STRUCTURE AND PHYLOGENETIC DIVERSITY OF VASCULAR PLANTS IN COMMUNITIES OF SEASONALLY DRY TROPICAL FOREST IN COLOMBIA

Presentado por
María Claudia González

Requisito parcial para optar al título de
Magister en Biología aplicada

Directora
Mailyn Adriana Gonzalez, Ph.D
Directora del laboratorio de genética de la conservación
Instituto de investigaciones Alexander von Humboldt

Co-director
Ericsson Coy-Barrera, Ph.D
Director del laboratorio de química bioorgánica
INQUIBIO

UNIVERSIDAD MILITAR NUEVA GRANADA
Facultad de Ciencias Básicas y Aplicadas
Programa de Maestría en Biología Aplicada
2018
Cajicá, Colombia

Acuérdate de mirar hacia cielo,

eso expande los límites de la mente

y nos recuerda que somos una

pequeña parte del inmenso universo,

que está siempre en movimiento

- Budismo -

AGRADECIMIENTOS

Este trabajo no hubiese sido posible sin la participación de muchas personas (tal vez muchas que no conozco), y han contribuido durante todo el proceso, desde el trabajo en campo hasta los análisis posteriores de los datos, principalmente al Instituto Humboldt, por apostarle a la realización de este proyecto.

Quiero agradecerle de manera muy especial a Mailyn, por haberme dado la oportunidad de realizar este trabajo bajo su dirección, por el apoyo y la confianza depositada en mí, por la paciencia y todo lo aprendido a su lado, por haberme enseñado el mundo de la ecología genética de comunidades y la importancia de esto en el estudio y entendimiento de nuestros ecosistemas, campo en el cuál quiero seguir guiando mis futuros estudios, gracias por haberme abierto un espacio en el grupo de genética del cuál siempre me he sentido parte. A Ghislaine por haberme acompañado y enseñado las técnicas en el laboratorio. Henry, Elkin y Paola quienes siempre estuvieron dispuestos a escuchar con paciencia las dudas que iban surgiendo en el proceso y ayudarme a resolverlas, con ustedes aprendí el manejo de muchas herramientas desconocidas para mí, y a cómo interpretar los resultados de los análisis desarrollados, espero en un futuro podamos llevar a cabo trabajos en conjunto.

Al profesor Ericson quién aceptó ser mi coordinador, aunque no tuviera mucho conocimiento en el tema, fueron muy importantes sus consejos al momento de tomar decisiones que fueron decisivas para la continuidad de este proyecto.

A Alma, Mari, Juli y Roy gracias por su invaluable amistad por apoyarme y animarme en los momentos de preocupación. Finalmente, y no por ser menos importante, a mi familia en Cali quienes me brindaron un hogar durante mi estadía en esa ciudad. A mis padres Rocío y Gonzalo, a mis hermanos Felipe y Carlos a Sebas y a Fede, quienes han apoyado y guiado cada uno de los pasos en mi vida, muchos de mis logros son por ustedes y para ustedes.

CONTENT

FIGURE LIST

TABLE LIST

ABSTRACT

Seasonally dry tropical forest (SDTF) is one of the most threatened ecosystems worldwide. With the aim to provide new insights into the conservation of this ecosystem, I characterized plant diversity (alpha and beta) and phylogenetic structure of SDTF of fifteen permanent plots located across the best-conserved remnants of this ecosystem in Colombia. I built a megatree phylogeny for 373 species an assessed the evolutionary distinctiveness of the communities as well as the phylogenetic diversity with two metrics mean pairwise distance (MPD) and mean nearest taxon distance (MNTD) that quantify basal and terminal evolutionary history of communities. Phylogenetic alpha diversity was found to be coupled with species diversity; however, species with high evolutionary distinctiveness where unevenly distributed. Further, I tested if patterns of phylogenetic diversity, could be explained by eleven environmental variables but none of these were good predictors of phylogenetic alpha diversity. Instead, for beta taxonomic and phylogenetic diversity four environmental variables explained up to 11% of species turnover. Phylogenetic structure was assessed with Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI), finding that 5 plots presents phylogenetic clustering meaning that coexisting species are more closely related than expected by chance; on the other hand, 9 plots showed a random assembly of species. The observed patterns were not associated with a particular region suggesting that despite the current fragmentation of SDTF in Colombia, lineages are not restricted to particular regions. Finally, to test if the resolution of the phylogeny employed could bias the observed patterns of phylogenetic structure, for a subset of six plots I generated a well-resolved phylogeny based on DNA sequences of the plastid region rbcLa and compared the results obtained with the megatree. There were no differences for patterns of MPD, however for MNTD the rbcLa phylogeny elucidated more structure indicating that well-resolved phylogenies are important when employing terminal measures of phylogenetic diversity. In summary, phylogenetic structure does not reflect the extreme environmental heterogeneity of SDTF.

However, the results suggest that deterministic and neutral processes act together in the processes of community assembly, highlighting the need to integrate more areas covered by SDTF to better understand the ecological processes acting in the community assembly.

INTRODUCTION

Understanding the processes that allow the coexistence of many species in a community is still a main target in ecology. In that sense, characterizing the evolutionary relationships of co-occurring species may shed light into their assembly process (Webb 2000, Cavender-Bares et al. 2009). Indeed, once dispersal limitation is overcoming, community assembly results both from the filtering of species by abiotic and biotic factors leading to particular patterns of phylogenetic structure (Losos 1996, Webb et al. 2002, Chase 2003, Kraft et al. 2007, Cavender-Bares et al. 2009, Mittelbach and Schemske 2015). Commonly, community diversity and structure has been characterized from a taxonomic perspective (species number and their relative abundance). However, assessing phylogenetic diversity allows to better understand the uniqueness of evolutionary lineages in a community and the ecological process dominating community assemblage (Forest et al. 2006, Faith 1992). A considerable number of studies have addressed the phylogenetic diversity and structure of temperate and tropical rain forests (Webb 2000, Cavender-Bares et al. 2004, Kembel and Hubbell 2006). However, only few studies have tackled the phylogenetic structure of tropical dry forests (Swenson and Enquist 2009, Freiro-Moro et al. 2015). In this study I sought to characterize patterns of species and phylogenetic diversity and structure in remnants of tropical dry forest in Colombia to understand community assembly and to provide insights for their future planning.

The seasonally dry tropical forest (SDTF) is characterized by a dry season of at least 3 months with precipitation below than 100 mm and evapotranspiration above 250 mm. Owing to this hydric deficit, species that live in this ecosystem have particular physiological and morphological adaptations that allow their survival and reproduction (Pennington et al. 2000, Portillo-Quintero et al. 2010, Sánchez-Azofeifa et al. 2005, Pizano et al. 2014). Globally, SDTF currently occupies an area close to 1,048,700 Km^2 (Miles et al. 2006, Portillo-Quintero and Sánchez-Azofeifa 2010) and

is distributed in south Sub-Saharan Africa (including Madagascar) (13.1%), south of Asia (3.8%) and north Australia (26.4%) (Pennington et al. 2000, FAO 2001, Reinaldo et al. 2011, Sunderland et al. 2015). In Mesoamerica, it has an extension of 12.5%, located in Mexico along the Pacific coast and Guanacaste in the north of Costa Rica. Nevertheless, it is in South America where SDTF reaches its largest extension with at least 54.2% (Figure 1) (Pennington et al. 2000, Miles et al. 2006, Reinaldo et al. 2011). Although SDTF currently has a highly fragmented distribution, it is believed that during the last glacial period (18.000 – 12.000 BP) in South America it reached its mayor extension, along the west of Brazil, Argentina and Paraguay and in more disjunct areas through the inter-Andean valleys of Peru, Bolivia, Ecuador and Colombia (Barneby 1991, Prado and Gibbs 1993, Barneby and Grimes 1996, Pennington et al. 2000).

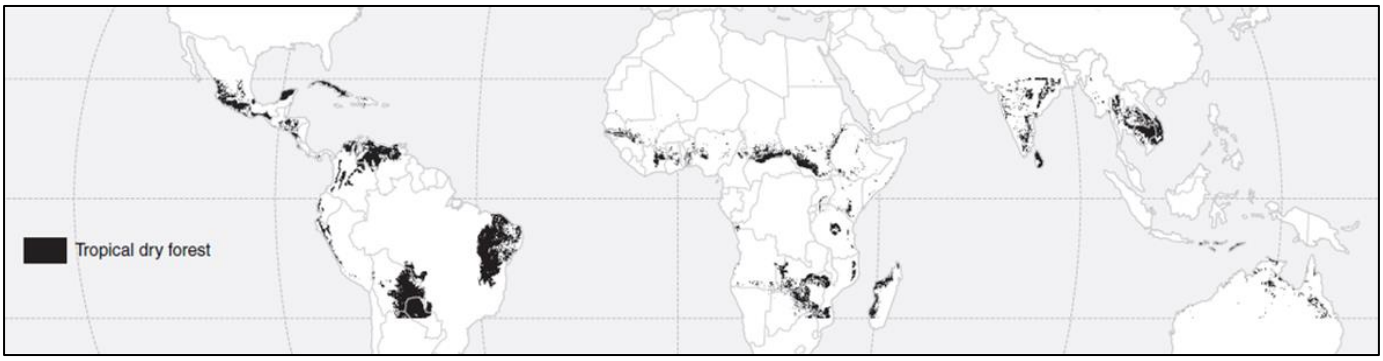

Figure 1. Global distribution of tropical dry forest in the year 2000, modified from Miles et al. (2006).

Globally, SDTF is considered the most threatened ecosystem with less than 10% of its original extension remaining in many countries (Banda et al. 2016, Portillo-Quintero and Sánchez-Azofeifa 2010, Sunderland et al. 2015). Particularly in Colombia, only 8% of the original extension of eight million hectares remains (Pizano and García 2014) (Figure 2). A recent study on species turnover across the neotropical dry forests identified 12 floristic groups highlighting the importance of considering such differences for conservation (Banda et al. 2016). In Colombia, the authors documented two floristic groups; the first one covers Central America and the north

of South America, and the second corresponds to the inter-Andean valleys. However, in a specific study for Colombia, Pizano et al. (2014) identified three floristic groups corresponding to the Caribbean coast, the inter-Andean valleys and the plains of Orinoquía. These studies are the first steps to understand the biogeographic history of the SDTF in Colombia which remains elusive.

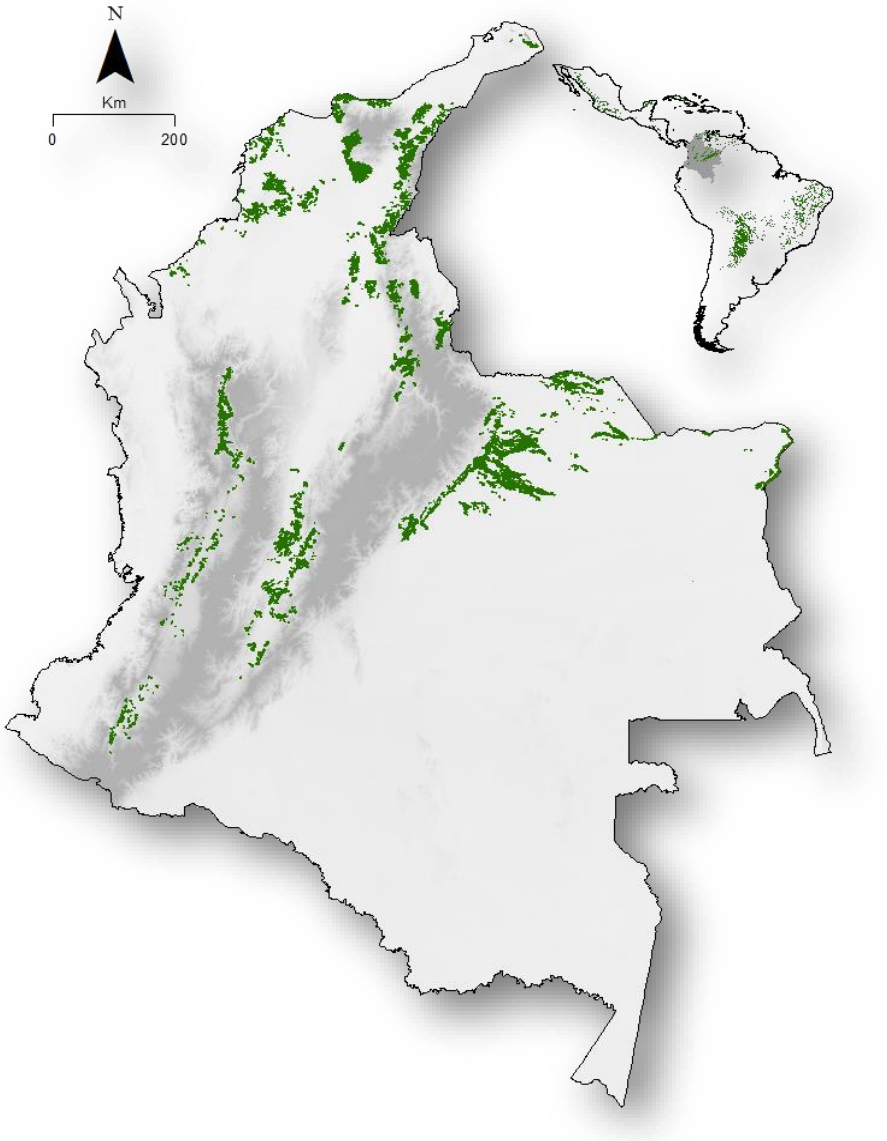

Figure 2. Current distribution of tropical dry forest in Colombia (modified from Pizano and García 2014)

Characterizing the evolutionary composition and turnover of SDTF communities should allow for a better understanding of the historical link of current floristic nodes and the dominant ecological processes that determine community assembly in this ecosystem. Community assembly results from the dispersion of species from a regional pool that may further be filtered through abiotic conditions and biotic interactions determining colonization and coexistence in a local site (Webb 2000, Kraft et al. 2007). Assuming that species more closely related in their evolutionary history are more ecologically similar, two major patterns of phylogenetic community structure have been described. For instance, under abiotic pressures, it is expected that species sharing similar ecological traits will coexist leading to a pattern of evolutionary clustering (Webb 2000, Cavender-Bares et al. 2004, 2006). On the other hand, when resource constraints increase competition among closely related species, their coexistence is limited (MacArthur and Levins 1967) leading to a community composed of distantly related species or phylogenetically overdispersed (Webb et al. 2002, Chase 2003, Cavender-Bares et al. 2009). However, another scenario supposes that species are ecologically equivalent and their presence and abundance in a community will be conditioned solely by their dispersion capacity and abundance in the regional pool (Hubbell 2001). This last scenario considers that local communities result from a random sampling from the regional pool.

Given the patchy distribution of SDTF, it has been argued that this biome is dispersal limited (Pennington 2006). Besides, given the differences on environmental and edaphic conditions throughout this ecosystem, (González-M et al. 2018) one could expect to find phylogenetic clustering in local communities where their assemblage would be composed of closely related species that share their tolerance to particular environmental conditions. The main objective of this study was to characterize the phylogenetic diversity and structure of fifteen permanent plots of SDTF present in five geographic regions of Colombia to understand the processes underlying their assemblage, in order to provide insights for dry forest conservation.

The increasing availability of DNA sequences and molecular phylogenies has allowed for the incorporation of the phylogenetic perspective in community ecology studies. Still, the majority of these studies assess phylogenetic metrics from "megatree" phylogenies that resolve up to the family or genera level a resolution that may bias the detection of structure patterns (Webb & Donoghue 2005, Gonzalez et al. 2010, Coronado et al. 2014). This limitation is stronger in neotropical rich flora like in Colombia, where less than 5% of the species have any genetic information available (González et al. 2014). Therefore, with the aim of testing if phylogenetic resolution may influence structure patterns, I built a well-resolved phylogeny with DNA sequences for a subset of six of the studied plots and compared the results with a megatree phylogeny.

METHODS

Study sites and sampling collection

This study was carried out in fifteen permanent plots distributed across five geographic regions in Colombia where the seasonally dry tropical forest (SDTF) is found (Table 1). The plots are one hectare each and were established between 2013 and 2015 by the "Alexander von Humboldt Biological Resources Research Institute" (IAvH) through the tropical dry forest network (Figure 3).

Table 1. Plots included in this study including region, species richness, and the number of trees with dbh>5 cm (density).

Plot names	Region	Species richness	Density
Colorados	Caribbean	73	1265
Macuira		32	809
Matitas		27	1381
Plato		32	887
Tayrona		42	1062
Cotove	Cauca Valley	19	694
Tamesis		15	221
Vinculo		43	1747
Tuparro	Llanos	62	551
Caparrapi	Magdalena Valley	62	67
Cardonal loma		42	2008
Cardonal plana		44	1341
Jabiru		35	1256
Tambor		68	567
Taminango	Patia Valley	4	623

Figure 3. Geographical location of plots included in this study. Plots are colored according to the region where they are located as follows: Caribbean in red, Magdalena in orange, Cauca in yellow, Llanos in green and Patia in grey. The triangle plots were used for the construction of a phylogeny with molecular sequences.

In each plot all vascular plants with diameter at breast height (dbh) ≥ 5 cm were marked and identified by botanical specialists in the herbarium (Figure 4). Overall, for the 15 plots, 373 morphospecies were recorded, from these 67% were identified to species level and 33% to genus

level. For further DNA analysis, small pieces of leaves of all species were collected, that were dried and stored in silica gel (Gonzalez and Quintero 2017) (Figure 5).

Figure 4. Establishment of permanent plots between 2013 and 2014 by the IAvH.

(a)

(b)

Figure 5. (a) Marking and census of all plants with DBH ≥ 5 cm. (b) Collection and storage of leaves for genetic analyses.

Phylogenetic trees

I generated two phylogenetic hypotheses using the phylomatic platform (http://phylodiversity.net/phylomatic/) (Webb et al. 2008). The first phylogeny comprises the total species present in the fifteen plots and the second phylogeny comprises the species present in a subset of six plots (Cotove, Matitas, Jabiru, Vinculo, Taminango and Tuparro). The base megatree used was R20120829 that is based on the APG III (Angiosperm Phylogeny Group 2009) phylogenetic classification of flowering plants from which I pruned 373 species as the regional pool of the 15 plots and 165 species as the regional pool of the 6 plots (Webb and Donoghue, 2005). Branch lengths for both phylogenetic trees were calculated with the BLADJ algorithm available in Phylocom v.4.2 (Webb et al. 2008) and based on 30 nodes ages from Magallón et al. (2015). BLADJ fixes dates of the given nodes to the megatree uniformly between known nodes on the phylogeny getting a phylogeny with branches calibrated in millions of years (Webb et al 2004).

A third phylogeny was comprising for the same subset of 165 species present in the subset of six plots (Cotove, Matitas, Jabiru, Vinculo, Taminango and Tuparro). That were constructed with molecular sequences, as follow:

DNA extraction, PCR amplification and sequencing

Around 25 mg of leaf tissue was used for DNA extraction; tissue was homogenized with liquid nitrogen. DNA was extracted following Ivannova (2008) protocol modifying the lysis buffer with CTAB 1% + PVP buffer and carrying incubation for one hour at 65°C. For the amplification of the first portion of the plastid rbcLa gene the primers: *1f: 5'*-ATGTCACCACAAACAGAAAC-3'; *r724: 5'*-TCGCATATGTACCTGCAGTAGC-3' were used. PCR mix contained 0,2 ml GoTaqH 51 U/ml (Promega), 2 µl of 0X buffer, 1 µl of 10 mM for each primer, 0,4 µl of dNTP 10 mM, 2 µl of ADN

and H_2O for a final volume of 20 µl. The PCR products were examined at 1,5% of agarose gel and 1% of TBE, and then were visualized using an ultraviolet trans-illuminator. The PCR program was 94 °C for 1 min, followed by 5 cycles of 30 s at 94 °C, 40 s at 50 °C and 2 min at 72 °C, continuing with 35 cycles of 30 s at 94 °C, 40 s at 54 °C, 2 min at 72 °C, with a final step of 5 min at 72 °C.

I obtained rbcLa sequences for 92 species that were deposited on the Barcode of Life Data System platform, in the project "Tropical Dry Forest Colombia" (BSTIH) and that correspond to the first's sequences for these species in global databases. The remaining 73 sequences were downloaded from GenBank (Table S1). Samples without genetic information at species level were replaced by a congener reported to be found in other tropical dry forests (http://www.dryflor.info/). Sequences obtained were assembled and edited in Geneious9, then aligned using MUSCLE of MEGA 7 packages (Kumar et al. 2016). Finally, the alignment was finally manually adjusted.

Phylogenetic analyses were conducted with Bayesian methods, implementing a GTR model (General Time Reversible) according to the results obtained from jmodeltest-2.1.10 (Posada 2010). Dating analyses were conducted with a relaxed molecular clock and the Yule model for calibration, in the software BEAST v2.0 (Drummond and Bouckaert 2014). I included a sequence of *Amborella trichopoda* (Amborellaceae) as the outgroup. For the calibration, 30 points were used from Magallón (2015) with a normal distribution (Ho 2007, Sauquet 2013), 30 million generations were run and 50% of the initial trees sampled in each run were discarded as burn-in. The results of the BEAST analyses were evaluated in TRACER v1.6 (Rambaut et al., 2014). In which was verified that the effective sample sizes of estimated parameters were greater than 100. Finally, the TreeAnnotator v2.4.3 software (Drummond and Bouckaert 2014) was used to generate a consensus tree.

DIVERSITY METRICS

1. Taxonomic and phylogenetic alpha diversity

Taxonomic alpha diversity was calculated with a Shannon's diversity (H') index that incorporates species abundance (Shannon 1948).

$$\text{Hsb} = \sum_{i=1}^{s} \text{pi} \log \text{pi}$$

Were *S*, is the number of species in the assemblage and *pi* is the relative abundance of the *ith* species.

The indices for assessing phylogenetic alpha diversity were the mean pairwise distance (MPD) and the mean nearest taxon distance (MNTD), both accounting for species abundances. MPD calculates the mean phylogenetic distance between all possible pairwise combinations of individuals of a local assemblage, and it is considered a basal measure of relatedness. MNTD quantifies the mean phylogenetic distance for each individual to its nearest relative in a plot; is more sensitive to variations towards the tips of the topology so it is considered a terminal relatedness measure (Webb 2002, Webb et al. 2008, Kembel et al. 2010, Tucker et al. 2016).

$$MPD = \sum_{i} \sum_{j} dij \; Pi \; Pj$$

And

$$MNTD = \sum_{i=1}^{s} [d_{i\,min} * Pi]$$

Where:

- *dij* = Phylogenetic distance between two species *i* and *j*.
- *di min* = Distance of a given specie *(i)* to its nearest neighbor relative in the assemblage.

- *pi*/*pj* = Probability to draw an individual of specie *i* from the assemblage (measured as a relative abundance).
- *S* = Species richness.

A Person correlation test with a confidence interval of 95% was performed to evaluate the correlation of phylogenetic diversity measurements (MPD and MNTD) with taxonomic diversity (Shannon index). To assess the relationship between environmental variables and phylogenetic diversity, (MPD and MNTD) I ran a linear model using the following variables: mean annual temperature (MAT, ºC), total annual precipitation (AP, mm), potential evapotranspiration (PET, mm) determined as the sum of monthly potential evapotranspiration using the Thornthwaite equation (Thornthwaite 1948), total precipitation during the three driest months (<100 mm·month-1) (TPdriest, mm), total of annual rainy days (Ard, no.), isothermality (Isoth, %) analyzed as the large day-to-night temperatures oscillations relative to annual oscillations (Bio3, O'Donnell and Ignizio 2012), aridity index (Aridity) calculated as the TAP/PET ratio (Zomer et al. 2008), the average number of rain days monthly (M.rainyd); the average number of rain days in a year (A.rainyd), number of dry months (Drymonths) and mean rainfall of the driest month (P.drymonth). These environmental variables were determined using the National Climatic Source, which include 2.046 weather stations in Colombia with monthly data of mean temperature, total precipitation and rainy days (~90 m spatial resolution, http://institucional.ideam.gov.co/jsp/1769) (González-M et al. 2018).

The possible collinearity between the selected eleven variables was verified using the variance inflation factor (VIF) for linear models, which measures the proportion by which the variation of a regression coefficient is inflated in presence of other explanatory variables. Variables with a VIF > 5 were eliminated. To identify which of the variables selected in the previous step were better

predictors for phylogenetic diversity, I applied the Akaike information criterion (AIC) and chose those with lowest values.

Evolutionary distinctiveness (ED)

This index represents the amount of evolutionary history kept by each lineage (Redding and Mooers 2006). To calculate this index, each branch is divided in segments (between nodes or to the root) called EDGE. To each EDGE an ED score is assigned which represents the timespan it covers (in millions of years) divided by the number of species at the end of the subtree it forms. Finally, the ED for each species corresponds to the sum of EDGES values for all branches from which the species descend (Pavoine et al. 2005, Redding et al. 2008).

$$ED_{(T,i)} = \sum_{e \in q(T,i,r)} \left(\lambda e \, . \frac{1}{S_e} \right)$$

Where e is a branch of length λ in the set $S(T,i,r)$ connecting species i to the root r and Se is the number of species that descend from edge e.

2. Taxonomic and phylogenetic beta diversity

Beta diversity represents the turnover of species among assemblages and quantifies the similarity or dissimilarity between two communities, (Whittaker 1960, Jaccard 1912, Simpson 1943) whereas phylogenetic beta diversity assesses the turnover of phylogenetic composition of assemblages, being a measurement related to evolutionary history (Graham and Fine 2008).

Taxonomic beta diversity (TBD) and phylogenetic beta diversity (PBD) between all pairwise communities was calculated with the indices Sorensen and PhyloSor respectively (Bryant et al. 2008), defined as:

$$Sorensen_{ij} = \frac{S_{ij}}{(S_i + S_j)\frac{1}{2}}$$

And

$$Phylosor_{ij} = \frac{BL_{ij}}{(BL_i + BL_j)\frac{1}{2}}$$

Where

- S_{ij} = Number of taxa common to both communities (*i* and *j*).
- S_i and S_j = Number of exclusive species in each community (*i* and *j*).
- BL_{ij} = Total length of branch lengths shared between communities *i* and *j*.
- BL_i and BL_j = Total branch lengths exclusive for each community *i* and *j*.

These indices ranged from 0 to 1. Values close to 1 indicate that two communities are completely different with regard to their taxonomic and phylogenetic composition; on the opposite, lower values (near 0), means that the two communities have the same set of species or evolutionary lineages.

A cluster analysis (for Sorensen and PhyloSor indices) was performed with the algorithm UPGMA (Unweighted Pair Group Means Algorithm). Based on the distance matrices (Sorensen and Phylosor) the algorithm begins by grouping the plots and calculates the average distance between each group, generating a new distance matrix which is compared with the original distance matrix, then the process is repeated until it finds the closest distances to the original matrix. (Kreft and Jetz 2010, Moreno Saiz et al. 2012, Holt et al. 2013).

Generalized dissimilarity modeling

Generalized dissimilarity model (gdm) is a nonlinear statistical approach, used to evaluate the contribution of the geographical distance and environmental variables to explain the species and lineages turnover among plots (Ferrier et al. 2004, 2007, Roseaur et al. 2013, Fitzpatrick et al. 2013). Different from the lineal models, such as Mantel test, the gdm has two principal advantages. First, it assumes a non-linearity in the measure of compositional dissimilarity and any gradient (the linear relation between the increasing of dissimilarity and the environmental or spatial scale). Second, it assumes that there is a different turnover rate in the entire range of each variable (Ferrier 2004, 2007). I used the eleven environmental variables (described above) as environmental predictors, and generated two gdm with Sorensen and PhyloSor dissimilarity between pairs, creating a site-by-site distance matrix (dissimilarity).Finally, the model was fitted, as described in Ferrier et al. (2007) in the R package "*gdm*" (Manion et al. 2017) available at https://www.rdocumentation.org/packages/gdm/versions/1.3.3. Then, the variables that were selected by the model were plotted. In the graph, the height of the curve represents the relative importance of each variable, while the amount of turnover is represented by the increase in the slope.

PHYLOGENETIC COMMUNITY STRUCTURE

Two metrics of phylogenetic community structure were calculated: the Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI). These indices are based on a comparison of observed values of MPD and MNTD respectively, relative to a null model distribution under which species position in the phylogeny are randomized. Positive values mean that the taxa in a sample are less related than expected by the random model (phylogenetic overdispersed) and negative values suggest that species are more closely related than expected by chance (phylogenetic clustering) (Webb 2000, Cavender-Bares et al. 2004, 2006, Webb et al. 2008, Vamosi et al. 2009). The observed phylogenetic distances for each sample were compared to the distribution of 9999 null communities. I used the "richness" null model, in which the abundances are assigned at random and the richness does not change within communities. The model assumes that the species of the regional species pool (all species present in the plots considered) are equally able to colonize any community (Gotelli 2000, Hardy 2008, Mouquet 2012, Miller et al. 2016). Phylogenetic structure indices were calculated using the R package "*Picante*" (Kembel et al. 2010) available at http://cran.at.r-project.org/web/packages/picante/.

$$NRI = \frac{\mathrm{MPDobs} - \mathrm{MPDrnd}}{\mathrm{std.MPDrnd}}$$

And

$$NTI = \frac{\mathrm{MNTDobs} - \mathrm{MNTDrnd}}{\mathrm{std.MNTDrnd}}$$

Where *obs* is the observed community, *rnd* is the random community and *std* is the standard deviation (Webb et al. 2002).

To determine if the phylogenetic structure metrics are affected by the resolution of the phylogeny employed, I compared the results of NRI and NTI obtained from the megatree and those with the phylogeny generated from DNA sequences of the plastid region rbcL-a, both for the plots of Tuparro, Matitas, Vinculo, Taminango, Cotova and Jabirú.

RESULTS

DIVERSITY METRICS

1. Taxonomic and phylogenetic alpha diversity

Taxonomic diversity shows a significant positive correlation with MPD (r = 0.89, p = 8.402 e-06; Figure 6) and a significant negative correlation with MNTD (r = -0.62, p = 0.01; Figure 6). Given the low number of species in the Taminango plot (only 4 species), the correlation was run without this plot. For MPD the positive correlation was kept significant (r = 0.80, p = 0.0006), but for MNTD the correlation disappeared (r = -0.11, p = 0.70).

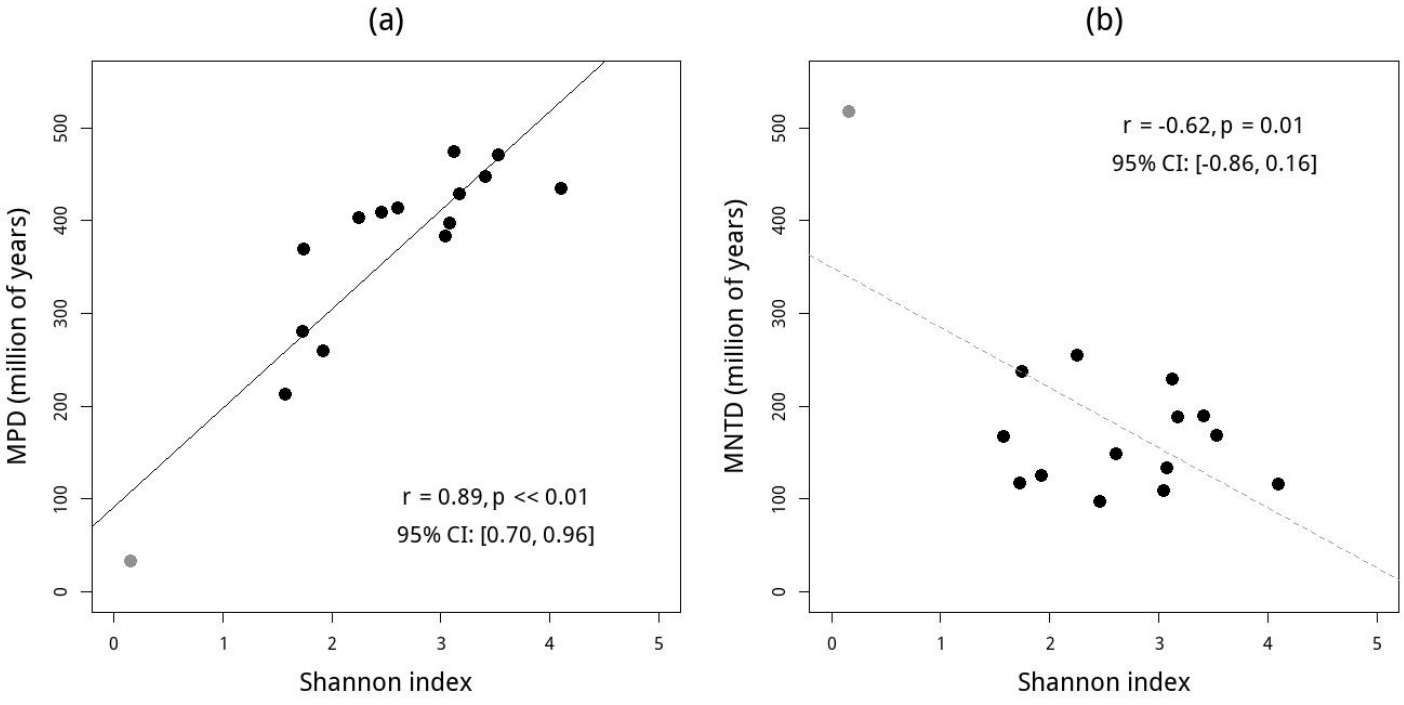

Figure 6. Correlation between taxonomic diversity (Shannon index) and (a) mean pairwise distance (MPD) and (b) mean nearest taxon distance (MNTD). Points correspond to permanent plots; grey point represents the Taminango plot.

In general, there was no correspondence between the results of MPD and MNTD. The top five plots with high MPD values correspond to Cardonal plana (474 Ma), Tambor (471 Ma), Tuparro (447 Ma), Caparrapi (434 Ma) and Colorados (429 Ma) (Figure 7.a) whereas for MNTD were Taminango (517 Ma), Matitas (254 Ma), Tamesis (236 Ma), Cardonal plana (229 Ma) and Tuparro (189 Ma) (Figure 7.b). Remarkably, Taminango had the lowest value for MPD and the highest value for MNTD (33 Ma and 517 Ma, respectively).

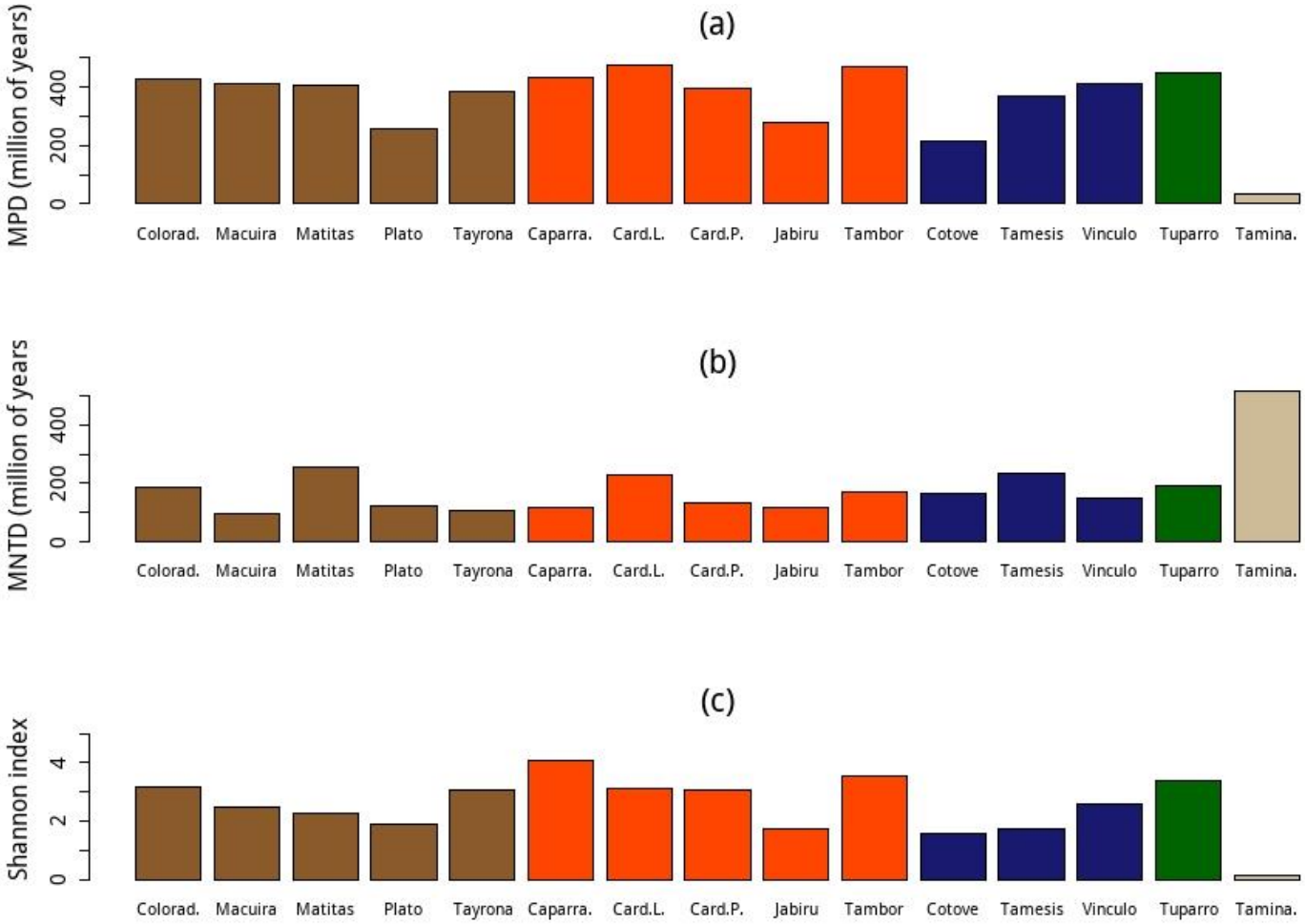

Figure 7. Bars represents the diversity index for each of the fifteen plots. (a) results for MPD, (b) results for MNTD and (c) results for Shannon index. Colors correspond to the region where each plot is located: Caribbean in brown, Magdalena in orange, Cauca in blue, Llanos in green and Patia in gray.

After performing the variance inflation factor the variables without collinearity were: MAT, Drymonths, Driests, Aridity, and A.rainyd. Then, with these five variables I ran the linear model for both metrics (MPD and MNTD), nevertheless none of the environmental variables fitted the model.

Evolutionary distinctiveness

I found that the species with the highest evolutionary distinctiveness are distributed unevenly across the phylogeny (Figure 8).

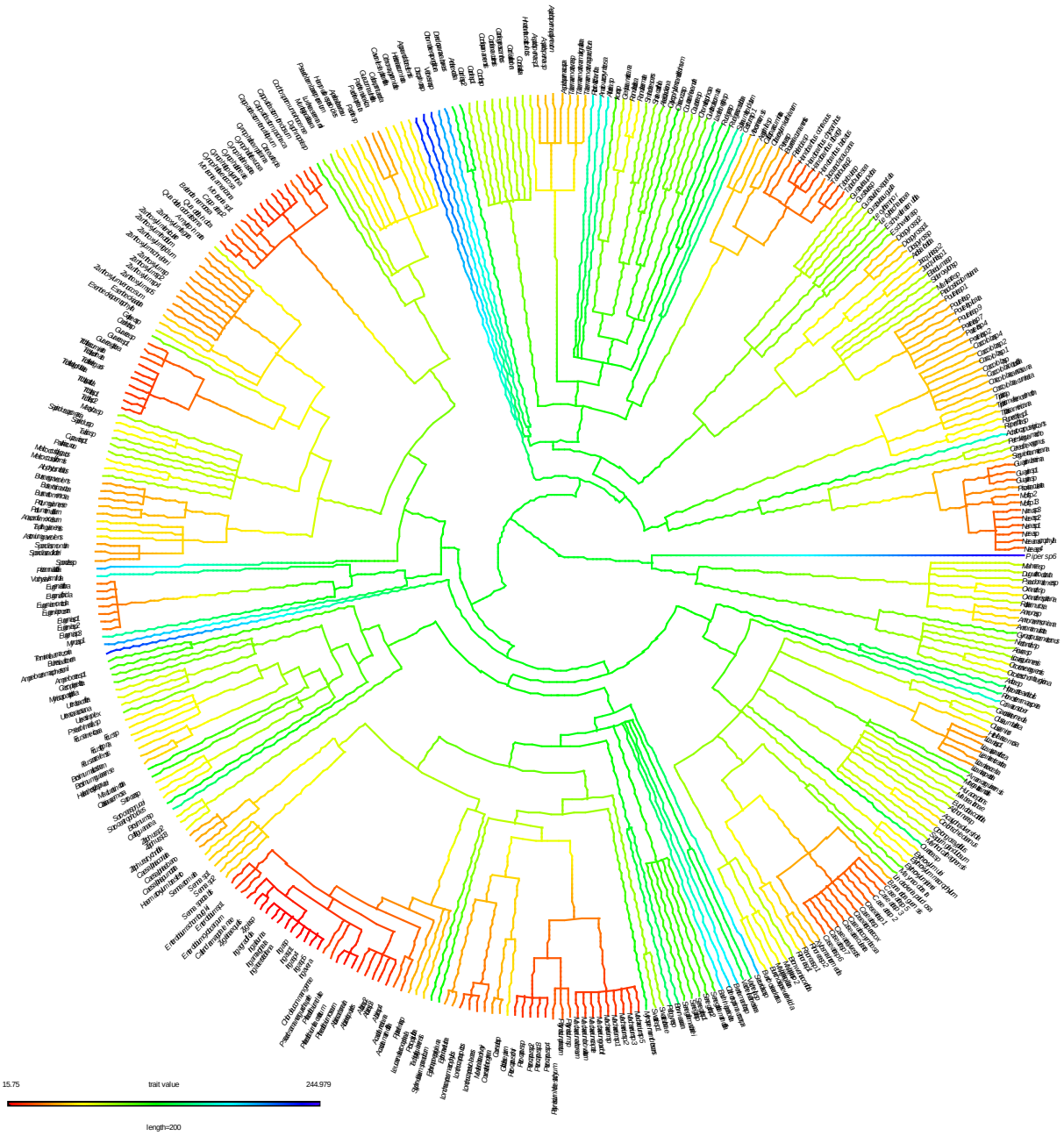

Figure 8. Evolutionary values (ED) distribution for each species. Blue edges represent species with highest values of ED whereas red represents species with lowest values of ED.

The fifteen species with the highest ED values are shown in Table 2**.** Colorados and Tuparro were the plots with more species out of these fifteen, each one with five, followed by Plato which had three. Of these fifteen species, *Agonandra brasiliensis*, *Bulnesia arborea*, *Picramnia latifolia*, *Securidaca sp*, *Bentamantha sp* and *Bauhinia petiolata* are typical of SDTF. Other species such as *Securidaca sp*, *Verbesina sp*, *Picramnia latifolia* and *Heisteria acuminata* (among others) are more characteristic of humid environments.

Table 2. The top 15 species values for evolutionary distinctiveness.

Species name	ED (millions of year)
Piper sp6	245
Heisteria acuminata	229
Agonandra brasiliensis	229
Bulnesia arborea	228
Discophora sp	222
Terminalia amazonia	203
Chromolaena perglabra	200
Verbesina sp	200
Picramnia latifolia	197
Securidaca sp	194
Clathrotropis macrocarpa	186
Bentamantha sp	186
Bauhinia petiolata	186
Alibertia sp	174
Amaioua corymbosa	174

2. Taxonomic and phylogenetic beta diversity

The UPGMA clustering based on the similarity of species composition and lineages groups most of the plots in the study to their biogeographic regions (Figure 9). Indeed, given that five major regions were represented in the study (Caribbean, Orinoquia, Patia, Magdalena and Cauca), a cut-off five group was established to test what groups were recovered. In terms of taxonomic distances, the Tamesis plot was distant from other plots of the Cauca, while the remaining Cauca and Magdalena plots were grouped together. For the phylogenetic distances, Tamesis remained as a single group but in this case, the Magdalena and Caribbean plots were grouped together. Both dendrograms (TBD and PBD) show more similarity of Colorados with the Magdalena than with other plots of the Caribbean region and recover Tuparro and Taminango as independent groups.

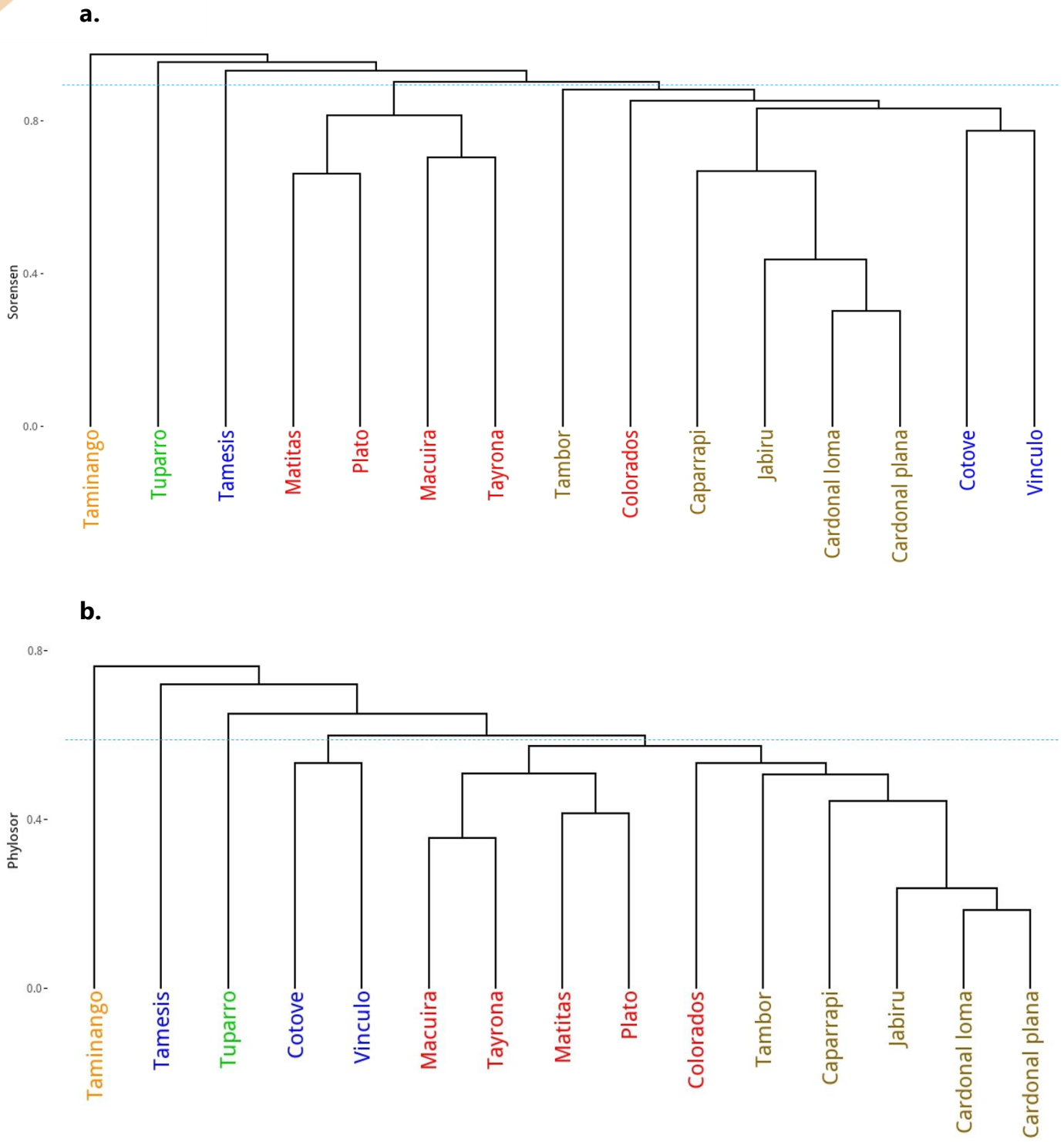

Figure 9 Dendrogram from UPGMA clustering calculated from Sorensen (a) and PhyloSor (b) distance matrices. The fifteen plots are colored according to their geographic region: Patia is in orange, Cauca is in blue, Caribbean is in red, Llanos is in green and Magdalena is in brown.

Generalized Dissimilarity Model

Of the 11 environmental variables tested as predictors of the phylobetadiversity four variables best fitted the model: 1. annual rainy/day (A.rainyd), 2. Aridity (Aridity), 3. total annual of precipitation (AP) and 4. mean annual temperature (MAT) (Figure 10). The relative importance for each variable were determined according to the height of the curve. The gdm model explained the 20.68% of turnover in community composition (TBD).

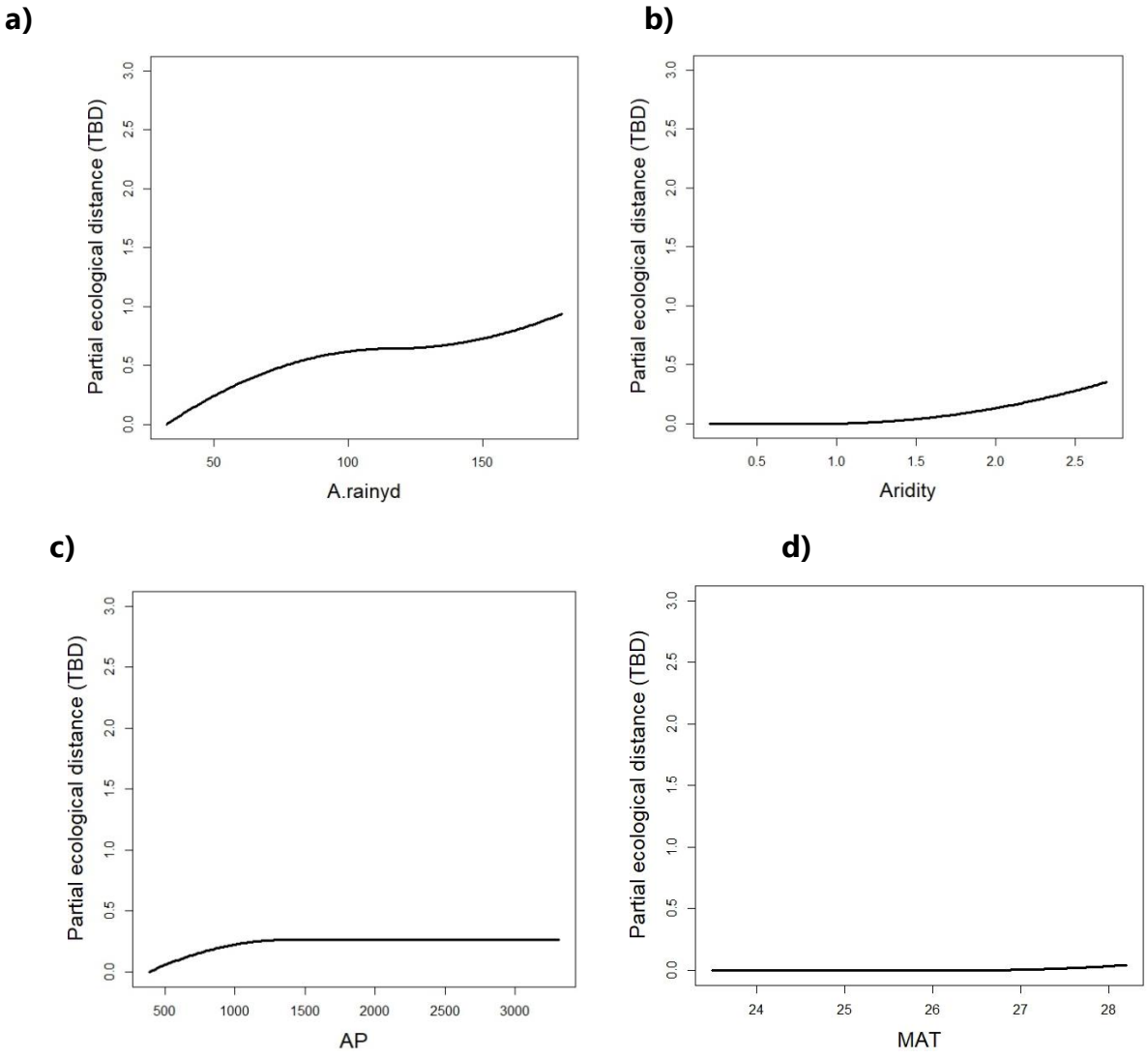

Figure 10. Environmental variables (a-d) selected as the best predictors for the Taxonomic Biodiversity Turnover (TBD). The total amount of turnover associated with each variable (holding all other variables constant) is represented by the maximum height reached by each curve. The shape of the curve indicates how the rate of turnover varies along the environmental gradient.

Four environmental variables and geographical distance explained 11.63% of PBD with the gdm. Those variables were in order of importance: month rainy/day (M.rainyd), annual rainy/day (A.rainyd), isothermality (Isotherm) and the driest month (P.drymonth) (Figure 11). Although both models share one of the four variables like best predictor, taking into account the percentages of variance explained for both models, PBD was less well predicted by the environment than TBD.

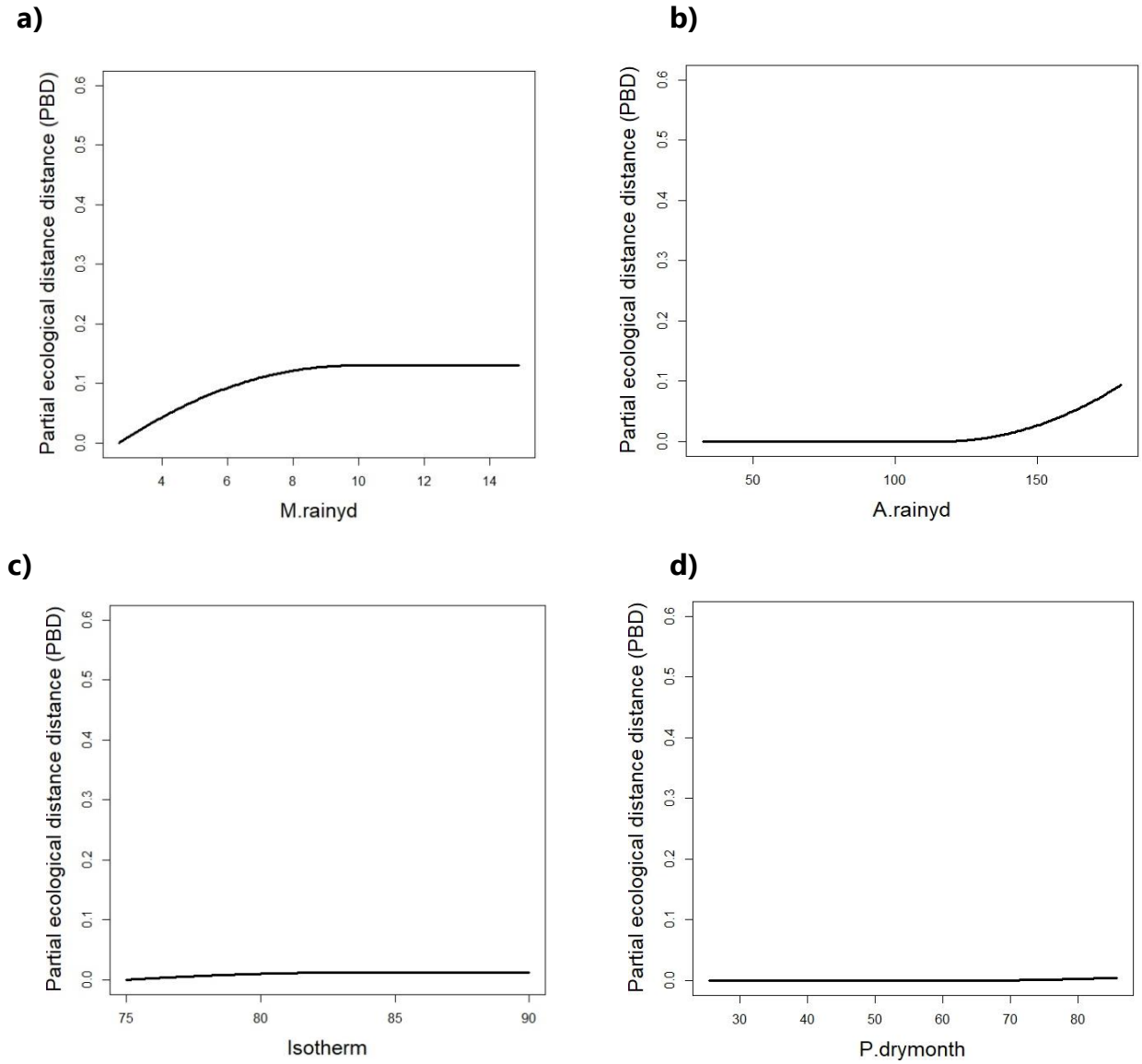

Figure 11. Environmental variables (a-d) selected as the best predictors for the phylogenetic dissimilarity of each site pair (PBD). The shape of the curve indicates how the rate of turnover varies along each environmental gradient.

The geographical distance was used in the adjustment of the model in both cases, but for TBD this variable had a greater influence (Figure 12).

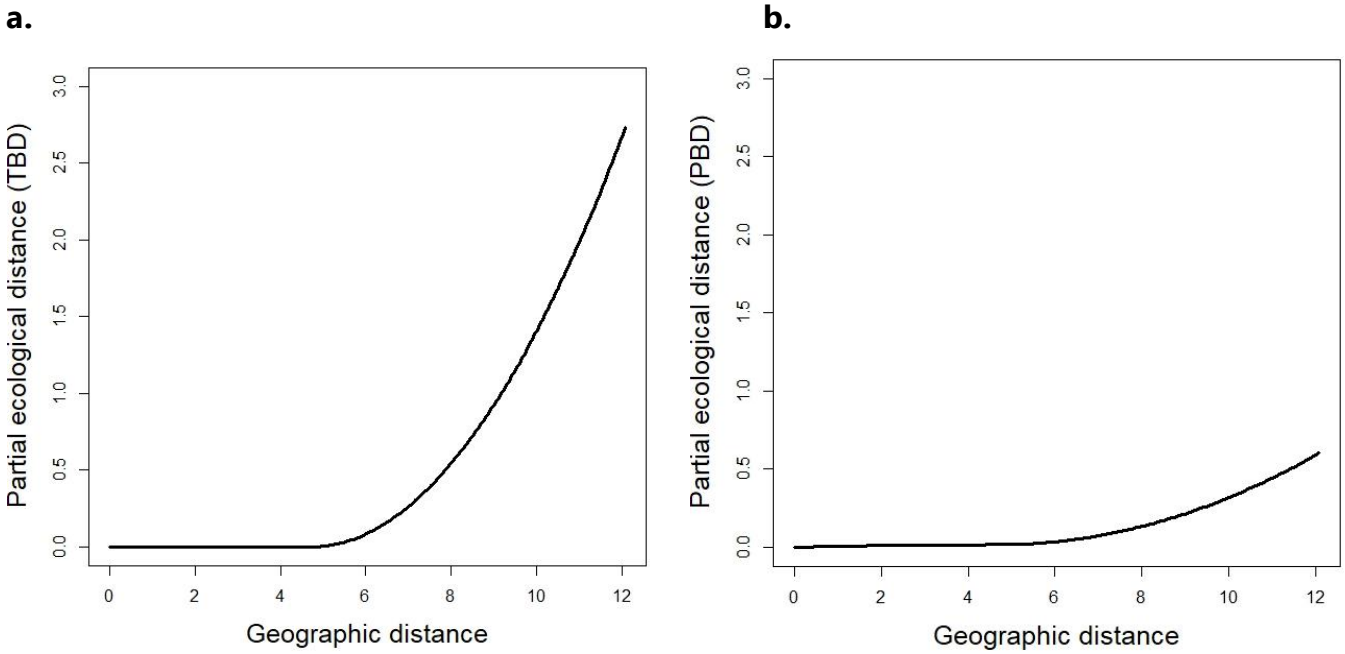

Figure 12. Fitted geographical variable of observed (a) taxonomic and (b) phylogenetic turnover for a Generalized Dissimilarity Model. The height reached by each function provides an indication of the total amount of turnover associated with this variable, and the slope indicates the variation of the rate of turnover along the geographical gradient.

PHYLOGENETIC COMMUNITY STRUCTURE

Out of the fifteen plots evaluated, five had significantly negative NRI values (Figure 13.a, p-value < 0.05), indicating clustering, and one plot (Cardonal plana) presented significant positive value indicating overdispersion (p-value < 0.05). The remaining nine plots display a structure not different from a random distribution. Figure 13.b show the values for NTI. Out of the fifteen plots: Macuira, Tayrona, Caparrapi and Cardonal loma showed phylogenetic clustering (p-value < 0.05) whereas the remaining plots did not differ significantly from the expectation from a random assemblage. The plots Tayrona, Caparrapi and Cardonal loma exhibit a cluster pattern for both indices, NRI and NTI.

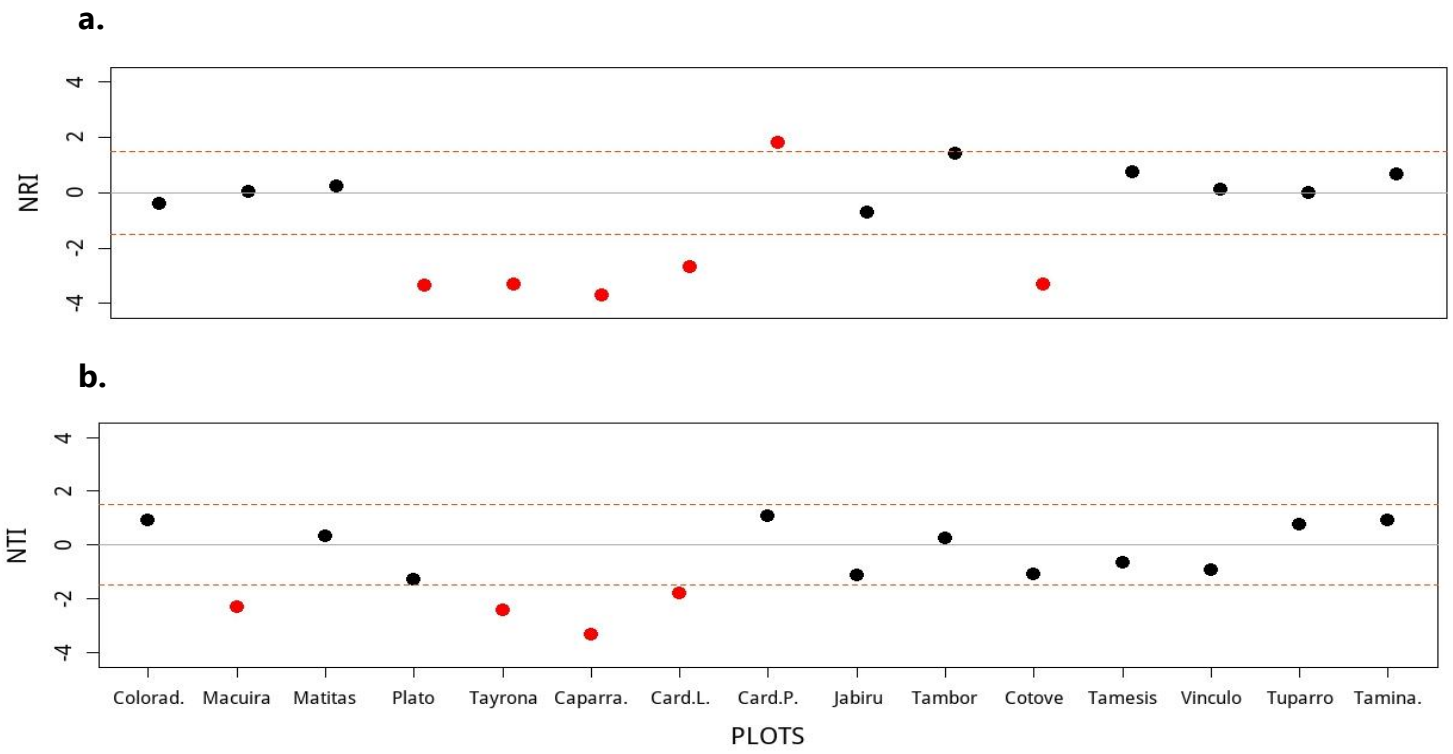

Figure 13. Values for NRI (a) and NTI (b) are displayed and those significantly different from random are indicated with red dots. Red dots above the dashed line indicate overdispersion (p>0.95) whereas red dots below the dashed line indicate clustering (p<0.05). Black dots correspond to plots with a phylogenetic structure not different from random.

When comparing phylogenetic structure patterns, for a subset of six plots, based on the rbcLa phylogeny (Figure S1) and the megatree phylogeny (Figure S2), I found a high correlation for NRI index while for NTI the correlation was not significant (Figure S3).

For NTI values (Figure 14) the Tuparro plot appears to be clustered based on the rbcLa phylogeny but not different from random based on the megatree phylogeny. On the other hand, for NRI consistently Cotove showed a cluster pattern while the other plots present a random pattern. Nonetheless, for NTI, Tuparro presented a significant cluster pattern based on the rbcLa phylogeny but a random pattern based on the megatree.

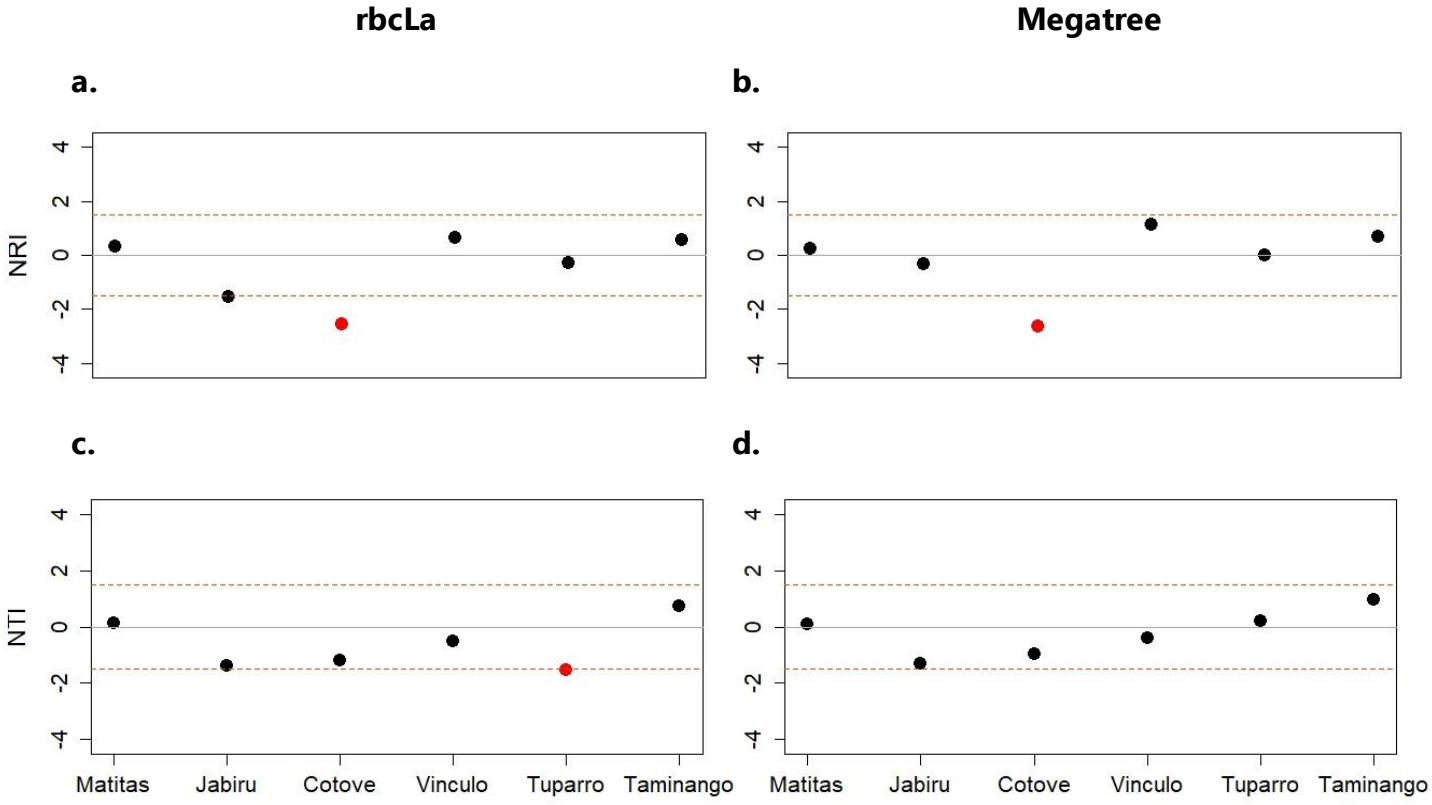

Figure 14. Net Relatedness Index (NRI) and Nearest Taxon Index (NTI) for a subset of six plots. Panels (a) and (c) correspond to the indices calculated with the rbcLa phylogeny and (b) and (d), to those calculated with the megatree. Red dots above the dashed indicate overdispersion (p-value < 0.05) and below the dashed line indicate clustering (p-value < 0.05).

DISCUSSION

The recognized threatened status of SDTF has driven recent research on the biological diversity associated to this ecosystem, although most of the research has focused mainly on taxonomic diversity. This study is among the first to characterize phylogenetic diversity and structure throughout the distribution of STDF in Colombia (Linares et al. 2011, Dryflor 2016, Pizano et al. 2016, González et al. 2014). Our results show a clear correlation between taxonomic diversity and phylogenetic diversity measured as MPD. This strong and positive correlation has been described previously in a mountain ecosystem in South Korea (Chun and Lee 2018). However, this result contrast with other studies showing that phylogenetic and taxonomic diversity is decoupled in plant communities (Forest et al. 2007, Chave et al. 2007). Interestingly, taxonomic diversity appeared to be negatively correlated with phylogenetic diversity measured as MNTD. This could be explained by the fact that MNTD is a terminal measure on the phylogeny and with increasing diversity and increasing unevenness in species abundances, MNTD tends to be lower. Our results are similar to those found by Coronado et al. (2015) in the Amazonian forest, where they found a negative correlation between species richness and MNTD. It is important to recall that the measures used in this study were weighted according to species abundance, which can result in different patterns when compared to presence/absence assessments (Cadotte et al. 2010, Gonzalez et al. 2010, Tucker et al. 2016).

Considering that SDTF is widely characterized by a marked seasonality and rainfall regime (Pennington et al. 2000, 2009, Portillo-Quintero and Sánchez-Azofeifa 2010), I tested whether phylogenetic diversity could be explained by eleven environmental variables but did not find a relationship between them. However, other studies have found that environmental variables are related to phylogenetic diversity patterns such as the study of Chun and Lee (2018) in a mountain ecosystem in South Korea. My results differ from the study of González et al. (2014) who found

a relationship between decreasing phylogenetic diversity with increasing temperature in different forests in Colombia; nonetheless their study included plant communities from different ecosystems and therefore their regional pool was wider and more heterogeneous.

Regarding Beta diversity, I found that environmental variables and geographic distance explain up to 20.68% of species turnover and 11.63% of phylogenetic Beta diversity. Interestingly, the environmental variables that best fit the models for TBD and PDB are not the same. In particular, TBD was partially explained by annual rainy/day (A.rainyd), Aridity (Aridity), total annual of precipitation (AP) and mean annual temperature (MAT); whereas PBD is best explained by month rainy/day (M.rainyd), annual rainy/day (A.rainyd), isothermality (Isotherm) and the driest month (P.drymonth). On the other hand, the results, are congruent with recent findings by González-M et al. (2018), who evaluated species turnover in 6 regions of SDTF in Colombia and showed that environmental variables explain up to 13% of species turnover, highlighting the large proportion of variance that remains to be explained. Species and phylogenetic beta diversity (TBD and PBD) were also related to geographic distances; still, this relationship is stronger in TBD (figure 7). Our results are similar to the study by González et al. (2014), who found a weak relationship between spatial distance and phylogenetic turnover.

Based on cluster analyses Tamesis and Tuparro presented the greater species and phylogenetic turnover when compared to other plots; they also correspond to areas with the highest rates of precipitation (González-M et al. 2018, Alvaro Idarraga personal communication). The difference in the composition of these two plots might be explained by the presence of more transitional species from humid environments. Also, is important to highlight that Tuparro is find isolated of the rest of the plots because two reasons, the first is the long distance between him and the nearest plot (the plots if the Magdalena valley) and that the eastern cordillera act as a geographical barrier. Other interesting result is that in the clustering analyses I consistently found

a greater proximity between the Caribbean and Inter-Andean valleys plots rather than with Tuparro, this is consistent with previous studies that have shown greater aggregation in terms of species composition between these two regions (Linares et al. 2011, DRYFLOR 2016). It is worth highlighting that the Colorados plot, from the Caribbean region, appears grouped with the plots of the Magdalena Valley. One plausible explanation for this is that Colorados is located on the western side of the Magdalena River that could represent a geographic barrier, limiting the dispersion of species with the rest of the Caribbean region but also this river might facilitate the migration of species from the SDTF of the Magdalena towards the Northern region (Pennington 2000).

Previous studies have identified species aggregations related in different regions where SDTF is present in the country; for example, DRYFLOR (2016) defined two floristic groups in the interior of Colombia: Inter-Andean valleys and the Caribbean region, coinciding with those defined by Pizano et al. (2014) and González et al. (2018) in a local scale, where they additionally identified the region of the Llanos (Orinoquía). Although in this study such strong relationships were not found, there was a tendency for plots to group according to the floristic regions where there are located. In relation to clustering by lineage turnover, I observe a greater turnover between the Magdalena and the Caribbean plots than between inter-Andean (Magdalena and Cauca Valley) plots; this may be a consequence of mountain ranges acting as geographical barriers that limit dispersion.

Despite the considerable species turnover observed between plots of SDTF in this study, the dominant phylogenetic structure did not differ from a random distribution. This reflects that lineages are shared between regions and that the effect of fragmentation is not yet detectable in evolutionary patterns. However, this does not exclude the presence of unique lineages such as the ones scored with the highest ED values; plots that hold these species, such as Colorados

and Tuparro, are of special concern for conservation. On the other hand, Phylogenetic community clustering was found in 5 out of 11 plots, indicating that species within these communities are more closely related to each other than expected in a random distribution (Webb 2000, Cavender-Bares et al. 2004, 2006). Clustering was not linked to plots of a particular region and in our dataset, was also not explained by environmental conditions. Remarkably, plots belonging to the same region such as Cardonal loma and Cardonal plana present completely opposite patterns of phylogenetic structure (Figure 13). Indeed, while clustering was observed in Cardonal loma overdispersion was found in Cardonal plana, the difference in abundance between these two plots (2008 individuals in Cardonal loma and 1341 in Cardonal plana) may be causing the contrast between the structure patterns. It is worth mentioning that although the communities belong to the same ecosystem, local conditions like the climatic, edaphic and topology may vary, Pennington et al. (2009) state that in some inter-Andean valleys, precipitation can change radically over very short distances. In a study conducted by González et al. (2014) on a larger scale with a regional pool that included more humid ecosystem species, they found cluster patterns associated to environmental variables mainly related to precipitation. In another study carried out in Caatinga (Freiro-Moro et al. 2015) when evaluating phylogenetic structure patterns, they clearly found a strong phylogenetic aggregation for NRI associated with edaphic characteristics, specifically those associated with water availability, reflecting greater environmental niche conservatism.

Community phylogenetic studies traditionally have used megatree to infer the relationship between the species in a community assemblage, this is an important tool especially when molecular information is not available or is limited; nevertheless, it has been argued that the estimates of the phylogenetic relationships among the species found within communities are not necessarily accurate (Davies et al. 2012). I tested if the phylogenetic structure patterns could be affected by the resolution between a megatree phylogeny (Webb & Donoghue 2005) and

molecular based phylogeny comparing the results of phylogenetic structure for 6 plots. I found that for the NRI index there is a positive correlation between the structure results (Figure 14 a-b), whereas for NTI I observe that there is no correlation (Figure 14 c-d). This difference observed in the results for NTI is due to the fact that since this is a measure associated with the terminals of the phylogeny it will be strongly influenced by the resolution. Hence, the importance of using more resolute phylogenies for example with the inclusion of other molecular markers such as ITS (nuclear gene), which has been reported as one of the molecular markers that achieves greater discrimination between species (Gonzalez et al. 2009, Cräutlein et al. 2011, Kang et al. 2017). Similar to this study, Kress et al. (2009) found significant differences in the evolutionary and ecological inferences when comparing phylogenies with different resolution; they concluded that there is a significant loss of information in the tips of the phylogeny when these are less resolved.

This study focused on the best-preserved fragments of SDTF in Colombia, the 95% of this ecosystem is highly intervened, and just 5% of the remnants are under protection. Because the fragmentation is very recent in evolutionary terms, the incorporation of population genetics would allow a better interpretation of the affectation by the perturbation regarding the phylogenetic structure of the ecosystem. For a better understanding of how different evolutionary responses may arise from environmental changes, it would also be interesting to integrate the results from this study with functional traits, that would entail to a better understanding of the community assembly processes.

It is worth noting some general limitations of this study are that we selected the individuals with a DBH > 5 cm, without including the individuals among 2.5 cm y 5 cm of DBH. I worked with the plots located in the remnants better conserved and having into account the high human intervention, it is not known with certainty how could be varying the composition in that parts

less conserved. Finally, and since these studies are based on phylogenetic relationships, the incorporation of DNA sequences with more molecular markers, offers the possibility of a better delimitation between species and better resolved phylogenetic hypotheses, as well as a greater precision in the inferences on community assembly process.

REFERENCES

ANGIOSPERM PHYLOGENY GROUP – APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society. 161(2):105-121. http://dx.doi.org/10.1111/j.1095-8339.2009.00996.x.

Beaulieu, J.M., Ree, R.H., Cavender-bares, J., Weiblen, G.D. and Donoghue, M.J. 2012. Synthesizing phylogenetic knowledge for ecological research. Ecology 93:4-13.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C-H., Xie, D., Suchard, M.A., Rambaut, A. and Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology, 10(4).

Bryant, J.A., Lamanna, C., Morlon, L., Kerkhoff, A.J., Enquist, B.J. and Green, J.L. 2008. Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. PNAS 105: 11505–11511.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. and Bazzaz, F.A. 2004. Phylogenetic Overdispersion in Floridian Oak Communities. Ame Nat. 163:823-843

Cavender-Bares, j., Keen A. and Miles, B. 2006. Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. Ecology 87:109-122.

Cavender-Bares, J., Kozak, K.H., Fine, P.V. A. and Kembel, S.W. 2009. The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715.

Chase, J.M. 2003. Community assembly: when should history matter? Oecologia 136:489–498.

Cräutlein, M., Korpelainen, H., Pietiläinen, M. and Rikkinen, J. 2011. DNA barcoding: a tool for improved taxon identification and detection of species diversity. Biodiversity and Conservation 20: 373–389.

Chun J.H. and Lee C.B. 2018. Diversity patterns and phylogenetic structure of vascular plants along elevational gradients in a mountain ecosystem, South Korea. Journal of Mountain Science. 15(2):280-295.

Davies, J.T., Kraft, J.N., Salamin, N. and Wolkovich, E.M. 2012. Incompletely resolved phylogenetic trees inflate estimates of phylogenetic conservatism. Ecology 93(2): 242–247.

Díaz, J.M. 2006. Bosque Seco Tropical Colombia. Banco de occidente, I/M Editores. Cali, Colombia.

Donoghue, M.J. 2008. A phylogenetic perspective on the distribution of plant diversity. PNAS 105: 11549–11555.

Drummond, A.J. and Rambaut, A. 2007. BEAST Bayesian evolutionary analysis by sampling trees. Evolutionary Biology 7:214.

DRYFLOR B. R. K. et al. 2016. Plant diversity patterns in neotropical dry forests and their conservation implications Science 353:1383–7.

Faith, D.P. 1992. Conservation evaluation and phylogenetic diversity. Biological conservation 61:1-10.

FAO. 2001. FRA 2000: Global Ecological Zoning for the Global Forest Resources Assessment 2000. Final Report. Food and Agriculture Organization of the United Nations. Rome. http://www.fao.org/docrep/006/ad652e/ad652e00.htm.

Ferrier, S., Manion, G., Elith, J. and Richardson, K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity Distrib. 13:252–264.

Fitzpatrick, M.C., Sanders, N.J., Normand, S., Svenning, J.C., Ferrier, S., Gove, A.D. and Dunn, R.R. 2013. Environmental and historical imprints on beta diversity insights from variation in rates of species turnover along gradients. Proc. R. Soc. B: Biol. Sci., 280 (2013), p. 20131201.

Forest, F., Grenyer, R., Rouget, M., Davies, J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proches, S., van der Bank, M., Reeves, R., Hedderson, T.A.J. and Savolainen, V. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445:757-760.

Freiro Moro, M., Silva, I.G., Soares de Araújo, F., Lughadha, E.C., Meagher, T.R. and Martins, F.R. 2015. The Role of Edaphic Environment and Climate in Structuring Phylogenetic Pattern in Seasonally Dry Tropical Plant Communities. PLoS ONE 10(3):1-18.

Gerhold, P., M. Pärtel., J. Liira, K. Zobel. and A. Prinzing. 2008. Phylogenetic structure of local communities predicts the size of the regional species pool. Journal of Ecology 96:709–712.

Gilbert, G., Laurance, W.F., Leigh Jr, E.G. and Nascimento, H.E.M. 2006. Can Neutral Theory Predict the Responses of Amazonian Tree Communities to Forest Fragmentation? The american naturalist 168:204-317.

González-Caro, S., Umaña, M.N., Álvarez, E., Stevenson, P.R. and Swenson, N.G. 2014. Phylogenetic alpha and beta diversity in tropical tree assemblages along regional scale environmental gradients in northwest South America. Journal of Plant Ecology 7:145–153.

Gonzalez, M.A., Roger, A., Courtois,E.A., Jabot, F., Norden, N., Paine C.E.T., Baraloto, C., Thébaud, C. and Chave,J. 2010. Shifts in species and phylogenetic diversity between sapling and tree communities indicate negative density dependence in a lowland rain forest. Journal of Ecology 98:137–146.

Gonzalez, M.A y Quintero, L. 2017. Plantas. pp 11. En: Gonzalez M.A., Arenas-Castro H. (Eds). Recolección de tejidos biológicos para análisis genéticos. Instituto de Investigaciones de Recursos Biológicos Alexander von Humboldt. Bogotá, D.C., Colombia. 33 pp.

González-M, R., García, H., Isaacs, P., Cuadros, H., López-Camacho, R., Rodríguez, N., Pérez, K., Mijares, F., Castãno-Naranjo, A., Jurado, R., Idárraga-Piedrahíta, A., Rojas, A., Vergara, H. and Pizano, C. 2018. Disentangling the environmental heterogeneity, floristic distinctiveness and current threats of tropical dry forests in Colombia. Environ. Res. Lett. 13 (2018) 045007.

Gotelli, N.J. 2000. Null model analysis of species co-occurrence patterns. Ecology, 81(9):2606–2621.

Graham, C.H. and Fine, P.V.A. 2008. Phylogenetic beta diversity linking ecological and evolutionary processes across space and time. Ecology Letters 11: 1265–1277.

Ivanova, N.V., Fazekas, A.J. and Hebert P.D. 2008 - Semi-automated, Membrane-Based Protocol for DNA Isolation from Plants. Plant Mol Biol Rep. 26:186–198.

Hardy, O. 2008. Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. Journal of Ecology 96:914–926.

Ho, S.Y.W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J. Avian Biol. 38:409-414.

Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves,G.R., Jønsson, K.A., Nogués-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldså, J. and Rahbek, C. 2013. An Update of Wallace's Zoogeographic Regions of the World. Science, 339:73-78.

Hubbell, S.P. 2005. The Unified Neutral Theory of Biodiversity and Biogeography. Paleobiology 31(2):122–132.

Jaccard, P. 1912. The distribution of the flora in the alpine zone. New phytologist 11(2):32

Johnson, M.T.J and J.R., Stinchcombe. 2007. An emerging synthesis between community ecology and evolutionary biology. Trends in Ecology and Evolution 22:250-257.

Kembel, S.W. and Hubbell, S.P. 2006. The phylogenetic structure of a neotropical forest tree community. Ecological Society of America 87(7):86-99.

Kraft, N.J.B., Cornwell, W.K., Webb, C.O. and Ackerly, DD. 2007 - Trait evolution, community assembly, and the phylogenetic structure of ecological communities. Am. Nat 170:271-283.

Kreft, H. and Jetz, W. 2010. A framework for delineating biogeographical regions based on species distributions. Journal of Biogeography, 37:2029–2053.

Kruskal, J.B. 1964. Nonmetric multidimensional scaling: A numerical method. Psychometrika. 29:115-129.

Kumar S., Stecher, G. and Tamura, K. 2016 - MEGA7_Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33(7):1870–1874.

Kembel S.W. 2009. Disentangling niche and neutral influences on community assembly_assessing the performance of community phylogenetic structure tests. Ecology Letters 12: 949–960.

Kembel, S.W., Cowan, P.D., Helmus, MR., Cornwell, W.K., Morlon, H., et al. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Koleff, P., Gaston, K.J. and Lennon, J.K. 2003. Measuring beta diversity for presence–absence data. Journal of Animal Ecology 72:367–382.

Lennon, J.J., Koleff, P., Greenwood, J.J.D. and Gaston, K.J. 2001. The geographical structure of British bird distributions: diversity, spatial turnover and scale. Journal of Animal Ecology 70:966–979.

Linares-Palomino, R., Cardona, V., Hennig, E I., Hensen I., Hoffmann, D., Lendzion J., Soto D., Herzog, S.K. and Kessler, M. 2009 Non-woody life-form contribution to vascular plant species richness in a tropical American forest Plant Ecol. 201:87–99

Linares-Palomino R, Oliveira-Filho A.T and Pennington R.T. 2011. Neotropical seasonally dry forests: diversity, endemism and biogeography of wood plants Seasonally Dry Tropical Forests: Ecology and Conservation ed RDirzo, H S Young, HA Mooney and G Ceballos (Washington, DC: Island Press) pp 3–21

Losos, J. 1996. Phylogenetic perspective on community ecology. Ecology society of America 77:1344-1354.

Macarthur, R. and Levins, R. 1967. The Limiting Similarity, Convergence and Divergence of Coexisting Species. The American Naturalist 101:377-385.

Magallón, S., Góomez-Acevedo, S., Sánchez-Reyes, L.L. and Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. 207:437–453.

Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Mokany, K. and Fitzpatrick, M.C. 2017. gdm: Generalized Dissimilarity Modeling.

Miller, E.T., Farine, D.R. and Trisos, C.H. 2016. Phylogenetic community structure metrics and null models: a review with new methods and software. Ecography 40: 461–477.

Miles, L., Newton, DeFries, A.C., Ravilious, R.S., Simon, I.B., Kapos. V. and Gordon, J. 2006. A global overview of the conservation status of tropical dry forests. Journal of Biogeography, 33:491–505.

Moreno, Saiz, J.C., Donato, M., Katinas, L., Crisci, J.V. and Posadas, P. 2012. New insights into the biogeography of south-western Europe: spatial patterns from vascular plants using cluster analysis and parsimony. Journal of Biogeography, 40:90-104.

Mouquet, N., Devictor, V., Meynard, C., Munoz, F., Bersier, L.F., Chave, J., Couteron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold1, M., Mouillot, M., Münkemüller, T., Pavoine, S., Prinzing1, A., Rodrigues, A.S.L., Rohr, R.P., Thébault1, E. and Thuiller, W. 2012. Ecophylogenetics: advances and perspectives. Biol. Rev. 87:769–785.

Mittelbach, G.G. and Schemske, D.W. 2015. Ecological and evolutionary perspectives on community assembly. Trends Ecol. Evol. 30:241–47.

Muneepeerakul, R., Bertuzzo, E., Lynch, h.J., Fagan, W.F., Rinaldo, A. and Rodriguez-Iturbe, I. 2008. Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin Nature 453:220-222.

Münkemüller, T., de Bello, F., Meynard, C. N., Gravel, C. N., Lavergne, S., Mouillot, D., Mouquet, N. and Thuiller,W. 2012. From diversity indices to community assembly processes: a test with simulated data. Ecography 35: 468–480.

Ndiribe, C., Salamin, N. and Guisan, A. 2013. Understanding the concepts of community phylogenetics Charlotte. Evolutionary Ecology Research 15: 1–16.

Olson, D. M. and Dinerstein, E. 2002. The Global 200: Priority ecoregions for global conservation. Annals of the Missouri Botanical Garden 89(2):199-224.

Pavoine, S., Ollier, S. and Dufour A.B. 2005. Is the originality of a species measurable?. Ecology Letters 8: 579–586.

Pennington, R. T., Prado, D.E. and C.A. Pendry. 2000. Neotropical seasonally dry forests and Quaternary vebetation changes. Journal of Biogeography 27:261–273.

Pennington, R.T., Richardson, J.E. and Lavin, M. 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. New Phytologist 172:605-616.

Pizano, C., y H. García. 2014. El bosque seco tropical en Colombia.

Pizano C et al. 2014. Las plantas de los bosques secos de Colombia Bosque seco tropical en Colombia ed C Pizano and H García (Bogotá: Instituto de Investigación de Recursos Biológicos Alexander von Humboldt) pp 49–93.

Portillo-Quintero, C.A. and Sánchez-Azofeifa, G.A. 2010. Extent and conservation of tropical dry forests in the Americas. Biological Conservation 143:144–155.

Prado, D.E and Gibbs, P.E. 1993. Pattern of species distribution in the Dry Seasonal Forest of South America. Annals of the Missouri Botanical Garden, 80, 902–927.

Rambaut, A., Suchard. M.A., Xie, D. and Drummond, A.J. 2014. Tracer version 1.5.0. [WWW document] URL http://beast.bio.ed.ac.uk/Tracer.

Redding, D.W., Hartmann, K., Mimoto, A., Bokal, D., DeV, M. and Mooers A.Ø. 2008. Evolutionarily distinctive species often capture more phylogenetic diversity than expected. Journal of Theoretical Biology 251:606–615.
Redding, D.W. and Mooers, A.Ø. 2006. Incorporating Evolutionary Measures into Conservation Prioritization. Conservation Biology 20(6):1670–1678.

Roseaur, D., Laffan, S.W., Crisp, M.D., Donnellan, S.C. and Cook, L.G. 2009. Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. Molecular Ecology (2009) 18, 4061–4072.

Rosindell, J., Hubbell, S.P. and Etienne, R.S. 2001. The unified neutral theory of biodiversity and biogeography at age ten. Trends in Ecology and Evolution, 26:340-348.

Sauquet, H. 2013. A practical guide to molecular dating. (Comptes rendus) C. R. Palevol 12:355–367.

Sánchez-Azofeifa, G.A., Quesada, M., Rodríguez, J.P., Nassar, J.M., Stoner, K.E., Castillo, A., Garvin, T., Zent, E.L., Calvo-Alvarado, J.C., Kalacska, M.E.R., Fajardo, L., Gamon, J.A. and Cuevas-Reyes, P. 2005. Research Priorities for Neotropical Dry Forests. BIOTROPICA 37:477–485.

Shannon, CE. 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27:379–423;623–56.

Simpson, G.G. 1943. Mammals and the nature of continents. American Journal of Science 241:1–31.

Stephen P. Hubbell, S.p., He, F., Condit, R., Borda-de-Agua, L., Kellner, J. and Steege H. 2008. How many tree species are there in the Amazon and how many of them will go extinct? PNAS 105:11498-11504.

Sunderland, T., Apgaua, D., Baldauf, R., Blackie, C., Colfer, A.B., Cunningham, K., Dexter, H., Djoudi, D., Gautier, D., Gumbo, A., Ickowitz, H., Kassa, N., Parthasarathy., Pennington, R.T., Paumgarten, F., Pulla, S., Sola, P., Tng, D., Waeber, P. and Wilmé, L. 2015. Global dry forests: a prologue. International Forestry Review 17:1-9.

Swenson, N.G. and Enquist, B.J. 2009. Opposing assembly mechanisms in a Neotropical dry forest: implications for phylogenetic and functional community ecology. Ecology 90: 2161-2170.
Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. and Zimmerman, J.K. 2007. The problem and promise of scale dependency in community phylogenetics. Ecology 88: 1770–1780.

Thornthwaite. C. W. 1948. An approach toward a rational classification of climate. Geogr. Rev. 38: 55 - 94.

Tucker, M.A., Cadotte, M.W., Carvalho, S.S., Davies, T.J., Ferrier, S., Fritz, S.A., Grenyer, R., Helmus, M.R., Jin, L.S., Mooers, A.O., Pavoine, S, Purschke, O., Redding, D.W., Rosauer, D.F., Winter, M. and Mazel, F. 2016. A guide to phylogenetic metrics for conservation, community ecology and macroecology. Biol Rev, 92: 698–715.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. and Webb, C.O. 2009. Emerging patterns in the comparative analysis of phylogenetic community structure. Molecular Ecology 18: 572–592.

Webb, C.O. 2000. Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. The American Naturalist 156:145-155.

Webb, C.O. 2000. Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. Am. Nat. 156:145–55.

Webb, C.O., Ackerly, D.D., McPeek, M. and Donoghue, M.J. 2002. Phylogenies and Community Ecology 33:475-505.

Webb, C.O. and Donogue, M.J. 2005. Phylomatic tree assembly for applied phylogenetics. Molecular Ecology Notes 5:181–183.

Webb, C.O., Ackerly, D.D. and Kembel. S.W. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 24:2098–100.

Whittaker, R.H. 1972. Evolution and measurement of species diversity. 21(2/3):213-251.

SUPPLEMENTARY MATERIALS

Table S1. Sequences downloaded from GenBank

Original Species name	GenBank Species name	Accesion number
Amaioua corymbosa		JQ626322
Amanoa guianensis	*Amanoa caribaea*	AY663561
Anacardium excelsum		GQ981661
Apeiba tibourbou		AJ233145
Belencita nemorosa		KU739570
Bunchosia pseudonitida	*Bunchosia armeniaca*	Z75274
Bursera simaruba		KJ773325
Bursera tomentosa		JQ590985
Caesalpinia coriaria		AY904380
Caesalpinia ebano	*Caesalpinia pulcherrima*	U74190
Casearia corymbosa	*Casearia decandra*	JQ626222
Casearia praecox		JQ593981
Casearia sp1	*Casearia guianensis*	GQ981688
Casearia sp7	*Casearia nitida*	JX664038
Casearia sylvestris		AF206746
Cecropia peltata	*Cecropia environmental*	KF270163
Ceiba pentandra		JX987571
Cereus hexagonus	*Cereus fernambucensis*	AY875240
Chiococca alba		KJ594151
Chloroleucon mangense		KU176144
Chomelia spinosa		JQ593643
Cochlospermum orinocense	*Cochlospermum vitifolium*	JQ591114
Cordia gerascanthus		JQ590899
Cynophalla amplissima	*Capparis spinosa*	AY167985
Enterolobium schomburgkii		JQ626149
Erythrina poeppigiana	*Erythrina velutina*	JX856697
Erythroxylum macrophyllum		JQ594783
Eugenia biflora		KJ082297

Eugenia monticola		JQ592953
Euphorbia cotinifolia		JN249286
Ficus zarzalensis	*Ficus obtusifolia*	GQ981740
Forsteronia spicata		JQ590223
Guarea glabra		U39085
Guazuma ulmifolia		KF724295
Haematoxylum brasiletto		KJ468097
Heisteria acuminata		GQ981760
Hirtella racemosa		KX180069
Inga gracilifolia		JQ626130
Lecythis chartacea		JQ626228
Licania excelsa	*Licania tomentosa*	L11193
Licania micrantha		JQ626165
Licania parvifructa		JQ898749
Licaria guianensis		GQ428569
Melicoccus bijugatus	*Melicoccus pedicellaris*	JQ626266
Myrcia sp1	*Myrcia fallax*	JQ625851
Myrospermum frutescens		JQ591961
Ocotea schomburgkiana	*Ocotea puberula*	KF561959
Oxandra espintana		AY319066
Pachira nukakica	*Pachira aquatica*	AY328178
Pereskia guamacho		AY875242
Pithecellobium lanceolatum		JQ591978
Platymiscium sp	*Platymiscium trifoliolatum*	KF436469
Pouteria plicata	*Pouteria venosa*	JQ413830
Prosopis juliflora		KF471677
Protium guianense		JQ625777
Pseudolmedia sp	*Pseudolmedia laevigata*	KX640875
Pterocarpus sp4	*Pterocarpus rohrii*	GQ981862
Quadrella odoratissima		KU739608
Sapium glandulosum		AY794841
Senegalia macbridei	*Senegalia bonariensis*	KX640832
Sideroxylon sp	*Sideroxylon foetidissimum*	KJ773891

Sorocea trophoides	*Sorocea saxicola*	KX640884
Spondias mombin		JQ590140
Talisia sp	*Talisia hexaphylla*	JQ625755
Tapirira guianensis		JQ626278
Trichilia oligofoliolata	*Trichilia martiana*	JQ592754
Trichilia pallida		JQ626046
Trichostigma octandrum		KJ594544
Triplaris melaenodendron		JQ593542
Zanthoxylum fagara		KJ773993
Zanthoxylum rhoifolium		KF561971
Ziziphus sp3	*Ziziphus mauritiana*	KR530242
Zygia inaequalis	*Zygia longifolia*	JQ592098

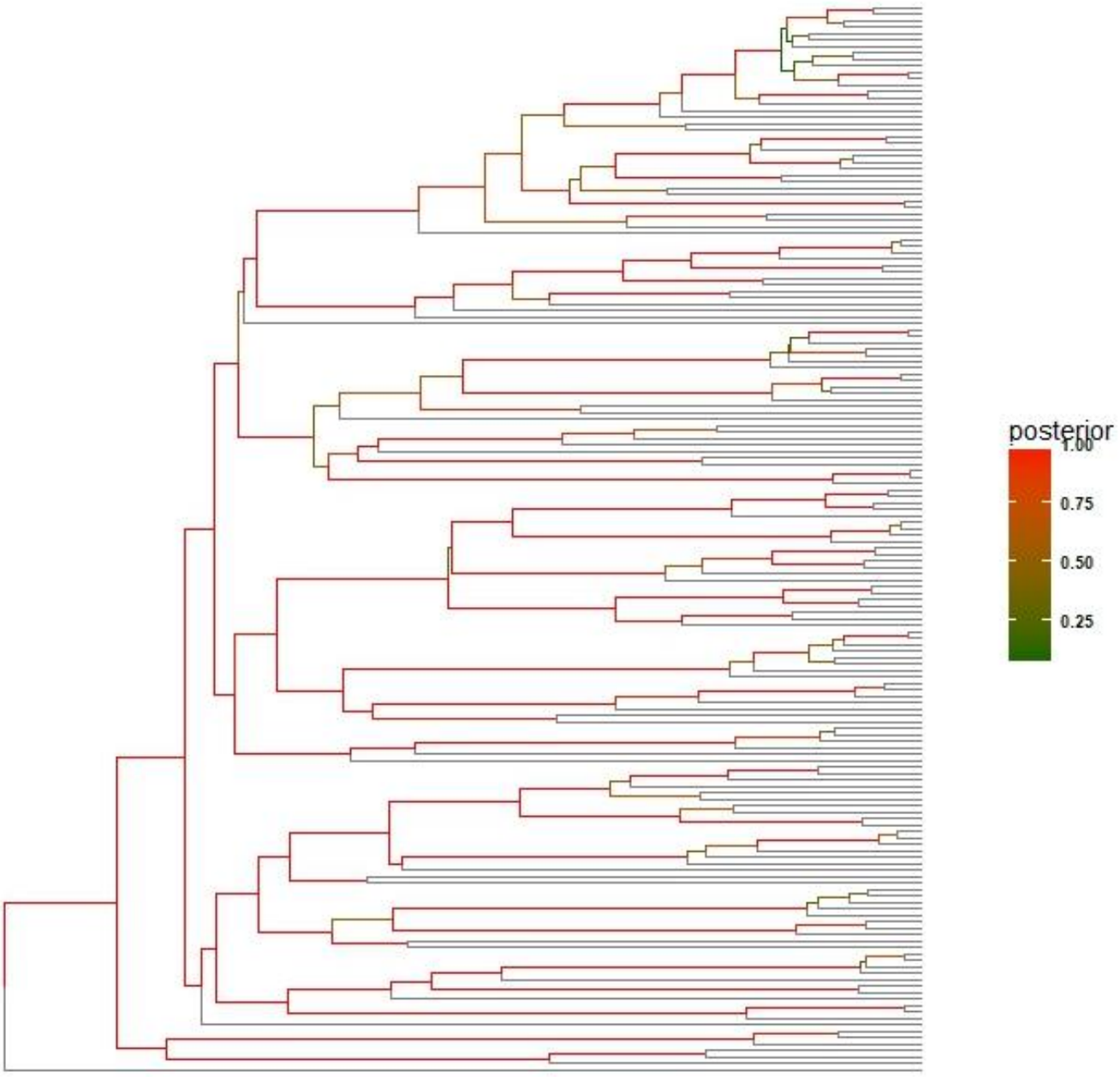

Figure S1 Phylogenetic tree according to Bayesian analysis with molecular sequences of rbcLa. The colors of the branches represent the "posterior" probability distribution. Red branches represent the highest confidence estimate of the evolutionary relationships, with a 95% the highest posterior density.

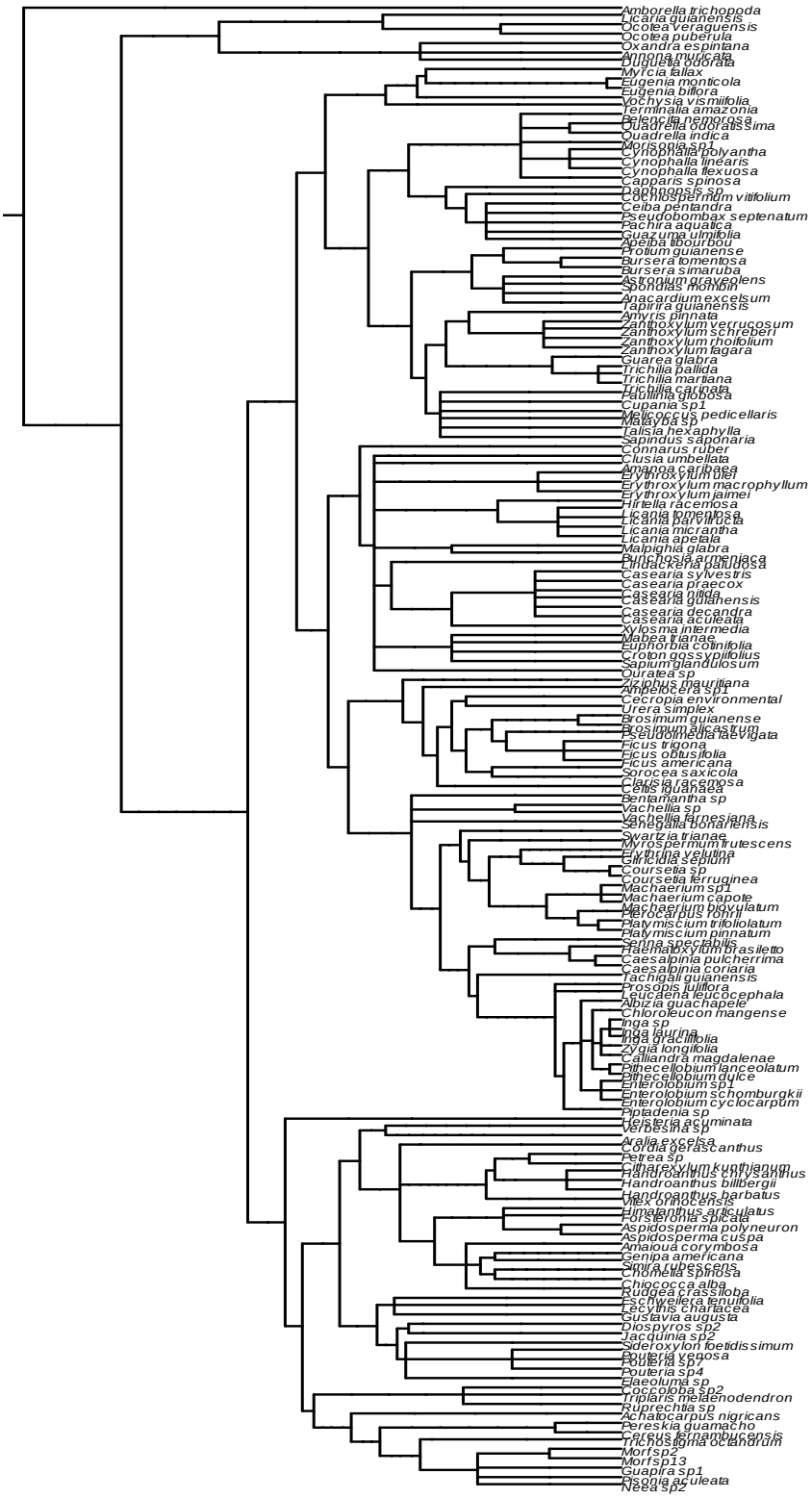

Figure S2 Phylogeny constructed with Phylomatic tool.

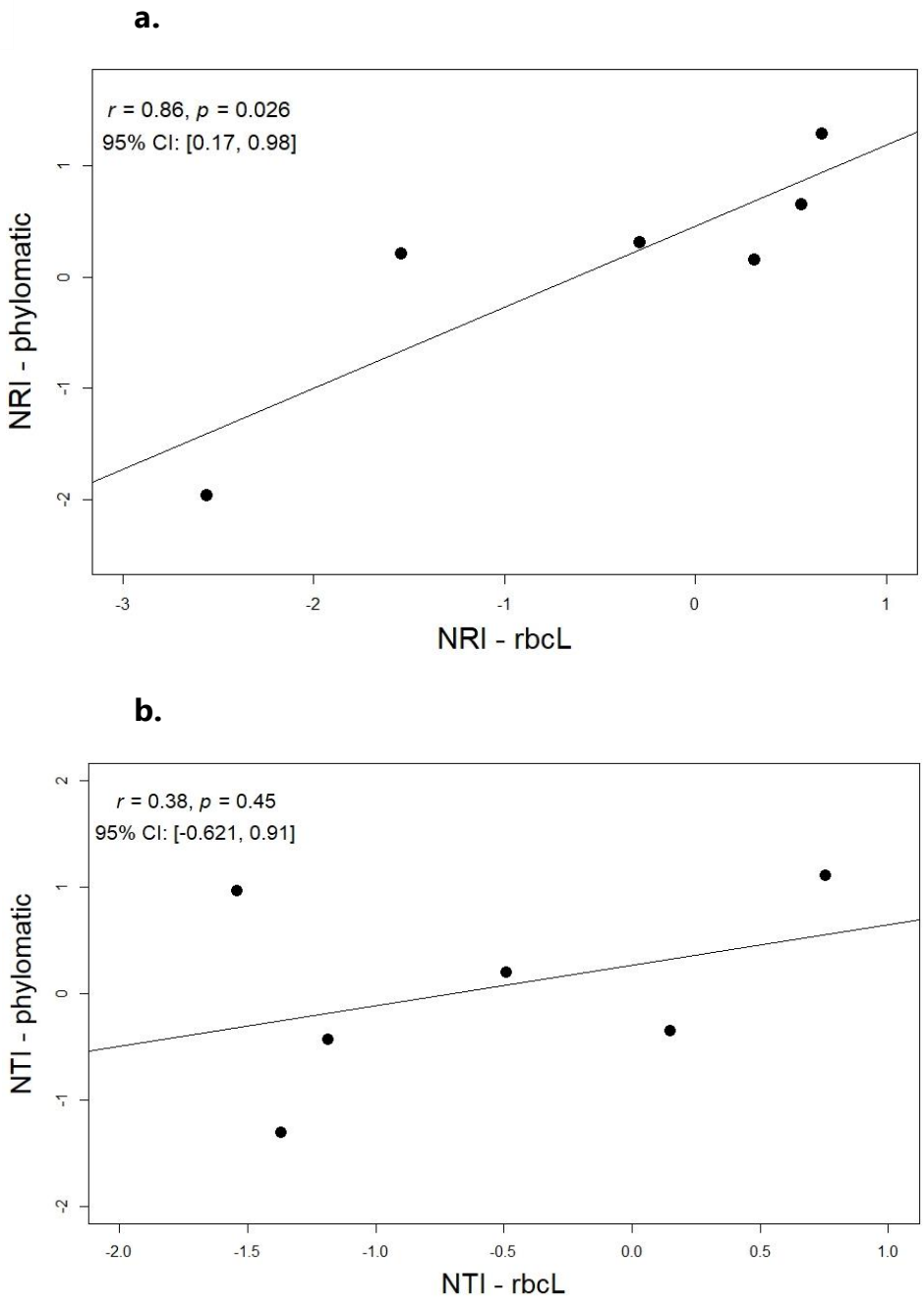

Figure S3 Correlation between phylogenetic structure index: a. NRI and b. NTI, based on the rbcLa phylogeny and the megatree phylogeny.