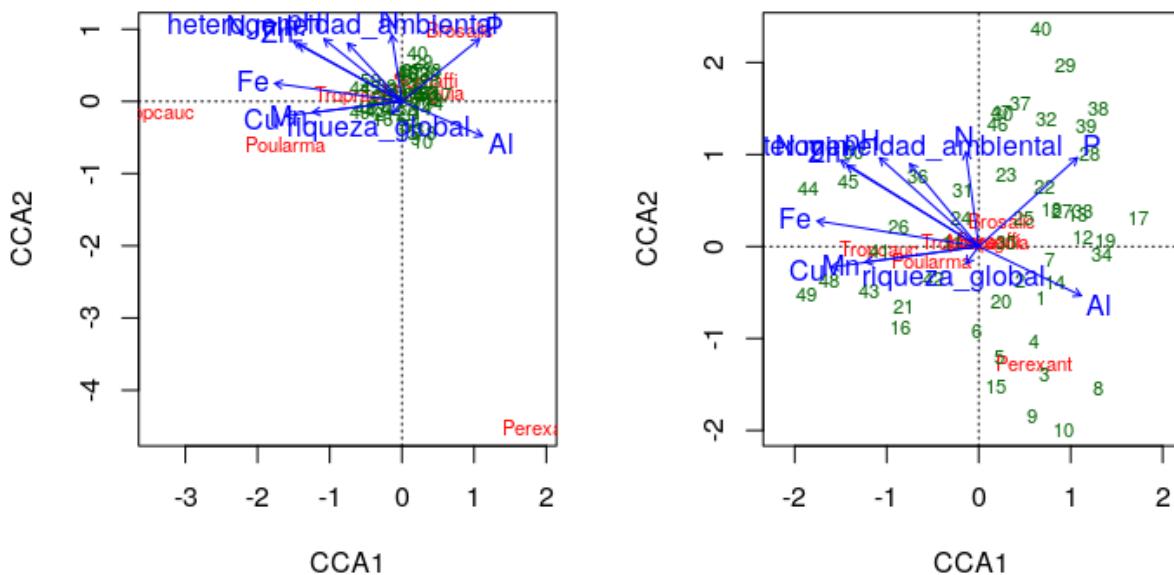


Figure 67: mapa de la isla Barro Colorado RDA especies no raras variables

## Mapa de CCA especies ~ var. selec4, escala

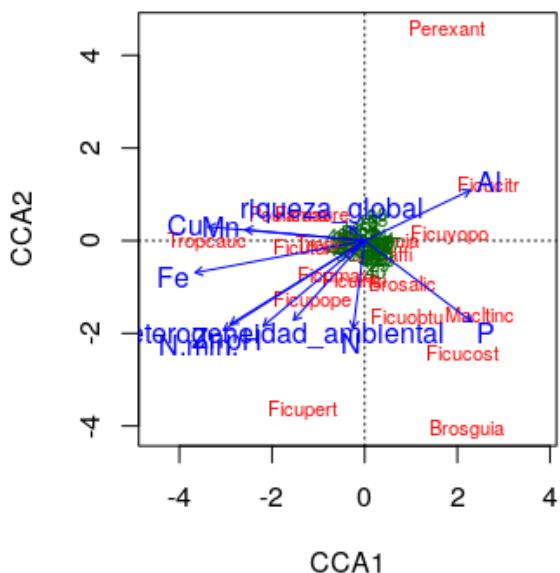
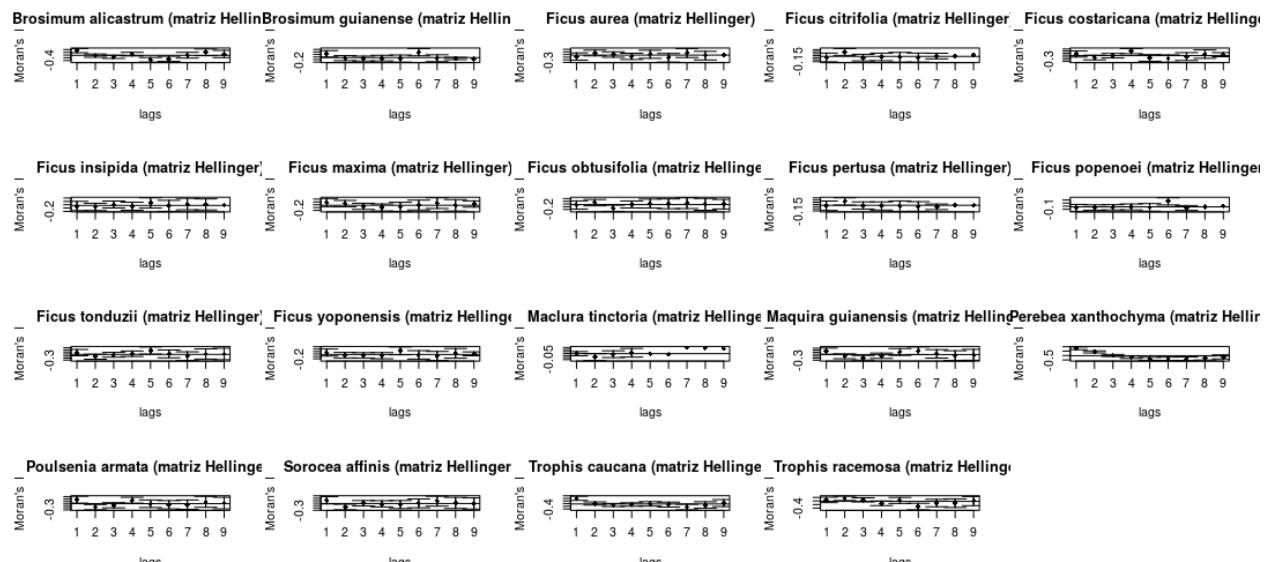


Figure 68: mapa de la isla Barro Colorado RDA digrama de especies variables



# Script reproducible

### 6.1 Análisis exploratorio de datos. Riqueza y abundancia

```
#' ---
#' título: 'Análisis exploratorio de datos. Riqueza y abundancia'
```

```

#' autor: "JR"
#'fecha: "13 de octubre, 2020"
#' salida: github_document
#' ---

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

#'### Imprimir datos en pantalla (impresiones parciales con head)
head(censo_morac)
head(mc_morac)
bci_env_grid # No necesita imprimirse parcialment

#' ### También podemos usar
# Requiere que se haya cargado ya la colección tidyverse
censo_morac %>% tibble
mc_morac %>% tibble

#' ### Lista de especies
sort(colnames(mc_morac))

#' ### Número de sitios, tanto en matriz de comunidad como en ambiental
# Verifica que coinciden
nrow(mc_morac) #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_morac)
sort(specnumber(mc_apcyn_melic_saptc)) # Ordenados ascendentemente
summary(specnumber(mc_morac)) # Resumen estadístico

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

#' ### Abundancia por especie
sort(colSums(mc_morac))
summary(colSums(mc_morac))# Resumen estadístico

```

```

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

#' ### Abundancia de toda la comunidad
sum(colSums(mc_morac))

#' ### Una tabla para el manuscrito, es necesario asignarle nombre
#' Para esto, usaré la colección "tidyverse"
abun_sp <- (censo_morac) %>%
  group_by(Latin) %>%
  count() %>%
  arrange(desc(n))
abun_sp

#' ### Un gráfico para el manuscrito
#' Gráfico de mosaicos de la abundancia por especie por cuadros
abun_sp_q <- crear_grafico_mosaico_de_mc(mc_apcyn_melic_saptc, tam_rotulo = 6)
abun_sp_q

```

## 6.2 Análisis exploratorio de datos. Mapas de variables ambientales

```

#' ---
#' title: "Análisis exploratorio de datos. Mapas de variables ambientales"
#' author: "JR"
#' date: "25 de octubre, 2020"
#' output: github_document
#' ---

#' ### Cargar paquetes
library(mapview)
library(tidyverse)
library(sf)
library(RColorBrewer)

#' ### Cargar datos
load('biodata/matriz_ambiental.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
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')

```

### 6.3 Análisis de agrupamiento (cluster analysis). Parte 1: agrupamiento jerárquico

```
#' ---
#' title: "Análisis de agrupamiento (cluster analysis). <br> Parte 1: agrupamiento jerárquico"
#' author: "JR"
#' date: "11 de noviembre, 2020"
#' output: github_document
#' ---

knitr::opts_chunk$set(fig.width=12, fig.height=8)

#' ## Preámbulo

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

#' ### Cargar datos
#'
load('biodata/Moraceae.Rdata')
mi_fam <- mc_morac
#'
#' ## Características de las técnicas de agrupamiento
#'
#' Las técnicas de agrupamiento se clasifican según los algoritmos que emplean y el orden de ejecución
#'
#' - Secuenciales o simultáneos.
#' - Por aglomeración o por división. En referencias en español encontrarás "aglomerativos" y "divisivos".
#' - Monotéticos o politéticos.
#' - Jerárquicos o no jerárquicos.
#' - Probabilísticos o no probabilísticos.
#' - Restringidos o no restringidos.
#'
#' ## Agrupamiento jerárquico
#'
#' El agrupamiento jerárquico (AJ) es una técnica de agrupamiento secuencial que consiste en la fusión de conjuntos.
#'
#' Dentro del AJ es frecuente usar un enfoque aglomerativo, lo cual implica aplicar algoritmos sencillos.
#'
#' El AJ aglomerativo dispone de varios algoritmos de resolución del agrupamiento por pares, que varían dependiendo del criterio de agrupamiento.
#'
#' Normalmente, en el análisis de agrupamiento nos interesa agrupar sitios en función de sus descriptores.
#'
#' Dado que los cuadros en BCI están autocorrelacionados espacialmente, violamos el supuesto de independencia entre observaciones.
#'
#' ### Agrupamiento "aglomerativo" por enlace simple
#'
```

```

#' Este método utiliza, como criterio de enlace para agrupar sucesivamente pares de objetos, la
#'
#' Para aplicar este método, debes transformar la matriz de comunidad utilizando alguno de los m
#'
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 'd
#'
(cl_single <- hclust(mi_fam_norm_d, method = 'single'))
#'
# Finalmente, el dendrograma a continuación:
plot(cl_single, labels = rownames(mi_fam), hang = -1,
      main = "Sitios de BCI según composición de especies de Moraceae\nEnlace simple a partir de
              xlab = 'Sitios', ylab = 'Altura')
#'
#' #### Agrupamiento "aglomerativo" por enlace completo
#'
# En este caso, el criterio de enlace para agrupar sucesivamente pares de objetos es la menor s
#'
(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 Moraceae\nEnlace completo a partir d
              xlab = 'Sitios', ylab = 'Altura')
#'
#' #### Agrupamiento "aglomerativo" por enlace promedio
#'
# En este caso, el criterio de enlace para agrupar sucesivamente pares de objetos es el promedi
#'
# Así, dependiendo de si se media o centroide, o si se ponderan o no las distancias originales,
#'
# Sólo crearé el 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 Moraceae\nUPGMA a partir de matriz d
              xlab = 'Sitios', ylab = 'Altura')
#'
#' #### Agrupamiento por el método de Ward de varianza mínima
#'
# Se basa en los mismos supuestos y criterios de la regresión lineal por mínimos cuadrados, sim
#'
(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 Moraceae\nMétodo de Ward a partir de",
      xlab = 'Sitios', ylab = 'Altura')

6.4 Análisis de agrupamiento (cluster analysis). Parte 2: Interpretación y comparación de resultados
#'
#' ---  

#' title: "Análisis de agrupamiento (cluster analysis). <br> Parte 2: Interpretación y comparación de resultados"
#' author: "JR"
#' date: "11 de noviembre, 2020"
#' output: github_document
#' ---  

knitr:::opts_chunk$set(fig.width=12, fig.height=8)  

#' ## Preámbulo
#'  

#'  

#' #### Cargar paquetes
#'  

library(vegan)
library(tidyverse)
library(broom)
library(cluster)
library(gclus)
library(pvclust)
library(sf)
source('biodata/funciones.R')
#'  

#' #### Cargar datos
#'  

load('biodata/Moraceae.Rdata')
mi_fam <- mc_morac
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
#'  

#' [En el script anterior] (aa_analisis_de_agrupamiento_1_jerarquico.md) realicé los dendrogramas
#'  

#' La primera evaluación de los dendrogramas NO debe venir de la mano de sofisticados análisis n

```

```

#'
#' Evita concentrar tu vista en grupos extremadamente pequeños; comienza analizando el árbol des
#'
#' Anota tus impresiones, para que las compares con los resultados que posteriormente obtendrás;
#'
#' Para la exploración visual, generaré los 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))
#'
#' En mi caso, exceptuando el dendrograma generado por medio del enlace simple, detecto al menos
#'
#' - Un grupo pequeño, compuesto por los sitios 1, 42, 12, 21, 11, 2 y 16.
#' - Un "grupo" heterogéneo y grande, conformado por 25, 31,..., 26,..., 35,..., 34,...,32, 17,..
#'
#' Además de los grupos anteriores, detecto elementos que no forman grupos, es decir, sitios que
#'
#' ## Elegir método y número de clústers
#'
#' Existen varios criterios para elegir un dendrograma idóneo, como por ejemplo, los gráficos ti
#'
#' ### Seleccionar método de agrupamiento por correlación cofenética
#'
#' La correlación cofenética implica conocer la distancia cofenética, y esta última se entiende m
#'
#' Usando la lista de objetos de clústers, calcularé la 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)
})
#'
#' Habrás notado que, tanto UPGMA como enlace completo, tienen valores altos de correlación cofe
#'
#' ### Elegir número de clústers
#'
#' Elegiré UPGMA como método de agrupamiento y determinaré cuántos grupos son idóneos de acuerdo

```

```

#'
#' El objetivo de la función 'calcular_anchuras_siluetas' está implícito en su nombre, y requiere
#'
#' 1. Las anchuras promedio para cada partición, excepto para las particiones 'i=1' y 'i=50', por
#'
#' 2. Número óptimo de grupos. Haré los cálculos para UPGMA y Ward, y luego explico en qué consisten
#'
#' Para UPGMA:
#'
anch_sil_upgma <- calcular_anchuras_siluetas(
  mc_orig = mi_fam,
  distancias = mi_fam_norm_d,
  cluster = lista_cl$cl_upgma)
anch_sil_upgma
#'
#' El objeto 'anchuras_siluetas' de la lista 'anch_sil_upgma' te muestra un vector con los promedios
#'
#' Igualmente, el objeto 'n_grupos_optimo' te indica cuál es el número óptimo de clústers a crear
#'
#' Haré el gráfico de dendrograma, aunque nota que en este caso primero reordenaré los sitios con
#'
u_dend_reord <- reorder.hclust(lista_cl$cl_upgma, mi_fam_norm_d)
plot(u_dend_reord, hang = -1)
rect.hclust(
  tree = u_dend_reord,
  k = anch_sil_upgma$n_grupos_optimo)
#'
#' Ahora compararé el dendrograma con el mapa de calor en un mismo gráfico, colocando los dendrogramas
#'
heatmap(
  as.matrix(mi_fam_norm_d),
  Rowv = as.dendrogram(u_dend_reord),
  symm = TRUE,
  margin = c(3, 3),
  col = rev(cm.colors(4)))
)
#'
#' En general, hay dos grupos, uno grande y otro pequeño, y parece haber un tercero en el mapa de calor
#'
#' Mostraré el resultado para Ward:
#'
anch_sil_ward <- calcular_anchuras_siluetas(
  mc_orig = mi_fam,
  distancias = mi_fam_norm_d,
  cluster = lista_cl$cl_ward)
anch_sil_ward
#'
#' En este caso, el valor máximo, que ocupa la posición número 2, no se diferencia mucho del de

```

```

#'
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)))
)
#'
## Nótese que este dendrograma hace más sentido que el sugerido por UPGMA. En cualquier casos, c
#'
## ### Evaluación mediante remuestreo por *bootstrap* multiescalar
#'
## Con suerte, un agrupamiento aplicado a datos muestrales reflejará los patrones naturales de o
#'
## Sin embargo, los datos de BCI también tienen sesgo, pues se usa un DAP de corte para decidir
#'
## No obstante, aun con todas sus bondades, los datos censales carecen de una fortaleza: no refl
#'
## Por estas razones, los análisis de agrupamientos realizados hasta este punto, reflejan tanto
#'
## El método de remuestreo por *boostrap* multiescalar está implementado en el paquete 'pvclust'
#'
## La función 'pvclust' devolverá un dendrograma Enriquecido, que incluirá los valores de AU y BP
#'
## - Rectángulos de borde azul, para todos aquellos grupos que resulten con valores de AU>0.91 e
#'
## - Líneas inferiores rojas, que resaltan aquellos grupos (o subgrupos) que obtuvieron AU>0.95.
#'
## Ten presente que, al realizar remuestreo por *bootstrap* multiescalar, cada corrida puede arr
#'
## #### 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
#'
#' Detecto algunos patrones consistentes en cuanto a grupos de sitios según composición de las e
#'
#' - Tanto en UPGMA como en Ward, detecté al menos dos o tres grandes grupos. Con el primer métod
#'
#' - En el caso específico del dendrograma Ward, 'pvclust' atomizó los sitios en demasiados grup
#'
#' ##### ¿Cómo declaro los grupos de sitios?
#'
#' Para conservar las clasificaciones de grupos de sitios anteriores, crearé un vector con el id
#'
#' UPGMA:
(grupos_upgma_k2 <- as.factor(cutree(lista_cl$cl_upgma, k = 2)))
#'
#' En este caso, los sitios 1 y 2 pertenecen al grupo 1, los sitios 3 al 6 pertenecen al grupo 2
#'
table(grupos_upgma_k2)
#'
#' Nota lo desiguales que son estos grupos, un efecto esperado dado el alto grado de autocorrelaci
#'
#' Ward:
#'

```

```

(grupos_ward_k3 <- as.factor(cutree(lista_cl$cl_ward, k = 3)))
table(grupos_ward_k3)
#
#' Guardaré estos vectores en archivos para reutilizarlos en *scripts* posteriores:
#
saveRDS(grupos_upgma_k2, 'grupos_upgma_k2.RDS')
saveRDS(grupos_ward_k3, 'grupos_ward_k3.RDS')
#
#' Evita usar este, y cualquier otro procedimiento, de manera mecánica. En tu caso, quizá tengas
#
#' En el próximo *script*, aprenderás a comparar este resultado con las variables ambientales. T

```

## 6.5 Análisis de agrupamiento (cluster analysis). Parte 3: Grupos (clústers), variables ambientales y mapas

```

#' ---
#' title: "Análisis de agrupamiento (cluster analysis). <br> Parte 3: Grupos (clústers), variabl
#' author: "JR"
#' date: "15 de noviembre, 2020"
#' output: github_document
#' ---

```

```
knitr::opts_chunk$set(fig.width=12, fig.height=8)
```

```

#' ## Preámbulo
#
#' ### Cargar paquetes
#
library(mapview)
library(tidyverse)
library(sf)
library(RColorBrewer)
source('biodata/funciones.R')
#
#' ### Cargar datos
#
load("~/unidad-0-asignacion-99-mi-manuscrito-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2) #Importante, tener en cuenta los desiguales tamaños de los grupos
grupos_ward_k3 <- readRDS('grupos_ward_k3.RDS')
table(grupos_ward_k3)
#
#' ### Paletas
#
rojo <- colorRampPalette(brewer.pal(8, "Reds"))
rojo_inv <- colorRampPalette(rev(brewer.pal(8, "Reds")))
colores_grupos <- brewer.pal(8, "Accent")

```

```

#'
#' ## Explorar efectos
#'
#' ### Pruebas de igualdad de promedios de las variables entre 2 grupos
#'
#' Para evaluar homogeneidad de promedios usaré las pruebas *t* (medias), basada en la distribuc
#'
#' Primero crearé un objeto que permita realizar tanto las pruebas como los diagramas de cajas.
#'
(m_amb_upgma_k2 <- bci_env_grid %>%
  select_if(is.numeric) %>% select(-id) %>%
  mutate(grupos_upgma_k2) %>%
  st_drop_geometry() %>%
  pivot_longer(-grupos_upgma_k2, names_to = "variable", values_to = "valor"))
#'
#' A continuación, las pruebas:
#'
m_amb_upgma_k2 %>%
  group_by(variable) %>%
  summarise(
    p_valor_t = t.test(valor ~ grupos_upgma_k2)$p.value,
    p_valor_w = wilcox.test(valor ~ grupos_upgma_k2, exact = F)$p.value) %>%
  arrange(p_valor_t) %>%
  print(n=Inf)
#'
#' Interesa observar las variables que obtuvieron valores de p<0.01. Reitero que, en mi caso, mi
#'
#' Por otra parte, este es un buen momento para "revisitar" tus análisis exploratorios de datos
#'
#' Los gráficos:
#'
m_amb_upgma_k2 %>%
  group_by(variable) %>%
  ggplot() + aes(x = grupos_upgma_k2, y = valor, fill = grupos_upgma_k2) +
  geom_boxplot() +
  scale_fill_brewer(palette = 'Accent') +
  theme_bw() +
  theme(legend.position="none") +
  facet_wrap(~ variable, scales = 'free_y')
#'
#' Mapas:
#'
mapa_upgma_k2 <- mapView(
  bci_env_grid %>% mutate(grupos_upgma_k2),
  layer.name = 'Grupos (2) UPGMA',
  alpha.regions = 0.6,
  map.types = 'OpenTopoMap',
  legend = T,

```

```

col.regions = colores_grupos[1:2],
zcol = 'grupos_upgma_k2') %>%
addStaticLabels(label = bci_env_grid$id) %>%
leaflet::setView(
  lng = -79.85136,
  lat = 9.15097,
  zoom = 15)
mapa_upgma_k2
mapa_upgma_k2 %>% mapshot(
  file = 'mapa_upgma_k2.png',
  remove_controls = c("zoomControl", "layersControl", "homeButton")
)
#
#'
#' Mapa de una de las variables donde se presentó efecto de su promedio (p<0.01), en este caso,
#'
mapa_zn <- mapView(
  bci_env_grid,
  layer.name = 'Zinc',
  alpha.regions = 0.6,
  map.types = 'OpenTopoMap',
  legend = T,
  col.regions = rojo,
  zcol = 'Zn') %>%
addStaticLabels(label = bci_env_grid$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_k3 <- bci_env_grid %>%
  select_if(is.numeric) %>% select(-id) %>%
  mutate(grupos_ward_k3) %>%
  st_drop_geometry() %>%
  pivot_longer(-grupos_ward_k3, names_to = "variable", values_to = "valor"))
#'
## Pruebas, en este caso ANOVA (evalúa homogeneidad de medias; no se cumplen muchos de los supue
#'
m_amb_ward_k3 %>%
  group_by(variable) %>%

```

```

summarise(
  p_valor_a = oneway.test(valor ~ grupos_ward_k3)$p.value,
  p_valor_k = kruskal.test(valor ~ grupos_ward_k3)$p.value) %>%
arrange(p_valor_k) %>%
print(n=Inf)
#'
#' Gráficos:
#'
m_amb_ward_k3 %>%
  group_by(variable) %>%
  ggplot() + aes(x = grupos_ward_k3, y = valor, fill = grupos_ward_k3) +
  geom_boxplot() +
  scale_fill_brewer(palette = 'Accent') +
  theme_bw() +
  theme(legend.position="none") +
  facet_wrap(~ variable, scales = 'free_y')
#'
#' Mapas:
#'
mapa_ward_k3 <- mapView(
  bci_env_grid %>% mutate(grupos_ward_k3),
  layer.name = 'Grupos (3) Ward',
  alpha.regions = 0.6,
  map.types = 'OpenTopoMap',
  legend = T,
  col.regions = colores_grupos[1:3],
  zcol = 'grupos_ward_k3') %>%
  addStaticLabels(label = bci_env_grid$id) %>%
  leaflet::setView(
    lng = -79.85136,
    lat = 9.15097,
    zoom = 15)
mapa_ward_k3
mapa_ward_k3 %>% mapshot(
  file = 'mapa_ward_k3.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_ph <- mapView(
  bci_env_grid,
  layer.name = 'pH',
  alpha.regions = 0.6,
  map.types = 'OpenTopoMap',
  legend = T,
  col.regions = rojo_inv,
  zcol = 'pH') %>%

```

```

addStaticLabels(label = bci_env_grid$id) %>%
leaflet::setView(
  lng = -79.85136,
  lat = 9.15097,
  zoom = 15)
mapa_ph
mapa_ph %>% mapshot(
  file = 'mapa_ph.png',
  remove_controls = c("zoomControl", "layersControl", "homeButton")
)
#'

```

## 6.6 Análisis de agrupamiento (cluster analysis). Parte 4: Especies indicadoras, especies con preferencia por hábitats

```
#' ---
```

```
#' title: "Análisis de agrupamiento (cluster analysis). <br> Parte 4: Especies indicadoras, especies con preferencia por hábitats"
#' author: "JR"
#' date: "15 de noviembre, 2020"
#' output: github_document
#' ---
```

```
knitr::opts_chunk$set(fig.width=12, fig.height=8)
```

```

#' ## Preámbulo
#'
#' ### Cargar paquetes
#'
library(indicspecies)
source('biodata/funciones.R')
#'
#' ### Cargar datos
#'
load('biodata/Moraceae.Rdata')
mi_fam <- mc_morac
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k3 <- readRDS('grupos_ward_k3.RDS')
table(grupos_ward_k3)
#'
#' ## 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_k3 <- multipatt(
  x = mi_fam,
  cluster = grupos_ward_k3,
  func = 'IndVal.g',
  max.order = 2,
  control = how(nperm = 999))
summary(iva_ward_k3, indvalcomp = TRUE)
colSums(mi_fam)
(p_ward_adj <- p.adjust(iva_ward_k3$sign$p.value))
(iva_ward_boot <- strassoc(
  X = mi_fam,
  cluster = grupos_ward_k3,
  func = "IndVal.g",
  nboot = 1000))
#'
#' ## Análisis de especies con preferencia por hábitat mediante el coeficiente de correlación bi
#'
#'
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_k3 <- multipatt(
  mi_fam,
  grupos_ward_k3,
  func = "r.g",
  max.order = 2,
  control = how(nperm = 999))
summary(phi_ward_k3)
colSums(mi_fam)
(phi_ward_boot <- strassoc(
  X = mi_fam,
  cluster = grupos_ward_k3,
  func = "r.g",
  nboot = 1000))

```

## 6.7 Técnicas de ordenación. Parte 1: Ordenación no restringida. PCA, CA y PCoA

```

#' ---
#' title: "Técnicas de ordenación. <br> Parte 1: Ordenación no restringida. <br> PCA, CA y PCoA"
#' author: "JR"
#' date: "21 de noviembre, 2020"
#' output: github_document
#'
#' ---
knitr::opts_chunk$set(fig.width=12, fig.height=8)

#' ## Preámbulo
#'
#' ### Cargar paquetes
#'
library(vegan)
library(tidyverse)
library(sf)
library(mapview)
source('biodata/funciones.R')
#'
#' ### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_morac
(colnames(mi_fam) <- make.cepnames(colnames(mi_fam)))
(df_equivalecias <- data.frame(
  nombre_original = colnames(mc_morac),
  colnames(mi_fam)))
bci_env_grid %>% tibble

```

```

grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k3 <- readRDS('grupos_ward_k3.RDS')
table(grupos_ward_k3)
#'
#' ## Ordenación
#'
#' La ordenación se basa en los mismos principios que la medición de asociación (similaridad) y
#'
#' A diferencia del análisis de agrupamiento, o como complemento de éste, el análisis de ordenación
#'
#' El análisis de ordenación puede ser no restringido (o simple) y restringido (o 'canónico'). Es
#'
#' Las principales técnicas de ordenación no restringida son análisis de componentes principales
#'
#' #### Análisis de componentes principales (PCA)
#'
#' Es el método tradicional basado en vectores propios que comúnmente se aplica a datos cuantitativos
#'
#' ##### PCA aplicado a datos ambientales
#'
#' Para aplicar PCA a datos ambientales, es necesario que todas las variables sean numéricas y "
#'
#' Dado que se requiere que las variables de entrada sean exclusivamente numéricas, el primer paso
#'
#' ¡IMPORTANTE! Haré esta demostración sólo con las variables de suelo, **pero puedes (y debes)
```

`#'`

```
#' A partir de los datos de suelo, la función 'rda', de 'vegan' realizará los siguientes pasos:
#'

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)
#'
#' Para agilizar la producción de scripts analíticos de referencia, trasladaré las explicaciones
#'
#' En el vídeo asociado, explico el significado de:
#'
#' - Inercia, *Inertia*
#' - Valores propios, autovalores, *Eigenvalues*
#' - Escalamiento, *Scaling*
#' - Puntuaciones de "especies", *Species scores*
#' - Puntuaciones de "sitios", *Site scores*
#'
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))
#
#' Observa que estoy generando un agrupamiento basado en los datos de suelo. No estoy comparando
#
#' Luego calculo las puntuaciones de los sitios para usarlas luego como coordenadas de los puntos
#
(puntuaciones <- scores(env_suelo_pca, display = 'wa', scaling = 1))
#
#' Luego creo el gráfico base, coloco los puntos sobre el gráfico usando las puntuaciones, les c
#
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
)
#'
#' Es razonable que el análisis cluster y el biplot muestren patrones consistentes, puesto que s
#'
#' Si hago lo mismo, pero usando mi análisis de agrupamiento anterior (*scripts* "aa_analisis_de
#'
# (mi_cluster_anterior <- grupos_upgma_k2)
(mi_cluster_anterior <- grupos_ward_k3)
(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
)

```

```

#'
#' Esto podría significar que las tendencias/patrones de mi matriz de comunidad (cuadros de 1 Ha
#'
#' ##### 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,
  ncpes = 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))
#'
#' Si intentáramos realizar el PCA a datos de comunidad no transformados, no recogeríamos apropi
#'
# Las especies que contribuyen mucho a los ejes 1 y 2 del PCA (aquellas cuyos vectores sobresal
#'
# Evaluaré el ajuste del PCA de datos de comunidad a datos ambientales, mediante la función 'en
#'
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)
#'
# Comento los resultados en el vídeo asociado. También probaré ajuste con todas las numéricas d
#'
# NOTA: te recomiendo probar otros métodos de selección de variables, como por ejemplo, usando
#'
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)
#'
#' Comento los resultados en el vídeo asociado.
#'
#' ¿Cuándo o a qué datos aplicar PCA?
#'
#' - PCA no es especialmente sensible a datos muy desviados de la normalidad.
#' - Como toda técnica, PCA tiene limitaciones.
#' - Las variables deben ser dimensionalmente homogéneas (unidades comparables o adimensionales).
#' - No usar en matriz transpuestas (no hace sentido la covarianza entre objetos).
#' - Es posible usar PCA con dato de presencia/ausencia, en cuyo caso, la matriz de comunidad de
#' - Las relaciones entre variables se miden por ángulos, no por proximidad de las puntas de los
#'
#' #### 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('Perexant', 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 Perexant")

```

```

)
plot(mi_fam_ca,
      scaling = 2,
      main = "CA, escalamiento 2, sin Thevetia ahouai")
par(mfrow = c(1, 1))
#
#' Análisis de coordenadas principales (PCoA)
#
#' Las técnicas de ordenación anteriores preservan la distancia euclídea entre los objetos. Si n
#
#' La función que realiza el PCoA en '{vegan}' es 'cmdscale' (de *Classical (Metric) Multidimens
#
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)

```

## 6.8 Técnicas de ordenación. Parte 2: Ordenación restringida o 'canónica'. RDA, CCA

```
#' ---
```

```
#' title: "Técnicas de ordenación. <br> Parte 2: Ordenación restringida o 'canónica'. <br> RDA,
```

```
#' author: "JR"
```

```
#' date: "21 de noviembre, 2020"
```

```
#' output: github_document
```

```
#' ---
```

```
knitr:::opts_chunk$set(fig.width=12, fig.height=8)
```

```
#' ## Preámbulo
```

```
#'
```

```

#' ### Cargar paquetes
#'
library(vegan)
library(tidyverse)
library(sf)
source('biodata/funciones.R')
#'
#' ### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_morac
(colnames(mi_fam) <- make.cepnames(colnames(mi_fam)))
(df_equivelencias <- data.frame(
  nombre_original = colnames(mc_morac),
  colnames(mi_fam)))
bci_env_grid %>% tibble
#'
#' ## Ordenación restringida
#'
#' Con este conjunto de técnicas de ordenación, las tendencias detectadas en un conjunto de datos
#'
#' Las principales técnicas de ordenación restringida son análisis de redundancia o RDA (siglas
#'
#' #### Análisis de redundancia (RDA)
#'
#' En el análisis de ordenación simple o no restringida, la matriz de datos (de comunidad o ambiente)
#'
#' RDA combina la regresión y el análisis de componentes principales. El RDA es una regresión lineal
#'
#' - Cada variable de respuesta (e.g. matriz de comunidad) es ajustada (por regresión) a la matriz de
#'
#' - Se prueba significancia estadística de la relación entre las variables de respuesta y explicativas
#'
#' - Si la prueba resulta significativa, se realiza un PCA de la matriz de valores ajustados. Esto
#'
#' - Luego se usa la matriz de vectores propios canónicos para calcular dos tipos de puntuaciones
#'
#'     - Usando la matriz de valores ajustados para obtener una ordenación en el espacio de las
#'
#'     - Usando la matriz de valores centrados de las variables de respuesta para obtener una ordenación
#'
#' Mostraré a continuación ejemplos aplicados a mi familia asignada. Para agilizar, trasladaré los
#'
#' Nota (de nuevo): recuerda que los datos de BCI están autocorrelacionados espacialmente, por lo tanto
#'
#' ##### Ejemplo usando las 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$', ignore.case = F))
env_suelo %>% tibble
mi_fam_hel_rda_suelo <- rda(mi_fam_hel ~ ., env_suelo)
summary(mi_fam_hel_rda_suelo)
#
#' ¿Qué partes del resumen debes mirar?
#
#' - La varianza, particionada, de la cual se muestra tanto la inercia (absoluta) como la proporcional
#
#' - Los valores propios (*eigenvalues*) y su contribución a la varianza. Existen *eigenvalues* y *eigenvectors*
#
#' - Las puntuaciones de sitios (*site scores*) y de especies (*species scores*) tienen el mismo signo
#
#' - Las puntuaciones restringidas de sitio (en '{vegan}' se usa la etiqueta *Site constraints*)
#
#' - Finalmente, el resumen presenta las puntuaciones para el *biplot*, que más adelante usaré en el análisis
#
#' Tal como comenté arriba, es importante calcular un R2 insesgado, siendo este más útil que el R2 sesgado
#
RsquareAdj(mi_fam_hel_rda_suelo)$adj.r.squared
#
#' Normalmente, el R2 insesgado es mucho más bajo que el sesgado, porque se relativiza con respecto a la varianza total
#
#' Otro diagnóstico importante es la determinación de los factores de inflación de la varianza (VIF)
#
vif.cca(mi_fam_hel_rda_suelo)
#
#' Variables con valores por encima de 10 deben ser examinadas y, desde una posición conservadora, se deben considerar
#
#' Finalmente, la representación del modelo se realiza en un *tripplot*, que es un gráfico enriquecido con información adicional
#
#' Escalamiento 1:
#
plot(mi_fam_hel_rda_suelo,
      scaling = 1,
      display = c("sp", "lc", "cn"),
      main = "Tripplot 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"
)
#
#'
#' Paso la interpretación de este gráfico y la del siguiente al vídeo correspondiente.
#
#' 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"
)
#
#'
## ##### Ejemplo usando las matriz ambiental con variables seleccionadas
#
#'
## El RDA anterior mostró que las variables de suelo son útiles para predecir la matriz de comunidades
#
## - Hay mucha colinealidad entre ellas.
#
## - No se probó mejorar el modelo añadiendo otras variables, además de las de suelo.
#
## Crearé una matriz ambiental con las variables que resultaron significativas en el ajuste *pos
#
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)
#'
#' Haré el plot del escalamiento 2 para comprobar gráficamente asociación entre variables sin la
#'
plot(mi_fam_hel_rda_selec,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de RDA especies ~ var. selec, escalamiento 2"
)
#'
#' Tal como comenté arriba, variables con valores VIF por encima de 10, deben ser examinadas. En
#'
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"
)
#'
#' Nota que las posiciones rotaron, no así la asociación entre la mayoría de las variables. 'B'
#'
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"
)
#'
#' Finalmente, la coordenada 'UTM.EW' tiene un alto valor VIF, por lo que es preferible excluirlo
#'

```

```

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"
)
#'
#' Las variables 'N.min.' y 'Zn' ahora presentan valores VIF aceptables. Habrás notado que es to
#'
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)

```

```

#'
#' Escalamiento 1
#'
plot(mi_fam_no_raras_cca_selec4,
      scaling = 1,
      display = c("sp", "lc", "cn"),
      main = "Triplot de CCA especies no raras ~ var. selec4, escalamiento 1"
)
#'
#' Escalamiento 2
#'
plot(mi_fam_no_raras_cca_selec4,
      scaling = 2,
      display = c("sp", "lc", "cn"),
      main = "Triplot de CCA especies no raras ~ var. selec4, escalamiento 2"
)

```

## 6.9 Análisis de diversidad alpha

```

#' ---
#' title: "Análisis de diversidad. <br> Parte 1: Diversidad alpha"
#' author: "JR"
#' date: "29 de noviembre, 2020"
#' output: github_document
#' ---

knitr::opts_chunk$set(fig.width=12, fig.height=8)

#' ## Preámbulo
#'
#' ### Cargar paquetes
#'
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-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_morac

```

```

bci_env_grid %>% tibble
grupos_upgma_k2 <- readRDS('grupos_upgma_k2.RDS')
table(grupos_upgma_k2)
grupos_ward_k3 <- readRDS('grupos_ward_k3.RDS')
table(grupos_ward_k3)
#'
#' ## Diversidad alpha
#'
#' **La tentación de medir la diversidad mediante un único número, ha atrapado a muchos investigadores**. La riqueza de especies, una cifra de uso común en estudios ecológicos, es un elemento muy sencillo y fácil de calcular.
#'
#' El término **"diversidad biológica"** se asume que fue acuñado a principios de los 80, y **se refiere a la diversidad de especies**.
#'
#' **Biodiversidad** es de factura más reciente, y es una contracción de "biológica-diversidad".
#'
#' La definición de biodiversidad del PNUMA es: **variabilidad entre organismos vivos** de todos los tipos.
#'
#' **Hubbell** (2001) ofrece una definición más adaptada a la práctica actual y mucho más restrictiva.
#'
#' En estas dos últimas acepciones (Hubbell y Magurran), la diversidad biológica puede dividirse en tres componentes principales:
#'
#' #### La diversidad de especies como un único número
#'
#' Usaré la notación *q* para designar el número de especies o riqueza de especies, y *n* para númer
#'
#' ##### Riqueza de especies y rarefacción
#'
#' Magurran (2004) distingue entre **densidad de especies**, que equivale al número de especies en un área dada.
#'
#' Para asegurar la comparabilidad entre sitios, se han propuesto distintos métodos. Uno es la rarefacción, que consiste en tomar muestras aleatorias de un sitio y calcular la diversidad en función de la cantidad de especies que se encuentran.
#'
#' 
#'
#' ##### Componentes de la diversidad de especies basada en abundancia: riqueza y equidad
#'
#' Asumamos que un sitio de muestreo es una variable cualitativa, y cada especie es un "estado". Los estados están ordenados por su abundancia.
#'
#' 
#'
#' Desde el punto de vista ecológico, la **H** tiene dos propiedades importantes:
#'
#' - Crece al aumentar la riqueza de especies *q*.
#' - Crece con la uniformidad (=equidad o equitabilidad, es decir, qué tan bien repartida se encuentra la diversidad).
#'
#' Para una *q* dada, *H* asume su valor máximo ( $H_{\max}$ ) cuando todas las especies es
#'
#' 

```

```

#'
#' Un ejemplo ilustra lo anterior:
#'
#' - 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)
#'
#' Otros índices miden exclusivamente la homogeneidad de la abundancia, a lo que se denomina equidad
#'
#' Uno de los más usados es la **equidad de Pielou** o *J* (1966), que es la razón entre la entropía y la riqueza
#'
#' 
#'
#' Teóricamente, **este índice sólo debería aumentar cuando aumenta la homogeneidad de la abundancia**
#'
#' Todos los **índices de equidad** se relacionan con la forma de los modelos de abundancia de especies
#'
#' Otra medida común en trabajos de ecología numérica, pero que es inversa o contraria a la equidad
#'
#' 
#'
#' Este valor aumenta con la dominancia, por lo que realmente no mide diversidad, sino más bien la riqueza
#'
#' La riqueza de especies (*q*), la entropía de Shannon (*H*), la equidad de Pielou (*J*) y el índice de Simpson (*S*)
#'
#' 
#'
#' donde *a* es el orden de la medida de entropía (*a=0,1,2...*), la cual cuantifica la importancia relativa de las especies
#'
#' 
#'
#' Las tres primeras entropías de Renyi (*Ha*), donde *a=0, 1 y 2*, y los correspondientes índices de diversidad de Hill
#'
#' Por otra parte, los tres primeros números de diversidad de Hill tienen significados que conocemos
#'
#'  <br>
#' > Según Borcard et al., 2018.
#'
#' Bajo esta notación:
#'
#' - La equidad de Pielou (denominada también equidad de Shannon) equivale a *J=H1/H*

```

```

#'
#' Usando la misma aproximación, Hill propuso también otras ratios que, a diferencia de lo que o
#'
#' -  $E<sub>1</sub>=N<sub>1</sub>/N<sub>0</sub>$ * a la cual el propio Hill denominó como su versió
#'
#' -  $E<sub>2</sub>=N<sub>2</sub>/N<sub>0</sub>$ , a la cual Hill denominó como su versión de la
#'
#' Los números de diversidad y las ratios de Hill son menos sensibles a las matrices de comunida
#'
#' **Estas afirmaciones tienen implicaciones muy importantes desde el punto de vista ecológico,
#'
#' **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
#'
abundancia[abundancia == min(abundancia)]
abundancia[abundancia == max(abundancia)]
(rango_abun <- range(abundancia))
#'
#' Abundancia en el sitio más pobre
#'
abundancia[riqueza == min(riqueza)]
#'
#' Abundancia en el sitio más rico
#'
abundancia[riqueza == max(riqueza)]
#'
#' Riqueza en el sitio con menor abundancia
#'
riqueza[abundancia == min(abundancia)]
#'
#' Riqueza en el sitio con mayor abundancia
#'
riqueza[abundancia == max(abundancia)]
#'
#' 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 dic
#'
#' Aproximación básica:
#'
specpool(mi_fam)
specpool(mi_fam)[[1]]/specpool(mi_fam)*100
#'

```

```

#' Lista comprensiva de métodos (incluyendo recientes):
#'
#' - **Enfoques asintóticos. Estiman la riqueza de especies**:
#'   - Paramétricos:
#'     - Modelo homogéneo (estándar y MLE)
#'   - No paramétricos:
#'     - Chao1 y Chao1-bc
#'     - iChao1
#'     - Basados en "cobertura" o "completitud de muestra". ACE para datos de abundancia
#'     - Estimadores Jackknife (de primer y segundo órdenes)
#' - **Enfoques no asintóticos. Se utilizan para hacer rarefacción y extrapolación**:
#'   - Basados en tamaño de la muestra
#'   - Basados en "cobertura" o "completitud de muestra"
#'
#' ##### Enfoques asintóticos y no asintóticos aplicados a la matriz de comunidad combinada
#'
#' 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$no_asintoticos_rarefaccion_extrapolacion_grafico
#'
#' ##### Enfoques asintóticos y no asintóticos aplicados a una matriz de comunidad agrupada
#'
#' #' Generar matriz de comunidad agrupada según el método de Ward (tres grupos), procedente de
#'
mi_fam_k3 <- mi_fam %>%
  mutate(g=grupos_ward_k3) %>%
  group_by(g) %>%
  summarise_all(sum) %>%
  select(-g) %>%
  data.frame
mi_fam_k3 %>% rowSums %>% sort
mi_fam_k3_chao <- estimacion_riqueza_chao(
  mc = mi_fam_k3,
  n_raras = 10)
mi_fam_k3_chao$asintoticos_estimacion
mi_fam_k3_chao$no_asintoticos_rarefaccion_extrapolacion
mi_fam_k3_chao$no_asintoticos_rarefaccion_extrapolacion_grafico

```

## 6.10 Análisis de diversidad alpha

```
#' ---
#' title: "Análisis de diversidad. <br> Parte 2: Diversidad beta"
#' author: "JR"
#' date: "2 de diciembre, 2020"
#' output: github_document
#' ---

knitr::opts_chunk$set(fig.width=12, fig.height=8)

#' ## Preámbulo
#'
#' ### Cargar paquetes
#'
library(vegan)
library(adespatial)
library(dplyr)
library(tidyverse)
library(sf)
library(vegetarian)
library(mapview)
source('biodata/funciones.R')
#'
#' ### Cargar datos
#'
load("~/unidad-0-asignacion-99-mi-manuscrito-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_morac
bci_env_grid %>% tibble
#'
#' ## Diversidad beta
#'
#' En la concepción de Whittaker, la diversidad alfa es la local o de sitio, la diversidad beta
#'
#' La diversidad alfa la medimos usando métricas que ponderan comúnmente la riqueza y la equidad
#'
#' ### Diversidad beta con un único número
#'
beta_multiplicativa <- calcular_beta_multiplicativa(
  mc = mi_fam,
  orden = 0:20)
beta_multiplicativa
#'
#' Notar que, para esta matriz de comunidad, la diversidad beta multiplicativa disminuye en exceso
#'
#' ### 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)
#'
#' Utilizar el mapa de cuadros para identificar aquellos que contribuyen más a la diversidad bet
#'
mapa_cuadros

```

### 6.11 Análisis espacial de datos ecológicos. Autocorrelación

```

#' ---
#' title: "Análisis espacial de datos ecológicos. <br> Autocorrelación"
#' author: "JR"
#' date: "5 de diciembre, 2020"
#' output: github_document
#' ---
#'
knitr::opts_chunk$set(fig.width=12, fig.height=8)
#'
#' ## Preámbulo
#'
#' ### Cargar paquetes
#'

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

```

```

#' library(ape)
library(spdep)
library(ade4)
library(adegraphics)
library(adespatial)
load("~/unidad-0-asignacion-99-mi-manuscrito-GeografosigloXXV/biodata/Moraceae.Rdata")
load('biodata/matriz_ambiental.Rdata')
mi_fam <- mc_morac
mi_fam %>% tibble
bci_env_grid %>% tibble
#
# ## Preparar datos
#
# ### 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)
#'
#' Para aplicar la prueba Mantel a datos de comunidad, es necesario quitar la tendencia espacial
#'
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" rodeado de "y")',
                                       leyenda = F,
                                       anchuratitulo = 50,
                                       tamanotitulo = 10,
                                       fuenedatos = '\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,
                  fuenedatos = '\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$leyenda <- gtable_filter(ggplot_gtable(ggplot_build(lisamaps_mifam[[1]] + theme(
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam[1:8], nrow = 3)), lisamaps_mifam$leyenda,
grid.arrange(do.call('arrangeGrob', c(lisamaps_mifam[9:16], nrow = 3)), lisamaps_mifam$leyenda,
#'
 #' ##### Aplicado a abundancias de especies transformadas sin tendencia
#'
mi_fam_sintendencia_sf <- bci_env_grid %>% select %>% bind_cols(mi_fam_sin_tendencia %>% as.data
lisamaps_mifam_sintendencia <- sapply(
  grep('geometry', names(mi_fam_sintendencia_sf), invert = T, value = T),
  function(x) {
    m <- lisamap(objesp = mi_fam_sintendencia_sf[x],
                 var = x,
                 pesos = pesos_b,
                 tituloleyenda = 'Significancia ("x-y", léase como "x" rodeado de "y")',
                 leyenda = F,
                 anchuratitulo = 50,
                 tamanotitulo = 10,
                 fuentedatos = '\nhttp://ctfs.si.edu/webatlas/datasets/bci/',
                 titulomap = paste0('Clusters LISA de "', x, '''))
    # dev.new();print(m$grafico)
    return(m$grafico)
}, simplify = F
)
lisamaps_mifam_sintendencia$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

```

## Referencias

- Cardona, W., De Ulloa, P. C., & Kattan, G. (2007). Avispas no polinizadoras asociadas a *ficus andicola* (moraceae) en la cordillera central de colombia. *Revista Colombiana de Entomología*, 33(2), 165–170.
- Clement, W. L., & Weiblen, G. D. (2009). Morphological evolution in the mulberry family (moraceae). *Systematic Botany*, 34(3), 530–552.
- Fredericksen, T. S., Justiniano, M. J., Rumiz, D., McDonald, E., & Aguape, R. (n.d.). *Bibosi higuerón ficus spp. moraceae*.
- Magallanes, A. B., Rocha, Y. R., & Terán, F. A. (n.d.). *Distribución y abundancia de las especies arbóreas de la familia moraceae en la reserva ecológica del mineral de nuestra señora de la candelaria, cosalá, sinaloa*.
- Piedra-Malagón, E. M., Ramírez Rodríguez, R., & Ibarra -Manríquez, G. (2006). El género *ficus* (moraceae) en el estado de morelos, méxico. *Acta Botánica Mexicana*, (75), 45–75.
- Simmonds, J. A., Gómez, J. A., & Villalaz, J. (2002). Parametros fisico-quimicos y biologicos en

aguas circundantes al canal de panama. *Tecnociencia*, 4(1), 47–69.

Van Devender, T. R., Felger, R. S., Fishbein, M., Molina-Freaner, F. E., Sánchez-Escalante, J. J., Reina-Guerrero, A., & others. (2010). Biodiversidad de las plantas vasculares. *Diversidad Biológica de Sonora. Universidad Nacional Autónoma de México*. México, DF, México, 229–261.

Williams-Linera, G., & Meave, J. (2002). Patrones fenológicos. *Ecología Y Conservación de Bosques Neotropicales, RM Guariguata Y GH Kattan (Eds .)*. Libro Universitario Regional, San José, 591–624.