UNIVERSIDAD NACIONAL DE COLOMBIA

Influencia de la historia biogeográfica sobre la estructura de las comunidades de árboles neotropicales

The influence of biogeographic history on the structure and functioning of Neotropical tree communities

Jilbert Sebastián González Caro

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias, Departamento de Ciencias Forestales
Medellín, Colombia

II

The influence of biogeographic history on the structure
and functioning of Neotropical tree communities

2021

Influencia de la historia biogeográfica sobre la estructura de las comunidades de árboles neotropicales

The influence of biogeographic history on the structure and functioning of Neotropical tree communities

Jilbert Sebastián González Caro

Dissertation submitted in partial fulfillment for the degree of:
Ph. D in Ecology

Advisor:
Ph.D. Álvaro Javier Duque Montoya

Research line:
Tropical forest ecology
Research group:
Conservación, Uso y Biodiversidad

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias, Departamento de Ciencias Forestales
Medellín, Colombia

IV

The influence of biogeographic history on the structure
and functioning of Neotropical tree communities

2021

No ecological study can fail to benefit in some way from an understanding of the phylogenetics relationships of its taxa. C.O. Webb et al. 2002

Acknowledgments

First, I have to express my gratitude to advisor Prof. Álvaro Duque for his teachings, support, advice and his patience that made this process possible.

To Colciencias (Call 785) for the scholarship that financed my doctoral studies. At the Missouri Botanical Garden for stay in my internship. To the National University of Colombia, Medellín headquarters, and to the Department of Forest Sciences for the space and the moments.

Sebastian Tello, Kenneth Feeley, Jonathan Myers; Joost Duiveenvoorden and other co-authors who participated throughout the development of the different chapters.

To my friends and colleagues who supported me throughout the process.

To my family and Alejandra for their support and inspiration.

Abstract

The biogeographic history is an overlooked component of community assembly analysis and its implications on ecosystem functioning. In this dissertation, we incorporated the geological history of South America to assess hypotheses about the mechanisms that drive the community assembly of Neotropical tree communities. We combined phylogenetic and vegetation plots across the Amazon and the Andean regions to evaluate the role of historical dispersal and environmental gradients as drivers of the phylogenetic structure, diversity, and phylogenetic differentiation among tree communities:

1. We assessed the role of geology, soil conditions, and flooding as environmental drivers of the Amazon tree phylogenetic diversity relative to historical dispersal. We found that historical dispersal overrides environmental at large spatial scales, while environmental conditions drive community assembly at local scales.
2. We tested the dispersal of temperate tree clades to tropical highlands after the Andean uplift and their implications on the phylogenetic diversity and forest structure. We found that the colonization of temperate clades on the tropical Andes increases the phylogenetic diversity, which increases elevation in three independent datasets (at local scales elevational gradient, at the regional scale in plots across Colombia, and at the continental scale across the subtropical and tropical Andes). We also reported the influence of temperate clades as drivers of large carbon stocks in tropical highlands because these trees maintain large sizes in cold conditions.
3. We highlight the importance of dispersal of temperate clades in the formation of tropical Andean flora.

In summary, historical dispersal is a key mechanism overlooked in Neotropical tree community assembly.

Keywords: Amazonia, Andes, phylogenetic structure, phylogenetic similarity, historical dispersal, environmental gradients, South America biogeography.

Resumen

La historia biogeográfica es un componente que generalmente no se tiene en cuenta en el análisis de ecología de comunidades y en el funcionamiento de los ecosistemas. En esta disertación, incorporamos la historia geológica de América del Sur para evaluar hipótesis sobre los mecanismos que determinan el ensamblaje de comunidades arbóreas neotropicales. Nosotros combinamos información filogenética y parcelas de vegetación a lo largo de las regiones amazónica y andina andinas para evaluar el papel de la dispersión histórica y los gradientes ambientales como determinantes de la estructura filogenética, la diversidad y la diferenciación filogenética entre las comunidades de árboles:

1. Evaluamos el papel de la geología, las condiciones del suelo y las inundaciones como determinantes ambientales de la diversidad filogenética de los árboles del Amazonas relativo a la dispersión histórica. Encontramos que la dispersión histórica es más importante que el medio ambiente a grandes escalas espaciales como determinante del ensamble de comunidades, mientras que las condiciones ambientales impulsan el ensamblaje de la comunidad a escalas locales.
2. Probamos la dispersión de linajes arbóreos templados a las tierras altas tropicales después del levantamiento de los Andes y sus implicaciones en la diversidad filogenética y la estructura del bosque. Encontramos que la colonización de linajes templados en los Andes tropicales aumenta la diversidad filogenética, a lo largo del gradiente de elevación en tres conjuntos de datos independientes (en un gradiente de elevación a escala local, a escala regional en parcelas a lo largo de los Andes en Colombia, y a escala continental a lo largo de los Andes subtropicales y tropicales). También, encontramos que la presencia de los linajes templados determina la estructura de los bosques Andinos y su capacidad de almacenar carbono, porque estos árboles son capaces de mantener grandes tamaños en condiciones frías.
3. Destacamos la importancia de la dispersión de los linajes templados en la formación de la flora andina tropical.

En resumen, la dispersión histórica es un mecanismo clave que se pasa por alto en el ensamblaje de comunidades arbóreas neotropicales.

Palabras clave: Amazonía, Andes, estructura filogenética, similaridad filogenética, dispersión histórica, gradientes ambientales, biogeografía de Sur América.

Content

List of figures

List of tables

List of abbreviatures

Abbreviature	**Name**
AGB	Above ground biomass
AIC	Akaike information criterion
ANOSIM	Analysis of similarities
DBH	Diameter at the breast height
dbRDA	Distance based redundancy analysis
EDG	Elevational diversity gradient
INDVAL	Indicator value index
LDG	Latitudinal diversity gradient
MNTD	Mean nearest taxonomic distance
MPD	Mean phylogenetic distance
MZO	Multiple zone of origen
NMDS	Non-metric multidimensional scaling
NRI	Net Relatedness Index
NTI	Nearest Taxon Index
NWAmazonia	Northwest Amazonia
OTT	Out of the tropics
PCA	Principal component analysis
PCNM	Principal coordinates of a neighbor matrices analyses

Abbreviature	**Name**
PD	Phylogenetic diversity
PhyloSor	Phylogenetic Sorensen
PhyloSorNes	Nestedness component of the Phylogenetic Sorensen
PhyloSorRep	Replacement component of the Phylogenetic Sorensen
PhyloSorTotal	Total Phylogenetic Sorensen
QMD	Quadratic mean diameter
SEM	Structural equation modeling
SR	Species richness
TNC	Tropical niche conservatism
WD	Wood density

Introduction

In the last decades, plants' community ecology has been led by two contrasting viewpoints: the niche and the neutral theories. The niche theory is based on deterministic mechanisms, such as environmental sorting of species along environmental gradients and biotic interactions (e.g., competition or predation) (Tilman 1987; Leibold and Chase 2017). The neutral theory proposes community assembly to be driven by speciation and dispersal (Hubbell 2001; Chave 2004). These processes have recently been unified in a simple framework that proposes four main processes as a base of community assembly: speciation, dispersal, selection, and ecological drift (Vellend 2016). Under this perspective, community ecology highlights the role of evolutionary processes and historical contingency as key drivers of plant community assembly. Then, the understanding of issues such as historical dispersal, the impact of key innovations, and the effect of environment on diversification rates can help us to understand current biodiversity patterns (Ricklefs 2006). In this way, the integration of phylogenetic information into the study of ecological communities is paramount to better understand forest structure and function.

The new dimension that provides the inclusion of phylogenetic information of community ecology has derived in the measurement of evolutionary diversity in ecological communities. The pattern of structure of the phylogenetic relationship of the co-occurring species, known as phylogenetic community structure (Webb et al., 2002), has been widely used to assess the role of ecological processes on community assembly. The co-occurrence of distantly related species (i.e., phylogenetic overdispersion) was in principle proposed as the outcome of competitive exclusion (Webb et al. 2002; Cavender-Bares et al., 2004). The opposite pattern, phylogenetic clustering, was interpreted as the effect of habitat filtering, which selects for ecologically similar species (Webb et al. 2002; Cavender-Bares et al., 2004). These predictions are based on the assumption that evolutionary relationship reflects ecological similarity among species (i.e., strong niche conservatism). However, this framework has been criticized because these assumptions could not appropriately infer about the main processes driving community assembly (Mayfield and Levine 2010). The main argument has to do with the evolution of traits at different rates within the same clade, which can surpass the effect of other ecological mechanisms as drivers of the local phylogenetic structure (Gerhold et al., 2015). However, the relevance of the inclusion of the phylogenetic structure into the study of community ecology is that it helps to integrate the effect of diversification, historical dispersal, and other evolutionary processes along with the ecology and current features of plant communities (Gerhold et al., 2018).

South America has a complex geological history impacting their current biodiversity patterns (Hoorn et al., 2010). Since the separation of Gondwana 100My ago, South America was isolated from other landmasses between 80-98My. During this time, the Andean uplift was one of the most important geological events in the continent. This mountain uplift began ~60 My in the Southern part of the continent and was progressive northwards (Zeil 1979; Graham 2009). Then, different parts of the Andean mountain range have different ages; the Central Andes have been dated between 30-40My, while Northern Andes were ~10My or less (Horn et al. 2010). The Andean uplift produces a new evolutionary arena with a vast diversity of niches and new conditions that promote the Neotropical biota diversification (Gentry 1982). Simultaneously, the Amazon landscape was modified as a response of the Andean uplift. Due to tectonic dynamics, the Pebas lake drainage produces a new mosaic of environments in Western Amazonia (Horn et al. 2010). A huge soil condition arrangement composes this landscape mosaic is broadly classified in terra firme, flooded, and white-sand forests. Under these scenarios, we evaluated the effect of historical dispersal and environmental gradients as drivers of phylogenetic diversity patterns and ultimately community assembly.

In this dissertation, we integrated large-scale plot datasets and phylogenetic information to analyze the historical imprint of plant evolution into the neotropical Amazon and Andean tree communities. Our main goal was to evaluate the influence of historical dispersal and environmental gradients on the spatial pattern of phylogenetic diversity on these two biodiversity hotspots. Specifically, we assessed i) the influence of environmental heterogeneity and historical dispersal as drivers of the phylogenetic structure and diversity tree communities under a complex geological history of the Amazon basin (Chapter I and II); and ii) the role of historical dispersal across elevation and latitude shaping the phylogenetic composition and structure of tropical and subtropical Andean tree communities (Chapter III, IV and V).

Outline of the thesis

In the first chapter, we used 400 0.1-ha forest plots distributed in the Colombian Amazon to assess the effect of geological substrate (terra firme, flooded and white-sand forests) and spatial distance on the phylogenetic similarity of tree communities larger than 10 cm of diameter at the breast height (DBH). We used multiple regressions based on dissimilarity matrices to assess the effect of geological substrate and spatial distance. Also, we calculated the degree of dispersion of indicator species of each geological substrate in the whole phylogenetic tree built up on Amazon trees. We found that phylogenetic similarity is primarily driven by spatial distance (i.e., dispersal overrides geological effects on phylogenetic community composition), suggesting a dispersal role at evolutionary timescales as a fundamental process in the Amazon community assembly. Additionally, we identified indicator species in each geological substrate and found that white-sand indicator species were phylogenetically clustered compared to the other geological substrates suggesting that this habitat could be difficult to invade due to the restrictions imposed by bad drainage and low soils fertility (Cárdenas et al. 2017). In the second chapter, we assess the effect of flooding (i.e., anoxia levels) on the phylogenetic similarity of tree communities in a dataset of 85 0.1-ha forest plots (DBH ≥ 2.5 cm) distributed in the Western Amazon (Colombia, Ecuador, Perú). Here, we used three species pools that change in spatial scale to calculate the standardized effect sizes of phylogenetic similarity. (whole Western Amazon or regional, country level and habitat level). This procedure allows us to assess the scale-dependency of dispersal and environmental filtering on clade sorting between flooded and unflooded units. We used variation partitioning to disentangle the role of these variables. We found that spatial variation explains phylogenetic similarity relative to the regional species pool, but flooding increases their importance at country and habitat restricted species pools. Our results also highlight the importance of historical dispersal as a driver of Amazon community assembly at a regional scale and flooding (or other soil-related factors) shaping local Amazon community structure (González-Caro et al. 2020).

In the second section (Chapter III, IV, and V), we evaluate the effect of historical dispersal from tropical lowlands and temperate regions as a source of species of the montane tropical Andes. We tested two hypotheses: the tropical niche conservatism (TNC) and the multiple zones of origin (MZO) hypotheses; the latter is a proposal of this study. The TNC assumes that all plant clades are originated under wet tropical conditions and the colonization of new environments (e.g., seasonal or cold) is restricted due to physiological constrains. Only a few clades that evolve particular traits can reach these new areas. The TNC predicts that phylogenetic diversity decreases with both elevation and latitude (Wiens & Donoghue 2004; Kerkhoff et al., 2014). In contrast, the MZO hypothesis assumes multiple ecological zones of origin and multiple adaptations to each environment. This hypothesis predicts a mix of ecological zones of origin in new environments mediated by the dispersal capability, which results in an increase of the phylogenetic diversity. along elevational and latitudinal gradients in the tropical Andes.

To assess these hypotheses, we evaluated phylogenetic diversity patterns in three independent datasets at three different scales: i) a single spatially restricted elevational gradient of 20 0.25-ha forest plots located along the eastern flank of the Andes-Amazonia conjuction in Colombia (landscape scale; III chapter; Ramírez et al., 2019), ii) 90 0.25ha forest plots distributed across Colombian Andes (regional scale, IV fourth; González-Caro et al., 2020), and iii) 245 forest plots (ranging between 0.25-1.1ha) distributed along the Andes from Argentina to Colombia (continental scale; V chapter; González-Caro et al., in prep). In all cases, we found that phylogenetic diversity increases along elevation gradients. We also found that phylogenetic diversity increases with latitude in the continental dataset. We also used simulations to evaluate our predictions on the TCN and the MZO. Based on the evidence of these three datasets, we suggest that the immigration of temperate clades to the tropical Andes is an overlooked process in the tropical flora formation.

1 Chapter 1: Asymmetrical niche determinism across geological units shapes phylogenetic tree communities in the Colombian Amazonia

1.1 Introduction

The Amazon basin harbors one of the most diverse tree communities on Earth (Slik et al., 2015, Duque et al., 2017). Such a high diversity has been attributed to a complex geological history resulting from the Andean uplift and involving different tectonic movements, changes in the direction of rivers, and processes of sedimentation (Hoorn et al., 2010). Therefore, Amazon forests have evolved on highly variable geological conditions that change in terms of age (2.5 - 65my), soil fertility, and patterns of drainage from the eastern white-sand Guiana shield dominated systems (hereafter craton unit) to the western Andean foothills (ter Steege et al., 2006). The extent to which these geological differentiations can be translated into evolutionary processes that determine the floristic differences in the Amazon forests has been intensively debated (Fine & Kembel, 2011; Dexter et al., 2012, 2017; Guevara et al., 2016). Although different mechanisms, such as environmental filtering or dispersal ability, have been suggested as being important historical drivers of the current biota in Amazon forests, the relative importance that each plays in structuring the phylogenetic structure of tree communities remains controversial (Fine and Kembel, 2011; Fortunel et al., 2016; Dexter et al., 2017).

Under the environmental filtering view of Amazonian diversity, we expect abrupt changes in tree community composition across different geological units (ter Steege et al., 2006). A comparison of functional traits between contrasting geological units, such as the white-sands in French Guiana and the tertiary terra firme forests in Peru, indicates strong habitat filtering as the main driver of the tree community composition in these areas (Fortunel et al., 2014). These findings suggest that dispersal plays a weak role in shaping the species composition of tree communities across geological units, and that environmental conditions primarily determine the capability of tree species to establish and reproduce in specific environments (Fortunel et al., 2016). In cases such as this, we can expect a high phylogenetic turnover between geological units controlled by plant-habitat specialization due to a constrained evolution of the species' niches through time (niche conservatism sensu Wiens and Graham, 2005). High niche conservatism between geographically separate geological units may influence allopatric speciation (Guevara et al., 2016), which will produce a pattern of geographically structured phylogenies with a high abundance of closely related species associated with each environmental unit.

In contrast, under a dispersal-mediated view of Amazonian diversity, we assume that species movement has been a primary determinant of floristic composition and diversity independent of the geological features. In other words, according to this viewpoint, species distributions in the Amazon basin have historically been determined by the species' differential dispersal abilities and with only a relatively weak influence of geographical barriers (Dexter et al., 2012, 2017). If there is effective establishment of propagules of clades with origins in different geological units through time, the ecological traits associated to each geological unit (sensu Fortunel et al., 2014) should be labile through evolutionary history. If this is true, ecological speciation may have been an important driver of Amazon diversification (Misiewicz and Fine, 2014). Based on this scenario of evolutionary convergence, we will expect that the local coexistence of tree species will be geographically structured by distantly related species (Dexter et al., 2017), which may also be widespread across geological units (Pitman et al., 1999). We also expect larger (Wang et al., 2013), older geological units to support more diverse communities due to greater likelihood and time to accumulate species (Fine & Ree, 2006). This idea of Amazon forests structured by geological units floristically differentiated just by the capability of species to migrate and colonize new areas in geographical space (Hubbell, 2001) has received recent support at evolutionary timescales as well (Dexter et al., 2017).

In the Amazon, the oldest geological craton units with precambric origin are derived from the Guiana shield; however, we know that there are fewer species, genera, and families in these units than in the younger tertiary sedimentary terra firme and the quaternary floodplains (Duivenvoorden, 1995; Duque et al., 2002), contradicting the expectation of the time-integrated species-area effect (Fine and Ree, 2006). This phenomenon, named by Hoorn et al. (2010) as the craton paradox, is by itself a strong argument supporting niche conservatism and environmental filtering as important mechanisms that shape the phylogenetic structure of tree communities in the Amazon basin (Fine and Kembel, 2011; Guevara et al., 2016; but see Dexter et al., 2017). The low-nutrient content, seasonal anoxia, and bad drainage of soils in white-sand systems (Lips and Duivenvoorden, 1996) may increase the level of environmental stress, and thus promote the development of convergent adaptive strategies. Likewise, in Amazon floodplains, the reported dominant pattern of phylogenetic clustering for tree communities (Aldana et al., 2016; Umaña et al., 2012), gives additional support to the assumption of environmental filtering as one of the most important mechanisms determining the phylogenetic structure of tree communities across geologic units in the Amazon basin (Fine et al., 2005; Fortunel et al., 2016). Therefore, comparative regional assessments of the phylogenetic structure and turnover of tree communities that include the most contrasting geological units, such as craton, tertiary sedimentary plains, and quaternary floodplains, are needed to help unravel the role played by asymmetric environmental filtering across geologic units in determining the distribution of tree communities in the Amazon. This prediction implies that the degree of niche conservatism will change between different evolutionary and ecological scales, and as so, between geological units.

Here, we aim to evaluate the role of differences in substrate age and environmental conditions, as represented by the main geological units, in determining the phylogenetic structure and distribution of tree communities in the Colombian Amazon. The Colombian Amazon has a strong longitudinal gradient of geological formations and substrates, which varies from the western Andean mountains to the eastern craton dominated formations, crossing through tertiary sedimentary plains and floodplains. The overall geological variation in the Colombian Amazon reflects the variation found throughout the Amazon basin, allowing us to use it as a benchmark to assess the extent to which environmental filtering determines the evolution and distribution of Amazon tree communities more generally. It is important to note that the use of phylogenetic indices to infer the main structuring mechanisms of plant communities has previously been criticized (Cavender-Bares et al., 2009). In particular, it has been argued that using assemblages of either distantly or closely related species to disentangle the roles played by competition vs. habitat filtering is “out of date” in modern species coexistence theory (Mayfield and Levine, 2010). According to this argument, more fundamental problems can arise when we assume that increasing phylogenetic differences favor species coexistence. To bypass some of these concerns, we employ phylogenetic turnover and structure, focusing on the main theoretical frame introduced above, to assess the extent to which dispersal (represented by geographical distances) and environmental filtering (represented by geological units) determine tree community composition in the Colombian Amazon, but avoiding any reference to competition. In other words, we assume that environmental filtering is due to a combination of both species’ strict abiotic habitat requirements and competitive exclusion based on species’ differential competitive abilities in different settings (Mayfield and Levine, 2010).

The main research questions and associated hypotheses addressed in this study are:

i) To what extent do environmental filtering and dispersal limitation determine the phylogenetic composition of tree communities across geological units in the Colombian Amazon? We expect a high correlation of phylogenetic turnover with geology as well as a high phylogenetic clustering of indicator species in each geological unit, which would support environmental filtering as the main factor shaping the evolution of tree communities in the Amazon forests.

ii) Are there differences between geological units in the extent to which niche conservatism shapes the phylogenetic structure of indicator species (i.e., strong habitat specialists) in tree communities? We hypothesize that niche conservatism increases with the age of the geological unit. In geological units with greater environmental stress, such as craton and floodplains, we expect the associated species to have evolved similar ecological traits that allow them to be abundant and widely distributed within each geological unit. In other words, the indicator species associated to each geological unit will be phylogenetically clustered if ecological advantages (traits) evolved in a conserved fashion.

1.2 Methods

Study region

The Colombian Amazon covers an area of roughly 483,119 km^2, which represents 41% of the country´s continental territory, and 8.7% of the Amazon basin (Fig. 1). Although the Colombian Amazon includes the largest deforestation hotspot in the country, the natural forests are still well preserved in more than 70% of the region. The Colombian Amazon has the lowest population density (1.5 people km^{-2}), and harbors the highest portion of indigenous communities of the whole country. To date, 84% of the region´s territory is consolidated into protected areas including Indigenous Reserves and National Parks.

The Colombian Amazonia can be divided into three main geological units according to substrate age and origin: 1) The Amazonian craton, which is the oldest geological formation with its origins in the precambric age (> 600 Ma). The craton unit is structured by Precambrian rock formations of the Guiana Shield that primarily emerged in the North-East part of the region (Hoorn et al., 2010). 2) The Amazon tertiary basin infills, which were mainly deposited during the Miocene and are part of the northern extension of the Solimões intracratonic basin (Hoorn et al., 2010). The depositional process of the area filled in by sediments from the Guiana shield gave origin to the sand units informally known as Mariñame and Apaporis (Hoorn et al., 2010). On the contrary, the depositional process of clay sediments with Andean origin created the Pebas formation, which was part of the mega-wetland that covered the Amazon during the Late Miocene (Hoorn, 2006). Although there are clear differences in the soils derived from either the sand or clay sediments in terms of structure and fertility (Lips and Duivenvoorden, 1996), here we include both as the tertiary unit but split them as needed into Tertiary-Shield or Tertiary-Pebas geological units, respectively. Any white-sand systems enclosed within these tertiary systems were treated as cratonic due to the parental material they are located on (Duivenvoorden et al., 1995). 3) Alluvial formations, which were all considered with Quaternary origin mostly due to the observation from the field to be located on floodplains. Although in the alluvial formations there could be fluvial terraces that stem from Pliocene, Pleistocene or Holocene, they are assumed to be never flooded and therefore not included within our Alluvial definition of a Quaternary origin. Likewise, the alluvial formations can be described and differentiated in terms of their geologic origin as well, but in this study, all alluvial formations were treated as a single geologic unit framed by the rivers´ flood plains.

Plot inventory data

Between 1990 and 2010, we conducted floristic surveys of 412 0.1-ha plots across the Colombian Amazon region (Figure 1.1). To establish the plots, the direction of the trails along which the forests were entered was planned a priori on the basis of the interpretation of aerial photographs and satellite images (e.g. Duivenvoorden and Lips, 1993). Forests were visually examined to identify terrain units that were more or less homogeneous. In these units, rectangular plots (20-m wide x 50-m long or 10-m wide x 100 m long) were located and delimited by compass, tape, and stakes. All of the plots were mapped by GPS and were established in mature forests that did not show obvious signs of recent human intervention. In each plot, all trees and palms with a DBH (Diameter at Breast Height) ≥ 10 cm were recorded and collected. Each plot was assigned to one of the three main geological units (craton, tertiary, or alluvial); when needed, we also divided the tertiary geological unit into Tertiary-Sand or Tertiary-Pebas as explained above.

The identification of the botanical collections was performed at the COAH, COL, MO, and AAU herbaria. Within families or groups of closely allied families, specimens that could not be identified at the species level because of a lack of sufficient diagnostic characteristics were clustered into morphospecies on the basis of simultaneous morphological comparisons with all other specimens. Hereafter, the term species refers to both morphospecies and botanical species.

Phylogenetic similarity

We built a hypothesized phylogenetic tree for all tree species found in our Colombian Amazon plot dataset using phylomatic version 3.0 (Webb and Donoghue, 2005) (Newick file in Appendix S1). We included unidentified species as genus or family level polytomies depending on botanical identification. Our phylogenetic tree was made up of 1693 tips and 422 nodes (24% of the evolutionary relationships were resolved). The largest proportion of polytomies is distributed within genera and deep phylogenetic relationships can be considered relatively well-resolved. Thus, our phylogenetic analyses are primarily focused at the genus level, minimizing analytic problems with plot comparisons associated with the level of resolution of tree species identification.

We used the PhyloSorensen index to build a matrix of phylogenetic similarities between plots based on species' presence-absence data. The PhyloSorensen index sums branch length of shared clades between sites (Bryant et al., 2008) relative to the sum of branch lengths of both sites:

$$PhyloSorensen\ ij = \frac{BL\ ij}{0.5(BL\ i + BL\ j)}$$

where BL*ij* is the sum of branch length shared between plots *i* and *j*, and BL*i* and BL*j* are the sum of branch length of unique tips within plots *i* and *j*, respectively. Thus, if shared and unique branch length is balanced, the index takes a value near to 1. In contrast, if unique branch length exceeds shared branch length the index is close to 0. This metric was estimated using the *phylosor.query* function available in the *PhyloMeasures* (Tsirogiannis and Sandel, 2015) package for R.

Statistical analyses

To define the compositional pattern of phylogenetic variation across the geological units, we used the hierarchical Ward´s clustering method (Cayuela et al., 2006) built from the phylogenetic similarity matrix of the plots. To find the cutoff point that maximizes the phylogenetic compositional differences between groups of plots, we used an iterative method based on the R-statistic obtained from an analysis of similarities (ANOSIM) (Legendre and Legendre, 2012). The ANOSIM method evaluates whether differences within groups are comparable with differences between groups. The method employed here systematically assessed the increment in the R-statistics of ANOSIM in relation to the increment in the length and number of branches in the Ward's clustering dendrogram until the R-value asymptoted. Furthermore, we used a chi-squared test to evaluate the congruence between Ward's dendrogram-based clusters and geological units. The chi-squared test allows us to evaluate if proportions are distributed randomly or the dendrogram clustering is constrained to specific geological units. The chi-squared test was performed on a contingency table that uses the geological units and dendrogram groups as categorical variables and the estimated number of plots for each combination.

We used the Mantel test (Legendre and Legendre, 2012) to evaluate the influence of geology in determining the phylogenetic turnover of trees in the Colombian Amazon. To run the Mantel test, the PhyloSorensen distance matrix was converted into a distance by applying 1 – similarity. This analysis allows us to assess the influence of geology on the phylogenetic tree turnover to understand the expected deterministic effects of environment (e.g. geology) on the pattern of tree distributions. For the Mantel analysis, we used the PhyloSorensen distance between plots and a presence-absence matrix of geological units. The presence-absence distance matrix of geology includes 0 when two plots were located on the same geological unit and 1 otherwise.

Likewise, we used the Mantel test to evaluate the influence of geographical distances between plots (straight line distances in km) in determining the phylogenetic turnover of trees in the Colombian Amazon. In this study, we used the geographic distances to account for the hypothesized effect of spatially structured biological processes, such as dispersal limitation (Condit et al., 2002). For this analysis, we used the PhyloSorensen similarity between plots and the log-transformed geographic distance between plots (Fine & Kembel, 2011). The distance-decay approach (Nekola & White, 1999; Tuomisto et al., 2003) was then applied to visualize the relationship between the phylogenetic similarities and the log-transformed spatial distances. The approximately linear distance-decay along the logarithm of geographical distance can be used as an explicit prediction (and test) of Hubbell's neutral theory (Morlon et al., 2010). Finally, we used the partial Mantel test to assess the effect from either geology or geographical space after controlling for the effect of the other. Mantel and partial Mantel tests were performed using the library *vegan* (Oksanen et al., 2013) for R.

Indicator species and Phylogenetic clustering

In order to explore species associations with geological units, we used the indicator value index (INDVAL) proposed by Dufrene and Legendre (1997). This method incorporates the relative abundance and relative frequency of individual species across habitats (e.g. geological units) to test the degree of habitat associations based on a null model assuming a random distribution of each species among habitat types. The significance is based on the distribution of the observed value relative to 1000 iterations of the random model. We used 0.01 as the confidence limit to define indicator species. Based on the set of indicator species of particular conditions that prevailed in site groups (Legendre and Legendre, 2012), we evaluated the role of niche conservatism for geological unit.

We expected that high niche conservatism along evolutionary history results in high habitat specialization and potentially habitat dominance. Taking into account the assumptions stated in the introduction section, phylogenetic niche conservatism can be used to predict phylogenetic clustering of co-occurring species under particular habitat conditions (Wiens and Graham, 2005). To assess niche conservatism, we conducted the following analysis: first, we calculated a phylogenetic mean pairwise distance (MPD) between indicator species within each geological unit. Next, we randomly chose the same number of species from a regional species pool that did not include the subset of indicator species and calculated their MPD. This procedure was repeated 1000 times and the standardized observed values were used to build the expected distribution. Standardized values were calculated as the observed value minus the mean random value divided by the standard deviation of the random distribution. We then multiplied these standardized values by -1 to obtain a Net Relatedness Index (NRI), a common phylogenetic community structure metric (Webb, 2000). Negative values of NRI indicate phylogenetic overdispersion and positive values indicate phylogenetic clustering. Significance was determined by ranking the observed value on the null distribution divided by the number of replicates. This procedure was run in the library *picante* (Kembel et al., 2010) for R.

Phylogenetic clustering/overdispersion patterns are dependent on null model and species pool definition (Lessard et al., 2012). Therefore, in order to test phylogenetic dispersion of indicator species within each geological unit, we used three different species pool definitions. Our first species pool definition included all of the species present in our dataset after having removed the species indicator set. Phylogenetic clustering could indicate habitat filtering of conserved traits imposed by geological units. The second species pool was made up of all species present within each geological unit. This species pool definition allows us to evaluate if indicator species of a particular geological unit are more clustered than expected by chance in relation to other species that occur within that particular unit. Within geological units, phylogenetic clustering of indicator species may indicate conservatism of evolutionary advantages (i.e defense to herbivory) that should facilitate ecological specialization through environmental filtering. On the contrary, phylogenetic overdispersion of indicator species within geological units may indicate converging ecological traits on the phylogeny along the evolutionary history. Finally, we used a species pool including all indicator species with the aim of evaluating whether ecologically indicator species in a particular geological unit are phylogenetically clustered in regards to all indicator species found across geological units.

1.3 **Results**

Across all 412 inventory plots, we tallied 27,192 individual trees belonging to 1,683 species from 84 plant families. Overall, 84.9% of individuals were identified to the species level, 9.1% to the genus level and 2.1% to the family level. The remaining 3.9% of individuals that were not identified at least to the family level were removed from the community matrix and all subsequent analyses.

In all plots, there are 358 singleton species (21.2%) and 206 doubletons (12.24%). In contrast, Oenocarpus bataua was the most abundant species with 703 individuals, followed by Oxandra polyantha (583 individuals), Euterpe precatoria (543 individuals), Socratea exorrhiza (459 individuals) and Micrandra sprucei (427 individuals). The five most diverse families were Fabaceae (249 species), Annonaceae (100 species), Lauraceae (96 species), Rubiaceae (95 species) and Chrysobalanaceae (75 species). On average, there were 29.4 species (range = 1 - 67) per plot and 66 individuals (range = 10 - 284) per plot. Across geological units, the number of individuals was significantly higher in the alluvial formation than in the tertiary, while species richness was significantly higher in the tertiary than in the craton and alluvial geological formations (Table 1.1). There were no differences in the mean phylogenetic distance between geological units (Table 1.1).

Based on the ANOSIM R-statistic, we differentiated seven phylogenetic groups primarily associated with geology (Figure 1.2). The plots located on the craton unit were divided into two main groups, one of them located in the middle of the tertiary forests. The chi-squared test showed a significant correlation between geological units and phylogenetic clustering (X2 = 10.21; $p < 0.001$), which suggests that the phylogenetic composition of each plot was strongly associated with a specific geological unit. However, none of the clusters was made up exclusively of plots from only one geologic unit.

When the entire dataset derived from plots located in all geological units was analyzed, the phylogenetic composition was significantly correlated with geology (Mantel r = −0.16). Furthermore, the correlation between phylogenetic composition and geology decreased (Mantel r = −0.09) when we split the tertiary unit into Tertiary-Shield and Tertiary-Pebas. When the alluvial formation was excluded, the Mantel correlation increased from −0.16 to −0.29 (Table 1.2), suggesting that phylogenetic similarity varies strongly between pairs of alluvial plots across the whole Colombian Amazon. Likewise, when the entire dataset was considered, phylogenetic composition had a stronger correlation with the spatial configuration of the plots (Mantel r = −0.25) than with geology; the Mantel correlation increased again from −0.25 to −0.39 (Table 1.2) when the alluvial plots were removed (Table 1.2). When the effect of either geology or geographical distance was accounted for, the correlation between phylogenetic similarity of tree communities and the remaining explanatory factor (geographic distance or geology, respectively) were in all cases still significant (Table 1.2). However, the stronger correlation observed with geographic distances than with geology suggests an active effect of plot distance that tended to increase with plot proximity up to an average distance of 20 km between plots (Figure 1.3).

According to the indicator value index (INDVAL), there were 241 indicator species in the three geological units (Table 1.3) that represent 14% of all species registered in the plots. Of these indicator species, 127 were found in the craton, 66 in the tertiary, and 48 in the alluvial unit (represent 25.6%, 5.0%, and 4.5% of the total of species found in each unit, respectively). In all cases analyzed, the craton indicator species were phylogenetically clustered. In contrast, indicator species from tertiary and alluvial units showed a clear, but statistically insignificant, trend towards phylogenetic overdispersion (Table 1.4).

1.4 **Discussion**

The results of our study give support for both environmental filtering and dispersal as being important drivers of the phylogenetic turnover and structure of tree communities in the Amazon basin. However, our first hypothesis about the expected dominant role of plant-habitat specialization being the most important mechanism determining tree phylogenetic assemblages in the main geological units was partially rejected. The extent to which the geographic distance between plots explained the phylogenetic turnover of tree communities, used to account for dispersal limitation (Eiserhardt et al., 2013; Morlon et al., 2010), surpassed that of geology. Dispersal limitation appears to actively drive phylogenetic turnover at distances equal to or lower than approximately 20 km; beyond this distance, abiotic filters have a significant complementary effect. This finding suggests that most of the lineages have had sufficient time to migrate and establish into adjacent geological units (Pitman et al., 1999; Dexter et al., 2017), contradicting the claims that geological conditions are the dominant factors determining the distribution of tree communities across northwestern Amazonia (Pitman et al., 2008).

The strong influence of spatial distance in determining the phylogenetic similarity between tree communities illustrates the fundamental role that dispersal has had in shaping the evolutionary patterns of tree communities in the Colombian Amazon (Dexter and Chave, 2016; Dexter et al., 2017; Fine & Kembel, 2011). This result is contrary to expectations based on the age of the geological formation and the strong role of the transition between Guiana to Andean systems (Hoorn et al., 2010). This finding suggests that niche conservatism is weak in Amazon trees - at least at the community level. Weak niche conservatism among geological units could be expected if the metacommunity of any local community is composed by the complete regional species pool (Dexter et al., 2017), which may make the development of specific traits labile through time (Silvertown et al., 2006). However, our study is mostly focused on large trees, which may mask the stronger abiotic filtering effects of habitat specialization in earlier developmental stages of trees as seedlings (Fortunel et al., 2016). Therefore, we cannot completely rule out the role played by environmental filtering as a more active complementary mechanism of dispersal in shaping the phylogenetic structure and turnover of Amazon tree communities.

The smaller but significant role played by geology in determining the phylogenetic structure of Amazon tree communities may in large part be attributable to physiological filters imposed on plant species by the stressful environmental conditions mostly derived from the seasonal low availability of oxygen in soils and very low nutrient contents, typical of the craton unit (Duivenvoorden et al., 1995; Fortunel et al., 2016). For example, based on the total number of species found in each geological unit, we found a larger portion of indicator species in the craton (25.6%) than in the tertiary (5%) and alluvial (4.5%) geological units. The significant phylogenetic clustering found in the craton unit confirms and emphasizes the role played by niche conservatism in shaping the evolutionary development of specific traits, such as herbivore defenses (Fine et al., 2005), that enable some particular species to succeed in this geological unit. The significant trend of phylogenetic clustering for indicator species in the craton unit persisted independently of the species pool definition employed (regional, habitat or only indicator species). Furthermore, the separation of the plots located on the craton into two different groups according to the cluster (Figure 1.1) suggests an apparently slow evolution of the species' niches and supports niche conservatism, likely due to a strong influence of an allopatric process of speciation as a main driver of diversification in the white-sand formations (Guevara et al., 2016).

Our results about the influence of niche conservatism in the craton unit defined from the coexistence of closely related species differed of recent results obtained from better-resolved phylogenies of some specific genera (Dexter et al., 2017). Dexter et al. (2017) found a lack of geographic phylogenetic structure in the Inga, Swartzia, Protieae, and Guatteria lineages across different regions of the Amazon. However, in French Guyana and the Atlantic coastal rainforest of Brazil where soil characteristics, such as low fertility due to leaching, are similar to those found in our craton unit, some of these taxa were phylogenetically clustered. A tendency to find geographical phylogenetic structure associated to systems with very low nutrient contents or drainage anomalies gives support to our claims of higher environmental filtering in the craton unit than in the other environments due to resource limitation. Although the use of phylogenetic clustering of only indicator species to infer environmental filtering as the main mechanism determining the composition of tree communities in the craton is debatable (Cavender-Bares et al., 2009; Mayfield and Levine, 2010), the evidence from other studies showing a trend towards geographic phylogenetic structuring in white sands (Guevara et al., 2016, Dexter et al., 2017) support our findings. Nonetheless, we acknowledge the need of additional experimental and observational studies based on well-resolved phylogenies to more conclusively test the idea that the craton unit operates as a dissimilar environment and drives closely related species to similar stabilizing niche differences, thereby promoting species coexistence.

In contrast to the craton and contrary to expectations, the alluvial geological unit was the most widespread geological formation across the phylogenetic clusters defined here. This suggests that flooding may not be a strict factor in restricting the colonization of species from across diverse lineages. Overall, tree communities in flooded plots tended to be more phylogenetically similar to nearby unflooded forests than to more distant flooded plots. In other words, although we cannot deny the physiological stresses imposed by flooding, floodplains may be defined as successional forests where the bulk of the tree community composition is mostly composed of species that disperse in from the locally and regionally adjacent tertiary sedimentary plains (Terborgh and Andresen, 1998; Pitman et al., 2014). Thus, the alluvial formation appears to be a geologic unit that is evolutionarily "easy" to colonize, which explains the phylogenetic overdispersion found for the alluvial indicator species. Although previous studies reported a trend of phylogenetic clustering in flooded forests (Umaña et al., 2012, González-Caro et al., 2014, Aldana et al., 2016), its ecological interpretation, in both this and former studies, need to be treated with caution due to statistical insignificance of the results. However, in this study, we find that the indicator species reported in the alluvial unit belonged to the Arecaceae family, which agrees with previous reports made for this geologic unit in Amazon forests (Duque et al., 2002, ter Steege et al., 2013). In contrast, in other studies (i.e., Aldana et al., 2016), Fabaceae, a lineage well represented in most of the Amazon geologic units, was one of the indicators and dominant families in the alluvial formation. Historical-biogeographic differences between geographic regions that determine the floristic composition could explain the differences in the underlying processes that drive the tree phylogenetic structure of tropical forests in this particular ecosystem.

The indicator species showed contrasting patterns between geological units in terms of abundance and phylogenetic origin. The craton units are largely dominated by species from the Eudicot clade, particularly from the Rosid families. For example, Micrandra sprucei is the most abundant species and is recognized as being hyperdominant on podzolised soils of the white-sand formations (ter Steege et al., 2013). In addition, Fabaceae is within this clade and is the most representative taxa of our craton units. Species of this family are recognized for their mutualism with Rhizobium species that increase their ability to uptake nutrients under poor soil conditions (Adams et al., 2016). The high number of Fabaceae indicator species may therefore support the increased importance of environmental filtering as a main driver of tree diversity in the craton unit (Fortunel et al., 2014). Alluvial units are dominated by the three major angiosperm clades: Magnoliids, Eudicots and Monocots. The most abundant species in all Amazonia (ter Steege et al., 2013), Euterpe precatoria, and other palms such as Mauritia flexuosa, were likewise the most representative species in this geological unit. Some Magnoliid species from Annonaceae such as Oxandra polyantha, which was dominant in our study, has also been reported as dominant in black-water flooded forests (ter Steege et al., 2013). As stated before, this phylogenetic overdispersion of indicator species could be the result of a weak evolutionary restriction of flooding conditions and water stress for species establishment into these forests (Pitman et al., 2014). The co-occurrence of distantly related species inhabiting floodplains emphasizes the need for understanding the wide range of ecological tolerances that species have developed to tolerate flooding and the mechanisms that allow such a high diversity of trees in this geologic unit (Parolin et al., 2004). In the tertiary unit, the most dominant species, such as Socratea exorrhiza (Monocot) and Pseudolmedia laevis (Eudicot), correspond to different angiosperm clades in a similar way as in the alluvial units. In summary, the phylogenetic structure of indicator species suggest that evolutionary and ecological mechanisms differ between the three geological units, but with a marked difference between craton and the other two units. Therefore, we find support for our second hypothesis about an asymmetric effect of niche conservatism between geological units that differentially shapes the evolution of tree communities in the Colombian Amazon.

The asymmetric effect of phylogenetic clustering between geologic units proposes that the movement of propagules from tertiary and alluvial to craton may have been less frequent than between them through time. For example, Misiewicz and Fine (2014) showed that the white-sand – terra firme transition influences the genetic structure and morphological traits of Protium subserratum populations, suggesting that divergent natural selection affects immigration between soil types. In contrast, Dexter et al. (2012) showed that the terra firme – alluvial transition does not affect the genetic population structure of eight Inga species, which is line with our results. In this context, craton units may have operated as a historical source of species to tertiary and alluvial plots rather than vice-versa, a pattern that may have had consequences on the species diversification of Amazon tree communities. Dexter and Chave (2016) showed that genus range size and species richness are negatively related, indicating that species formation may depend upon range fragmentation and slow secondary contact, both processes directly related to dispersal limitation. In addition, they suggest that dispersal-related traits such as tree height are important drivers of the genus - species richness relationship found in different geological units in Amazon forests, indicating that intrinsic factors such as dispersal ability may be one of the most important drivers of tree diversification in the Amazon basin (Claramunt et al., 2012; Dexter and Chave, 2016; Dexter et al., 2017), as it is proposed in this study.

To conclude, our study suggests that from an evolutionary perspective, both dispersal and environmental filtering interact to shape the assembly of Amazon tree communities. The extent at which these two processes determine the phylogenetic structure of Amazon tree communities is asymmetrical between geological units. In the craton unit, environmental filtering appears as the dominant force due to the more restrictive conditions that may constrain seedling establishment (Fortunel et al., 2016). In the craton unit, the environmental restrictions have diminished the evolutionary opportunity of ecological success, resulting in a dominance of closely related species, a mechanism widely known as niche conservatism. The restrictions imposed by the craton for species establishment is directly associated with the magnitude and significance of statistical tests of beta diversity when this geologic unit is included. In contrast, in the younger and more diverse tertiary and alluvial units, a very active exchange of propagules is driving compositional patterns in these two geological units. In particular, our findings support the idea that alluvial units may have largely been colonized by habitat generalist species with few species specifically restricted to it (Pitman et al., 2014). This result suggests that there is a wide range of ecological strategies that enable species to inhabit floodplains, and that these strategies may have evolved in a convergent fashion across the angiosperm phylogeny. However, the lower amount of specialized species in floodplains may be more important in those areas subjected to more intensive and frequent inundations. This combination of contrasting mechanisms of speciation that has acted at a differential extent on the main geological units, and so along the evolutionary history of the Amazon basin, have left their fingerprint on the biogeographic patterns of the Amazon basin.

Table 1.1 Summary of the 412 plots. Total species and individual counts include identified and unidentified morphospecies. Parenthetical values indicate standard deviations. Letters show distinct groups based on Tukey honest significant difference. The mean phylogenetic similarity is based on the assessment of the mean Phylosorensen index estimated for all species present in each plot.

	Craton	Tertiary	Alluvial	Total
Number of plots	69	201	142	412
Number of individuals	4997	12419	9776	27192
Number of species	495	1312	1060	1683
Number of families	52	76	72	84
Mean species per plot	23.27^{C}	32.97^{A}	27.23^{B}	29.37
	(7.2)	(11.8)	(12.7)	(12.1)
Mean stems per plot	75.57^{AB}	73.94^{B}	82.41^{A}	66
	(25.7)	(19.4)	(43.9)	(31.1)
Mean PhyloSorensen	0.39	0.38	0.35	0.36
	(0.10)	(0.07)	(0.08)	(0.07)
Number of indicator species	127	66	48	241

Table 1.2 Mantel and Partial mantel tests results from phylogenetic similarity related with geological and spatial (log-transformed) distances. Analyses were run while including and excluding alluvial plots.

	Variable	R	P
Including Alluvial plots	geo	-0.16	0.001
	spa	-0.25	0.001
	geo, spa	-0.15	0.001
	spa, geo	-0.23	0.001
Excluding Alluvial plots	geo	-0.29	0.001
	spa	-0.39	0.001
	geo, spa	-0.17	0.001
	spa, geo	-0.33	0.001

Table 1.3 Ten most important indicator species per geological unit analyzed. The indicator value is the index from Legendre and Borcard analysis (see Methods). Relative abundance is the proportion of individuals of each species in the associated geological unit. N is the total number of individuals per species.

Geological unit	Indicator species	Family	Indicator value	Relative Abundance	N
Craton	*Clathrotropis glaucophylla*	Fabaceae	0.318	0.95	108
	Leopoldinia piassaba	Arecaceae	0.297	0.89	230
	Macrolobium limbatum	Fabaceae	0.215	0.78	67
	Caraipa longipedicellata	Calophyllaceae	0.209	0.76	142
	Protium divaricatum	Burseraceae	0.207	0.95	69
	Eperua purpurea	Fabaceae	0.202	1	97
	Micrandra sprucei	Euphorbiaceae	0.201	0.99	427
	Protium paniculatum	Burseraceae	0.196	0.9	50
	Mouriri grandiflora	Melastomataceae	0.181	0.96	36
	Glandonia williamsii	Malpighiaceae	0.171	0.98	32
Tertiary	*Socratea exorrhiza*	Arecaceae	0.315	0.76	448
	Pseudolmedia laevis	Moraceae	0.271	0.72	297
	Micropholis guyanensis	Sapotaceae	0.17	0.53	175
	Iryanthera ulei	Myristicaceae	0.163	0.55	263
	Crepidospermum rhoifolium	Burseraceae	0.158	0.8	80
	Dendropanax arboreus	Araliaceae	0.117	0.72	66
	Pseudolmedia laevigata	Moraceae	0.112	0.62	127
	Perebea xanthochyma	Moraceae	0.112	0.62	99
	Protium sagotianum	Burseraceae	0.104	0.74	57
	Protium apiculatum	Burseraceae	0.097	0.8	51
Alluvial	*Euterpe precatoria*	Arecaceae	0.225	0.59	535
	Mauritia flexuosa	Arecaceae	0.146	0.8	374
	Oxandra polyantha	Annonaceae	0.112	0.99	583
	Vatairea guianensis	Fabaceae	0.105	0.83	44
	Licania longistyla	Chrysobalanaceae	0.094	0.95	36
	Virola surinamensis	Myristicaceae	0.079	0.93	64
	Oxandra mediocris	Annonaceae	0.073	0.74	46
	Gustavia pulchra	Lecythidaceae	0.07	1	50
	Couratari oligantha	Lecythidaceae	0.067	0.96	42
	Zygia inaequalis	Fabaceae	0.066	0.72	37

Table 1.4 Test of the degree of habitat association based on a null model of the net relatedness index (NRI) from indicator species for each geological unit assuming a random distribution of each species among habitat types. The significance is based on the distribution of the observed value relative to 1000 iterations of the random model. We used 0.01 as the confidence limit to define indicator species. The NRI was estimated for three different species pool definition: 1: All species reported in our dataset. 2: All species in each geological unit. 3: All indicator species. *: $p < 0.1$; **: $p < 0.05$.

	1	2	3
Craton	1.54*	1.60**	1.56**
Tertiary	-0.76	-0.81	-0.77
Alluvial	-0.91	-0.98	-0.91

Figure 1.1 Map showing the location of study plots. Red points indicate craton plots, blue points indicate alluvial plots and green points indicate Tertiary plots.

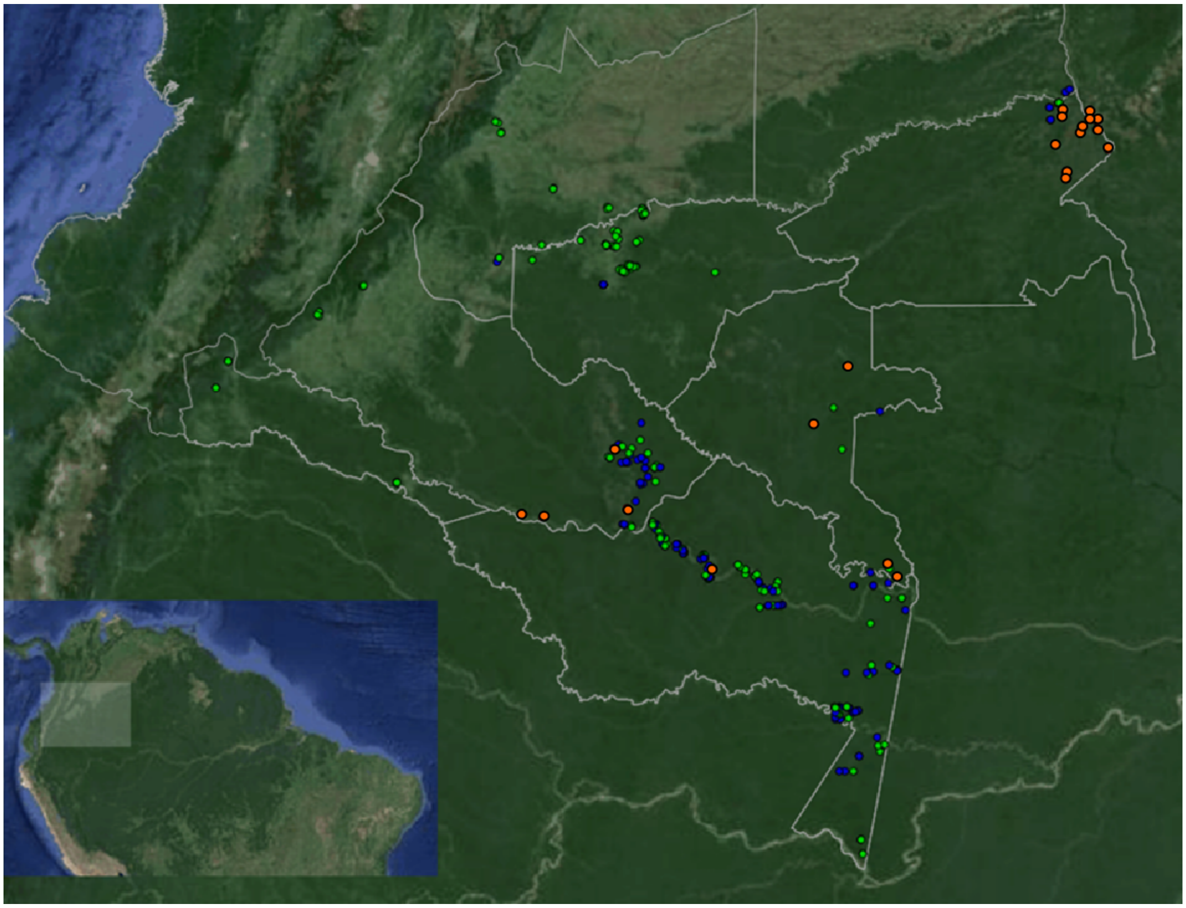


Figure 1.2 A) Analysis of similarity (ANOSIM) showing the trend of the correlation value (R). B) Dendrogram based on phylogenetic similarities between plots. Dotted line indicates significant groups based on a maximization of ANOSIM R statistics. Groups are colored according to the main geological unit associated with each group. Non-conforming geological units are indicated by coloring the terminal branches of the dendrogram. Red: craton plots; blue: alluvial plots; and green: Tertiary plots.

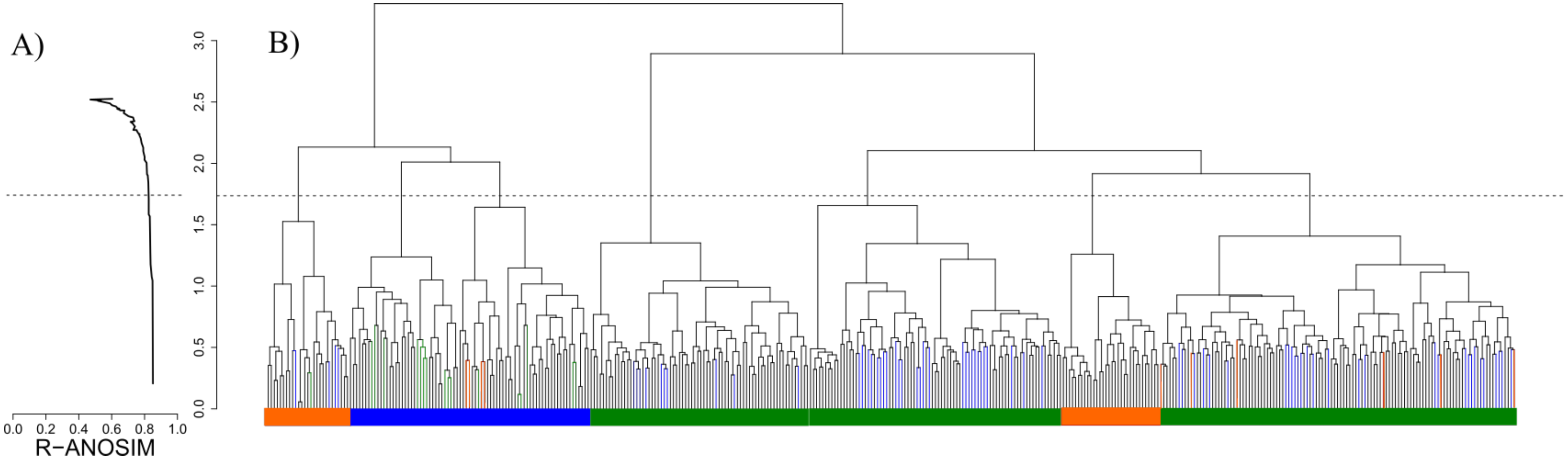
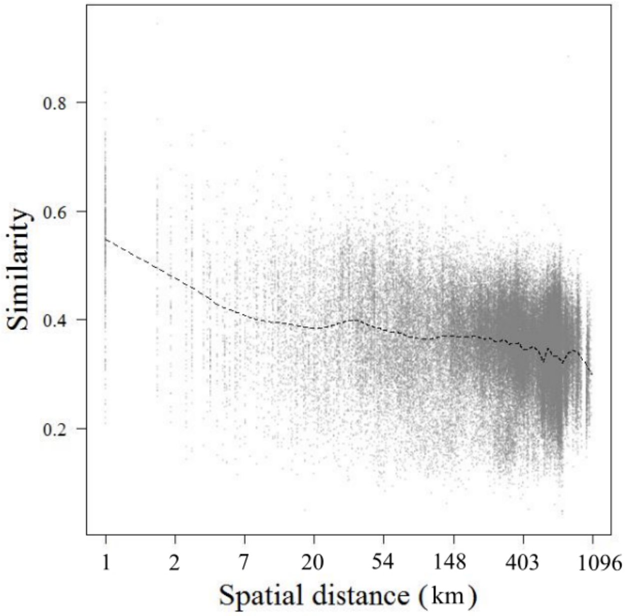


Figure 1.3 Observed relationship between the phylogenetic similarity and geographic distance (log transformed) for pairwise combinations of plots.



2 Chapter 2: Scale-dependent drivers of the phylogenetic structure and similarity of tree communities in northwestern Amazonia

2.1 Introduction

The complexity of tree communities results from the dynamic interplay of evolutionary and ecological processes that operate at different spatial scales (Ricklefs, 2004; Cavender-Bares, et al., 2009). In forests, the interaction between speciation and extinction (Ricklefs, 2004), as well as between habitat specialization and biotic interactions (Stropp et al., 2009), largely shapes the phylogenetic structure and distribution of tree communities at both local and regional scales (Cavender-Bares et al., 2009; Fine & Kembel, 2011). Nonetheless, the capability of lineages to disperse across large geographic areas (Dexter & Chave, 2016) plays a key role connecting distantly separated areas, and thus, adding new species into local communities (Cavender-Bares et al., 2009). The extent to which dispersal can shape tree phylogenetic communities largely depends on the frequency and age of geographic barriers across the region, which can constrain species distributions across different regions and habitats. Overall, the tradeoff between lineage age and the time of formation of physical barriers (e.g. big rivers) often shapes lineage distributions across geographical space (Ruokolainen et al., 2019). For lineages that originated before the presence of geographical barriers, dispersal may be paramount determining the phylogenetic structure of the whole metacommunity at large spatial scales (Dexter et al., 2017). On contrary, geographic barriers may constrain the spatial distribution of young lineages (i.e. those younger than emergence of geographical barriers), increasing the co-occurrence of closely related species within subregions separated by older geographical barriers (Ruokolainen et al., 2019).

The Amazon basin, which roughly harbors 15,000 tree species across ~6 million km2 (ter Steege et al., 2020), is covered by the largest and most diverse reserve of forests on Earth. Particularly in northwestern (NW) Amazonia, where local tree species richness achieves mean values up to 640 woody species per hectare (Duque et al., 2017), the interaction between dispersal and geographical barriers has direct consequences on how speciation and local selection shape the phylogenetic structure of plant communities (Fine & Kembel, 2011). On one hand, if the distribution of lineages is strongly influenced by historical dispersal, a homogeneous phylogenetic structure of tree communities would be expected.

Under that scenario, the phylogenetic diversity of any local community would be a random draw of the phylogenetic composition of the species pool of the metacommunity (Dexter et al., 2017). On the other hand, a dominant pattern of local phylogenetic clustering across regions separated by geographical barriers, such as big rivers, could be due to an extensive in situ speciation and high habitat specialization. This pattern would also lead to high turnover of lineages among regions separated by big rivers across the Amazon (Fine et al., 2014).

The influence of habitat heterogeneity and associated environmental constraints imposed by the endogenous properties of the main habitat types of Amazon forests, such as floodplains, swamps and terra firme (Duivenvoorden, 1995), may also act as local barriers to species distributions. In NW Amazonia, some studies have proposed that floodplains represent phylogenetically specialized habitats that differentiate tree communities at intermediate spatial scales (Aldana et al., 2016). Others, however, described NW Amazonian floodplains as habitats that can easily be invaded by lineages that come from adjacent habitats with contrasting environmental features (Terborgh & Andresen, 1998; Pitman et al., 2014; Cárdenas et al., 2017). Specifically, Terborgh and Andresen (1998) suggested that flooded and unflooded forests within a region would resemble each other more than either flooded or unflooded forests located in distant regions (i.e. dispersal overrides environmental filtering). In swamps, where few species dominate, poor soil drainage and anoxia have been shown to operate as important selection factors that make this habitat type more difficult to invade (ter Steege et al., 2019). The environmental filter associated to swamps proposes that closely related species could tend to coexist more often than expected by chance, which supposes high levels of phylogenetic turnover at both intermediate and large spatial scales. Regarding the largest and most diverse terra firme forests, which are located on sedimentary Tertiary plains, the lack of harsh conditions such as flooding along with an older age of this geological formation, should have facilitated the colonization of lineages with different evolutionary histories (Dexter et al., 2017). Therefore, assuming dispersal across NW Amazon as a frequent historical process, the terra firme forests might be expected to have a high representation of the main phylogenetic lineages at local scales. High phylogenetic diversity at a local scale will in turn increase phylogenetic evenness at the deepest level of the phylogeny, but will decrease phylogenetic turnover (β-diversity) at both subregional (i.e. country) and regional (i.e. NW Amazon) spatial scales. Improving our understanding on the concurrent effect of the spatial scale and the evolutionary history will help identifying the main drivers of tree community assembly in Amazonian forests (Fine & Kembel, 2011; Fine & Baraloto, 2016; Cárdenas et al., 2017).

Virtually all phylogenetic patterns currently known from NW Amazonian forests are based on information from large canopy trees (diameter at breast height -DBH- □ 10 cm) (e.g. Fine & Kembel, 2011; Honorio Coronado et al., 2015; Cárdenas et al., 2017). The inclusion of only canopy trees rules out more than a half of the total number of species expected in a plot (Duque et al., 2017). The main evolutionary features of typical understory lineages (e.g. Rubiaceae, Melastomataceae, and Piperaceae), usually differ from those of dominant canopy lineages (e.g. Fabaceae, Lecythidaceae, Sapotaceae, and Myristicaceae). Overall, small plant size in wet tropical forests has been associated with an increase in both habitat specialization (Duque, et al., 2002; Macia, 2011; Fortunel et al., 2016) and dispersal limitation (Dexter & Chave, 2016). Hence, including small-statured lineages in plot analyses can be expected to increase phylogenetic turnover at intermediate and large spatial scales, compared to analyses only based on large canopy trees. Assessing the extent to which distinct evolutionary and ecological processes shapes the phylogenetic structure and phylogenetic β-diversity of plant communities in relation to plant size at different spatial scales (Swenson et al., 2007), will significantly contribute to our understanding of the main determinants of the diversification and maintenance of the high diversity of Amazonian forests.

In this study, we aimed to assess the extent to which historical dispersal, flooding, geographical barriers, and soil fertility variation determine the phylogenetic structure of tree communities at intermediate and large scales. For this purpose, we used 85 0.1-ha plots (Trees and shrubs with DBH □ 2.5 cm) covering three of the main habitat types (terra firme, floodplains, and swamps) in three subregions of NW Amazonia (located in Colombia, Ecuador, and Peru). The main hypotheses evaluated were: H1) Historical dispersal overcome geographical barriers and played a key role on structuring the phylogenetic assembly of tree communities in NW Amazonia (sensu Dexter et al., 2017). Under this assumption, local phylogenetic community structure of tree communities will be a random draw of the regional species pool, and thus, we expect a low phylogenetic β-diversity regardless of spatial scale and stem size of the included lineages. H2) Geographical barriers have caused an evolutionary separation of ancestral communities imposing limitations to the dispersal of lineages across the region. A separation of subregions by geographical barriers proposes an increase in phylogenetic similarity between forest communities located closer geographically (the Terborgh and Andresen hypothesis), particularly if smaller-statured lineages are incorporated (Dexter & Chave, 2016). If this process is frequent along evolutionary time, this would imply that two communities located in contrasting habitat types but in the same subregion, would be more similar in their phylogenetic composition than two communities located in the same habitat type but in different subregions (Terborgh & Andresen, 1998; Cárdenas et al., 2017). H3) Abiotic environmental heterogeneity shapes the phylogenetic structure across NW Amazonian forests (Fine & Kembel, 2011). Under this scenario, we expect a high phylogenetic clustering in habitats prone to strong environmental filtering (i.e. flooded forests), especially if small-statured lineages are included (Duque et al., 2002; Comita et al., 2007; Fortunel et al., 2016). Two communities located in different subregions but in the same habitat would be more similar in phylogenetic composition than two communities in one subregion but in different habitats.

2.2 Metodology

Study site

The study was conducted in three different subregions in NW Amazonia: Metá-Chiribiquete (hereafter Metá), forming part of the middle Caquetá basin in Colombia; Yasuní, part of the Napo basin in Ecuador; and Ampiyacu, in the Maynas Province of Perú (Figure 2.1). The average temperature in all areas is around 25°C and annual precipitation varies around 3000 mm (based on worldclim database; Hijmans et al., 2005). All months show an average precipitation above 100 mm (Lips & Duivenvoorden, 2001).

Vegetation sampling and identification of botanical vouchers

A total of 85 0.1-ha plots were established: 35 in Metá-Chiribiquete (Duque et al., 2002; Duque, 2004) and 25 in both Yasuní (Romero-Saltos, Valencia, & Macía, 2001) and Ampiyacu (Grández et al., 2001) (Figure 2.1). Plots were rectangular (20 x 50 m) and delimited by compass, tape and stakes from a random starting point respect to topographic conditions (i.e. elevation and slope), with the restriction that the long side of the plot was parallel to the contour line. The plots were located in areas with relatively homogeneous soils and physiognomically uniform forest stands. In each plot, all shrubs, treelets, and trees with DBH ≥ 2.5 cm were numbered and measured with tape. Plots were located in forest that lacked signs of human intervention, except in some swamp plots in the floodplain of the Ampiyacu River in Peru, where few palms had been cut recently to harvest fruits from Mauritia flexuosa L.f. Plots were established at a minimum between-plot distance of 500 m and were mapped with GPS.

In each plot, at least one botanical collection of each morpho-species was collected. The nomenclature of families and genera followed Angiosperm Phylogeny Group version 4 (Chase et al., 2016). Within families or groups of closely allied families, specimens that could not be identified to species because of lack of sufficient diagnostic characteristics, were treated as morpho-species on the basis of simultaneous morphological comparisons with all other specimens in herbarium. Hereafter, we will mostly refer as to species for both morpho-species and species. Botanical identification took place at the herbaria COAH, QCA, QCNE, AMAZ, US MO, NY, and AAU. Vouchers of around 90% of the species and morpho-species described by each independent group were pooled together and compared in MO and AAU between 2000 and 2002.

Habitat type definition

Floodplain forests were those located in areas flooded by river water in times of high river water levels. Because river water levels vary in a 'seasonal' way (depending on local and regional rainfall), these forests can be seen as seasonally flooded forests. If not flooded by river water, soils in floodplain forests were well drained. Swamp forests are forests on soils that were permanently inundated due to poor drainage. These swamp forests could be located in or outside areas subjected to flooding by river water. Terra firme forests occurred in areas not affected by flooding by river water, where soil drainage conditions were good (see Duivenvoorden & Lips (2001) for additional details).

Soil characterization

In the center of each plot, using a soil auger, a soil sample was taken at a depth of 65-75 cm. For analyses, soil samples were dried at temperatures below 40°C, crumbled and passed through a 2-mm sieve. Total contents of Ca, Mg, K, Na, and P were determined by means of atomic emission spectrometry of a subsample of 100-200 mg from the sieved fraction, previously digested in a solution of 48% HF and 2M H2 SO4 (following Lim & Jackson, 1983). Total content of C and N was determined for the sieved fraction using a Carlo Erba 1106 elemental analyser. Soil analyses were done at the soil laboratory of the Institute for Biodiversity and Ecosystem Dynamics of the Universiteit van Amsterdam (for more details see Lips and Duivenvoorden, 2001).

Species pool randomization and hypotheses testing

We used three different species pool, each one associated with one of the main hypotheses stated above, as follows: i) To address the first hypothesis (H1; Dexter et al., 2017), which assumes that historical dispersal overrides geographical barriers (e.g. big rivers) at the entire NW Amazon scale, we applied a standardization procedure based on a null model that randomized the species composition of the entire dataset (85 0.1-ha plots). Hereafter we will refer to this procedure as the Z1 species pool randomization. To draw a null distribution based on 999 replicates, we used an algorithm that retains the species richness within each plot and the relative frequency of species occurrences, and randomize species identity for each community (Kembel et al., 2010). The standardization of the metrics employed to assess phylogenetic structure and turnover (see below) was obtained by subtracting the mean of the null model from the observed value, divided by the standard deviation (Webb, 2000). ii) To address the second hypothesis associated to the effect of geographical barriers as a cause of dispersal limitation (H2; Terborgh & Andersen, 1998), we randomized the species presence-absence across all plots of each subregion (i.e., countries) maintaining species richness and frequency in each plot, applying the same standardization procedure described above. In this way, the habitat effect within each subregion was ruled out in the null communities. Hereafter we will call this procedure the Z2 species pool randomization model. iii) To address the third hypothesis associated to the effect of habitat filtering (H3; Duque et al., 2002; Phillips et al., 2003; Tuomisto et al., 2003), we randomized the species presence-absence among plots located in one and the same habitat type (terra firme, floodplains, and swamps) across all three subregions, applying the same standardization procedure described above. Hence, in this null model the effect of spatial distance between subregions was ruled out within the habitat types. Hereafter it will be called the Z3 species pool randomization.

Growth form and life stage development

In order to assess the effect of plant size, which is associated with growth form (shrub or tree) and the stage of development (juvenile or adult), on determining changes in the magnitude of the responses to either environmental or spatial factors, we divided the dataset in four categories according to a size cut-off threshold, as follows: i) all individuals: includes all trees (juvenile and adults; DBH ≥ 2.5 cm) and shrubs in each plot. Shrub species were classified based on the BIEN database (Engemann et al., 2016). ii) Only trees (excluding shrubs). iii) Adult trees (DBH ≥ 10.0 cm). iv) Juvenile trees (DBH < 10.0 cm). Although only shrubs were also separated as an independent category, the high abundance of zeros and presence of unique species in many plots impeded the capacity to a comprehensive analysis of this growth form.

Local phylogenetic structure of communities

Using all species and morpho-species (hereafter species) that could be assigned to at least family taxonomically, we obtained a phylogenetic tree for our entire dataset based on the Phylomatic repository using the R20120803 backbone phylogenetic tree (Webb & Donoghue, 2005). We calculated the net relatedness index (NRI) and the nearest taxon index (NTI) to analyze the local phylogenetic community structure (Webb, 2000; Webb et al., 2002). The NRI is a standardized average mean phylogenetic distance between all pairs of coexisting species in a site. The NTI is a normalized measure of the average phylogenetic distance between each species and its mean nearest taxon distance. NTI tends to quantify the degree of clustering among terminal taxa, whereas NRI signals clustering at deeper levels in the phylogeny.

We further standardized the NTI and NRI values of each plot by subtracting the mean value of the null communities from the observed plot value and dividing the resulting difference by the standard deviation of the plot values drawn from the null model, following each one of the three species pool definitions (Z1, Z2, and Z3). The standardized values were then multiplied by -1 so that negative values indicated phylogenetic over-dispersion (i.e., more distantly related species than expected under the null model), and positive values phylogenetic clustering (i.e., more closely related species than expected under the null model) (see Methods in the Supplementary Information). These indexes were calculated using the “picante” package (Kembel et al., 2010) in R.

Phylogenetic similarity among communities

We calculated the phylogenetic similarity using the PhyloSorensen index (hereafter PhyloSor). This index estimates the length of the branches on the phylogenetic tree of shared lineages between two samples relative to the sum of the lengths of lineages not shared between the samples (Bryant et al., 2008; see Methods in the Supplementary Information). We standardized the phylogenetic similarity according to each one of the species pool definitions (Z1, Z2, Z3). Positive values of the standardized effect sizes of phylogenetic similarity represent greater divergence than expected by chance, while negative values represent greater similarity than expected by chance. The PhyloSor index was calculated for all possible pairs of plots using the “PhyloMeasures” package (Tsirogiannis & Sandel, 2015) in R.

Environmental and spatial explanatory variables of phylogenetic turnover

To characterize environmental variation, we applied soil cation contents and a binary anoxia factor. The latter factor was built up by an amalgamation of the flooding and swamp effect (anoxia level 1) and a terra firme effect (anoxia level 0). Flooding and swamp effects are considered as an environmental constraint for plant establishment due to the anoxia it creates mainly for the root component of trees. In addition, we used principal component analysis (PCA; Legendre & Legendre, 2012) to reduce soil variation on the seven continuous variables measured in soil samples (C, N, P, Ca, Mg, K and Na). These variables were scaled (i.e., mean equal to zero and variance to one) previous to the PCA because their units have different magnitudes. The two main principal components that explain most of the total variance were used as explanatory variables. PCA allowed interpreting soil variation in terms of a gradient of change across the whole region. This analysis was made using stats package in R.

The spatial variation was assessed by principal coordinates of a neighbor matrices analysis (PCNM; Legendre & Legendre, 2012). PCNM analysis was run using a threshold value or truncated Euclidean distance among the geographic coordinates of the plots, which returns an independent set of orthogonal axes that represent the spatial configuration of plots relative to themselves. As recommended, we only selected PCNM axes with positive eigenvalues that represent spatial associations among plots. The PCNM was run using the default option of the pcnm function in the "vegan" package (Oksanen et al. 2019) in R. In the case of irregular sample designs, the PCNM represent a series of irregular wavelengths. The magnitude of the wavelengths is associated with the grain-size or spatial scale at which the hypothetical or surrogate processes, such as geographical barriers or dispersal, acts. We then used a forward selection procedure to select the PCNM axes that were significantly associated with phylogenetic turnover. The selected PCNM axes, employed as surrogate of spatially structured processes at different scales, were PCNM1, PCNM2, PCNM3, PCNM4, and PCNM5 (Fig. S1). Large-scale processes, such as historical dispersal or geographic barriers, can be associated to the first two PCNM axes. On contrary, the effect of dispersal limitation tends to be better described by intermediate to small wavelengths, such as PCNM 3, 4, and 5.

Environmental and spatial determinants of phylogenetic similarity

We used a non-metric multidimensional scaling (NMDS) ordination procedure on the observed and standardized effect sizes of phylogenetic similarity to visualize differences among habitat types and subregions. To assess environmental and spatial effects on phylogenetic similarity, we used a distance-based redundancy analysis (dbRDA; Legendre & Legendre, 2012) to partition the amount of variation of phylogenetic similarity explained by each set of explanatory variables. The dbRDA is a method to carry out constrained ordinations using non-Euclidean distance measures. In the dbRDA, the standardized PhyloSor matrix was internally transformed to principal coordinates, which were then used in a redundancy analysis (RDA). Prior to the analysis, the standardized PhyloSor values were converted to positive distances by scaling them between 0 and 1. Then, we ran dbRDA with only the spatial variables and applied a forward selection procedure to identify significant PCNM axis. The same procedure was applied for the soil PCA axes and anoxia. Finally, we ran a dbRDA including significant variables from both spatial and environmental variables, and we carried out a variation partitioning to evaluate the relative importance of each set of explanatory factors (Legendre & Legendre, 2012). For all models, we reported our results in terms of the adjusted R2 values. The analyses were run using the "vegan" package (Oksanen et al., 2019) in R.

2.3 **Results**

Structural patterns

The 30,150 individuals used in the analyses belonged to 2,266 morpho-species (1,859 fully identified), 592 genera, and 136 families. In total, 29,376 individuals were classified as trees (2,127 morphospecies) and 774 were shrubs (139 morphospecies). Plot species richness averaged 106 ± 46 (1 s.d.) and average stem density was 333 ± 145 individuals per plot (Table S2.1).

The proportion of significant NRI and NTI values relative to the null expectation varied among habitat types and subregions (Figure 2.2). NRI was not significantly different from the null expectations for the entire region, subregion, or habitat type (NRIZ1 = 0.35 ± 0.86; NRIZ2 = -0.19 ± 0.91; NRIZ3 = -0.27 ± 0.88). In contrast, around 40% of the NTI values were significantly clustered at each scale (NTIZ1 = 1.36 ± 0.81; NTIZ2 = 1.01 ± 0.91; NTIZ3 = 1.07 ± 0.90). NRI and NTI patterns were not significantly different across life stage of development (Figure 2.2, Figure S2.2-S2.3).

Pattern and drivers of phylogenetic turnover

The first PCA axis of the soil variables explained 42% and was positively correlated to total bases (Ca, Mg, K, and Na). The second PCA axis explained 31% and was negatively correlated to C, N, and P concentrations (Table S2.2). PCNM1 was highly correlated with longitude ($r = 0.98$, $p < 0.001$). PCNM2 and PCNM3 were correlated to latitude ($r = 0.80$, $p < 0.001$; $r = -0.47$, $p < 0.001$, respectively), while the others two PCNM axes represented smaller scales of spatial variability.

The scores of the first two NMDS axes (calculated on the basis of the PhyloSor index for all individuals using each of the three species pool definitions (Z1, Z2, Z3; Fig. 3 and Fig S4-S5)) were highly correlated (Figure S2.6), suggesting that the principal patterns of phylogenetic similarity were independent from the species pool sizes. The mean standardized effect size of phylogenetic similarity was negative for all randomization procedures (PhyloSorZ1 = -0.24 ± 1.24; PhyloSorZ2 = -1.40 ± 3.13; PhyloSorZ3 = -1.42 ± 3.17). The PhyloSorZ1 mean was distinctly higher than PhyloSorZ2 and PhyloSorZ3 ($F = 284.9$, $p < 0.001$; Figure S2.7). Overall, the PhyloSorZ1 means were lower in absolute values (closer to zero), which indicate a lower phylogenetic similarity between the plots if tested against null model Z1 than that compared to the other two null models (Z2 and Z3). In other words, using the whole of NW Amazonia as a reference (Z1) tends to drive the phylogenetic composition of plots to being random samples of the metacommunity.

Under the Z1 species pool definition, both anoxia and large-scale spatial processes (captured by PCNM1 and PCNM2) explained most of the overall variation of the standardized phylogenetic similarity for the four developmental stage categories. Under the Z2 and Z3 species pool definition, however, environmental factors, such as anoxia and PCA1 soils, became more important than spatially structured processes as drivers of the changes in the standardized phylogenetic similarity (Table 2.1). The total explained variation was always higher under the Z1 species pool definition, and lower for adult trees (DBH ≥ 10 cm) independent of the species pool definition applied. When we analyzed juvenile and adult trees, under the Z2 and Z3 species pool definition, the spatially structured processes were negligible in the former but more important in the later. Further, the inclusion of shrubs as an independent category of juvenile trees had a negligible effect in the RDA analysis (Table 2.1).

2.4 Discussion

Patterns of local phylogenetic structure in NW Amazonia

Our findings identify an important pattern of scale-dependency in the local phylogenetic structure in the tree communities of NW Amazonia. We found striking differences in the patterns obtained from NRI and NTI, which measures deep and tip phylogenetic levels, respectively. Since NRI emphasize patterns at deep levels of the phylogeny, the random NRI patterns we observed at the plot scale indicate high lability in ancestral traits in the NW Amazon region across habitat types. Therefore, at least at the deepest part of the phylogeny, we found support to our first hypothesis (H1), i.e. that local samples represent random draws from the regional phylogeny (Dexter et al., 2017). Contrary to NRI, NTI results showed an important proportion (3% to 41%) of plots were phylogenetically clustered at both the subregion and habitat type level. At both of these smaller spatial scales, the number of plots with a clustered phylogenetic tip-level pattern increased along with the size of the species pool (Figure 2.2). In other words, when we applied the Z1 randomization procedure, the removal of the sampling effect from the whole phylogeny at the scale of the entire NW Amazonia increased the capability to detect closely related species coexisting in a particular site. Thus, the NTI clustering trend suggests more recent events, such as the relatively recent change in the channel location of big and medium size rivers (Ruokolainen et al., 2019), have acted as effective geographic barriers and constrained plant dispersal. These findings also suggest that small peripheral populations isolated by distance could act as an important source of speciation (Hubbell, 2001). Taken together, our results indicate that at local scales, particular lineages are overrepresented in term of species and individuals, probably as result of selective pressures or competitive dominance. However, across habitats and at larger spatial scales, this pattern dissolves as other lineages with origin in different habitat types (e.g. terra firme) are also represented (Fine & Baraloto, 2016; Cárdenas et al., 2017).

Phylogenetic similarity across NW Amazonia

The phylogenetic similarity of tree communities in NW Amazonia was primarily accounted for by anoxia and large scale spatially structured processes, respectively. The separation of the individuals inhabiting the forest understory in different growth forms, such as juvenile trees and shrubs, did not substantially changed the relative importance of either environmental or spatial factors as drivers of the phylogenetic similarity pattern of tree communities (Table 2.1). However, both the spatial scale (randomization method applied) and the stage of development of plant communities, led the substantial differences in the extent to which either environment or dispersal shaped the phylogenetic assembly of tree lineages. Our results are similar to other studies that assessed the effect of habitat heterogeneity and spatial distance on the phylogenetic structure of Amazon tree communities (Fine & Kembel, 2011). However, to our knowledge, this is the first study that evaluates the effect of the stage of development on the phylogenetic structure of Amazon tree communities.

At the spatial scale of the entire NW Amazonia (Z1), the importance of anoxia and large scale spatially processes were almost the same, which rejects our expectations (H2) of an overwhelming influence of dispersal rather than environmental constraints on determining the phylogenetic similarity of tree lineages (sensu Terborgh & Andersen, 1998). These findings contradict the idea of Amazonian floodplains as an forest type that is easy to invade by lineages from adjacent habitat types (Terborgh & Andresen, 1998; Pitman et al., 2014; Cárdenas et al., 2017). The need of species to establish and survive in conditions of severe soil anoxia may have triggered the development of particular traits, such as a thick cuticle and thick outer epidermal walls to avoid fast rot when submerged (Parolin et al., 2004).

However, in NW Amazonia, the highly dynamic geological history have promoted historical and continuous river readjustments and formation of terrestrial mosaics (Hoorn et al., 2010) that seemed to have been overcome by some lineages but not for others, explaining the paired relative importance of both anoxia and historical dispersal at a regional scale.

The reduction of the species pool size (Z2 and Z3 species pool definitions) drove the relative importance of spatial processes to almost negligible values and pointed to anoxia as the overriding factor structuring the phylogenetic similarity patterns of tree communities at both the country and habitat type level (Table 2.1). These findings support our third hypothesis (H3), which suggest the effect of environmental filtering as the overriding mechanism that structures Amazon tree communities, but only when the overall size of the regional species pool is reduced (e.g. intermediate scales) (Duque et al. 2002, Phillips et al., 2003). Although more subtly, soil fertility also increased in relative importance in shaping phylogenetic similarity among communities when the species pool size was reduced (Table 2.1). Therefore, at more local and intermediate (i.e. subregion) scales, soil conditions also played a small but significant role on differentiating tree lineage assembly across NW Amazon (Phillips et al. 2003).

Under the Z2 and Z3 species pool definition, the inclusion of small trees (i.e., shrubs and juveniles) had a significant effect on determining the relative important of spatial processes on shaping the phylogenetic similarity of either adult or small trees. However, contrary to our expectations, environmental variation had a stronger effect on phylogenetic differentiation on adult than small trees. Although under the Z2 and Z3 species pool definition the total amount of explained variation of adult trees decreased, the proportional importance of the spatial variables became almost negligible (Table 2.1). Therefore, environmental filtering, primarily due to the limited tolerance to flooding in the earlier developmental stages of tree species and the local adaptation to soil conditions, played a key role on shaping the phylogenetic assembly of adult trees across NW Amazonia. When small trees were included (under Z2 and Z3), the relative importance of the spatial variables almost four-folded that of adult trees (Table 2.1). This finding suggests that dispersal limitation of small trees, probably many small-statured lineages (e.g. Rubiaceae), is also driving community assembly at local to intermediate scales (Dexter & Chave, 2016). Then, comparisons of lineages of different plant size can inform about the effect of dispersal on community assembly and diversification of Amazon tree lineages (Dexter & Chave, 2016).

Our results also highlight that phylogenetic differentiation within habitat types are significantly affected by variation in soil fertility throughout NW Amazonia (Tuomisto et al., 2003). For example, the within habitat type average of key elements for species distributions, such as P (Condit et al., 2013), was almost as high as that observed between habitat types within the same subregion. In the case of P in particular, the concentration of this element in Peruvian terra firme forests (11.32 ± 10.3 ppm) was two-fold higher that that in the same habitat type in Colombia (5.2 ± 1.5 ppm). A similar pattern was found for C, N, and other key soil elements (see Table S2.1). Such a high heterogeneity in soil cation concentrations support a need of including soils fertility as an explanatory variable in the models that aim to explain an understand the evolutionary patterns of tree species in Amazon forests. Therefore, we call for caution for assuming homogeneous features in soil properties within a similar habitat type at large geographical scales (Fine & Kembel, 2011; Cárdenas et al., 2017). Likewise, understanding the trade-off between soils fertility and biotic interactions at finer spatial scales should help to unravel the mechanisms that structure the evolutionary history and distribution of tree lineages in tropical forests (Baldeck et al., 2016).

In conclusion, the extent to which either environmental or spatial features shapes the local phylogenetic structure and phylogenetic similarity of tree communities in Amazon forests is scale dependent. Overall, the relative importance of environmental factors increases with both decreases in spatial scale and the size of the species pool. We emphasize on the need of additional studies that focus on the trade-off between biotic interactions and soil fertility at finer spatial scales than those employed here. In the face of global environmental changes, understanding how evolutionary history can shape the phylogenetic structure and turnover of Amazonian plant communities will identify opportunities to preserve this highly diverse but threatened ecosystem.

Table 2.1 Variation partitioning based on distance-based redundancy analysis (dbRDA) applied to the PhyloSor metric and their standardized values by the three species pool randomizations (Z1 = all data; Z2 = restricted by subregion; Z3 = restricted by habitat type) and four community definitions (all individuals of trees and shrubs, only trees, adult trees and juvenile trees). In the left panel we have the overall percentages of phylogenetic β-diversity explained by each set of variables associated to either the environmental or the spatial hypothesis. Env = pure environment; Spa= pure spatial; Over = spatially structured environmental variation. In the right panel we have the percentage of phylogenetic β-diversity explained by each one of the explanatory variables selected by the forward selection procedure. Anoxia: Flooding regimen. PCA1 and PCA2: axis derived from the Principal Component Analysis applied to soil variables. PCNM1, PCNM2, PCNM3, PCNM4 and PCNM5 are the spatial variables selected from the principal coordinates on neighbor matrices analysis.

Community	Null model	Env	Over	Spa	Total	Anoxia	PCA1	PCA2	PCNM1	PCNM2	PCNM3	PCNM4	PCNM5
Trees & Shrubs	Z1	22.51	6.18	27.74	56.43	21.49	2.84	--	16.69	6.89	3.69	3.41	1.43
	Z2	38.08	2.55	5.28	45.92	29.59	5.55	2.09	1.79	2.37	--	2.76	1.77
	Z3	36.33	2.40	5.77	44.46	25.83	5.43	1.79	4.08	2.70	1.26	3.67	1.65
Only trees	Z1	23.58	6.44	28.61	58.64	23.05	3.11	--	17.50	7.31	4.06	3.62	0.00
	Z2	38.80	2.58	5.34	46.73	30.22	5.47	1.95	2.07	2.45	--	2.76	1.81
	Z3	36.64	2.31	5.69	44.66	26.07	4.74	1.55	3.76	2.42	1.25	3.31	1.56
Large trees	Z1	24.12	5.85	22.44	52.31	24.14	2.02	--	15.61	5.09	2.98	2.47	--
(>10cm DBH)	Z2	30.62	1.62	1.23	33.47	25.98	2.96	1.69	1.33	--	--	1.51	--
	Z3	33.96	2.37	-1.39	34.95	25.11	3.02	1.56	1.94	1.13	--	2.19	--
Small trees	Z1	22.72	6.08	27.77	56.57	21.46	2.91	--	15.43	7.22	4.29	3.51	1.74
(<10cm DBH)	Z2	31.49	2.47	10.84	44.82	28.18	2.03	--	7.08	5.60	--	2.81	2.12
	Z3	31.63	1.90	10.51	44.06	25.30	--	--	6.71	4.41	--	3.76	1.79

Figure 2.1 Location of the three subregions of NW Amazonia recognized in this study. The rectangles denote the area where the plots were established.

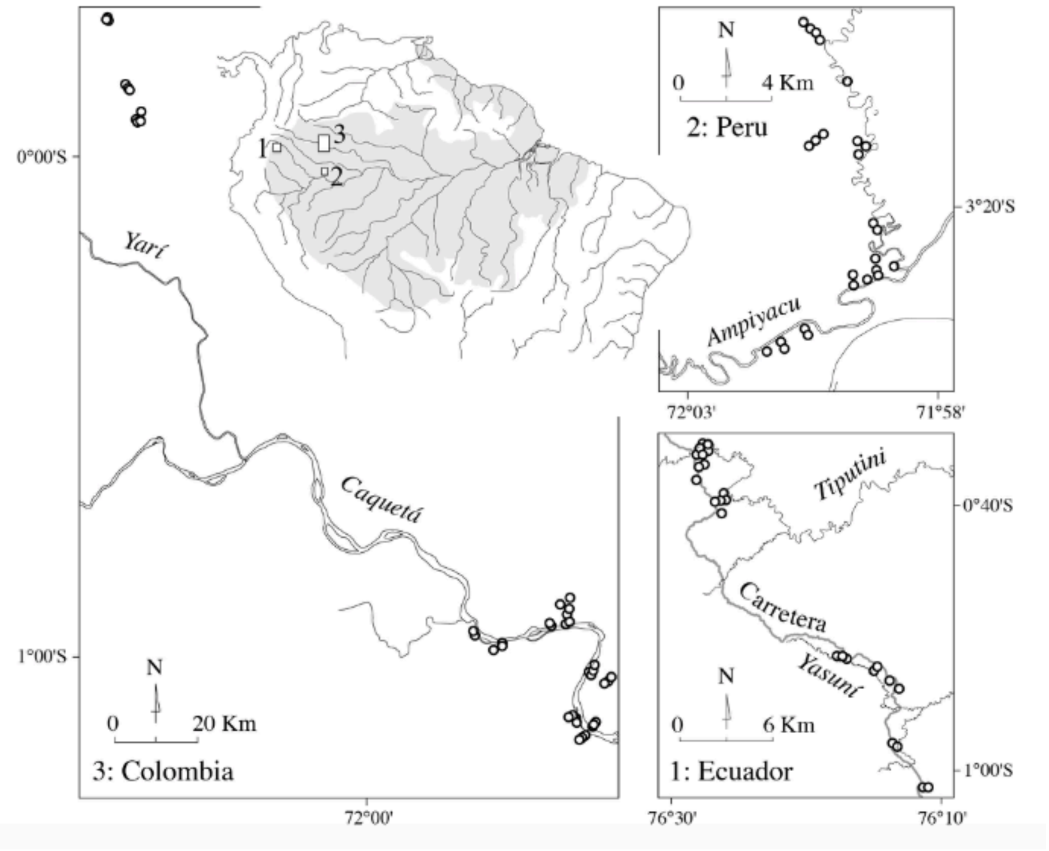


Figure 2.2 Proportion of plots that showed significant clustering (underdispersion) or evenness (overdispersion) in NRI and NTI values considering all individuals (DBH ≥ 2.5 cm) under the three species pool randomization models (Z1 = All data; Z2 = Restricted by subregion (effect of habitat type removed within subregion); Z3 = Restricted by habitat type (effect of subregion removed within habitat type)). Light gray represents the proportion of plots significantly clustered. Dark gray bar represents the proportion of plots with a significant evenness pattern. White bars represent the plots with either NRI or NTI values that were not significantly different from random. FP = Flood plain. SW = Swamp. TF = Terra Firme. Metá-Chiribiquete = 1. Yasuní = 2. Ampiyacu = 3.

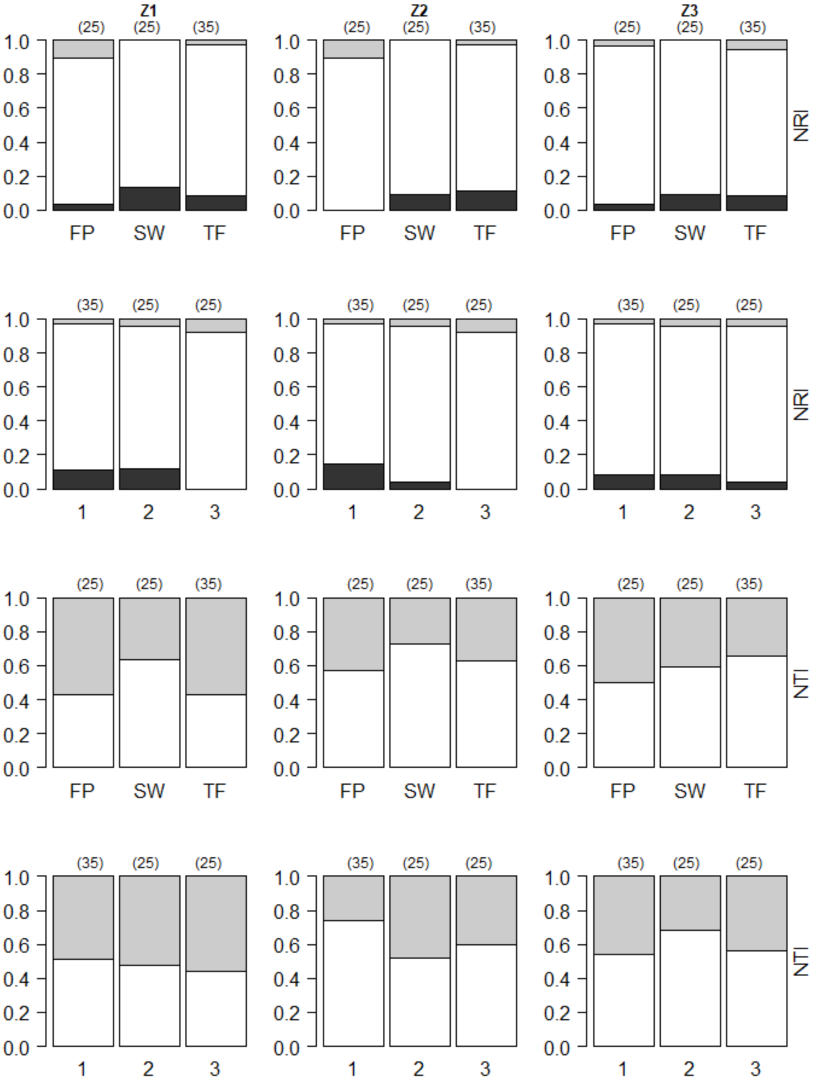
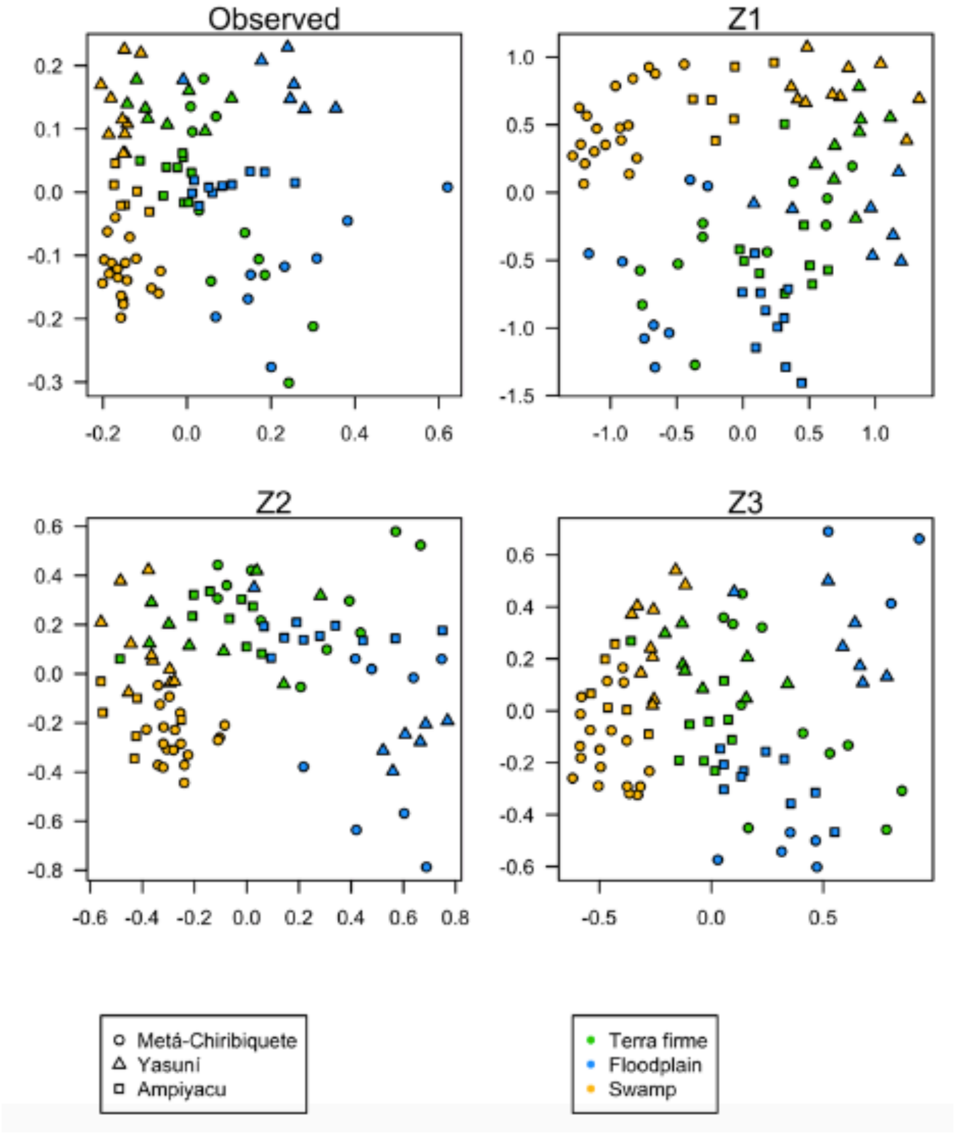


Figure 2.3 Non-metric multidimensional scaling ordination analysis of phylogenetic similarity tree communities including all individuals under different scenarios of species pools used in this study. Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by habitat type. Symbols denoted subregions: circle: Metá-Chiribiquete; triangle: Yasuní; square: Ampiyacu. Colors denoted habitat types: Flood plains: orange; Swamp: blue; Terra Firme: green.



3 Chapter 3: The influence of historical dispersal on the phylogenetic structure of tree communities in the tropical Andes

3.1 Introduction

Explanations for decreases of species richness with elevation often invoke different ecological and evolutionary hypotheses related to the variation in temperature (Laiolo et al. 2018) that mirror the latitudinal gradient (Stevens 1992). Among those, the Tropical Niche Conservatism (TNC) hypothesis emerges as one of the most common ideas explaining the decline in species richness from the tropics towards the poles. According to the TNC, tropical species originated in warm conditions (i.e., tropical lowlands) and mostly retain their ancestral climatic niches (Wiens & Donoghue 2004, Kerkhoff et al. 2014). As such, the TNC implies that the evolution of adaptations for cold tolerance is infrequent, which means that only a small number of lineages have developed traits required to inhabit thermal environments other than the wet tropical lowlands. This hypothesis can also be applied to elevation gradients in mountain ecosystems, such as the tropical Andes. In the younger Andean highlands, a lower time for species establishment may explain the lower richness of lineages as well as the increased tendency for closely related species to co-exist (Graham et al. 2009). The tendency for closely related species to co-exist is referred to as phylogenetic clustering (Webb et al. 2002), and the increase in clustering along elevational gradients is predicted by the TNC.

In contrast to the TNC, the Out of The Tropics (OTT) hypothesis proposes that extra-tropical regions have been colonized multiple times for different lineages since the Eocene when the extent of tropical climates was drastically reduced (Jablonski. et al. 2006, Qian & Ricklefs 2016). Thus, according to the OTT, the adaptations of species to colder high-elevation environments in tropical mountains may have also originated from multiple immigrations of different lineages from outside of the tropics. This process of historical dispersal through time is sometimes used to explain the frequent occurrence of extra-tropical lineages in high-elevation tropical montane systems (Gentry 1988, Hughes & Eastwood 2006). In tropical highlands, the co-existence of lineages with contrasting zones of origin (tropical lowlands versus extratropical) and evolutionary history will lead to a phylogenetically overdispersed distributional pattern (González-Caro et al. 2014, Qian & Ricklefs 2016).

Changes in phylogenetic composition (hereafter phylogenetic β-diversity) along elevational gradients will depend on which pattern of immigration prevails. Overall, phylogenetic β-diversity may be partitioned into nestedness and replacement (Baselga 2010). Phylogenetic nestedness is primarily due to changes in phylogenetic diversity and species richness. Under the TNC hypothesis, we expect nestedness to be the overriding pattern due to most species originating in the tropical lowlands and only a subset of lineages dispersing into highlands. In contrast, under the OTT hypothesis, we expect the replacement of major lineages (Leprieur et al. 2012) to be the overriding pattern of phylogenetic β-diversity due to different elevational bands being occupied by lineages with divergent evolutionary histories. The idea that larger turnover of lineages supports the OTT hypothesis comes from the immigration of lineages from extra-tropical regions is believed to have been more frequent in the highlands than in the lowlands (Gentry 1982; van der Hammen & Cleef 1983; Hooghiesmtra et al. 2006). The addition of clades associated exclusively with the highlands, such as Fagaceae, Cunnoniaceae, Proteaceae, and others, will increase phylogenetic turnover, in contrast to the expectation of the phylogenetic nestedness predicted by the TNC hypothesis. Understanding the extent to which either nestedness of turnover explains patterns of phylogenetic β-diversity along elevational gradients will help us to infer how historic dispersal of lineages shapes the current composition of tropical elevational gradients.

In this study, we integrate presence-absence data of tree species tallied in plots with phylogenetic information to assess the extent to which the TNC vs. OTT hypotheses shape phylogenetic structure and changes in phylogenetic composition of forest communities along an elevational gradient in the Andean-Amazon flank of Colombia. Our study had two specific goals. The first goal was first to assess changes of phylogenetic clustering of tree communities along the elevational gradient. We predicted a priori that phylogenetic clustering will increase with elevation due to changes in temperature, in accord with the TNC hypothesis. In contrast, if the OTT hypothesis prevails, phylogenetic overdispersion should increase with elevation. The second goal of our study was to assess the extent to which either phylogenetic nestedness vs. replacement shapes phylogenetic β-diversity along the elevational gradient. The expectation under the TNC hypothesis is that the species occurring in the highlands are derived from lowland lineages. This leads to the a priori predictions that highland communities will be phylogenetic subsets of lowland communities and that phylogenetic nestedness will be the main process determining the phylogenetic β-diversity along the elevational gradient. In contrast, the OTT hypothesis predicts that the multiple extratropical colonization events will increase the phylogenetic diversity in the highlands, leading to phylogenetic replacement along the elevational gradient. To test this hypothesis, we assess the prevalence of extra-tropical affiliated lineages in highland vs. lowlands tree communities.

3.2 Metodology

Study area

The study was conducted in the Putumayo and Nariño departments of Colombia. Mean annual temperatures in the region range from 10 °C to 26 °C, and mean total annual precipitation ranges from 2,301 to 2,968 mm. The study site is located in a remote region that is largely inaccessible and covered with mature forest with non-signs of recent or past human disturbances. We established 20 0.25 ha (5 ha total) permanent plots in the study area distributed along an elevational gradient from 750 to 2,802 m asl. The plots were randomly located in pairs along the trail at intervals of approximately 250 m elevation, following the sampling design used in other studies of tropical montane forests (Feeley et al. 2011). In each sample elevation, we set up two plots of 0.25 ha (50 m × 50 m), except in the two lowest elevations where we established three plots at each elevation (Figure S3.1).

Plot surveys

The establishment and delimitation of the plots were carried out following standard protocols using a precision compass, clinometer, and tape. The coordinates and elevation were recorded at each plot corner. In the field, plots were delimited with PVC tubes in the corners and polypropylene yellow ropes on the perimeter and parallel lines every 10 m; slope correction was applied so that plots sample 0.25 aerial hectares. Within each plot, all woody plants, including palms, with a diameter at breast height (DBH; measured at 1.3 m above the ground) greater than or equal to 10 cm, were measured, mapped and identified. Botanical samples were determined, and vouchers stored in the Herbario Amazónico Colombiano (COAH).

Phylogenetic data

We registered a total of 3,038 individuals excluding tree ferns. We were only able to assign taxonomic identification at any level to 2,853 individuals; 75.5% of these individuals were identified to species, 23% to genus and 1.5% to family level. The 6.1% of individuals without any taxonomic identification were excluded from subsequent phylogenetic analysis. From the combined species list, we created a phylogenetic tree using Phylomatic version 3.0 (Webb & Donoghue 2005), employing the phylogenetic backbone proposed by Zanne et al. (2014) as a reference tree. The resulting phylogenetic tree has 55% of polytomies. Individuals that were identified at genus or family levels were included as polytomies.

Species richness and phylogenetic structure

We estimated the species richness employing a rarefaction analysis to compare among them due to differences in the number of individuals. We used 100 individuals as the cutoff to draw 1000 random species number for each plot and estimate the confidence intervals.

We calculated two common indexes to quantify the phylogenetic dispersion between co-occurring species in each plot (Webb 2000, Webb et al. 2002): the net relatedness index (NRI) and the nearest taxon index (NTI). The NRI is a normalized average mean phylogenetic distance (MPD) between all pairs of coexisting species in a site and is calculated as:

$$NRI_{sample} = -1\ x\ \frac{MPD_{sample} - MPD_{null}}{sd(MPD_{null})}$$

The NTI is a standardized measure of the average phylogenetic distance between each species and their mean nearest taxon distance (MNTD), NTI, is calculated as:

$$NTI_{sample} = -1\ x\ \frac{MNTD_{sample} - MNTD_{null}}{sd(MNTD_{null})}$$

NTI quantifies the degree of clustering among terminal taxa and is independent of the phylogenetic clustering at deeper levels (NRI).

For the two indexes, we applied a standardization procedure based on a null model that randomized the species composition based on the entire dataset. Specifically, we used the independent-swap algorithm to draw a null distribution based on 999 replicates, which retains the species richness within each plot and the relative frequency of species occurrences, but changes species co-occurrences. For each randomization, the phylogenetic distances between the species assigned to each community were calculated. We then standardized the NTI and NRI by subtracting the random mean from the observed value and dividing by the standard deviation of the values drawn from the null model. The standardized values were then multiplied by -1 so that negative values of each index indicate phylogenetic over-dispersion (i.e., more distantly related species than expected under null model), and positive values indicate phylogenetic clustering (i.e., more closely related species than expected under the null model). These indexes were calculated using the package "picante" in R (Kembel et al., 2010).

Phylogenetic betadiversity

We used a PhyloSorensen index (hereafter PhyloSor) to construct a phylogenetic β-diversity matrix of the 20 plots based on the presence-absence of species. This index estimates the length of the branches of shared lineages between two sites relative to the sum of the lengths of lineages not shared between the sites (Bryant et al. 2008). The PhyloSor index is calculated as:

$$PhyloSorensen\ ij = \frac{BL\ ij}{0.5(BL\ i + BL\ j)}$$

where BLij is the sum of the length of the branches shared between sites i and j, and BLi and BLj is the sum of the length of the branches of each i and j, respectively. Subsequently, a null model distribution was drawn, and a similar standardization procedure as described above was applied. Positive values represent greater similarity than expected by chance, while negative values represent greater divergence than expected by chance. This metric estimates the replacement of lineages towards the terminal portions of the phylogeny; however, our phylogeny has a high percentage of polytomies (55%) at the shallow level. The PhyloSor index was calculated for all possible pairs of plots using the "PhyloMeasures" package in R (Tsirogiannis & Sandel 2015).

The PhyloSor index was decomposed based on the additive framework proposed by Baselga, (2010) and adjusted to the phylogenetic compositional change context proposed by Leprieur et al., (2012) to distinguish between phylogenetic diversity differences (nestedness) and phylogenetic replacement. The additive framework of PhyloSor controls for the differences in the phylogenetic diversity between two sites. These differences are the result of the extirpation of an entire lineage resulting in nestedness or the replacement of one lineage for other part of the phylogenetic tree. The differences in phylogenetic diversity are the result of contrasting diversification or immigration histories between the highlands and lowlands in our study system.

Finally, we employed a null model to control for the effect of the regional species pool on the observed phylogenetic β-diversity indexes. The null distributions were drawn from 999 randomizations of the presence-absence matrix and the estimation of each index. Then, we standardized the observed values relative to the null distribution. Based on these analyses, we obtained six phylogenetic β-diversity indexes: PhyloSorTotal, PhyloSorRep, PhyloSorNes and their standardized effect sizes.

Statistical analyses

We assessed the effects of elevation on NRI, NTI, and rarefied species richness to 100 individuals (SR) using bivariate Pearson correlation analysis. We expect a positive correlation between elevation and NRI and NTI, and a negative relationship between elevation and SR when TNC is the predominant assembly mechanism; we expect the opposite patterns when the OTT processes prevail (i.e., negative relationships between elevation and NRI and NTI, but still a negative relationship between elevation and SR). In addition, to looking at linear relationships, we used a quadratic model to assess potential non-linear relationships between SR and elevation.

We next employed a Mantel correlation test to analyze the relative importance of elevation in explaining the variation of the six previously described phylogenetic β-diversity indexes. We expected that the nestedness component of phylogenetic β-diversity will be strongly correlated with elevational differences when TNC is the predominant pattern.

We performed a sensitivity analyses to assess the effect of the resolution of the phylogenetic tree on the Mantel correlations. We used 1000 draws of hypothetical fully resolved dichotomies of the employed phylogenetic tree using the PDcalc R package (Rangel et al. 2015). Then, we calculate the correlation between the PhyloSor indexes described above and elevation. We used the 1000 bootstrapped correlation indexes to estimate the confidence intervals (quantiles 0.05 and 0.95) of the Mantel correlation.

Finally, we tested the hypothesis that temperate-affiliated lineages will be more common in highlands tree communities than in the lowlands by assigning a hypothetical origin to each family following Heywood et al. (2007). For each family, we then estimated the mean elevation range along the elevational gradient and used a t-test analysis to assess for differences in the mean elevation range between tropical and temperate affiliated species.

3.3 **Results**

The 2853 individuals used in the analyses belonged to 464 species, 205 genera, and 71 families (Table 3.1). The three families with the highest number of genera were Rubiaceae (23 genera), Fabaceae (13), and Lauraceae (11). 31 families (44.4%) had only one genus, and eleven families (15.2%) had two genera. The genera with the highest number of species were Miconia (27 spp.), Inga (24), Ocotea (10), and Pouteria (9). The most abundant species were Piptocoma discolor (60 individuals), Wettinia maynensis (59), Clusia sp. 3 (58), and Clethra fagifolia (45). We found an average species richness (SR) of 51 ± 14.2 (rarefied average SR100 = 40.8 ± 9.9) species per each 0.25 ha plot and an average stem density of 607.6 ± 117.9 individuals per hectare. The total number of individuals and species richness per plot were not significantly correlated ($r = 0.32$, $P = 0.17$).

Although the rarefied species richness showed a linear and negative significant correlation with elevation ($r = -0.72$, $P < 0.001$; R2 = 0.52, AIC= 139.1), the most parsimonious model was a unimodal model with a peak in richness at approximately 1600 m asl (R2 = 0.68, AIC = 132.6) (Figure 3.1A). The net relatedness index (NRI) and the nearest taxon index (NTI) were both significantly correlated with elevation (rNTI = -0.73, $P < 0.001$; rNRI = -0.60, $P < 0.001$). NRI averaged -0.07 ± 1.00 and NTI averaged in -0.36 ± 1.11. In addition, the species assemblages in plots above 2000 m asl were consistently phylogenetically overdispersed and significantly different from the null model mainly at the most extreme elevations (Figure 3.1B and 3.1C).

Phylogenetic β-diversity was negatively associated with differences in elevation in all cases (Figure 3.2). In other words, communities separated by large elevational gaps were more dissimilar in composition (greater β-diversity) than plots at nearby elevations. The observed and standardized effect size of PhyloSorTotal showed the highest Mantel correlation value with differences in elevation (r = -0.82, P <0.001; r = -0.72, P <0.001, respectively). The standardized effect size of PhyloSorNes phylogenetic nestedness component had the lowest correlation with elevation differences and showed a tendency to have low values throughout the entire elevational gradient (Figure 3.2F). The latter finding indicates a weak contribution of nestedness in structuring phylogenetic β-diversity in these communities. In contrast, the observed and standardized PhyloSorRep phylogenetic replacement had a higher correlation with elevation differences in a similar fashion that PhyloSorTotal (Figure 3.2B and 3.2E). According to the sensitivity analysis, the Mantel correlation between the observed PhyloSor values and elevation was not significantly different from those calculated from randomly resolved polytomies of the observed phylogenetic tree (Figure 3.3).

Finally, the tropical-affiliated families occurred at significantly lower average elevations than temperate-affiliated families (t = -21.18, P < 0.001; Figure S3.2).

3.4 Discussion

In contrast to the predictions of the TNC hypothesis, phylogenetic clustering does not increase with elevation. This suggests that multiple dispersal events through time have defined the biogeographic connections among tropical and extratropical regions and lineages (Gentry 1982, Qian & Ricklefs 2016). Likewise, the observed changes of phylogenetic β-diversity along this elevational gradient indicate low overlap of the major clades between lowland and highland tree communities. Low species richness and phylogenetic overdispersion in the highlands suggests limited but effective immigration of extra-tropical linages adapted to the colder conditions that occur at high elevations (Hughes & Eastwood 2006). In other words, highland Andean tree communities comprise a mixture of species originating from the lowland-tropics and species originating from extra-tropical temperate zones that are pre-adapted to overcome the physiological limitations imposed by high-elevation conditions, such as low temperatures and cloudiness (Girardin et al., 2010; Donoghue & Edwards, 2014). Therefore, our findings support the OTT hypothesis and highlight the important role that biogeographic dispersal plays in determining the structure of tree communities in tropical mountain ecosystems.

Although there was not a strong pattern in the phylogenetic structure of lowland tree communities, the influence of locally dominant and highly diverse linages, such as Fabaceae (see Gentry 1982), may help to explain the phylogenetic clustering observed in some sites, similar to what has been reported in other studies (González-Caro et al. 2014; Honorio-Coronado et al. 2015). In contrast, the mix of different evolutionary histories due to the influx of extra-tropical lineages can explain the phylogenetic overdispersion observed in most of the highland sites. Even though our system showed a “humped” rather than a linear change of species richness with elevation, there was a clear systematic decline in species richness above ~1800m asl. We suggest that the reduction of diversity in the highlands compared to lowlands is largely due to the limited rate that the thermal niches of species originated in tropical lowlands can evolve to highland conditions. However, explicit tests of the impact and rate of the dispersal flux and genetic exchange among tropical lowlands, highlands, and extratropical regions are still needed (e.g. Antonelli et al., 2018).

In the tropical mountains, similar to what is found along latitudinal gradients (Stevens 1992), the decrease of the niche width with the increase of temperature would promote phylogenetic overdispersion in lowlands and phylogenetic clustering in highlands, as predicted under the TNC hypothesis (Kerkhoff et al. 2014). This pattern has been observed in other taxa in tropical Andean mountains (e.g., Hummingbirds; Graham et al., 2009). However, we did not find evidence of this pattern in our study of tropical trees. Indeed, our findings showed a tendency towards phylogenetic overdispersion at higher elevations (Figure 3.3B and 3.3C), which contrasts with the expected pattern of change in the phylogenetic structure of tree communities under the TNC. The observed tendency for trees to be phylogenetically overdispersed in these tropical highlands is in agreement with other studies carried out along broader elevational gradients (González-Caro et al., 2014; Qian & Ricklefs 2016), and lends support for the OTT hypothesis as the underlying mechanism (Qian & Ricklefs 2016).

Alternative hypotheses have suggested increases of habitat filtering and deterministic processes as major causes of the decrease in species richness with elevation (Mori et al. 2013, Tello et al. 2015). This habitat-filtering hypothesis predicts a systematic increase in phylogenetic clustering along the elevational gradient – the opposite of what we observed. Other alternative ecological processes, such as facilitation or priority effects, could increase in importance along with in elevation and potentially lead to phylogenetic overdispersion at higher elevations. For example, in the younger tropical highlands, the relative recent migrations of divergent lineages may have facilitated the entrance of other taxa with similar requirements but with different evolutionary histories (Valiente et al. 2006; Fukami 2015), thereby enhancing phylogenetic diversity. These hypotheses require testing.

Phylogenetic replacement was the most important component of the phylogenetic β-diversity among our study plots. According to the standardized effect sizes, we estimate that plots with elevational differences of approximately 1000 m or more (corresponding to ~ 5.7°C difference in mean annual temperatures) harbor tree communities that differed significantly in phylogenetic composition (Fig. 3.2D and 3.2E). As above, this finding suggests that in the lowlands the in situ diversification of dominant clades may have played a paramount role on shaping the phylogenetic assemblages of tree communities (Honorio-Coronado et al. 2015). In the highlands, historical dispersal appears to lead the system to have a pattern of lineages with contrasting evolutionary histories occupying different elevational bands. Thus, the elevational range (i.e., thermal niche) of Andean tree communities can be seen as the result of the interplay of biogeography and lineage evolution. Understanding the proportion to which either dispersal of lineage evolution shapes the tree communities in Andean mountains could be used to help predict the likely impacts that global warming will have on species' thermal niches, distributions and forest function (Feeley et al. 2011; Duque et al. 2015; Fadrique et al. 2018).

Assuming that tropical Andean tree communities are composed of a mix of species with divergent origins from the tropical lowlands (e.g., Amazonia) and extra-tropical zones (e.g., the Southern Andes or North America), we can expect different elevations to be primarily inhabited by lineages pre-adapted for each particular condition as expected under the OTT hypothesis. The conservation of the thermal niche through the evolutionary history of species could explain why the observed phylogenetic β-diversity as well as the dominance of extra-tropical species in the highlands was higher than expected by chance. Indeed, the hypothesis of dominance of tropical lineages in the lowlands versus temperate lineages in the highlands is supported by the comparisons of the elevational distribution tropical vs. temperate families (Figure S3.2). High but differential colonization rate of lineages along the slope of the mountains according to their origin and location could also explain the dominance of the phylogenetic overdispersion observed in the upper part of the elevational gradient.

The analysis of sensitivity showed that the observed correlation between phylogenetic β-diversity and elevation remain significant even after randomly assuming better-resolved dichotomies. In fact, under the assumed better-resolved phylogenies, the correlation between phylogenetic β-diversity and elevation tended to increase (Fig. 3.2A). It is a remarkable result if we take into account that the fully identified 464 species tallied in this study represent ~ 18% of the total number of tree species (2640 spp.) reported for the whole Colombian tropical Andes (Bernal et al. 2019). However, additional work focusing on improving tree phylogenies in Andean mountains are urgently needed. An ancestral reconstruction of the ecological zone of origin based on robust molecular phylogenies will help to improve our understanding of the roles played by the evolutionary mechanisms and migration patterns on determining tree communities and species coexistence in tropical montane ecosystems (Ricklefs 2006, Kerkhoff et al. 2014).

In conclusion, this study suggests that the colonization of tropical mountains by extra-tropical lineages (i.e., the OTT hypothesis) has influenced the modern distribution of tree species along the elevational gradient of the Andean-Amazon transition, and perhaps in other tropical mountain systems throughout the world. In addition, the interplay between biogeography and the evolution of the thermal niche emerges as a critical mechanism determining regional tree diversity in the tropical mountains. Species' thermal niches, defined by different morphological and physiological characteristics, may constrain species movements and transitions between contrasting environments and elevational bands (Zanne et al. 2014; Fadrique et al. 2018). For this reason, a young tropical mountain community can be the result of the immigration of different pre-adapted lineages and subsequent diversification within the limits of their niches (Donoghue & Edwards 2014). Since restricted transitions between biomes can be one of the most critical processes in the distribution of angiosperms (Donoghue & Edwards 2014), at least a portion of the flora along an elevational gradient of the tropical Andes may have migrated from areas at other latitudes and elevations but with similar temperatures (Qian & Ricklefs 2016). This contribution to our understanding of the width of the phylogenetic thermal niche of tree communities will add to our understanding of the need for species to migrate in order to track their preferred thermal conditions and avoid extinction due to climate change (Feeley et al. 2011, Duque et al. 2015; Fadrique et al. 2018).

Table 3.1 Description of plots used in this study. Longitude (X), Latitude (Y), Number of individual trees (N; ha-1), species richness (SR.; 0.25 ha-1), genus richness (GR, 0.25 ha-1), family richness (FR, 0.25 ha-1), Net Relatedness Index (NRI) and Nearest Taxon Index (NTI).

Elevational Band	X	Y	Elevation (masl)	N	SR	FR	GR	NRI	NTI
1	-80.63	-3.77	772	512	40	25	31	0.47	1.76

1	-80.64	-3.78	749	640	66	27	52	-0.72	-0.91
1	-80.65	-3.74	764	496	49	25	40	1.33	0.68
2	-80.67	-3.75	941	424	53	30	46	-0.77	-0.30
2	-80.66	-3.74	972	512	55	30	48	-1.18	-0.57
2	-80.65	-3.75	960	516	54	31	42	-0.10	-0.36
3	-80.67	-3.72	1119	696	72	31	56	-0.63	-0.95
3	-80.68	-3.72	1065	576	51	24	41	-0.80	-1.12
4	-80.69	-3.72	1412	748	77	32	56	1.07	-0.39
4	-80.70	-3.71	1414	528	51	27	37	1.13	0.17
5	-80.70	-3.72	1642	608	74	30	53	1.50	0.95
5	-80.70	-3.71	1663	640	58	23	43	1.68	-1.41
6	-80.72	-3.69	1848	652	52	25	36	0.99	2.18
6	-80.72	-3.68	1903	824	50	28	40	-0.73	-1.59
7	-80.72	-3.68	2111	728	48	25	37	-0.40	0.15
7	-80.73	-3.68	2083	860	43	20	28	-0.43	-1.27
8	-80.73	-3.68	2364	556	30	19	22	-0.73	-1.21
8	-80.73	-3.68	2314	592	41	19	26	-1.06	-0.28
9	-80.74	-3.68	2647	592	36	19	27	-0.71	-0.60
9	-80.75	-3.67	2802	452	20	10	13	-1.44	-2.28

Figure 3.1 Variation across elevation in rarefied species richness to 100 individuals (A). NRI (B) and NTI (C). The gray circles are significant phylogenetic structure indexes relative to the null model.

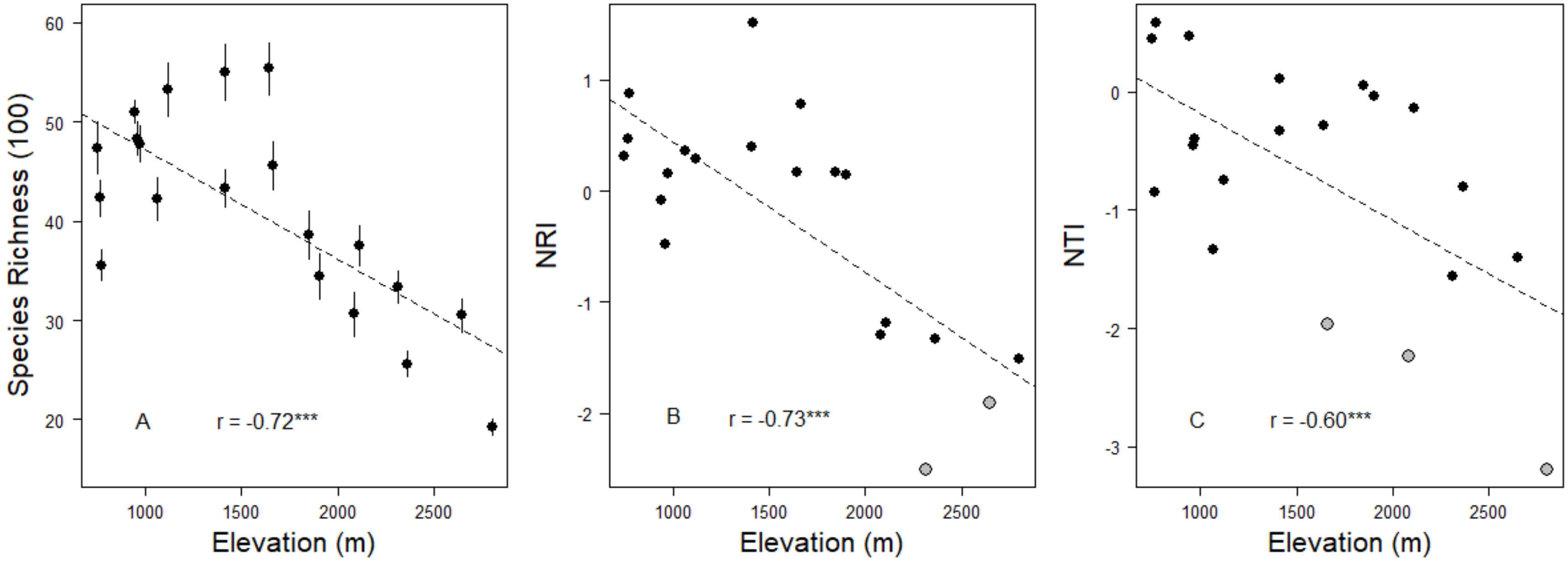


Figure 3.2 Variation in phylogenetic similarity indexes with respect to elevational differences between plots. Observed PhyloSor$_{\text{Total}}$ (A). PhyloSor$_{\text{Nes}}$ (B). PhyloSor$_{\text{Rep}}$ (C) and standardized effect size of PhyloSor$_{\text{Total}}$ (D). PhyloSor$_{\text{Nes}}$ (E). PhyloSor$_{\text{Rep}}$ (F). The gray circles indicate significant standardized effect size with respect to null expectations.

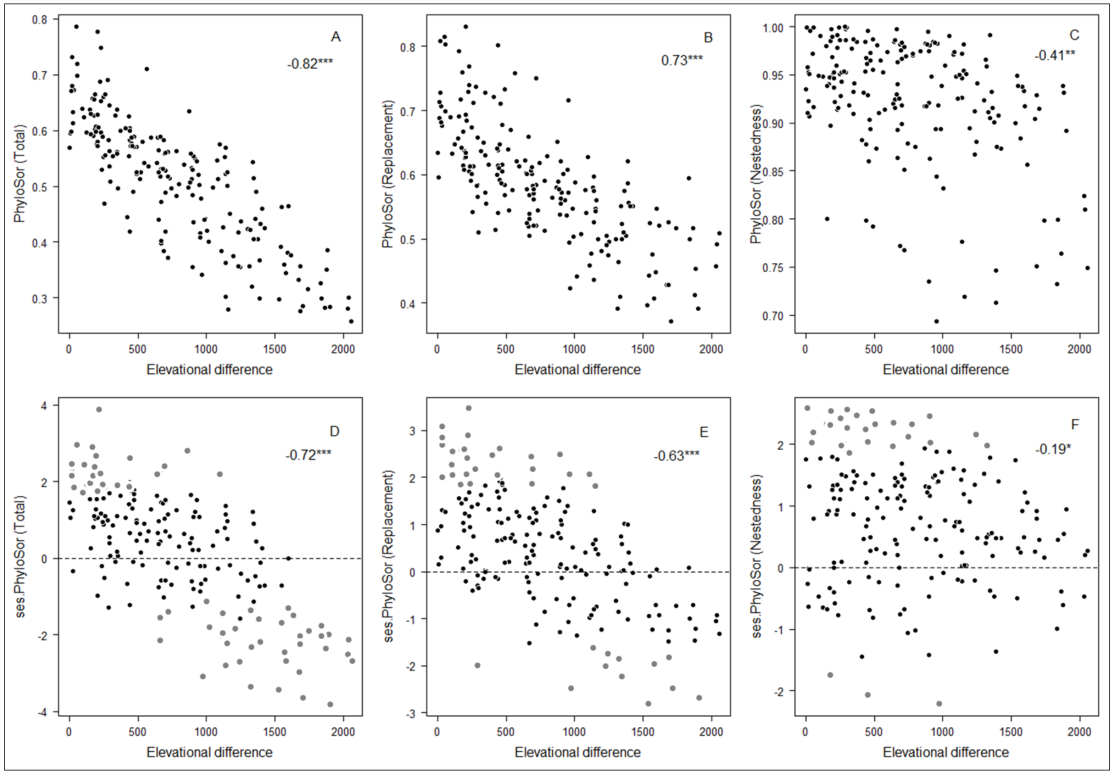
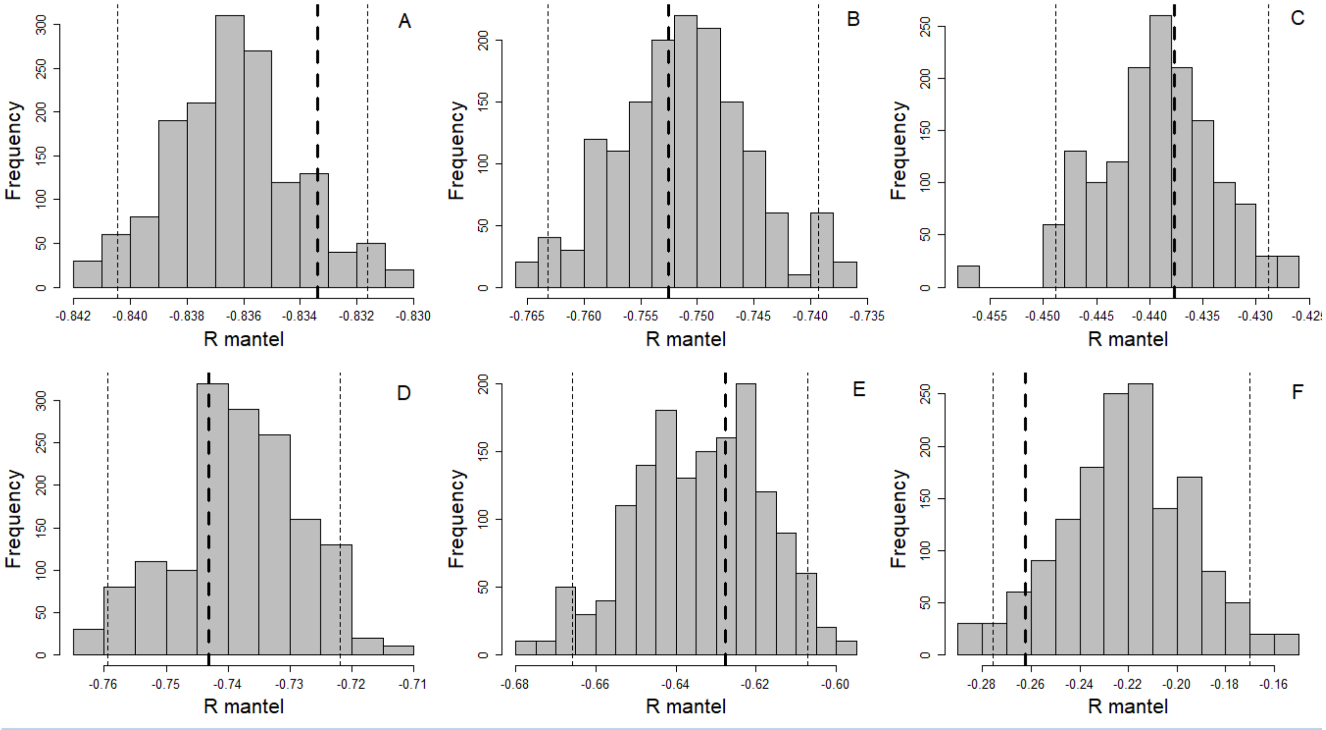


Figure 3.3 Distribution of mantel test parameter for 1000 runs resolving randomly the phylogenetic tree polytomies of correlation among Observed PhyloSor$_{\mathrm{Total}}$ (A). PhyloSor$_{\mathrm{Nes}}$ (B). PhyloSor$_{\mathrm{Rep}}$ (C) and standardized effect size of PhyloSor$_{\mathrm{Total}}$ (D). PhyloSor$_{\mathrm{Nes}}$ (E). PhyloSor$_{\mathrm{Rep}}$ (F) and elevation. The dashed dark line indicates the observed value with the raw phylogenetic tree. The dashed thin lines indicate the confidence intervals of the null distribution.



4 Chapter 4: The legacy of biogeographic history on the composition and structure of Andean forests

4.1 Introduction

Humboldt & Bonpland's (1807) classic work in Chimborazo Mountain linking the physical and biotic attributes of tropical mountains is a cornerstone of modern ecology and biogeography (Wake et al. 2009). One of Humboldt' most striking observations was the almost-dogmatic decrease in species richness and tree height at higher elevations. This observation lead Humboldt & Bonpland (1807) to hypothesize that elevation and temperature are primary drivers of forest community assembly and structure. These ideas formed the foundation for modern expectations of a systematic decline in tree height and aboveground biomass (AGB) with elevation, as well as subsequent theories for the importance of this pattern in predicting changes in the global carbon cycle under climate change (Enquist et al. 2009; Olson et al. 2018).

In Neotropical forests, a decrease in net primary productivity due to lower temperatures and nutrient availability is believed to drive a decline in AGB along elevational gradients (see Table 3 in Girardin et al. 2014; Unger et al. 2012). In southeastern Peru, a decline in AGB with elevation was associated with a reduction of tree height in colder areas (Girardin et al., 2014) due to the decrease of light, nutrient availability, and temperature within the cloud immersion zone (Girardin et al., 2010). However, a contrasting pattern was recently reported from elevational gradients in Africa (Cuni−Seanchez et al. 2017, Marshall et al. 2012), Papua New Guinea (Venter et al. 2017) and South America (Peña et al. 2018), where AGB did not decline monotonically with elevation. These contrasting results suggest that complex abiotic and biotic factors may influence forest AGB stocks and dynamics along elevational/thermal gradients in the tropics (Rapp et al. 2012, Peña et al. 2018).

The biogeographic origin of tropical Andean tree species could be one important factor influencing the distribution of tree height and AGB along elevational gradients. Some studies have hypothesized that, after the Andean uplift, tropical lowland clades were the primary source of species colonizing the highlands (Gentry 1982; Qian & Ricklefs 2016). Under this scenario, we could expect a higher colonization success of small-statured species because of their greater resistance to embolism due to freezing (Zanne et al. 2014) due to the narrower vessel diameters (Zanne et al. 2014; Olson et al. 2018). Indeed, the large-diameter vessels of large tropical lowland trees may have prevented their movement into the cold tropical highlands (e.g. Olson et al. 2018). If vessel size did indeed act as a filter to successful colonization, we would expect highland forests to be dominated by closely-related (sensu Webb et al. 2002), and mostly small-statured, species (Gentry 1982). In other words, this scenario predicts that average tree height - and thus the AGB of forests - and the phylogenetic diversity of forests should decrease with elevation.

The tropical Andes, however, have also been colonized by clades that originated in temperate regions (Gentry 1982; van der Hammen and Cleef 1983; Hooghiemstra et al. 2006). Many of these temperate immigrant clades include large-statured trees (e.g. Quercus humboldtii and Colombalanus excelsa; van der Hammen and Cleef 1983) whose ancestors may have evolved adaptations to maintain hydraulic safety under extremely cold and seasonal temperate climates since the Eocene (Larjavaara 2014). The past colonization of the tropical highlands by large-stature temperate clades could cause modify the expected systematic decline of tree height and AGB along elevation (Culmsee et al. 2010; Peña et al. 2018). Moreover, we expect the contribution from temperate immigrant clades to the total AGB of forests to be greater in the highlands than in the lowlands because of the greater climatic similarity between high-latitude and tropical high-elevation systems (Körner 2007; Feeley and Stroud 2018). Likewise, the mixing of clades from the tropical lowlands with temperate immigrants should produce more phylogenetic diversity in the highlands than in the lowlands (Ramírez et al. 2019).

The historical immigration of temperate-affiliated clades into tropical montane systems may modify the expected decrease of AGB with elevation in tropical forests (e.g. Venter et al. 2017). For example, in Indonesia and Colombia, Culmsee et al. (2010) and Peña et al. (2018), respectively, found that AGB did not co-vary with elevation largely due to the contribution of the Fagaceae family. Fagaceae is a temperate-affiliated tree family in which species richness increases at higher latitudes (Heywood et al. 2007; Culmsee et al. 2010). Adaptations to cold may have facilitated Fagaceae's colonization and dominance in the highlands of Indonesia and the northern regions of South America. Since Fagaceae is comprised mostly of large-statured tree species, its dominance at high elevations may therefor explain why AGB does not decrease with elevation in some tropical mountain systems (Culmsee et al. 2010; Peña et al. 2018). The immigration of other large-statured temperate-affiliated lineages may similarly help to explain patterns of AGB in other tropical montane forests.

The tropical Andes includes climates similar to the temperate zones where the largest species of trees occur around the world (Larjavaara 2014; Feeley and Stroud 2018). Furthermore, the highlands of the tropical Andes have been colonized multiple times by north- and south-temperate immigrants tracking to cold environments (Hooghiemstra and van der Hammen, 2004; Segovia and Armesto 2015). Accordingly, we hypothesized that the migration of some large-statured temperate-affiliated clades promotes an increase of the AGB stocks in high-elevation tropical forests (Figure 4.1). To test this hypothesis, we first evaluated how phylogenetic diversity changes along elevational gradients. We then used structural equation modelling (Shipley 2016) to disentangle the effects of tree size (i.e., tree height and diameter) and the biogeographic affinity of tree clades in explaining patterns of AGB along elevational gradients in northwest South America. Using data from intensive surveys of 0.25-ha plots (e.g. Schepaschenko et al. 2019) in native forests of Colombia from lowland Amazonia and Caribbean forests to the highland Andean cloud forests, we addressed the following questions: i) How does phylogenetic diversity (PD) of tree communities change with elevation? ii) How does AGB and tree size (height and diameter) change along the elevation gradients? iii) What is the contribution of temperate-affiliated species to changes of AGB along the elevational gradients? By answering these questions we hope to provide new insights into the biogeographic factors influencing patterns of forest structure in tropical montane forests.

4.2 Materials and Methods

Study sites and Plot surveys

Colombia has traditionally been divided into five major biogeographic regions that delineate the principal forest types and ecosystems: the Caribbean, Pacific, Andean, Orinoquia, and Amazon regions. The present study was conducted primarily in the Andean and Amazon regions, but some sites from the Caribbean region were also included. To assess the variation in biomass and carbon stocks along elevational gradients, we established 92 forest inventory plots (Figure S4.1). These plots were situated between 50 and 2928 m asl in six focal areas: the Antioquia province in Northwest Colombia, the Guaviare and Meta areas located in the Amazon region, the El Guavio forests, the Gamuez elevational gradient between Putumayo and Nariño, and the Regional Natural Park Corredor Biológico Puracé – Cueva de Los Guácharos in the province of Huila. Plots were each 0.25 hectare (2500 m2) to help maximize landscape coverage (i.e., Clark and Clark 2000; Chave et al. 2004).

In each inventory plot, all living trees with diameter at breast height (DBH) ≥ 10 cm were mapped, tagged, measured, and identified. Trunk diameters were measured at 130 cm above ground or 50 cm above buttresses and aerial roots when present. Multiple stems were measured and recorded separately but noted as being of the same individual. We also measured the total height (H, in m) of roughly 40% of the individuals selected at random within each plot. For small trees (height < 15 m), tree height was measured using a fiberglass telescoping pole (Hastings Measuring Sticks 3JF-108823); for taller trees, height was measured using a laser hypsometer (Nikon Laser 550). Overall, the heights of 7976 trees were measured for this study.

Vouchers were systematically collected for each unique species in each plot, as well as for individuals for which taxonomic identity was uncertain. Taxonomic identifications were made by comparing the voucher specimens with herbaria material and with the help of specialists in some groups. All of the vouchers are maintained at the University of Antioquia's Herbarium (HUA) and/or in the Herbario Amazónico Colombiano (COAH). In total, 73.3% of the trees were identified to species, 16.0% to genus, and 2.3% to family; the remaining 8.4% of trees are unidentified.

Phylogenetic diversity

To calculate phylogenetic diversity sensu stricto (PD; Faith 1992), we first generated a phylogenetic tree of hypothesized relationships using phylomatic for the complete species list as recorded across all 92 plots (excluding unidentified taxa). We used the bladj algorithm to date the phylogenetic tree by adjusting phylogenetic branch lengths to respective fossil ages (Webb & Donoghue 2005). The PD of each plot community was then calculated as the total sum of the phylogenetic branch lengths connecting the co-occurring species in each plot along the minimum spanning path to the root of the tree.

The observed PD was compared to a null distribution to control for the sampling effect of the regional diversity. The null model used an independent swap algorithm that maintained the frequency and richness of species in each plot while randomizing community composition (Gotelli 2000). The standardized effect size of the PD (PDz) was then calculated by subtracting the expected mean PD derived from the null distribution of 999 random draws to the observed PD value in each plot, divided by the standard deviation of the null distribution. This metric was estimated using the picante R package (Kembel et al. 2010).

To evaluate the possible effect of poor resolution in the phylogenetic tree generated from phylomatic, we used 1000 random fully resolved phylogenetic trees employing the bifurcatr function available in the PDcalc R package (Rangel et al. 2015) and estimated the associated uncertainty of the calculated PD in each plot.

Aboveground biomass estimation

The AGB of each tree (in kg) was estimated using a generalized allometric model developed from trees harvested from a broad range of tropical forests. The model we employed uses DBH (cm), total tree height (H in m) and wood density (oven−dry weight divided by fresh volume; WD in g cm-3) to estimate the AGB of each individual tree according to the following equation (Chave et al. 2014):

$$AGB = 0.0673 * (DBH2 * H * WD) 0.976$$

DBH was available for every tree. For trees with missing height measurements, we developed an allometric model based on our dataset of ca. 8000 measured tree heights to estimate H from DBH using a non-linear mixed effect model (Banin et al. 2012; Hulshof et al. 2015) with the following Weibull equation (Feldpausch et al. 2012):

$$H = a – b * \exp (-c * DBH)$$

where a represents the asymptotic maximum tree height, b is the difference between the maximum and minimum tree height and c is a curve fitting parameter. This model outperforms the power model to estimate tree height of large stems. We included plot identity as a random factor to assess environmental effects and to remove the possible correlation of H with elevation that can indirectly affect AGB estimations.

We obtained wood density (WD) of the species from a global database (Chave et al. 2009; Zanne et al. 2009). In the case of missing species-level values in the database or incomplete identifications, we used the genus-level average (29% of the trees), or a family-level average (5.1% of the trees). Finally, the WD of individuals without botanical identification was assigned as being the mean WD of their respective plot. We then calculated the stand-level AGB of each plot as the sum of the estimated aboveground biomass of all living trees. We report stand-level AGB values on a per hectare basis.

Biogeographic affinities

We assigned biogeographic affinities to each tree species based on family-level data from Heywood et al. (2007) following the categories used by Culmsee et al. (2010). These categories are defined as: i) tropical families (T) that are constrained to tropical regions; (ii) families centered in the tropics (CTrop) for which the majority of species occur in tropical areas but several species occur in subtropical and temperate regions; iii) families centered in the temperate zones (CTemp) for which the majority of species are located in the southern and/or northern temperate regions. iv) Other families (O) including angiosperm plant families that are cosmopolitan or that did not correspond to any of the previous classifications. For our analyses, we combined the T and CTrop classes into a single tropical (Trop) category. This approach could underestimate the number and contribution of temperate clades that have had high rates of radiation in the tropics after immigration (Kooyman et al. 2014).

Differences between tropical- and temperate-affiliated species

We used the quadratic mean diameter (QMD) to characterize tree size (DBH) at the plot-level. The QMD is described as follows:

$$QMD = \sqrt[2]{\frac{BA}{(k * n)}}$$

where BA is the stand basal area, n is the number of individuals, and k is a constant equal to 0.0000785 that transforms BA in square meters to QMD in square centimeters (Curtis and Marshall 2000).

In each plot, we defined three different biogeographic indexes of size. First, we calculated the biogeographic diameter index (BDI), which is defined as the log-transformed ratio between the QMD of only temperate–affiliated (QMDtemp) and only tropical-affiliated species (QMDtrop) (BDI = QMDtemp / QMDtrop). The BDI >0 when temperate-affiliated species are the dominant component of the plot, and BDI<0 when tropical-affiliated species dominate. We employed similar approaches for tree height (BHI = Htrop / Htemp) and abundance of individuals (BNI = Ntrop / Ntemp) in order to assess the biogeographic influence of either tropical- or temperate-affiliated species in structuring both H and N in each plot. Each of the biogeographic indexes was employed independently as a surrogate for the overall biogeographic affinity of each plot.

Statistical analyses

We used linear and quadratic models to assess the relationship of PDz and elevation. The Akaike information Criterion (AIC) was employed to select the most parsimonious model. We also evaluated the correlation between AGB and PDz.

To analyze the relationship between AGB stocks and PD along the elevational gradients, we used linear and quadratic models to assess the variation of quadratic mean diameter (QMD), mean tree height per plot (Hmean), the number of individuals (N), and their respective biogeographic index (BDI, BHI, or BNI) along the elevational gradient. The AIC was employed to select the best model.

To assess the contribution of temperate-affiliated species to changes of AGB along the elevational gradients, we used structural equation modelling (SEM) to evaluate the relative importance of biogeographic affinity (as assessed by the biogeographic indexes BDI, BHI, or BNI) in determining the variation of the AGB along the elevational gradient after controlling for other variables. The SEM used elevation as an exogenous variable (we used both linear and quadratic terms of elevation), and QMD and either BDI, BHI, or BNI as endogenous explanatory variables. In the SEM, the use of quadratic terms as explanatory variables was required to fulfill the assumption of normal distributions of the residuals. Prior to the analysis, all variables were standardized to mean zero (0) and variance one (1). The SEM was run in the lavaan package in R.

All analyses were performed using R 2.15.3 (R Development Core Team 2012).

Comparison with other datasets

In order to determine if our results from Colombia are generalizable to other parts of the tropics, we also analyzed the Gentry dataset of 139 0.1 ha plots (DBH ≥ 2.5 cm) from Central and South America (Phillips & Miller 2002) as well as the data available in Venter et al. (2017) for Papua New Guinea and tested for the influence of biogeographic affinities on trees species dominance and composition (Fig. S4.1). When analyzing the Gentry dataset, we used basal area to evaluate the contribution of temperate-affiliated lineages instead of AGB due to the lack of tree height information for each plot.

4.3 Results

In all 92 Colombian plots, there were a total of 20,060 trees (281 ± 37.7 trees per 0.25ha) belonging to 1,625 species or morphospecies. Phylogenetic diversity (PD) showed a "hump-shaped" relationship with elevation (Table 4.1; Figure 4.2A) and was positively related to AGB (Table 4.1; Figure 4.2B). The Weibull model we employed to estimate tree height (H) was $H = 32 - 2.1 * \exp(-0.94 * DBH)$ and the estimated H was weakly affected by elevation (marginal R2 = 0.68; conditional R2 = 0.73).

The overall mean total AGB in the plots was 255.3 ± 108.4 Mg ha-1. There were significant differences in the mean AGB across focal areas ($F = 19.39$; $P < 0.001$). However, these differences were due to the fact that the highest values of mean AGB recorded in the province of Huila were twice the mean AGB observed in the other provinces (Table S4.1). The quadratic model best described the relationship between AGB and elevation (Table 4.1). Overall, the highest AGB occurred in plots between 1,800 and 2,200 masl (Figure 4.3A).

The quadratic model was also best for describing the relationship between N and QMD with elevation (Table 4.1; Figure 4.3). The biogeographic indexes (BDI, BHI and BNI) all increased significantly with elevation (Table 4.1; Figure 4.3), indicating an increasing importance of temperate-affiliated species at higher elevations.

The structural equation model (SEM) indicated that elevation and AGB were negatively related after controlling for the other variables, but that the effect of elevation on AGB was lower than the effect of the QDM and biogeographic indexes. The three SEM models that used the different biogeographic indexes explained a similar proportion of the AGB variance (BDI = 81%; BHI = 83%; BNI = 83%). However, the model that used BDI had the highest AIC value, while BHI and BNI models had similar AIC values (Figure 4.4). Overall, QDM explained the largest proportion of variation in AGB. However, BHI showed a higher correlation with QDM than the other two biogeographic indexes, which points to the overriding importance of temperate-affiliated species in indirectly determining patterns of change in the AGB.

Across the 139 Neotropical plots, we found that 15.25% (6,986 m2) of the total basal area (45,803m2) was represented by temperate-affiliated species. In highland (>1500 m asl) plots (n = 19), the contribution of temperate species was 34.6%. The ten most important genera contributing to tree basal area in Gentry's highland plots were Podocarpus, Ocotea, Weinmannia, Colombobalanus, Quercus, Alchornea, Meliosma, Miconia, Otoba and Clusia. With the exception of Otoba, Miconia and Clusia, these genera are all hypothesized to be temperate-affiliated lineages. In Papua New Guinea (see Table 3; Venter et al. 2017), 45% of the largest (tallest) species were temperate−affiliated in the plots ≥ 2,000 m asl. For example, the tallest species in the Papua New Guinea highlands belonged to the Nothofagaceae and Cunoniaceae families, both of which are typical families from temperate regions.

4.4 Discussion

We found a humped-shaped relationship between the phylogenetic diversity (PDz) of forests and elevation. This pattern is consistent with the phylogenetic diversity vs. elevation relationships observed in other Neotropical mountain ecosystems (González−Caro et al. 2014; Qian and Ricklefs 2016, Ramírez et al. 2019) and lends support for the hypothesis that tropical highlands of communities contain species of mixed biogeographic affinities (i.e. tropical and temperate). Overall, we observed a strong influence from both Godwanian and Laurasian biogeographic clades on the structure of the Neotropical highland tree communities (Gentry 1982; van der Hammen and Cleef 1983). The positive correlation between PDz and AGB (Figure 4.2B) indicates that the evolutionary history of tree communities may play an important role in shaping changes in both forest diversity and forest structure along tropical elevational gradients. Increases in PDz at intermediate elevations may be due to the mixing of highland floras that have a high proportion of temperate-affiliated species with lowland floras mostly dominated by tropical-affiliated species. This suggests that complementarity effects are a potential key determinant of the observed increase in AGB between 1,800 and 2,200 masl. Based on these results, we contend that the biogeographic history of the Andean mountains' flora is strongly influencing the average size and AGB of tree communities.

In contrast to the decrease in AGB stocks with elevation reported in some Neotropical montane forests (Girardin et al. 2010, 2014), we found a humped relationship between AGB stocks and elevation with peak biomass occurring >1,500 m asl (Figure 4.3A). This pattern is in accord with recent findings from Papua New Guinea (Venter et al. 2017) where highland forests were found to have high AGB due to a high abundance and dominance of large-statured tree species, as well as with findings from Indonesia (Culmsee et al. 2010) and Colombia (Peña et al. 2018) where there was no relationship between AGB and elevation. The observed non-negative trend between AGB and elevation was primarily due to the high dominance of trees belonging to Fagaceae at higher elevations. In this study, the dominance of Colombobalanus excelsa and Quercus humboldtii, both belonging to the Fagaceae family, explained the higher AGB stocks found in highlands (> 1,500 masl; 401.4 ± 247.4 kg ha-1) than in lowlands (< 1,500 masl; 227.7 ± 77.8 kg ha-1). The endemic C. excelsa and the Q. humboldtii tree species accounted for more than 25% of the total AGB in all plots and 60% of the AGB in highland plots (> 1500 m asl). C. excelsa is an extreme case of hyperdominance for the tropics. When C. excelsa is included, a total of 40 species (2.5%) represented

~50% of the total AGB across all plots (C. excelsa account for 2,046.2 Mg of 11,606.1 Mg in our dataset, 17.6%); when C. excelsa is excluded, 125 species (7.69%) represented 50% of the total AGB. In the case of Q. humboldtii, it has been shown that this clade colonized South America after the uplift of the Andes (Raven and Axelrod, 1974; Gentry, 1982; Hooghiemstra and van der Hammen, 2004). In contrast, C. excelsa is believed to have a tropical origin (Manos and Standford 2001). However, it is possible that both species maintained their ancestral tolerances for cold environments, allowing them to contribute to high AGB stocks in highland tropical forests.

The increase in the abundance and tree size (tree height and diameter) of temperate-affiliated species with elevation (Figure 4.3E – G) appears as a key factor controlling the observed non-linear relationship between AGB and elevation. In contrast to the temperate-affiliated species, the tropical-affiliated species decrease in size and dominance along the elevational gradient (Figure 4.3F). This finding agrees with the suggested decrease of tree size of tropical clades along the elevational gradient due to hydraulic constraints restricting the establishment of tropical large-stature trees in cold environments (Olson et al. 2018). Thus, in the Neotropics, in the absence of large-statured tree species from temperate clades, such as C. excelsa and Q. humboldtii, the negative relationship between AGB and elevation would likely prevail (Girardin et al. 2010). These allometric differences between the two contrasting biogeographic groups (tropical and temperate-affiliated species) indicate different ecological strategies and responses to changes in temperature. Overall, we propose that large-statured temperate-affiliated species that have migrated into the tropical highlands should have smaller vessel diameters than their counterparts from lowlands, increasing hydraulic safety and helping these species avoid embolism in the cold tropical highlands (Zanne et al. 2014; Olson et al. 2018). This is a new physiological/evolutionary hypothesis for tropical elevational gradients that requires further study.

Some authors have proposed the existence of temperate maritime climatic bands, located between 7.0 °C and 15.4 °C, where forest carbon storage can reach maximum values (Larjavaara 2014). Although our results agree with the idea of the existence of optimal climatic conditions for high AGB stocks for some specific clades (e.g., Fagaceae), the dominance of large-statured temperate trees in tropical Andean highland forests may be due primarily to physiological adaptations to cold that are conserved from their ancestors, a mechanism widely known as niche conservatism (Wiens & Graham 2005). In addition to the physiological constraints of vessel diameters (Olson et al. 2018), some studies have suggested that shallow soils may prevent tree species from achieving large statures on steeply sloped terrains (Venter et al. 2017). Although we lack adequate data on soil depth in the study plots, the relationship between the slope, which can be used as a possible surrogate of soil depth, and AGB was not significant (r = 0.09, P = 0.45). Likewise, the correlation between soil pH and AGB was not significant in a subset of 60 plots for which we have soil data (r = 0.13, P = 0.32). Indeed, differences in the AGB among plots with similar pH's were primarily explained by the percentage of temperate- vs. tropical-affiliated species (Figure S4.2).

The comparison of our results with additional datasets from other similar studies, such as those published by Gentry in the Neotropics (Phillips & Miller 1993) and Venter et al. (2017) for Papua New Guinea, support our hypothesis that small-statured species in highlands (e.g. Miconia and Clusia) is a common feature of clades originated in tropical lowlands. In contrast, in both these independent datasets, immigrants from temperate regions contributed significantly to the increase of tree size at higher elevations. For example, in Gentry´s plots located in highlands (>1,500 ma asl) the most dominant genera, e.g., Quercus and Weinmannia, are temperate-affiliated clades widely distributed in extra-tropical regions. In the case of Colombobalanus and other relatives belonging to the Fagaceae family that originated in the tropics, their ancestral adaptations to the cold may have favored their evolution and development primarily in the high mountains (Culmsee et al. 2010). Understanding the likely development of the evolutionary traits that promotes the successful performance of select species in the Neotropical mountains could help us to understand the physiological basis that allows some clades with temperate origin to maintain a large size in in the tropical highlands.

In conclusion, we find support for the hypothesized legacy of biogeographic affinity on the modern composition and structure of tropical mountain forests. Historical dispersal of temperate clades helps to explain the high AGB stocks observed in some Andean highland forests. Temperate-affiliated species that migrated into tropical montane forests from the northern or southern hemispheres may have conserved the thermal tolerances evolved by their cold-adapted temperate ancestors, enabling them to maintain high performance and carbon stocks in the cold tropical highlands. In particular, the contribution of the Fagaceae family to high AGB stocks in Colombian mountain forests (Peña et al. 2018; see also Culmsee et al. 2010 for Indonesia) suggests that further research is needed to improve our understanding of the fingerprints left by biogeographic history on these under-studied forests (Cavender-Bares et al. 2016).

Table 4.1 Models of the linear and quadratic relationships between elevation and standardized phylogenetic diversity (PDz), quadratic mean diameter (QMD), mean tree height per plot (H), number of individuals (N), biogeographic diameter index (BDI), biogeographic height index (BHI), biogeographic abundance index (BNI) and aboveground biomass (AGB). AIC = Akaike information criterion. Explained variance (Exp.Var(%)) = (Null deviance – Residual deviance) / Null deviance. SW P-Value = P value of the Shapiro Wilcoxon test to normality of the residuals.

Response Variable	Model	AIC	Exp.Var(%)	P-value	SW P-Value
PDz	elevation	362	27.9	>0.001	0.11
	elevation^2 + elevation	330	49.7	>0.001	0.25
QMD	elevation	265.9	1.2	>0.001	>0.001
	elevation^2 + elevation	265.1	1.9	>0.001	0.12
H	elevation	264.2	9.8	0.18	>0.001
	elevation^2 + elevation	258.5	49.7	0.009	0.21
N	elevation	242	22.9	>0.001	0.11
	elevation^2 + elevation	241.8	24.8	>0.001	0.11
BDI	elevation	348.9	6.8	>0.001	0.34
	elevation^2 + elevation	351	6.8	>0.001	0.11
BHI	elevation	282.2	50.2	>0.001	0.25
	elevation^2 + elevation	289.1	51.3	>0.001	0.11
BNI	elevation	288.1	50.1	>0.001	0.27
	elevation^2 + elevation	289.3	51.6	>0.001	0.11

Response Variable	Model	AIC	Exp.Var(%)	P-value	SW P-Value
AGB	PD	253.1	13.1	>0.001	0.10
	elevation	261.1	5.2	0.04	0.001
	$elevation^2$ + elevation	250.4	17.4	>0.001	0.33

Figure 4.1 (A) A hypothetical scenario of colonization by mainly by small-stature trees from the tropical lowlands that are pre-adapted to cold and seasonal condition in the highlands following the Andean uplift. The expectation is that tree size and aboveground biomass (AGB) decrease with elevation. (B) A hypothetical scenario in which large-statured trees colonized the tropical highlands from the temperate regions. The assumption in this scenario is that the temperate trees are adapted to cold conditions and maintain their large size after colonizing the tropical highlands. This scenario predicts that tree size and AGB do not change systematically alongside elevation.

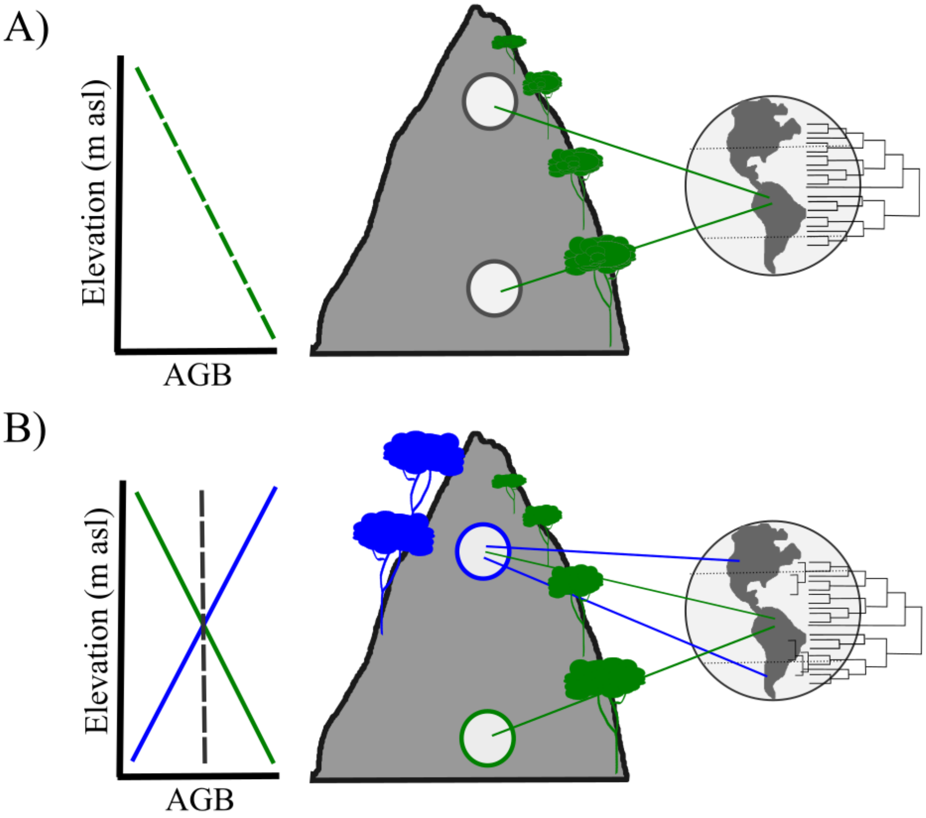


Figure 4.2 (A) Variation of standardized phylogenetic diversity (PDz) with elevation in Colombian forests, and (B) Correlation between phylogenetic diversity and aboveground biomass (AGB). The length of the bars represents confidence intervals in PDz. Confidence intervals of the PDz for each plot were calculated based on the sensitivity analysis used to assess the effects of the phylogenetic tree resolution. The continuous line represents the model (linear or quadratic) that best explain the relationship with elevation; dashed lines represents the calculated 95% confidence intervals of the model fit.

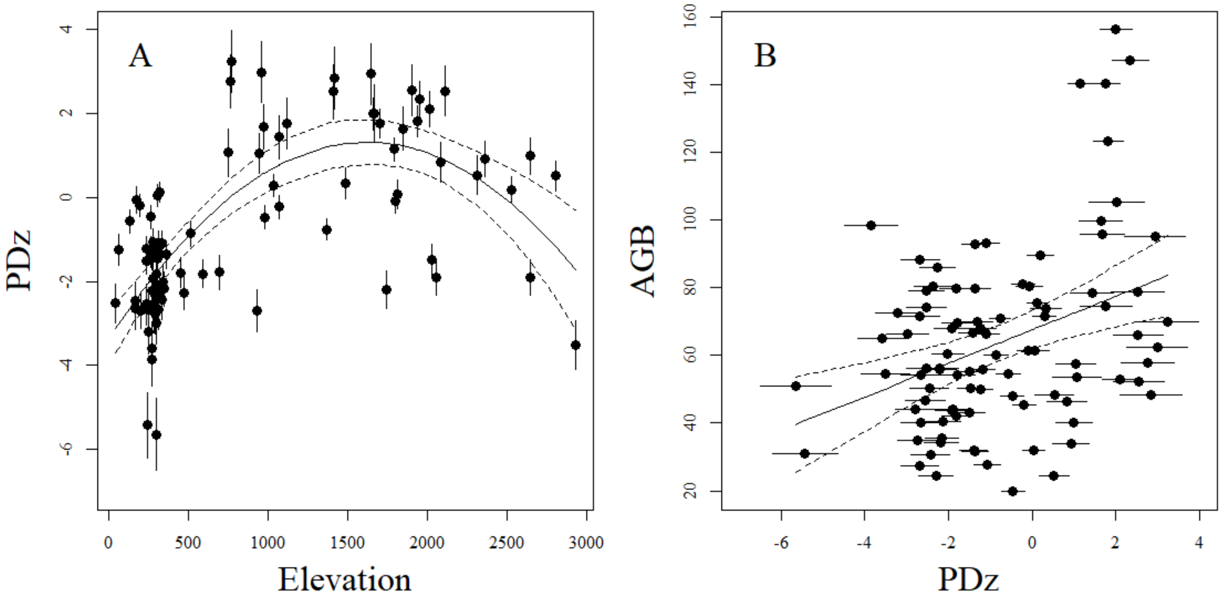


Figure 4.3 Relationships between (A) Aboveground biomass (AGB), (B) quadratic mean diameter (QMD, (C) mean tree height per plot (H), (D) number of individuals (N), (E) biogeographic diameter index (BDI), (F) biogeographic height index (BHI), and (G) biogeographic abundance index (BNI), and elevation. Lines are shown for significant models; dashed lines indicate confidence intervals.

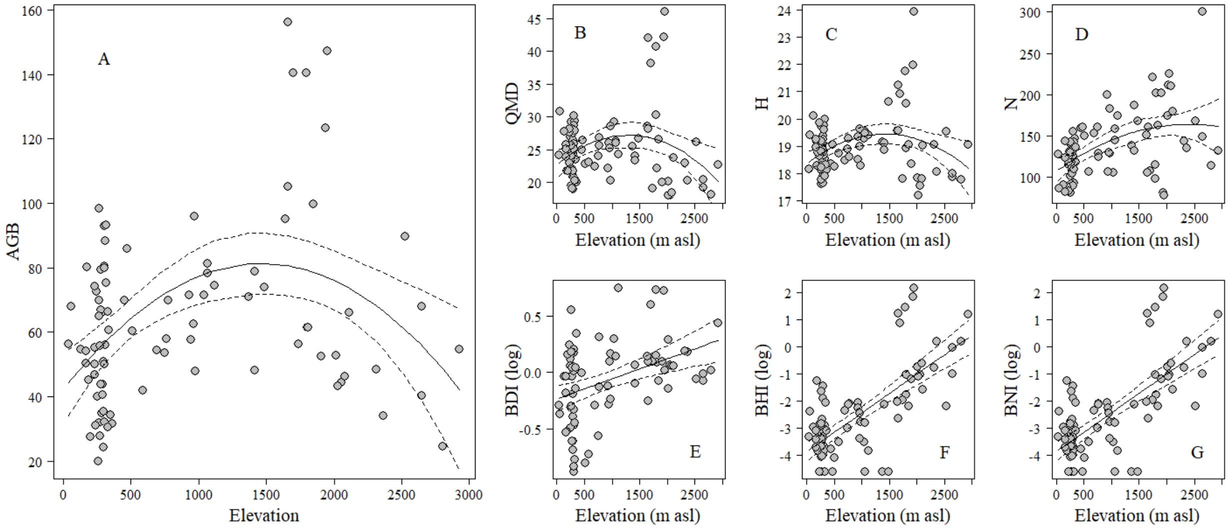
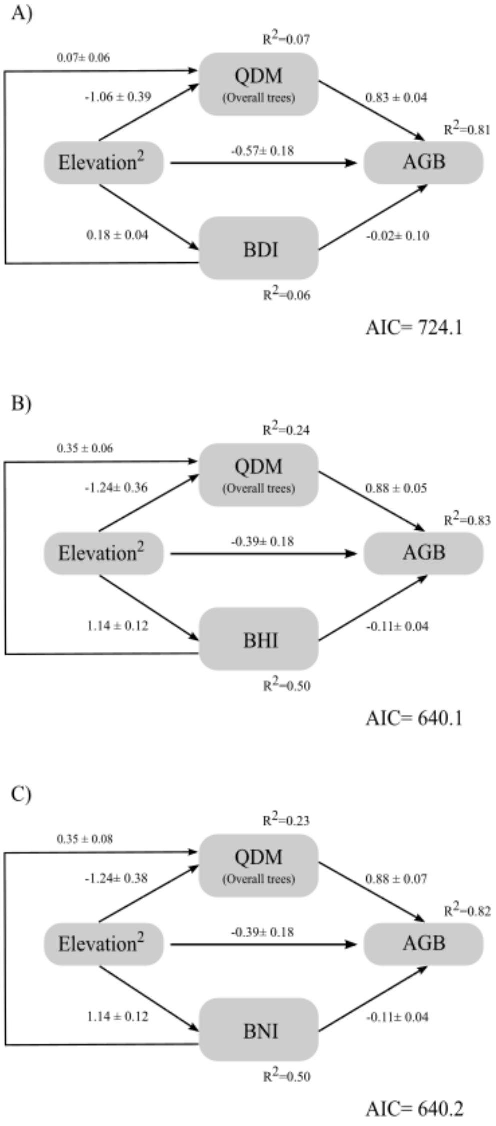


Figure 4.4 Structural equation models (SEM) used to explain aboveground biomass (AGB) as a function of elevation, quadratic mean diameter for overall trees and (A) biogeographic diameter index, (B) biogeographic height index and (C) biogeographic abundance index. Standardized coefficients with standard errors are indicated over each arrow. R^2 values above each endogenous variable indicate the amount of variance explained.



5 Chapter 5: Historical assembly of Andean tree communities

5.1 Introduction

The latitudinal and elevational gradient of species diversity have drawn ecologists' attention for centuries, but we still lack a consensus about the underlying structuring mechanisms of the claimed reduction of species richness and diversity either from the tropics to the poles or from lowlands to highlands (Mittelback et al., 2007; Laiolo et al., 2018; Etienne et al. 2019). Since Humboldt, the spatial variation of climate along continents and mountains has been widely recognized as a key driver of changes in species diversity (Körner and Spehn 2019; Kreft and Jetz, 2007). However, diversity anomalies associated with differences in species richness among regions with similar environmental conditions, contradicts the idea that climate by itself shapes most of the changes in diversity across the geographical space (Latham & Ricklefs 1993; Ricklefs and He, 2016). That said, both the latitudinal and the elevational diversity gradients (LDG and EDG, respectively) could also emerge as the outcome of a historical balance among processes such as diversification and migration (Wiens and Donoghue 2004, Etienne et al., 2019). To move forward on our understanding about the LDG and EDG we must take into account the biogeographic history of clades and ecosystems, a commonly overlooked keystone of the theory on patterns of regional and global diversity (Ricklefs 2004).

The most prominent hypothesis aiming to explain the evolutionary pattern of vascular plant diversification and species migration along both the LDG and the EDG seems to be the well-known Tropical Niche Conservatism hypothesis (TNC; Wiens and Donoghue 2004, Kerkhoff et al. 2014). The TNC proposes that the higher diversity at low latitudes results from the low chance that clades originated in tropical regions have had to settle and evolve after migrating to the colder temperate zone (Kerkhoff et al. 2014). That said, the tropical lowlands must be the ecological zone of origin of tree species, which capability to disperse and settle in contrasting climates (e.g. temperate/cold/dry) is constrained by a high clade retention of the environmental preferences of their tropical ancestors (strong niche conservatism). These features suppose an implicit low rate of niche evolution and the reduction of the adaptive shift rate across the LDG. Under this scenario, the TNC hypothesis predicts tree communities in both the temperate and the tropical highland zones to be structured by a handful of closely related species with tropical origin that generates a decrease of the phylogenetic diversity from the tropics poleward (Algar et al., 2009; Kerkhoff et al., 2014) and from lowlands to highlands (Graham et al. 2009).

The South America region, which has been recognized as one of the most diverse areas on earth (Myers et al. 2000), provides a rich setting for testing hypotheses about the historical assembly of plant communities across latitudinal and elevational gradients (Qian & Ricklefs 2016). After the separation of South America from other landmasses nearly ~100My, the continent remained isolated and was dominated by a tropical biome triggered by the increase of temperatures during the early Eocene (Jaramillo et al. 2006). Fossil records of tropical-like floras found in the southern Patagonia dating from ~60My, support the purported past broad latitudinal expansion of tropical conditions (Jaramillo et al. 2006, Jaramillo & Cárdenas 2013). Later, during the late Eocene (~35 My), climate cooling caused a contraction of the tropics triggering extinction and shrinking the spatial distribution of many tropical clades (Jaramillo et al. 2006). This tropical contraction may have provided opportunities for cold-restricted lineages to arise, diversify, and spread across the Neotropics. In South America, the migration of cold-temperate clades originated in the southern part of the continent into warmer and wet regions located at low latitudes was favored by the Andes' uplift, the most extensive mountain range on earth. This mountain uplift began nearly ~80My in the south of the continent, while the Central Andes began their uplift ~45My and the Northern Andes ~10My ago (Hoorn et al. 2010). The emergence of the Andes created a very pleasant corridor along highlands, with environmental conditions such as coldness and seasonality, climatically similar to those typical of temperate biomes (Segovia et al. 2020). At the same time, the closing of the Isthmus of Panama ~15My connected South America and North/Central America (Montes et al. 2015), providing new opportunities for species migration north-southwards (van der Hammen and Cleef, 1983, Bacon et al. 2015).

This quite alterative climatic history in South America has surely left its imprint on the evolution and historical migration of many of the clades that actually shape the spatial patterns of species assembly across Neotropical forests. Climatic and geological changes, such as the cooling during the late Eocene and/or the Andean mountain uplift (Hoorn et al., 2010), may have promoted the purported mix of species with tropical and temperate origin in the mountains of the tropical zone (González-Caro et al. 2014; Qian & Ricklefs 2016; Ramírez et al., 2019, González-Caro et al. 2020, Segovia et al. 2020). Likewise, in South America, although most of the tree clades may have been originated in Gondwana and diversified isolated over long periods, the tropical biome expansion during the early Eocene (Jaramillo et al. 2006) may have generated a contraction of the temperate biome that reduced the opportunity of diversification for temperate clades. However, the temperate alike climatic remnant areas towards the southern of South America, helped to maintain very old clades, which are still represented as long branches in the phylogenetic tree (Segovia et al. 2020). These very old surviving clades, such as Nothofagus (~80My), favored the subsequent mixing of clades with different origin at southern latitudes, and thus, the increase of the phylogenetic diversity far from the tropical zone (Rezende et al., 2017). These predictions of an increase of the phylogenetic diversity towards highlands in the tropics, but a reduction from south towards the equator, is defined here as the multiple ecological zones of origin hypothesis (MZO), which is expected to play a key role on shaping both the latitudinal and elevational phylogenetic diversity gradient in South America. Therefore, in contrast to TNC, under the MZO hypothesis, we expect a reduction of the phylogenetic diversity towards the tropics due to a higher accumulation of younger clades, but an increase upward in the mountains towards the tropical band due to the mix of older and younger clades.

Here, we combined simulations and an empirical approach to assess the predictions of the Tropical Niche Conservatism (TNC) and the Multiple Zones of Origin (MZO) hypotheses (Figure 5.1) as determinants of the LDG and EDG of the phylogenetic diversity of tree species along the subtropical and tropical Andes. First, we used computer simulations to evaluate the theoretical predictions of both the TNC and the MZO on the relationship between the mean pairwise phylogenetic distance (MPD; Webb 2000) among co-occurring species in permanent plots and the LDG and EDG. Then, we assessed the same relationship between MPD and the LDG and EDG on a dataset of 245 forest plots distributed across the subtropical and tropical Andes from Argentina to Colombia. Overall, under the TNC hypothesis, we expect a negative relationship between MPD and the LDG/EDG; in contrast, under the MZO hypothesis, we expect a positive relationship between MPD and the LDG/EDG (Figure 5.1).

5.2 Methods

Simulations

The simulations assume that tree species belong to, or were originated in, any of the following four regions: Tropical Lowlands (TrL), Tropical Highlands (TrH), Temperate Lowlands (TeL), and Temperate Highlands (TeH). We also assume that TrH, TeL, and TeH have similar climatic conditions (i.e., cold and seasonal) and differ from TrL (i.e., warm and wet). Under the TNC hypothesis, we assumed a unique origin of all species in TrL, from which species can then get dispersed onto TrH and TeL (movements between different climates), while TeL species can also get dispersed onto the TeH (movements between similar climates). On contrary, under the MZO hypothesis, the model assumes that tree species were originated in both TrL and TeL, and can disperse onto TrH and TeH, respectively. However, in both models, we assumed that TeH can in turn get dispersed onto TrH and vice versa (Figure 5.1).

In our model, as a starting point, we used the pbtree function available in the phytools R package to simulate a phylogenetic tree. This function requires parameterizing speciation and extinction rates to define the lineage accumulation rate through time and the final number of species in the resultant phylogenetic tree. We used 0.5 as the speciation rate value, 0.05 as the extinction rate value, and 10000 as a final number of species to generate the phylogenies in our simulations. A value of 0.5 for speciation means that 1 species would generate another species in 2 million years. Likewise, and extinction rate of 0.05 means a species would go extinct each 20 million years (after Ricklefs 2007). We generated 1000 phylogenetic trees for further simulations. The extinct branches were removed from the simulated phylogeny employing the drop.extinct function from the geiger R package.

To define the presence of each species in any of the four regions across the phylogeny (TrL, TrH, TeL, TeH), we used the sim.history function available in phytools R package. This function simulates discrete traits based on an ancestral trait state (here represented by the region of origin), a phylogenetic tree, and a square symmetric matrix that contains the probability of transition among trait states (i.e., regions). To define the region of origin, we used TrL as the region of origin in the function sim.history to simulate TNC. To simulate MZO, first, we simulated two phylogenetic trees in which the ancestral state is either TrL or TeL, which were then joined to build a unique representative phylogeny. The tropical lineage was fixed to initiate with 75% of the total species richness, while the temperate lineages initiated with the remaining 25% of the total species richness. The transition matrix's diagonal includes the probability (range between 0 and 1) of one lineage to retain or stay in its ancestral region. In this model, we used a constant probability value of 0.5 in the main diagonal, which represents a moderate conservatism probability to stay in the ancestral region. The off-diagonal elements of the transition matrix represent the probability of dispersal among regions. The probability of dispersal between regions with different climates (TrL and the others) was defined as 0.25. The probability of dispersal between regions with similar climates (TrH, TeL and TeH) was defined as 0.5, which is similar to the probability to remain in the same region and means twice the probability of dispersal between different climates.

Then, we randomly generated 10 local communities for each region (TrL, TrH, TeL, TeH) by bootstrapping 10% of species in each run or independent generation of the phylogenetic tree, which means the reproduction of 1,000 independent sets of 40 local communities. In each run, and for each local community, we calculated the standardized effect size of the mean pairwise phylogenetic distance (MPDses; Webb 2000). The MPD is the average of all possible combinations of phylogenetic distance pairs within a community. This metric was standardized relative to a null distribution to control by the effect of the species richness (see Supplementary Methods), and calculated using the picante (Kembel et al., 2010) R package. Likewise, each one of the 1,000 times we generated the 40 local communities (10 by region), the 40 MPDses values were independently regressed against latitude and elevation using the Ordinary Least Squares method. All variables were standardized to zero mean and one standard deviation before applying the regression using the scale function available in R. The average of the slope (the intercept is not significantly different than zero) after the 1000 draws was employed to evaluate the overall pattern of change of the MPDses along the LDG and EDG. Finally, we assessed if the results of the simulations for each hypothesis were consistent with our theoretical expectations (Figure 5.1). For TNC, we expected the regression coefficient of both elevation and latitude to be negative, which means a decrease in MPDses with both elevation and latitude. In contrast, for MZO, we expected the regression coefficients to be positive, which mean an increase in MPDses with both elevation and latitude. We used a t-test to assess whether the simulated coefficients of latitude and elevation for each hypothesis differs from zero. We also used the Kolmogorov-Smirnov test to compare the distributions of the coefficients of latitude and elevation generated by either the TNC or the MZO hypotheses.

All code was written in R (R Core Team 2018) and is available in the supplementary information.

Empirical analysis

The empirical analysis was conducted using tree census data collected from 245 forest inventory plots located that covered a latitudinal geographical range from 7.1°N (Colombia) to 27.8°S (Argentina), a longitudinal range from 79.5° to -63.8° W, and an elevation range from 150 m asl to 3511 m asl. The mean annual temperature (MAT) of plots ranged from 7.3 to 23.8 °C (mean = 16.7 ± 4.1 °C; mean ± SD) and mean annual precipitation (MAP) of the plots ranged from 608 to 4313 mm y-1 (mean = 1405.0 ± 623.9 mm y-1). The number of plots sampled in each country was: Argentina = 46, Bolivia = 40, Peru = 22, Ecuador = 21, and Colombia =116 (Figure 5.2). Plot size varied from 0.25 to 1 ha, with a cumulative sample area of 156.5 ha. In each plot, we tagged, mapped, measured, and collected vouchers of all trees and palms with stem diameter at breast height (DBH, 1.3 m) ≥ 10 cm.

To homogenize and validate species names of vascular plants recorded in each plot, we submitted the combined list from all plots to the Taxonomic Name Resolution Service (TNRS; http://tnrs.iplantcollaborative.org/) version 3.0. Any species with an unassigned TNRS accepted name or with a taxonomic status of ‘no opinion’, ‘illegitimate’, or ‘invalid’ was manually reviewed. Families and genera were changed in accordance with the new species names. If a full species name was not provided or could not be found, the genus and/or family name from the original file was retained. We registered a total of 125.670 individuals, from which we assigned a taxonomic identification at any level to 93.9% of them (45.6% were identified to species, 27.3% to genus, and 21% to family). The 6.1% of individuals without any taxonomic identification were excluded from subsequent analysis.

Using the plot data, we used Phylomaker V.2. to build a phylogenetic tree based on the Smith and Brown (2018) phylogenetic backbone. The phylogenetic tree was built employing all of species reported in our database that were identified at the family level at least. We used the bifurcating algorithm available in the PDcalc R package (Rangel et al. 2015) to solve the polytomies at the tip level. This procedure was repeated 1000 times to estimate confidence intervals in further analyses. We used this phylogenetic tree to estimate MPDses for each plot. We included in our analysis both angiosperms and gymnosperms. In order to test our hypotheses associated with the predictions of TNC and MZO (see above), we used a linear model that included MPDses as the dependent variable, and either latitude or elevation as explanatory variables. However, due to the use of different plot sizes, before assessing the relationship between MPDses and latitude/elevation, we partialled out the effect of plot size, which was negatively associated with the MPDses.

All analyses were run in R package (R Core Team, 2018).

5.3 **Results**

Simulations of the tropical niche conservatism and multiple zones of origin hypotheses

For the TNC hypothesis, as expected, we found a decrease of MPDses across latitude and elevation (Figure 5.3A-B, respectively). A large proportion of the simulations showed a significant negative relationship between MPDses and latitude (86.1% of simulations, mean = -0.43) significantly different than zero ($t = -21.68$, $p < 0.001$). Likewise, a large proportion of the simulations showed a significant negative relationship between MPDses and elevation also significantly different from zero (80.1%, mean = -0.33, $t = -16.08$, $p < 0.001$). For the MZO hypothesis, also as expected, we found an increase of MPDses across latitude and elevation (Figure 5.3 C-D, respectively). A large proportion of the simulations showed a significant positive relationship between MPDses and latitude significantly different than zero (88.1%, mean = 0.53, $t = 12.31$, $p < 0.001$) as well as a large proportion of significant positive relationship between MPDses and elevation also significantly different from zero (85.8%, mean = 0.36, $t = 13.14$, $p < 0.001$). For both hypotheses, the effect of latitude was stronger than the effect of elevation (Figure 5.3). The coefficients of the regression of MPDses against latitude and elevation were significantly different between TNC and MZO hypotheses coefficient distribution of both hypotheses differs significatively (Dlat = 0.78, $p < 0.001$; Delev = 0.70, $p < 0.001$). Then, our theoretical predictions for both TCN and MZO hypotheses (Figure 5.1) were supported by the computer simulations (Figure 5.3).

Phylogenetic diversity along elevation and latitude in Andean tree communities

As predicted by the MZO hypothesis, the standardized effect size of the mean pairwise phylogenetic distance (MPDses) increased with latitude and elevation (Figure 5.4). The regression of MPDses against latitude, after partialling out the plot size effect, showed a coupled positive increase (slope = 0.11; F = 17.62; R2 = 0.07; $p < 0.001$) (Figure 5.4A). Similarly, the regression of MPDses against elevation, after partialling out the plot size effect, was also positive and significant (slope = 0.45; F = 30.10; R2 = 0.12; $p < 0.001$) (Figure 5.4B).

5.4 **Discussion**

Our results shed strong support for the predictions of the multiple zones of origin hypothesis (MZO) as a primary structuring mechanism of the species assemblages along the latitudinal and elevational gradient in tropical and subtropical Andean forests. The consistent increase of the phylogenetic diversity (assessed by the MPDses) with latitude and elevation, confirms the historical mix of clades with different origin (i.e., tropical – temperate) as an important component of the actual variation in species richness and composition of tropical mountain ecosystems (Gentry 1982; Ramírez et al. 2019, Linan et al. under review). In the tropical and subtropical Andes, the immigration of very old clades that evolved in the meridional part of South America, such as Nothofagus, Drymis, Podocarpus, and Weinmania, as well as the immigration of septentrional clades, such as Quercus, have played a key role on determining the structure and functioning of montane forests (González-Caro et al., 2020). Therefore, the MZO hypothesis appears as an important paradigm to improve our understanding of the latitudinal and elevational gradient of tree phylogenetic diversity, and thus, the astonishing and highly endangered woody diversity harbored by the Andean region (Myers et al. 2000, Malizia et al. 2020).

Previous studies had emphasized on the importance of the tropical niche conservatism hypothesis (TNC) as a main determinant of the community assembly along the Andean corridor (Segovia et al. 2020). The TNC is an evolutionary model widely invoked to explain the latitudinal diversity gradient (LDG) assuming a single zone of origin (i.e., tropical conditions) and infrequent species dispersal to other latitudes with contrasting climates, such as subtropical or temperate, due to strong niche conservatism (Wiens & Donoghue 2004). Under the TNC, the expected reduction of the phylogenetic diversity along with the increase of elevation results from the purported infrequent colonization from lowland clades, and predicts speciation alongside elevation (i.e., ecological speciation) as the primary driver of species accumulation along mountainsides. However, adaptive radiation of tree clades alongside elevation on the tropical Andes seems to have been less frequent than expected, which by itself seems to be insufficient to explain the assembly of tree communities across the tropical Andes (Ramírez et al., 2019; Linan et al. under review).

The increase of the phylogenetic diversity from tropical to subtropical Andes may be largely due to the expansion of the tropics in the early Holocene (~60My) and the subsequent contraction later on (~35My), which favored the mixing of floras originated under different climatic regimes (Jaramillo et al. 2006, Jaramillo & Cárdenas 2013). Our findings of an increase of the phylogenetic diversity southwards in the American continent differed of those previously found northwards, where a reduction of the phylogenetic diversity was reported (Kerkhoff et al., 2014). One likely explanation for this difference was the relatively recent connection between south and north America through the Panama Isthmus closing around ~15My (Montes et al., 2015), which prevented an earlier northward expansion of the already very diverse Amazon/Andean forests. Within the wholly isolated South America before the Panama Isthmus closing, a more active historical flux of propagules exchange was highly benefited by the Andean uplift (Gentry 1982; Graham et al., 2009; Hoorn et al., 2010), a process that paralleled the formation of the current Neotropical forests during the Paleogene (Carvalho et al. in press). That said, in South America, the tropical-temperate connection was persistent along geological time, facilitating migration among the contrasting tropical and temperate environments (Jaramillo & Cárdenas 2013).

The increase of the phylogenetic diversity along elevation, pinpoints to an evolution of the ancestral niche of pre-adapted temperate (no tropical) originated species, which by definition contradicts the expected low rate of niche evolution of tropical clades purported to control the eco-evolutionary forces that structure the current communities of tree species upslope in the Andean mountains (i.e., TNC). By saying this, we don't aim to deny the existence and importance of the TNC hypothesis as a mechanism that helps to explain the patterns of diversification both inside and out of the tropics (Wiens & Donoghue 2004). Here, we rather emphasize on the need of acknowledging the role played by historical patterns of immigration not only from the tropics towards the temperate zone, but also in the opposite direction (Qian and Ricklefs 2016; González-Caro et al., 2020), on shaping the assembly of tree communities along climatic gradients. Our simulated and empirical findings bring out the need of changing the magnified interpretation of the actually quite low liability of tropical traits evolution as the main driver of the latitudinal and elevational gradient of species diversity (Algar et al., 2009; Kerkhoff et al., 2014).

An open question would be to define the main evolutionary traits that characterize functional differences between tropical and temperate originated clades. It seems that along the elevational gradient, tropical originated clades have to reduce total mass and size to survive in the colder and more stressful, less productive highlands. On the contrary, for temperate originated clades the pre-adaptations to harsher climates, such as tropical highlands, do not necessarily compromise size (Culmsee et al., 2010; González-Caro et al., 2020). These differences in size show that temperate species may have physiological adaptations to coldness that help them to have metabolic rates similar to those typical of lowland species (Peña et al. 2018). For example, we hypothesize that temperate originated clades may have a narrower vessel diameter than their tropical counterparts, which decreases the risk of cavitation and embolism. To test this hypothesis, we re-analyzed the dataset available in Olson et al. (2018) including the species' biogeographic origin as a covariate in their model of plant height – vessel diameter. We found that temperate originated species had narrower vessels than tropical ones at the same height (Fig. S5.1). This result suggests physiological and morphological differences between tropical and temperate clades that allow temperate tree species to maintain large size in cold conditions. Advancing our understanding on the evolutionary differences between clades that could confer them adaptive advantages to certain climatic conditions (e.g., tropical and temperate) can help to predict the response of tropical mountain communities to climate change.

In South America, the gymnosperms (e.g., Aracauriaceae) were a dominant feature during the Maastrichtian, which may have left its fingerprint on the composition of the subsequent neotropical forest formation after the massive Cretaceous extinction event (Carvalho et al. in press). A high abundance of gymnosperms could distort or enhance the length of the branches in the phylogenetic tree, and thus, the assessment of the phylogenetic diversity. Thus, to assess the robustness and validity of our results, we reduced our taxonomic scope to only angiosperms. However, our results using only angiosperms remained very much the same than those employing gymnosperms + angiosperms (Figure S5.2). These findings show that the effect of gymnosperms was not the cause of the pattern of increase of the phylogenetic diversity along with latitude or elevation. (Figure S5.2). Then, our claims about the MZO hypothesis as an underlying mechanism of the tree community assembly along the latitudinal and elevational gradient in Andean forests is well supported.

The phylogenetic diversity patterns are influenced by the phylogenetic tree's shape, which ultimately depends on the speciation and extinction rates of clades (Ricklefs 2007). We used different combinations of speciation and extinction rates in our computer simulation to assess different phylogenetic tree shapes' effect in our conclusions. We did not find a signal of the effect of speciation and extinction rates on the patterns of phylogenetic diversity along latitude and elevation in both hypotheses, TNC and MZO (Figure S5.3). Then, these phylogenetic diversity patterns emerge from dispersal among regions rather than by the effect of differences on speciation or extinction rates. These results were consistent with recent theoretical and empirical studies. Etienne et al. (2019) used a simple simulation model to assess the effect of speciation, extinction, dispersal, and area on LDG. They found a little effect of speciation and extinction promoting LDG. Recent studies on plants and vertebrates showed that diversification rates (i.e., speciation minus extinction) are similar across latitude or increase poleward (but see Schluter and Pennell 2017, Rabosky et al., 2015; Igea & Tanentzap, 2020). Then, our findings of the increment of phylogenetic diversity along latitude and elevation could be robust to variations on diversification rates.

In conclusion, the MZO hypothesis and the tropical–temperate connection formed by the climatic similarity between Southern temperate and Tropical highland regions along evolutionary history may have played a paramount role on structuring the Andean flora. The phylogenetic diversity increases alongside elevation and latitude found in both the simulated and the empirical data does not support the TNC hypothesis as a universal or exclusively structuring mechanism of the latitudinal and elevational gradients. The tropical-temperate mixing of floras alongside latitude and elevation can represent contrasting ecological strategies to respond to climate change that should be considered in future assessment of the response of Andean tree communities to climate change (e.g., Duque et al. 2015, Fadrique et al. 2018). We argued that a hypothesis that explicitly includes the historical context of communities, such as the MZO hypothesis, could advance our knowledge of the mechanisms that shape tree community assembly from local to regional scales (Ricklefs 2006).

Figure 5.1 Two hypotheses for the historical assembly of communities across elevations in temperate and tropical latitudes. Tropical niche-conservatism hypothesis (A-C) and Multiple Zone of Origin hypotheses (D-F). Under both hypotheses, niche conservatism limits adaptive shifts and colonization among regions with different environmental conditions (thin arrows) and facilitates colonization among regions with similar environmental conditions (thick arrows). However, the hypotheses differ in whether colonization results from a single (tropical) or multiple (tropical + temperate) ecological zones of origin. Under the TNC hypothesis, clades originate in tropical lowlands (TrL) and have limited colonization of temperate lowlands (TeL) and tropical highlands (TrH). Once clades have evolved adaptions to temperate lowlands, they can colonize temperate highlands. This hypothesis predicts that phylogenetic diversity decreases with elevation and from tropical to temperate latitudes. Under the MZO hypothesis, clades originate in both tropical lowlands and temperate lowlands. Clades that evolved in tropical lowlands follow the same colonization history as in the first hypothesis. In addition, some clades originate in the temperate lowlands and readily colonize the temperate and tropical highlands but have limited colonization of tropical lowlands. This hypothesis predicts that phylogenetic diversity increases with elevation and from tropical to temperate latitudes. For simplicity, the two hypotheses assume identical rates of speciation and extinction among regions.

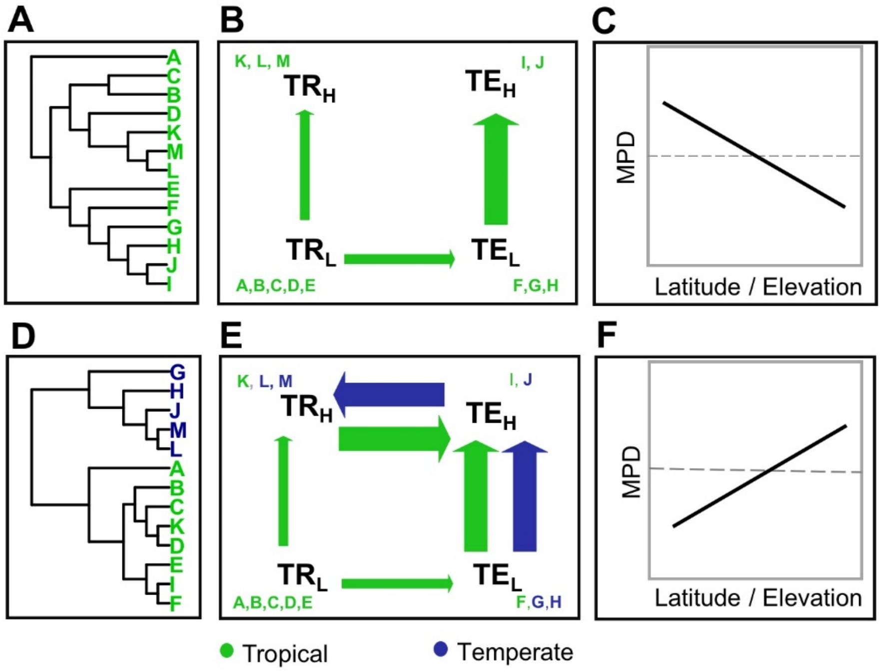


Figure 5.2 Distribution of Andean-forest plots used in this study represented by orange points along elevational, latitudinal, and temperature gradients. (A) Elevational profile of the Andes in South America (gray area). Squares represent sampled regions and orange circles represent forest plots. green colors represent warmer, more unseasonal sites and blue colors represent colder, more seasonal sites. Colors are based on the first axis of a principal component analysis of 100,000 random points excluding non-forested areas. (B) Spatial distribution of plots in South America. Gray areas represented mountains over 1000 masl and dark green showed forested areas.

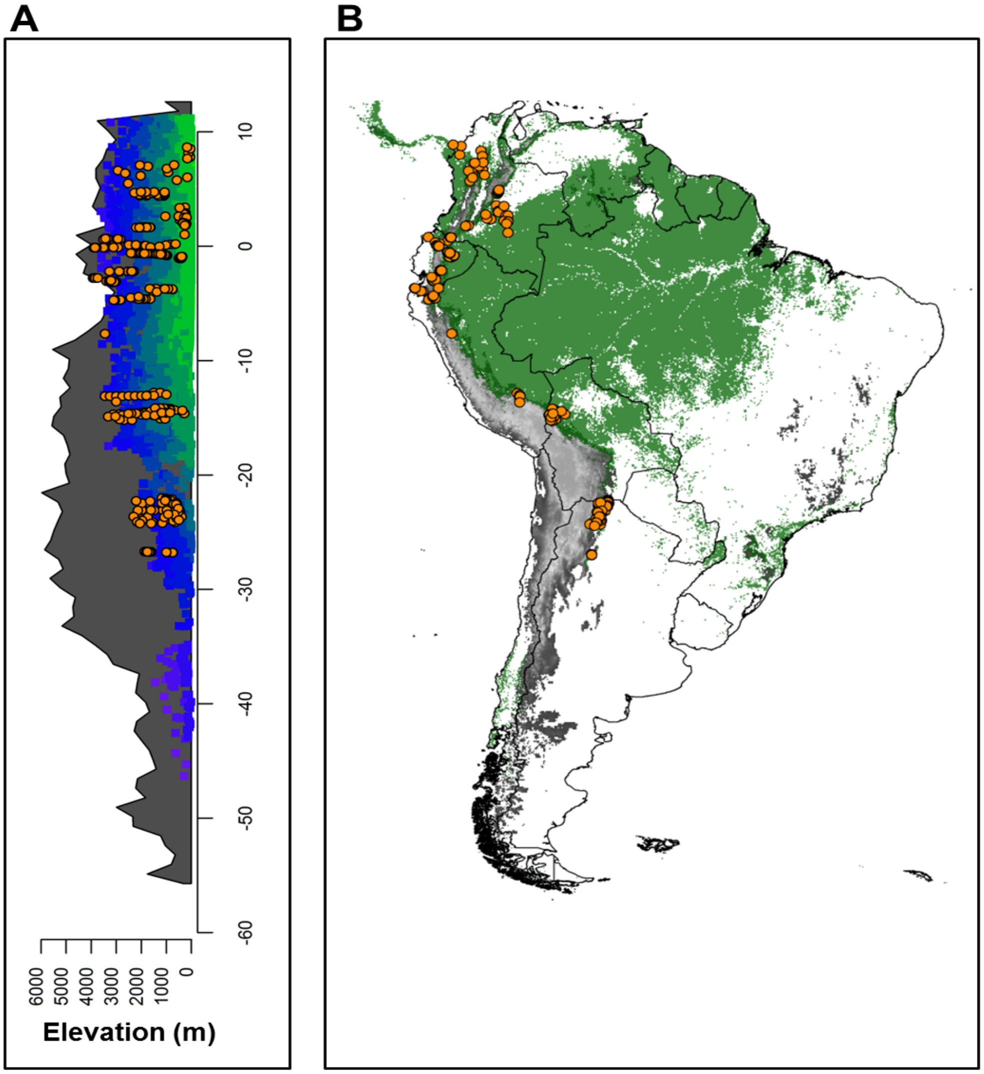


Figure 5.3 Variation of simulated standardized effect size of mean pairwise phylogenetic distance (MPD_{ses}) across latitude and elevation obtained from simulations of the Tropical Niche Conservatism hypothesis (TNC; A and B) and the Multiple Zone of Origin hypothesis (MZO; C and D), respectively. The black lines showed the main effect of the x-axis variable and gray lines represent simulations between 0.25 and 0.75 quantiles. The equation is based on the average of regression coefficients of 1000 simulations.

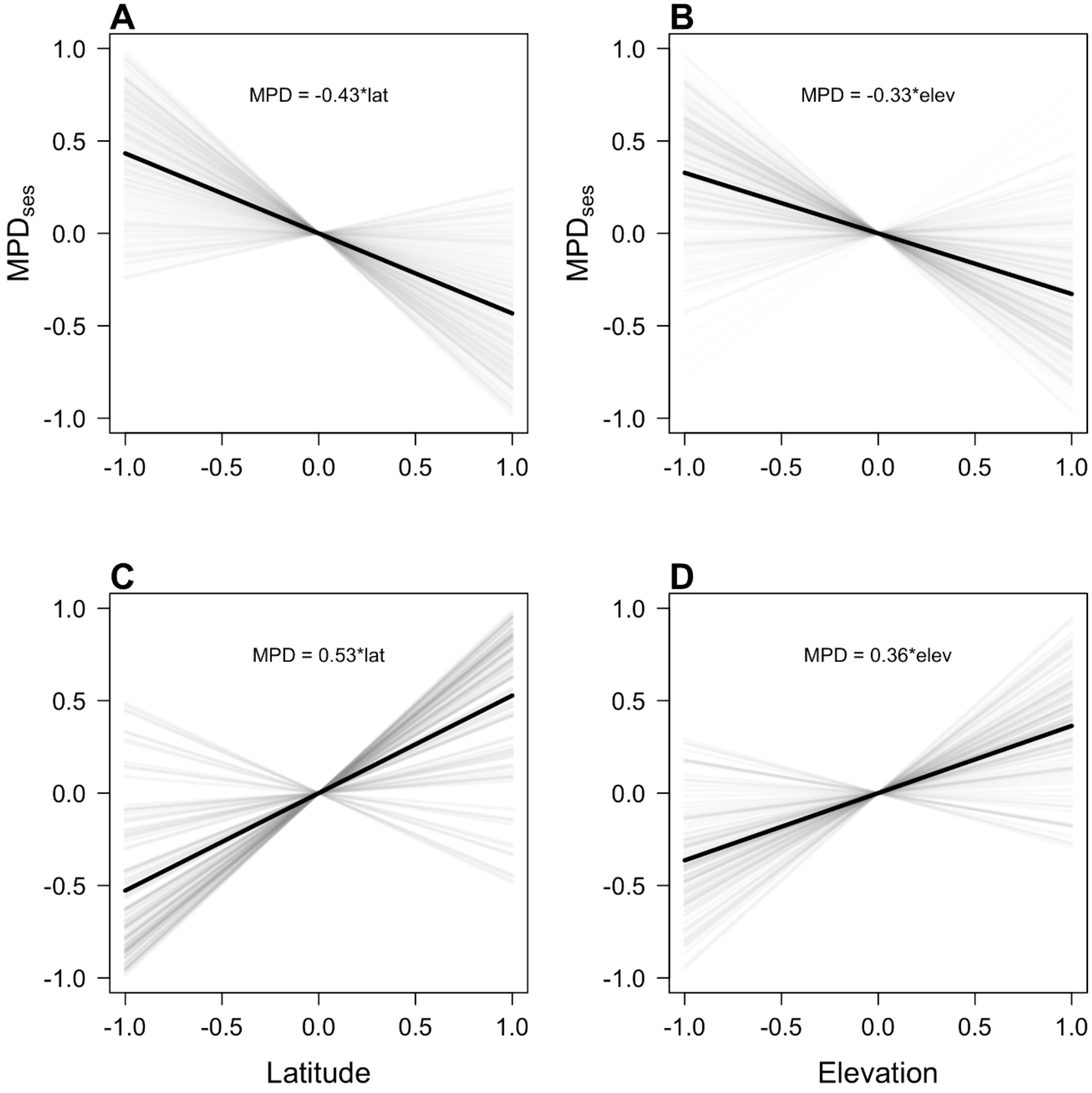
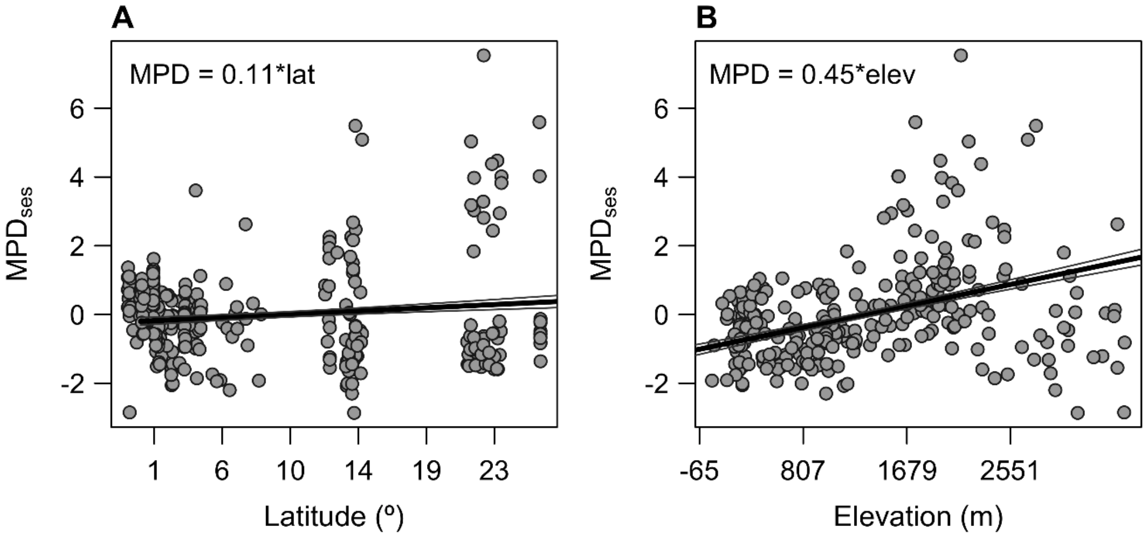


Figure 5.4 The standardized mean pairwise phylogenetic distance among co-occurring species in each plot (MPD_{ses}) after control by the effect of plot area (i.e., residuals of a regression between MPD_{ses} and plot area) along latitude (A) and elevation (B). The black lines showed the main effect of the variable and the gray lines indicates their confidence intervals.



6 Chapter 6: Conclusions

Vellend´s framework propose four main processes of community assembly: speciation, dispersal, selection, and ecological drift (Vellend 2016). These could be separated into two classes of processes that, from one side, add species to communities (speciation and dispersal), and from the other side subtract species from communities (selection and drift). Speciation is the main process to increase species number at a global scale. However, dispersal is also important and promotes species accumulation at local scales through the colonization of new environments and the mixing of different evolutionary histories, which increases opportunities of speciation. Therefore, dispersal is a key mechanism that can unify and facilitate the operation of the other three processes (Hubbell 2001, Claramaunt et al., 2012; Dexter et al., 2017, González-Caro et al., 2020). Based on this assumption, a simple equation to describe community assembly would be:

$$\text{Community assembly} = \text{Speciation} + \text{Dispersal} - \text{Selection} - \text{Drift}$$

Speciation plays a primary role at continental scales; then, it could be disregarded in local community studies. In contrast, community drift is determinant at a local scale, a spatial scale in which stochasticity matters. The impact of ecological drift and its consequences may be ruled out along at large spatial scales (Vellend 2016). Assuming that speciation is not correlated to species richness across worldwide (Schulter & Pennell 2017) and that ecological drift can be disregarded at large spatial scales, the effect of dispersal and selection and their interaction may drive community assembly across broad environmental gradients. We can assume two scenarios: i) Dispersal effects override selection effects; and ii) Selection effects override dispersal effects. In the first scenario, high dispersal maintains connections among local communities and homogenized community composition at least at deep phylogenetic levels, while selection (via habitat filtering or biotic interactions) shapes their distribution at local scales. This scenario can promote local adaptation, population divergence, and therefore, ecological speciation. In the second scenario, species arrive at one locality and selective pressures, such as habitat filtering or biotic interactions (e.g., competition), reduces the probability of establishment. The strong effect of selection may reduce the number of new species, and ultimately, the effect of ecological speciation (e.g., adaptative radiations across environmental gradients). Then, if dispersal overrides selection, the interaction can promote speciation across the metacommunity (Dexter et al., 2017; González-Caro et al., 2020). In contrast, if selection overrides dispersal, then, speciation could be reduced across the metacommunity and immigration from other metacommunities become an important process to add new species (Linan et al., under review; González-Caro et al., in prep). Thus, the trade-off between these two processes can contribute to ecological speciation and will shape the spatial patterns of biodiversity. In the two regional tree communities of the Neotropics analyzed, we found support for this model.

For the Amazon forests, Dexter et al. (2017) claims that the whole Amazon basin, defined as the metacommunity (i.e., all sites are connected and interact via dispersal over long periods), may homogenized the phylogenetic composition of local tree communities. We added evidence to this hypothesis evaluating in two independent datasets the phylogenetic similarity between local tree communities (Cárdenas et al., 2017; González-Caro et al., 2020). At the Western Amazon scale, phylogenetic similarity was primarily driven by the spatial variation among plots, which was associated to historical dispersal. In addition, we found that the relative importance of the environmental conditions, such as the geological substrate and anoxia, increases the non-random distribution of species at local scales in relation to the environmental axes evaluated. Based on these findings, we suggest that: i) historical dispersal is an active process across the Amazon basin; and ii) local conditions, such as edaphic heterogeneity, shape the spatial distribution of species. If one species spread and increases its geographic range, and then, the local conditions generates local adaptation or extinctions, this process could ultimately derive in multiple species. Under this view, dispersal overrides selection and could promote opportunities for ecological speciation.

The Andean uplift produced new environments that were occupied by tropical lowland species and temperate immigrant species (Ramírez et al., 2019; González-Caro et al., 2020; Segovia et al., 2020; Linan et al., under review). In this system, the frequency of dispersal from each ecological zone of origin was mediated by ecological barriers and the historical connectivity between biomes (González-Caro et al., in prep). One the one hand, the ecological barriers could be filtering out species with particular traits that able them to persist in the new environments. In the case of the

Andes, small lowland trees could have a greater chance to colonize the high mountains than relative large trees because they may also have a lower probability of embolism due to physiological adaptations, such as narrower vessels. However, some temperate clades evolved adaptations to maintain large size but small vessels relative to their size. Then, clades with large-sized trees may have developed particular adaptation to colonize highlands after the mountain uplift. For example, large-sized and specious clades originated in the Amazon, such as *Eschweilera* or *Brosimun* among others, are rare in the Andes. Under this view, selection overrides dispersal and could reduce speciation chances. On the other hand, temperate pre-adapted clades that colonize tropical Andes increases their speciation chance mediated by their dispersal ability. The historical connectivity of south and north temperate regions with the tropical Andes is different: while the southern part has been constantly connected, the Northern part was recently connected after the Panama Isthmus closing. Then, the colonization of both hemispheres into the tropical Andes is asymmetric, which means that pre-adapted clades from the Southern part of South America should be more frequent that North American immigrant because they may have had a longer period of connectivity and genetical interchange. It is exemplified by the distributions of the *Weinmmania* and *Quercus* genera in the tropical Andes. *Weinmmania* was originated in the Southern part of South America and accumulated many species in the tropical Andes, while *Quercus* is typical a North American immigrant that only have one species restricted to the Northern Andes (i.e., Colombia).

In summary, historical assembly models of tree communities should include multiple ecological zones of origin (mainly based on climatic reconstructions), the historical connectivity among ancestral ecological zones, and key ecological traits that allow colonization of new environments and increase diversification, to understand the actual spatial patterns of biodiversity. In addition, the inclusion of phylogenetic patterns in local communities such as the phylogenetic structure and phylogenetic similarity are useful to assess the role of historical dispersal across environmental gradients on shaping community assembly and diversification. Assessing the trade-off between dispersal and selection at large spatial and temporal scales can help us to disentangling community assembly mechanisms. Finally, we suggest that include explicitly the role of history on the analysis of community assembly is fundamental to understand broad scale diversity patterns (e.g., Latitudinal and elevational diversity gradients) and the potential response of species to environmental change.

7 Annex A

List of additional publications or manuscripts under review:

Martínez-Villa, J. A., **González-Caro, S.,** & Duque, Á. (2020). The importance of grain and cut-off size in shaping tree beta diversity along an elevational gradient in the northwest of Colombia. Forest Ecosystems, 7(1), 1-12.

Duque Á:, Peña M., Cuesta F., **González-Caro, S.,** Kennedy P., Phillips O., Calderón M., Blundo C., Carilla J., Cayola L., Farfán-Ríos W., Fuentes A., Grau R., Homeier J., Loza M., Malhi Y., Malizia A., Malizia L., Martinez-Villa J., Myers J., Osinga O., Peralvo M., Pinto E., Saatchi S., Silman M., Tello S., Terán A., Feeley K. Andean forests as globally important carbon sinks and future carbon refuges. (Accepted in Nature Communications)

Giraldo J., del Valle J., **González-Caro, S.,** Sierra C. Isotope signatures in tree rings reveal growth rhythms occurring in the least rainy season in a hyper-humid neotropical forest. (Under review in Trees)

Idárraga-Piedrahíta, Á., **González-Caro, S.,** Duque, Á., Jimenez, J., González, R., Parra, J.L., Rivera-Gutierrez, H. Drivers of beta diversity along a precipitation gradient in tropical forests of the Cauca River Canyon in Colombia. (Under review in Journal of Vegetation Science)

Linan, A. G., Myers, J. A., Edwards, C. E., Zanne, A. E., Smith, S. A., Arellano, G., ...**González-.Caro, S.**…, & Tello, S. J. (2021). The evolutionary assembly of forest communities along environmental gradients: recent diversification or sorting of pre-adapted clades?. bioRxiv, 2020-12. (Under review in New Phytologist).

8 Annex B

Chapter 2

Scale-dependent drivers of the phylogenetic structure and similarity of tree communities in northwestern Amazonia

Figure S2.1. Spatial variation of the principal coordinate neighbor axes estimated from the geographic coordinates of the plots used in this study. Plots are ordered from west to east in the x-axis.

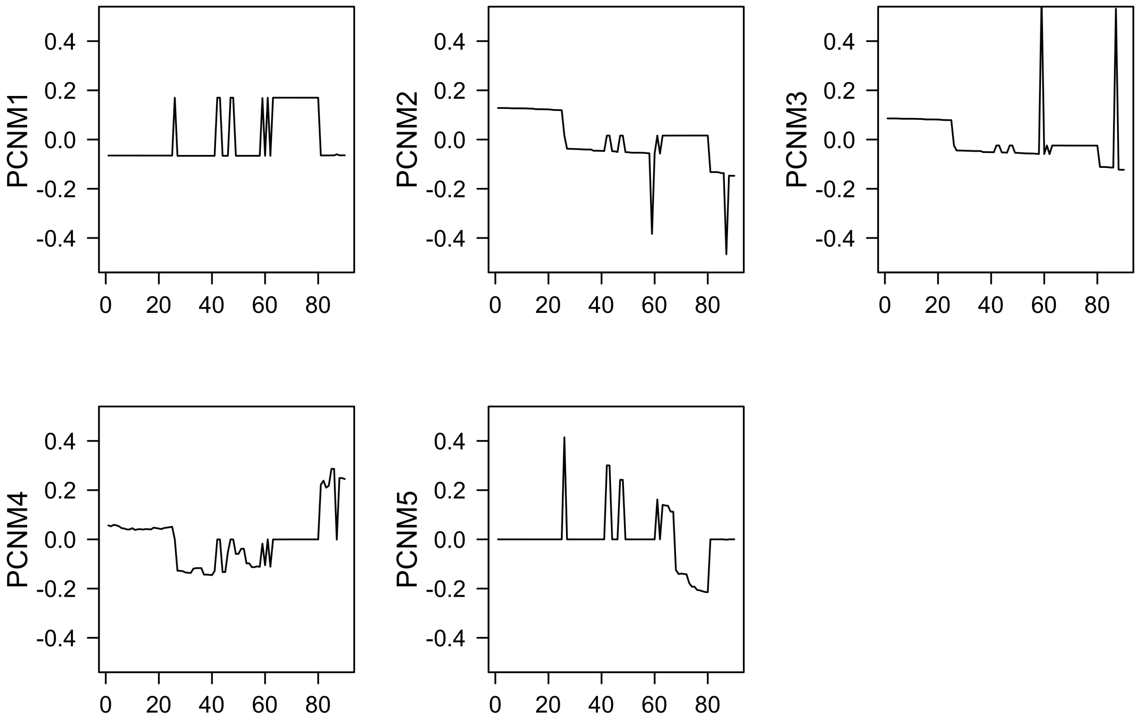


Figure S2.2. Proportion of NRI and NTI values calculated to tree communities of adult trees with that were significantly clustered or evenness under the three randomization models (Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit). Light gray represents the proportion of plots significantly clustered. Dark gray bar represents the proportion of plots with a significant evenness pattern. White bar represents the plots with either NRI or NTI values that were not significantly different from random. FP = Flood plain. SW = Swamp. TF = Terra Firme. Metá-Chiribiqute = 1. Yasuní = 2. Ampiyacu = 3. Numbers in parenthesis represent the total of number of plots evaluated.

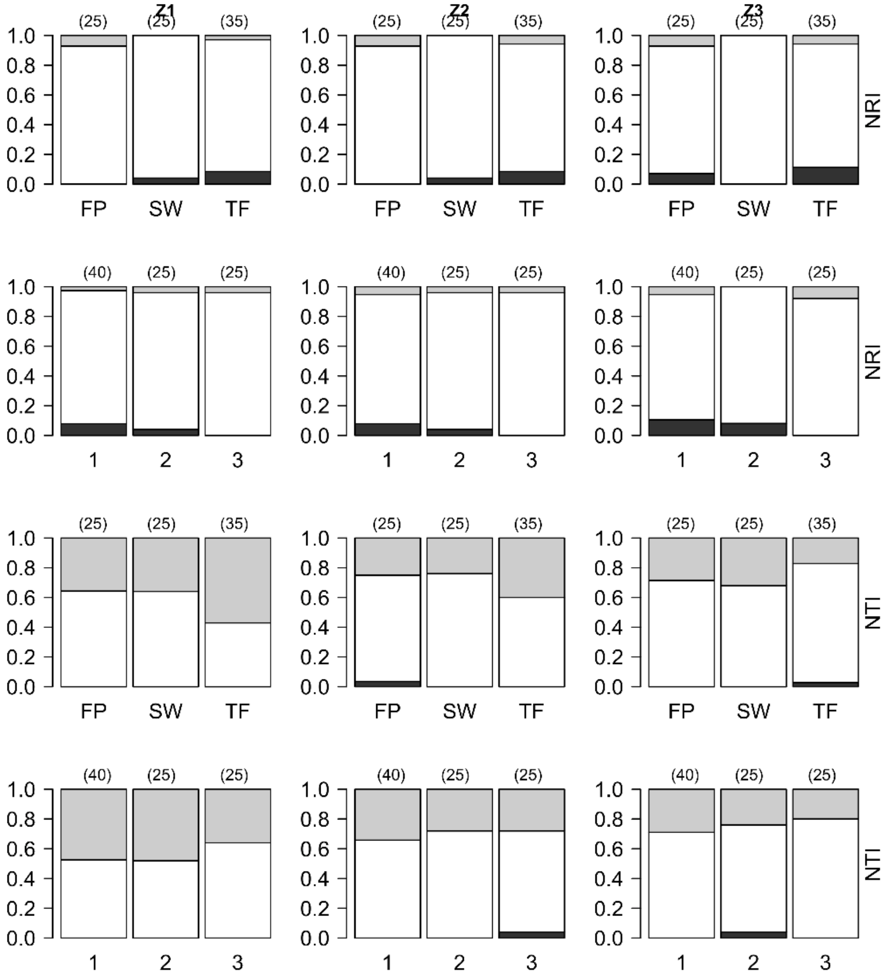


Figure S2.3. Proportion of NRI and NTI values calculated to tree communities of juvenile trees with that were significantly clustered or evenness under the three randomization models (Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit). Light gray represents the proportion of plots significantly clustered. Dark gray bar represents the proportion of plots with a significant evenness pattern. White bar represents the plots with either NRI or NTI values that were not significantly different from random. FP = Flood plain. SW = Swamp. TF = Terra Firme. Metá-Chiribiqute = 1. Yasuní = 2. Ampiyacu = 3. Numbers in parenthesis represent the total of number of plots evaluated.

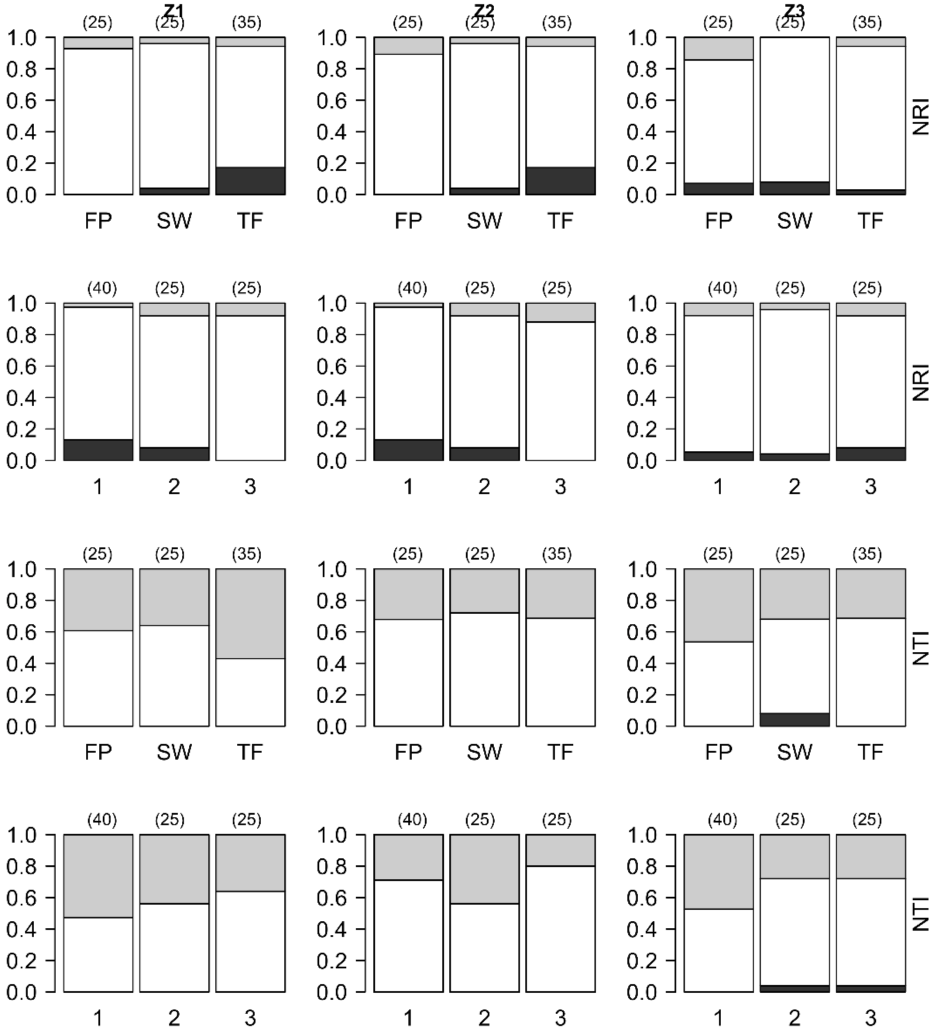


Figure S2.4. Non-metric multidimensional scaling ordination analysis of phylogenetic similarity tree communities including adult trees under different scenarios of species pools used in this study. Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit. Symbols denoted subregions: circle: Metá-Chiribiquete; triangle: Yasuní; square: Ampiyacu. Colors denoted landscape units: Flood plains: orange; Swamp: blue; Terra Firme: green.

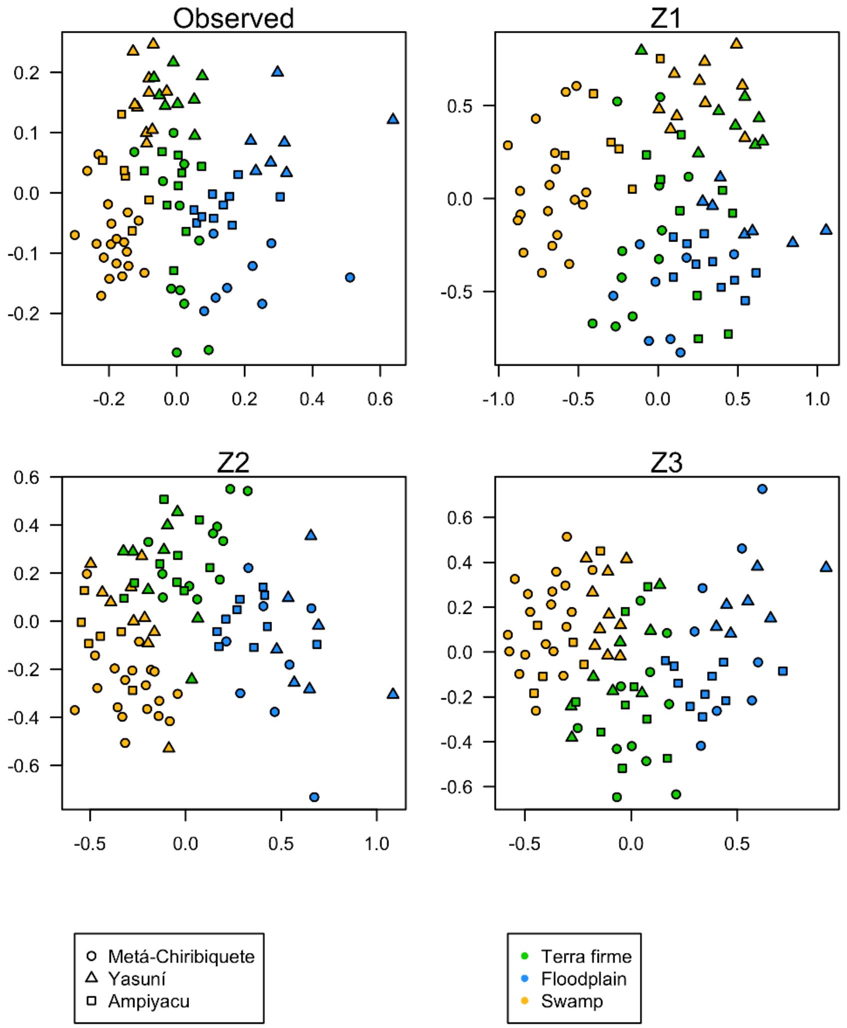


Figure S2.5. Non-metric multidimensional scaling ordination analysis of phylogenetic similarity tree communities including juvenile trees under different scenarios of species pools used in this study. Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit. Symbols denoted subregions: circle: Metá-Chiribiquete; triangle: Yasuní; square: Ampiyacu. Colors denoted landscape units: Flood plains: orange; Swamp: blue; Terra Firme: green.

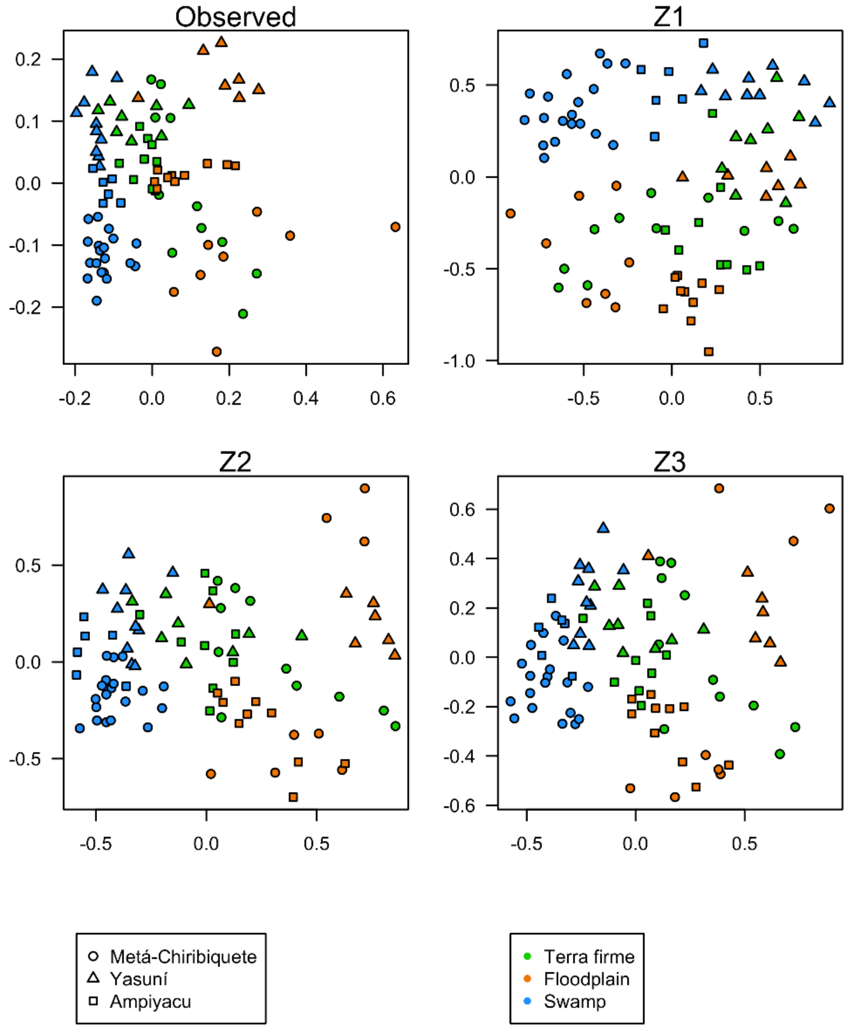


Figure S2.6. Pearson correlation coefficient between the the first and second axes of the non-metric multidimensional scaling ordination. Correlation values are assessed for standardized effect size of phylogenetic β-diversity among the three different species pools. The upper panel showed the Pearson correlation between each pair of axes. Species pool randomization used in this study are: Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit.

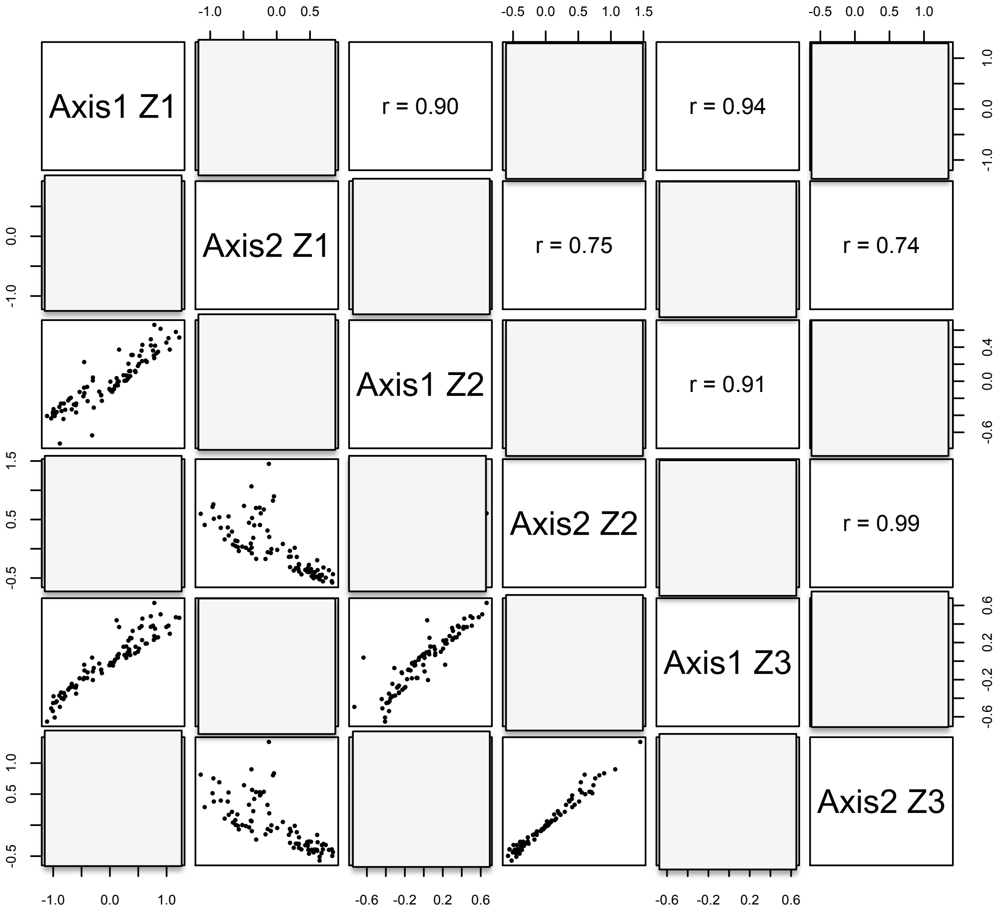


Figure S2.7. Comparison of standardized effect sizes of PhyloSor under the three species pool randomization methods employed in this study. Z1 = All data; Z2 = Restricted by subregion; Z3 = Restricted by landscape unit. Different letters indicate significantly different groups based on the Tukey honest significant differences test.

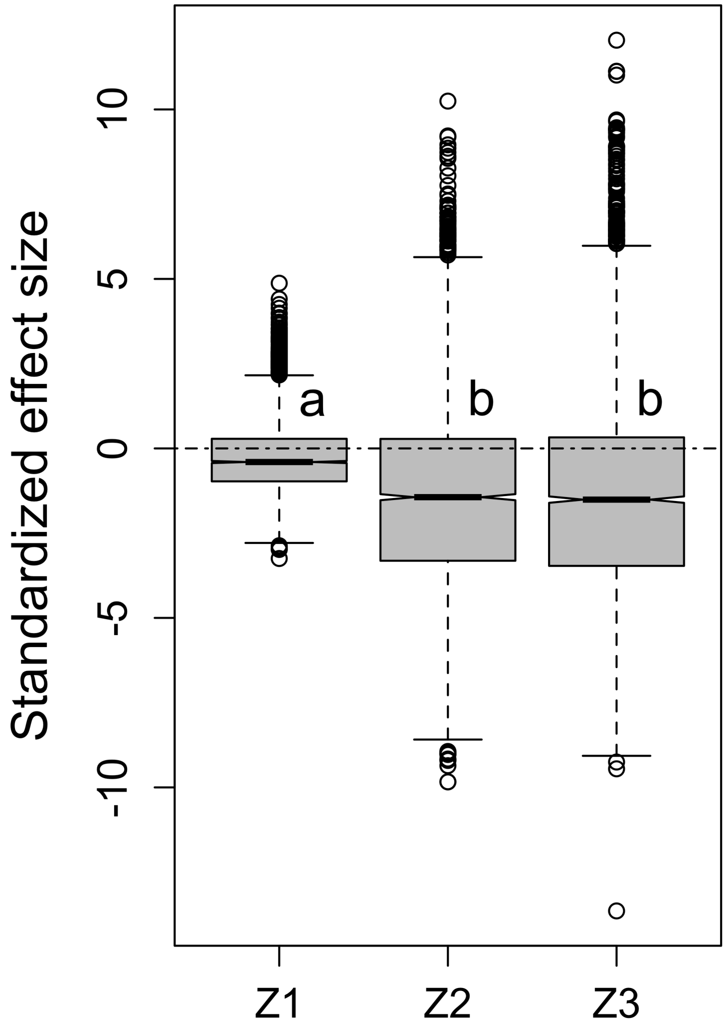


Table S2.1. Summary of the plot inventory and soil data for subregion and landscape included in this study. FP: Flood Plain. SW: Swamp. TF: Terra Firme. Plots: Number of plots. Stems: Total number of Stems. Adult trees: Total number of Adult stems. Juvenile trees: Total number of juvenile stems. Shrubs: Total number of shrub stems. S: Total Number of Species. G: Total Number of genera. F: Total number of families. Ind: Mean individual density per plot (0.1-ha). SR: Mean species richness (0.1-ha). Ca: Calcium (mmol/kg). Mg: Magnesium (mmol/kg). K: Potassium (mmol/kg). Na: Sodium (mmol/kg). P: Phosphorus (mmol/kg). C: Carbon (%). N: Nitrogen (%). C:N Ratio between Carbon and Nitrogen. Standard deviation in parenthesis. Lower case denoted significant differences among landscape units within each country. Upper case denoted significant differences among subregions.

	Metá-Chiribiquete				Yasuní				Ampiyacu			
	FP	SW	TF	Total	FP	SW	TF	Total	FP	SW	TF	Total
Plots	11	5	19	35	8	7	10	25	9	10	6	25
Stems	3068	3147	7223	13438	2229	842	3372	6443	2784	2912	1836	7532
Adult trees	619	1092	1294	3005	427	323	679	1429	497	463	383	1343
Juvenile trees	2449	2055	5929	10433	1802	519	2693	5014	2287	2449	1453	6189
Shrubs	73	58	123	254	78	61	106	245	72	63	110	245
S	472	205	1020	1349	490	231	673	886	492	347	501	814
G	252	128	373	462	251	151	319	361	240	180	230	325
F	85	49	105	120	75	58	86	91	79	61	75	88
Ind	278.90^{a}	629.4^{b}	380.15^{c}	399.52^{A}	278.62^{ab}	120.28^{a}	337.2^{b}	257.72^{B}	309.33	291.2	306	301.28^{B}
	(66.86)	(307.67)	(65.18)	(174.11)	(52.23)	(31.78)	(43.81)	(100.27)	(63.80)	(77.92)	(21.39)	(61.65)
SR	65.18^{a}	60.8^{a}	145.84^{b}	98.9^{A}	109.87^{a}	45.14^{b}	156.2^{c}	110.28^{AB}	116.55^{ab}	90.3^{a}	143.83^{b}	112.6^{B}
	(23.77)	(17.54)	(22.84)	(50.80)	(38.01)	(22.40)	(14.13)	(52.32)	(13.28)	(24.45)	(12.98)	(27.84)
Ca	77.3^{a}	4.72^{ab}	1.79^{b}	22.88	42.01	66.51	19.73	39.96	25.64	47.82	2.54	28.97
	(95.75)	(1.8)	(0.73)	(59.2)	(24.79)	(44.66)	(47.44)	(43.58)	(36.83)	(28.69)	(0.75)	(32.98)
Mg	169.38^{a}	83.15^{ab}	30.9^{b}	71.83^{A}	222.95	157.87	162.05	180.36^{B}	254.61^{ab}	289.87^{a}	127.01^{b}	238.09^{B}
	(164.66)	(34.81)	(21.68)	(106.98)	(28.14)	(60.29)	(79.78)	(66.47)	(121.99)	(70.09)	(40.15)	(106.97)
K	233.51^{a}	173.4^{ab}	59.99^{b}	114.52^{A}	222.95	164.07	155.5	179.48^{A}	367.77^{ab}	403.67^{a}	211.51^{b}	344.63^{B}
	(157.78)	(77.9)	(54.67)	(126.18)	(24.29)	(69.79)	(81.95)	(69.62)	(119.72)	(117.09)	(71.77)	(130.67)
Na	157.38^{a}	28.96^{ab}	14.54^{b}	53.93	132.02	106.45	80.77	104.36	158.38	93.28	25.57	100.47
	(190.93)	(12.67)	(14.62)	(116.95)	(41.03)	(123.11)	(68.27)	(80.72)	(160.74)	(76.94)	(4.96)	(116.27)
P	8.45^{a}	16.82^{b}	5.23^{a}	6.98^{A}	14.06^{ab}	23.58^{a}	9.85^{b}	15.04^{B}	11.65^{a}	21.98^{b}	11.32^{a}	15.7^{B}
	(4.52)	(8.19)	(1.5)	(5.74)	(3.79)	(9.85)	(2.41)	(7.96)	(3.74)	(11.88)	(10.3)	(10.35)
C	0.49^{a}	14.82^{b}	0.5^{a}	2.43	0.43^{a}	14.1^{b}	0.47^{a}	4.27	0.48	6.52	0.56	2.92
	(0.29)	(15.35)	(0.25)	(6.86)	(0.12)	(13.35)	(0.15)	(9.15)	(0.27)	(6.03)	(0.08)	(4.76)

N	0.04[a]	0.99[b]	0.05[a]	0.16	0.07[a]	0.74[b]	0.07[a]	0.26	0.05	0.34	0.06	0.17
	(0.02)	(0.7)	(0.02)	(0.39)	(0.02)	(0.67)	(0.02)	(0.46)	(0.02)	(0.28)	(0.01)	(0.22)
C:N	15.69	17.16	12.82	21.12	6.36	15.46	6.96	9.15	10.59	16.25	9.88	12.68
	(9.76)	(5.39)	(7.1)	(28.75)	(1.23)	(6.33)	(2.53)	(5.39)	(5.13)	(4.77)	(1.08)	(5.14)

Table S2.2. Loadings of the soil cations employed to assess the principal component analysis.

	Axis 1	Axis 2
Ca	0.43	0.03
Mg	0.51	0.09
K	0.46	0.08
Na	0.46	0.17
P	0.28	-0.44
C	0.05	-0.61
N	0.06	-0.60
C:N	-0.15	-0.09

Chapter three

The influence of historical dispersal on the phylogenetic structure of tree communities in the tropical Andes

Figure S3.1. The location of plots in the elevation gradient in the Eastern Cordillera of Colombia. Green points indicate the location of the plots.

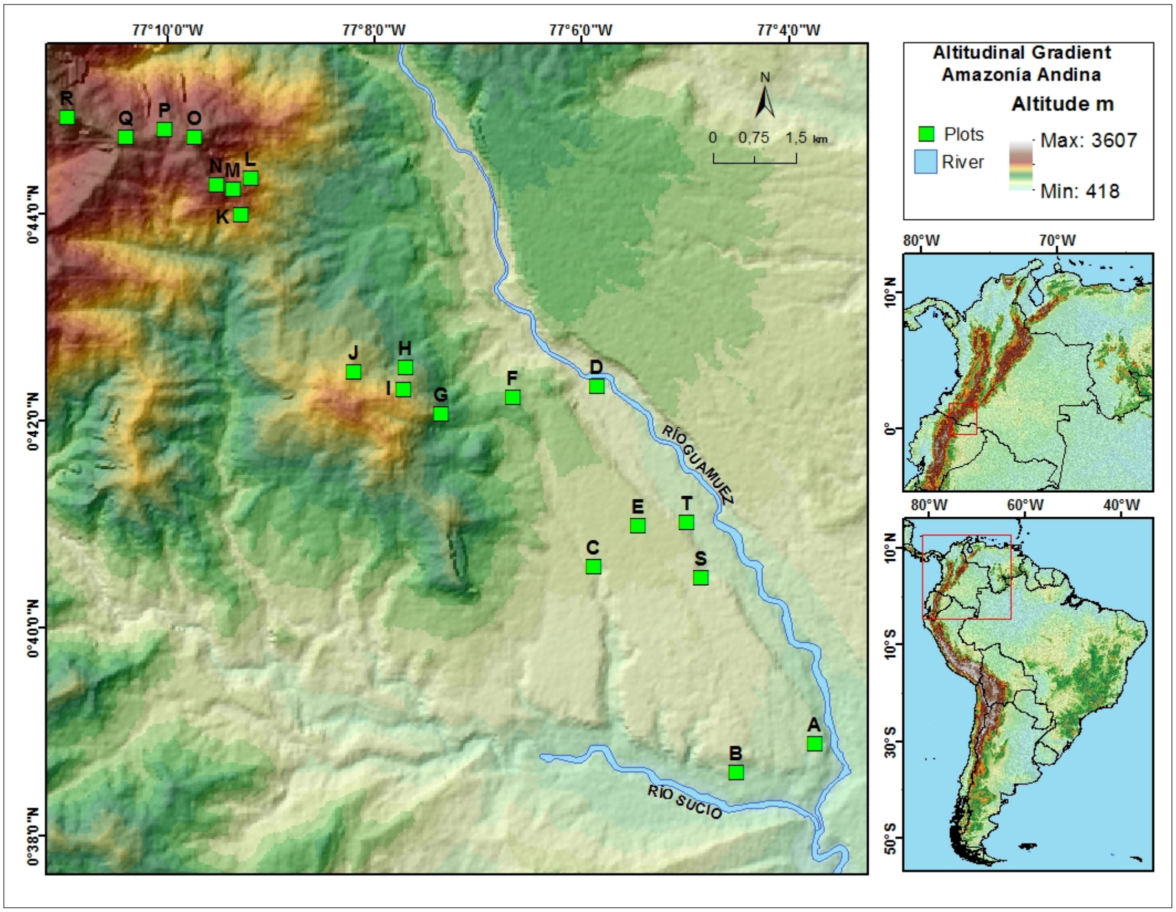
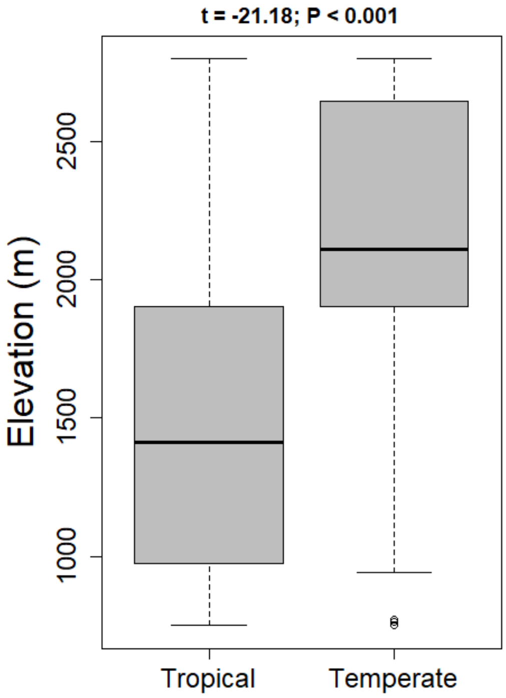


Figure S3.2. Elevational distribution of tropical vs temperate tree families classified based on Heywood et al. (2007). The t-test analysis showed significant differences in their elevational distribution.



Chapter 4

The legacy of biogeographic history on the composition and structure of Andean forests

Figure S4.1. Map of study sites denoted by the white squares and the code letter in the Table S1. The distribution of other plots used to comparison in the discussion of this study, Gentry´s plots: red circles, Culmsee et al. (2010): blue square, Venter et al. (2017): green square.

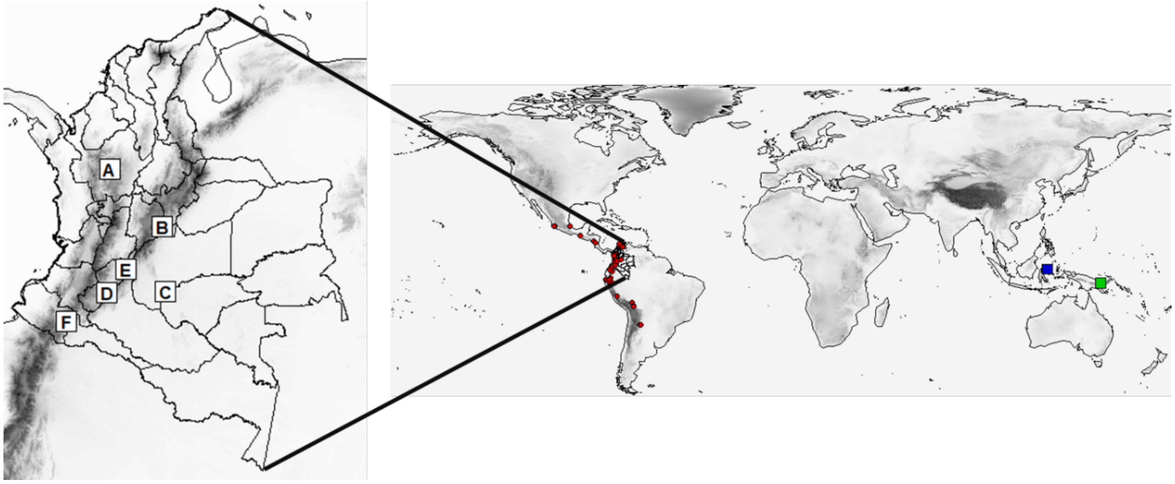


Figure S4.2. Relationship between AGB and pH in each plot based on data from 60 plots. The color showed the percentage of AGB contributed by temperate-originated species. The plots used do not include *Colombobalanus excelsus* that is the most important temperate-originated species that influence AGB stock variation.

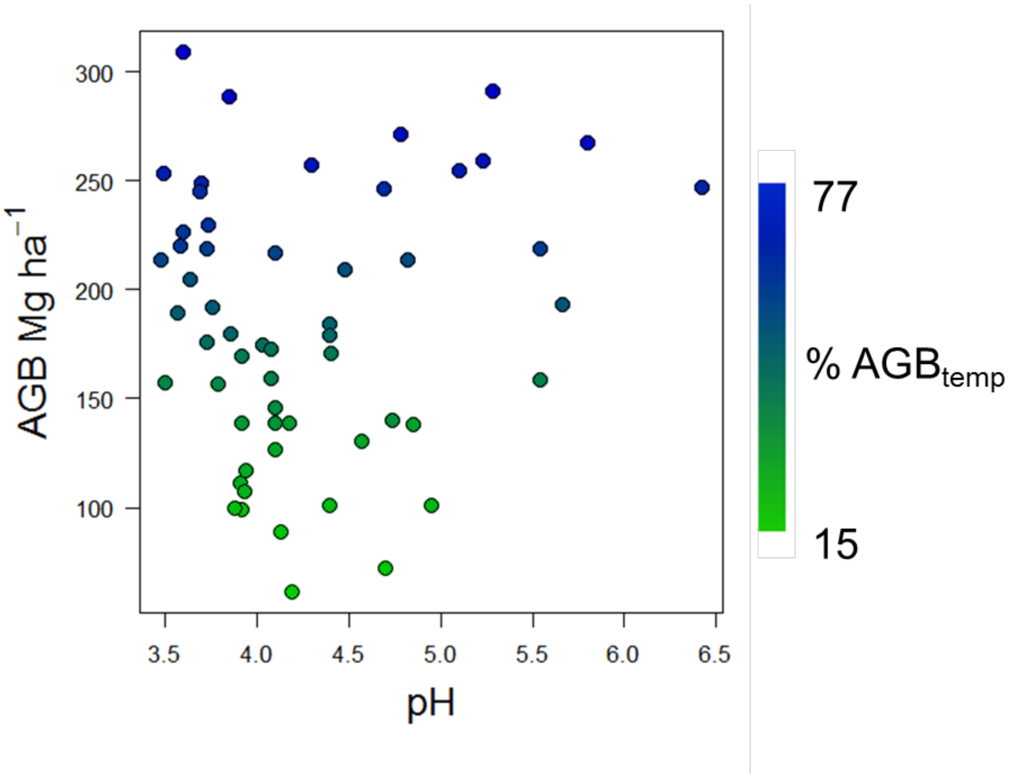
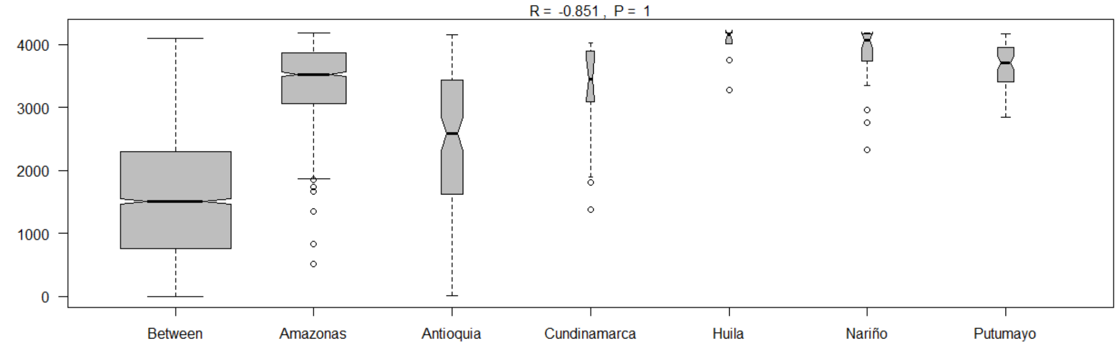


Figure S4.3. (A) Result of ANOSIM among biogeographic regions included in the analysis based on their differences in the phylogenetic composition (i.e. PhyloSorensen). (B) Phylogenetic similarity decay along elevational gradient. One potential point of concern with our results was that the observed differences in AGB between focal areas were primarily due to the high AGB in the Huila province, where C. excelsus was the dominant species. This could suggest that the significant relationship between AGB and PD is just due to this particular case, and that the observed pattern do esnot represent a more general pattern along the whole elevational gradient. To address this concern, we used ANOSIM to test for differences in phylogenetic turnover, evaluated by the PhyloSorensen index (Bryant et al. 2008), along the elevational gradient among focal areas and to evaluate the likely effect that individual focal areas have on phylogenetic compositional variation in the entire region. We found that phylogenetic turnover was higher within than among regions ($R = -0.85$, $P = 0.99$). This confirms that most of the variation in the phylogenetic composition is due to changes in elevation (or temperature) across the entire area ($R_{Mantel} = -0.61$, $P < 0.001$) rather than to differences between focal areas with contrasting biogeographic history.


R = -0.851 , P = 1
4000
3000
2000
1000
0
Between
Amazonas
Antioquia
Cundinamarca
Huila
Nariño
Putumayo

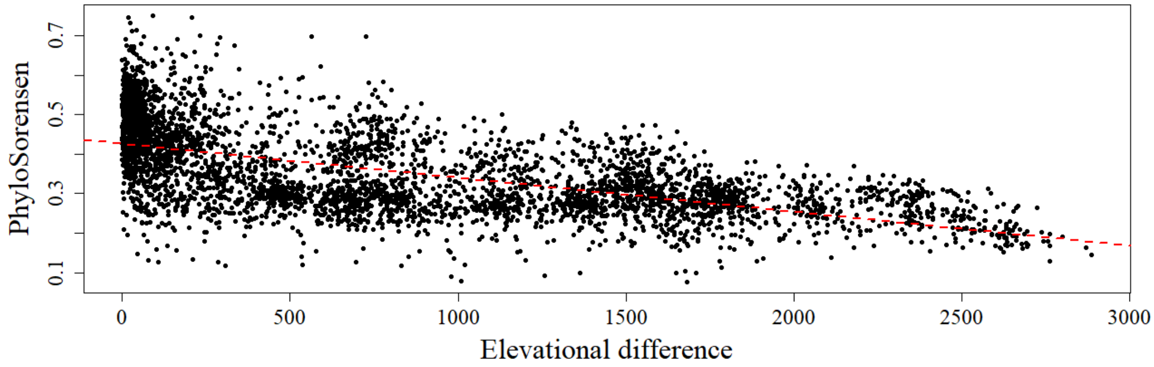

PhyloSorensen
0.7
0.5
0.3
0.1
0
500
1000
1500
2000
2500
3000
Elevational difference

Table S4.1. Average aboveground biomass estimation (AGB ± SD) and environmental characteristics of plots in Colombian provinces. Significance of differences ($P \leq 0.05$) in AGB between provinces was assessed by the Tukey Honestly Significant test. Values with the same letter represent means without significant differences. QDM = Quadratic mean diameter.

Code	**Province**	**N**	**Mean AGB (Mg ha^{-1})**	**QDM**	**Elevational range (m asl)**	**Rainfall (mm)**
A	Antioquia	16	199.6 ± 48.4^{a}	23.42 ± 2.12	41-2928	850-4250
B	Cundinamarca	16	283.7 ± 87.9^{a}	22.80 ± 2.65	1043-2260	3196-5679
C	Guaviare	18	222.4 ± 80.5^{a}	26.08 ± 2.96	190-349	2000-4000
D	Huila	15	503.6 ± 265.0^{b}	32.85 ± 9.53	1639-2204	1933-3273
E	Meta	11	222.4 ± 77.9^{a}	25.05 ± 2.88	235-1037	2000-4000
F	Putumayo	16	201.9 ± 75.2^{a}	25.46 ± 3.56	764-2802	2301-2968
	All	92	272.2 ± 117.3	25.92 ± 6.33	41-2928	850-5679

Table S4.2. Expected biogeographic affinities of Taxonomic families included in this study. These data were taken from Heeywood et al. 2007 Flowering plant families of the world.

Family (APG4)	Expected biogeographic affinity (Based on Heywood et al. 2007)
Acanthaceae	Tropical
Achariaceae	Tropical
Achatocarpaceae	Tropical
Actinidiaceae	Tropical
Adoxaceae	Tropical
Alzateaceae	Tropical
Amaranthaceae	Tropical
Anacardiaceae	Tropical
Annoceae	Tropical
Annonaceae	Tropical
Apocynaceae	Tropical
Aquifoliaceae	Temperate
Araceae	Tropical
Araliaceae	Tropical
Araliaceae	Temperate
Arecaceae	Tropical
Asparagaceae	Tropical
Asteraceae	Tropical
Berberidaceae	Tropical
Betulaceae	Tropical
Bignoniaceae	Tropical

Bixaceae	Tropical
Blechnaceae	Tropical
Boraginaceae	Tropical
Brunelliaceae	Temperate
Burseraceae	Tropical
Buxaceae	Tropical
Cactaceae	Tropical
Calophyllaceae	Tropical
Campanulaceae	Temperate
Cannabaceae	Tropical
Capparaceae	Tropical
Cardiopteridaceae	Tropical
Caricaceae	Tropical
Caryocaraceae	Tropical
Celastraceae	Temperate
Chloranthaceae	Temperate
Chrysobalanaceae	Tropical
Cleomaceae	Tropical
Clethraceae	Temperate
Clusiaceae	Tropical
Combretaceae	Tropical
Connaraceae	Tropical
Convolvulaceae	Tropical
Cornaceae	Tropical
Cunnoniaceae	Temperate

Dichapetalaceae	Tropical
Dipentodontaceae	Tropical
Ebenaceae	Tropical
Elaeocarpaceae	Tropical
Ericaceae	Temperate
Erythroxylaceae	Tropical
Escalloniaceae	Temperate
Euphorbiaceae	Tropical
Fabaceae	Tropical
Fagaceae	Temperate
Gentianaceae	Tropical
Gesneriaceae	Tropical
Goupiaceae	Tropical
Hamamelidaceae	Temperate
Hernandiaceae	Tropical
Hippocrateaceae	Tropical
Humiriaceae	Temperate
Hydrangeaceae	Tropical
Hypericaceae	Temperate
Icacinaceae	Tropical
Juglandaceae	Temperate
Lacistemataceae	Tropical
Lamiaceae	Tropical
Lauraceae	Tropical
Lauraceae	Tropical

Lecythidaceae	Tropical
Linaceae	Tropical
Loranthaceae	Tropical
Magnoliaceae	Temperate
Malpighiaceae	Tropical
Malvaceae	Tropical
Marantaceae	Tropical
Marcgraviaceae	Tropical
Melastomataceae	Tropical
Meliaceae	Tropical
Menispermaceae	Tropical
Monimiaceae	Tropical
Moraceae	Tropical
Muntingiaceae	Tropical
Myricaceae	Tropical
Myristicaceae	Tropical
Myrtaceae	Tropical
Nyctaginaceae	Tropical
Ochnaceae	Tropical
Olacaceae	Tropical
Oleaceae	Temperate
Opiliaceae	Tropical
Orchidaceae	Tropical
Passifloraceae	Tropical
Pentaphylacaceae	Temperate

Peraceae	Tropical
Petiveriaceae	Tropical
Phyllanthaceae	Temperate
Phytolaccaceae	Tropical
Picramniaceae	Tropical
Picrodendraceae	Tropical
Piperaceae	Tropical
Podocarpaceae	Temperate
Polygalaceae	Tropical
Polygonaceae	Tropical
Primulaceae	Temperate
Proteaceae	Temperate
Rhamnaceae	Tropical
Rhizophoraceae	Tropical
Rosaceae	Temperate
Rubiaceae	Tropical
Rutaceae	Tropical
Sabiaceae	Tropical
Salicaceae	Tropical
Santalaceae	Tropical
Sapindaceae	Tropical
Sapotaceae	Tropical
Simaroubaceae	Temperate
Siparunaceae	Tropical
Smilacaceae	Temperate

Solanaceae	Tropical
Staphyleaceae	Tropical
Stemonuraceae	Tropical
Styracaceae	Tropical
Symplocaceae	Temperate
Tapisciaceae	Tropical
Theaceae	Tropical
Thymelaeaceae	Tropical
Trigoniaceae	Tropical
Ulmaceae	Temperate
Urticaceae	Tropical
Verbenaceae	Tropical
Viburnaceae	Tropical
Violaceae	Tropical
Vitaceae	Tropical
Vochysiaceae	Tropical
Winteraceae	Tropical
Ximeniaceae	Tropical

Chapter five

Historical assembly of Andean tree communities

Figure S5.1. A) Relationship between plant height and vessel diameter based on the dataset from Olson et al. (2018). Black line indicates linear regression for all data employing the same model used by Olson et al. (2018). The biogeographic origin was a sinificat covariate in the model (F = 11.23; p < 0.001). The green line depicts the linear regression for tropical originated species and blue line for temperate originated species. B) Distribution of residuals of the linear regression model including all data between temperate and tropical species showed significant differences in the tree height expected by vessel diameter between biogeographic origins (t = 3.96, p < 0.001).

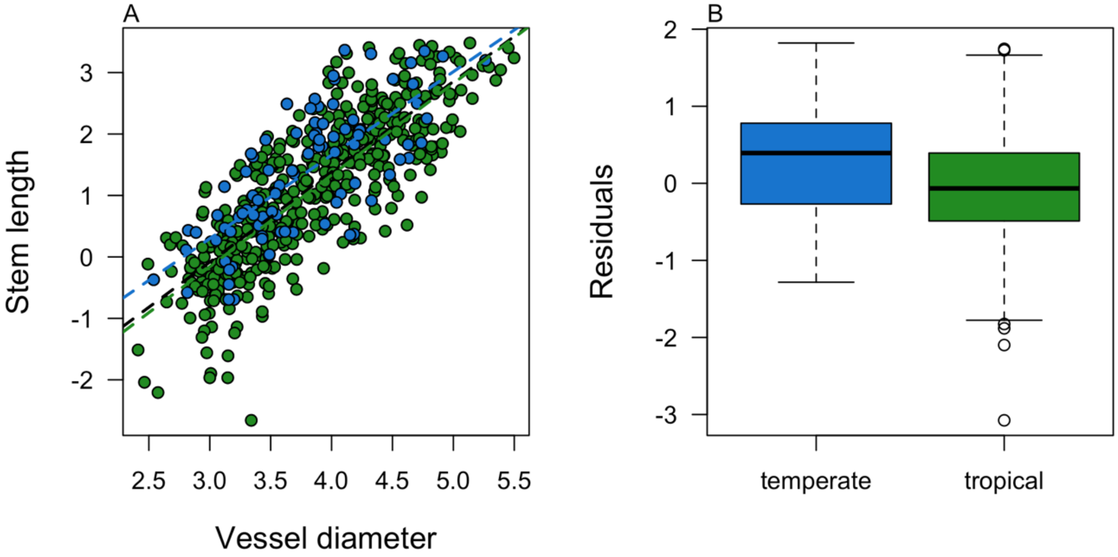


Figure S5.2. The standardized mean pairwise phylogenetic distance among co-occurring species in each plot (MPD_{ses}) after control by the effect of plot area (i.e., residuals of a regression between MPD_{ses} and plot area) along latitude (A) and elevation (B) including angiosperms alone. The black lines showed the main effect of the variable and the gray lines indicates their confidence intervals.

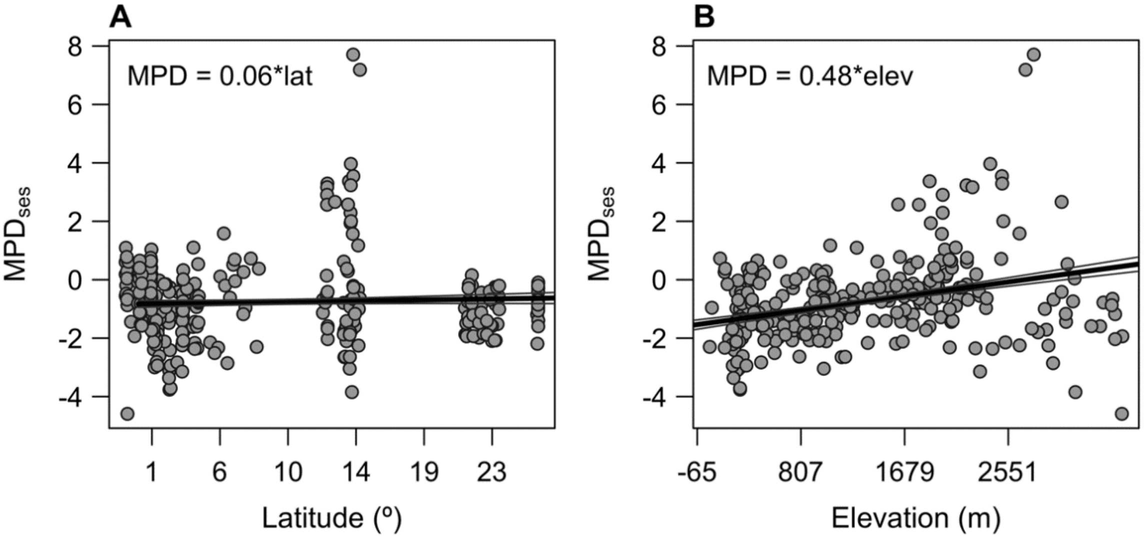
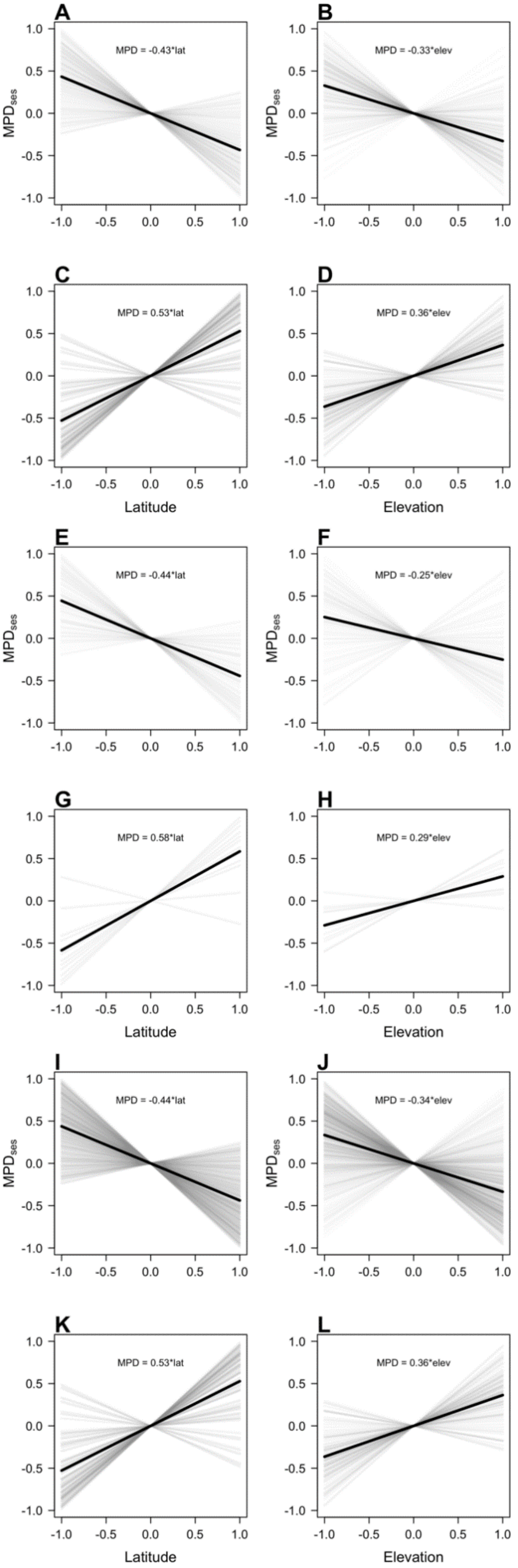


Figure S5.3. Variation of simulated standardized effect size of mean pairwise phylogenetic distance (MPD_{ses}) across latitude and elevation obtained from simulations of the Tropical Niche Conservatism hypothesis (left panels) and the Multiple Zone of Origin hypothesis (right panels). The black lines showed the main effect of the x-axis variable and gray lines represent simulations between 0.25 and 0.75 quantiles. The equation is based on the average of regression coefficients of 1000 simulations. Simulations using speciation rate = 0.5 and extinction = 0.1 (A, B, C, D), simulations using speciation rate = 0.75 and extinction = 0.05 (E, F, G, H) and simulations using speciation rate = 0.5 and extinction = 0.05 (I, J, K, L).


A
MPD = -0.43*lat
B
MPD = -0.33*elev
C
MPD = 0.53*lat
D
MPD = 0.36*elev
E
MPD = -0.44*lat
F
MPD = -0.25*elev
G
MPD = 0.58*lat
H
MPD = 0.29*elev
I
MPD = -0.44*lat
J
MPD = -0.34*elev
K
MPD = 0.53*lat
L
MPD = 0.36*elev
MPDses
Latitude
Elevation
1.0
0.5
0.0
-0.5
-1.0

9 References

Adams, M. A., Turnbull, T. L., Sprent, J. I., & Buchmann, N. (2016). Legumes are different: Leaf nitrogen, photosynthesis, and water use efficiency. Proceedings of the National Academy of Sciences, 113(15), 4098-4103.

Aldana, A. M., Carlucci, M. B., Fine, P. V., & Stevenson, P. R. (2017). Environmental filtering of eudicot lineages underlies phylogenetic clustering in tropical South American flooded forests. Oecologia, 183(2), 327-335.

Algar, A. C., Kerr, J. T., & Currie, D. J. (2009). Evolutionary constraints on regional faunas: whom, but not how many. Ecology letters, 12(1), 57-65.

Antonelli, A., Zizka, A., Carvalho, F. A., Scharn, R., Bacon, C. D., Silvestro, D., & Condamine, F. L. (2018). Amazonia is the primary source of Neotropical biodiversity. Proceedings of the National Academy of Sciences, 115(23), 6034-6039.

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., & Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proceedings of the National Academy of Sciences, 112(19), 6110-6115.

Banin, L., Feldpausch, T. R., Phillips, O. L., Baker, T. R., Lloyd, J., Affum-Baffoe, K., ... & Lewis, S. L. (2012). What controls tropical forest architecture? Testing environmental, structural and floristic drivers. Global Ecology and Biogeography, 21(12), 1179-1190.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. Global ecology and biogeography, 19(1), 134-143.

Bernal, R., S.R. Gradstein & M. Celis (eds.). 2019. Catálogo de plantas y líquenes de Colombia. Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá. http://catalogoplantasdecolombia.unal.edu.co

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J., & Green, J. L. (2008). Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences, 105(Supplement 1), 11505-11511.

Chave, J. (2004). Neutral theory and community ecology. Ecology letters, 7(3), 241-253.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a worldwide wood economics spectrum. Ecology letters, 12(4), 351-366.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., ... & Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global change biology, 20(10), 3177-3190.

Cárdenas, D., González-Caro, S., Duivenvoorden, J., Feeley, K., & Duque, A. (2017). Asymmetrical niche determinism across geological units shapes phylogenetic tree communities in the Colombian Amazonia. Perspectives in Plant Ecology, Evolution and Systematics, 28, 1-9.

Cavender-Bares, J., Ackerly, D. D., Baum, D. A., & Bazzaz, F. A. (2004). Phylogenetic overdispersion in Floridian oak communities. The American Naturalist, 163(6), 823-843.

Cavender-Bares, J., Kozak, K. H., Fine, P. V., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. Ecology letters, 12(7), 693-715.

Cavender-Bares, J., Ackerly, D. D., Hobbie, S. E., & Townsend, P. A. (2016). Evolutionary legacy effects on ecosystems: Biogeographic origins, plant traits, and implications for management in the era of global change. Annual Review of Ecology, Evolution, and Systematics, 47, 433-462.

Cayuela, L., Benayas, J. M. R., Justel, A., & Salas-Rey, J. (2006). Modelling tree diversity in a highly fragmented tropical montane landscape. Global Ecology and Biogeography, 15(6), 602-613.

Claramunt, S., Derryberry, E. P., Remsen Jr, J. V., & Brumfield, R. T. (2012). High dispersal ability inhibits speciation in a continental radiation of passerine birds. Proceedings of the Royal Society B: Biological Sciences, 279(1733), 1567-1574.

Comita, L. S., Condit, R., & Hubbell, S. P. (2007). Developmental changes in habitat associations of tropical trees. Journal of Ecology, 95(3), 482-492.

Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., ... & Hubbell, S. P. (2002). Beta-diversity in tropical forest trees. Science, 295(5555), 666-669.

Condit, R., Engelbrecht, B. M., Pino, D., Pérez, R., & Turner, B. L. (2013). Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. Proceedings of the National Academy of Sciences, 110(13), 5064-5068.

Culmsee, H., Leuschner, C., Moser, G., & Pitopang, R. (2010). Forest aboveground biomass along an elevational transect in Sulawesi, Indonesia, and the role of Fagaceae in tropical montane rain forests. Journal of Biogeography, 37(5), 960-974.

Cuni-Sanchez, A., Pfeifer, M., Marchant, R., Calders, K., Sørensen, C. L., Pompeu, P. V., ... & Burgess, N. D. (2017). New insights on above ground biomass and forest attributes in tropical montane forests. Forest Ecology and Management, 399, 235-246.

Curtis, R. O., & Marshall, D. D. (2000). Why quadratic mean diameter?. Western Journal of Applied Forestry, 15(3), 137-139.

Dexter, K. G., Terborgh, J. W., & Cunningham, C. W. (2012). Historical effects on beta diversity and community assembly in Amazonian trees. Proceedings of the national academy of sciences, 109(20), 7787-7792.

Dexter, K., & Chave, J. (2016). Evolutionary patterns of range size, abundance and species richness in Amazonian angiosperm trees. PeerJ, 4, e2402.

Dexter, K. G., Lavin, M., Torke, B. M., Twyford, A. D., Kursar, T. A., Coley, P. D., ... & Pennington, R. T. (2017). Dispersal assembly of rain forest tree communities across the Amazon basin. Proceedings of the National Academy of Sciences, 114(10), 2645-2650.

Donoghue, M. J., & Edwards, E. J. (2014). Biome shifts and niche evolution in plants. Annual Review of Ecology, Evolution, and Systematics, 45, 547-572.

Duivenvoorden, J. F., & Lips, J. M. (1993). Landscape ecology of the middle Caqueta basin: explanatory notes to the maps (+ maps, scale 1: 25,000). Tropenbos-Colombia.

Duivenvoorden, J. E. (1995). Tree species composition and rain forest-environment relationships in the middle Caquetá area, Colombia, NW Amazonia. Vegetatio, 120(2), 91-113.

Duque, A., Sánchez, M., Cavelier, J., & Duivenvoorden, J. F. (2002). Different floristic patterns of woody understorey and canopy plants in Colombian Amazonia. Journal of Tropical Ecology, 499-525.

Duque, A., Stevenson, P. R., & Feeley, K. J. (2015). Thermophilization of adult and juvenile tree communities in the northern tropical Andes. Proceedings of the National Academy of Sciences, 112(34), 10744-10749.

Duque, A., Muller-Landau, H. C., Valencia, R., Cardenas, D., Davies, S., de Oliveira, A., ... & Vicentini, A. (2017). Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots. Biodiversity and Conservation, 26(3), 669-686.

Engemann, K., Sandel, B., Boyle, B., Enquist, B. J., Jørgensen, P. M., Kattge, J., ... & Svenning, J. C. (2016). A plant growth form dataset for the New World. Ecology, 97(11), 3243-3243.

Enquist, B. J., West, G. B., & Brown, J. H. (2009). Extensions and evaluations of a general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences, 106(17), 7046-7051.

Eiserhardt, W. L., Svenning, J. C., Baker, W. J., Couvreur, T. L., & Balslev, H. (2013). Dispersal and niche evolution jointly shape the geographic turnover of phylogenetic clades across continents. Scientific reports, 3(1), 1-8.

Etienne, R. S., Cabral, J. S., Hagen, O., Hartig, F., Hurlbert, A. H., Pellissier, L., ... & Storch, D. (2019). A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. The American Naturalist, 194(5), E122-E133.

Fadrique, B., Báez, S., Duque, Á., Malizia, A., Blundo, C., Carilla, J., ... & Feeley, K. J. (2018). Widespread but heterogeneous responses of Andean forests to climate change. Nature, 564(7735), 207-212.

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

Feeley, K. J., Silman, M. R., Bush, M. B., Farfan, W., Cabrera, K. G., Malhi, Y., ... & Saatchi, S. (2011). Upslope migration of Andean trees. Journal of Biogeography, 38(4), 783-791.

Feeley, K. J., & Stroud, J. T. (2018). Where on Earth are the "tropics"?. Frontiers of Biogeography, 10(1-2).

Feldpausch, T. R., Lloyd, J., Lewis, S. L., Brienen, R. J., Gloor, M., Monteagudo Mendoza, A., ... & Phillips, O. L. (2012). Tree height integrated into pantropical forest biomass estimates. Biogeosciences, 9(8), 3381-3403.

Fine, P. A., Daly, D. C., & Cameron, K. M. (2005). The contribution of edaphic heterogeneiyt to the evolution and diversity of burseracear trees in the western Amazon. Evolution, 59(7), 1464-1478.

Fine, P. V., & Ree, R. H. (2006). Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. The American Naturalist, 168(6), 796-804.

Fine, P. V., & Kembel, S. W. (2011). Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. Ecography, 34(4), 552-565.

Fine, P. V., Zapata, F., & Daly, D. C. (2014). Investigating processes of neotropical rain forest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae). Evolution, 68(7), 1988-2004.

Fine, P. V., & Baraloto, C. (2016). Habitat endemism in white-sand forests: insights into the mechanisms of lineage diversification and community assembly of the Neotropical flora. Biotropica, 48(1), 24-33.

Fortunel, C., Paine, C. T., Fine, P. V., Kraft, N. J., & Baraloto, C. (2014). Environmental factors predict community functional composition in A mazonian forests. Journal of Ecology, 102(1), 145-155.

Fortunel, C., Valencia, R., Wright, S. J., Garwood, N. C., & Kraft, N. J. (2016). Functional trait differences influence neighbourhood interactions in a hyperdiverse Amazonian forest. Ecology letters, 19(9), 1062-1070.

Graham, A. (2009). The Andes: a geological overview from a biological perspective. Annals of the Missouri Botanical Garden, 96(3), 371-385.

Graham, C. H., Parra, J. L., Rahbek, C., & McGuire, J. A. (2009). Phylogenetic structure in tropical hummingbird communities. Proceedings of the National Academy of Sciences, 106(Supplement 2), 19673-19678.

Gentry, A. H. (1982). Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny?. Annals of the Missouri Botanical Garden, 69(3), 557-593.

Gentry, A. H. (1988). Changes in plant community diversity and floristic composition on environmental and geographical gradients. Annals of the Missouri botanical garden, 1-34.

Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Functional Ecology, 29(5), 600-614.

Gerhold, P., Carlucci, M. B., Procheş, Ş., & Prinzing, A. (2018). The deep past controls the phylogenetic structure of present, local communities. Annual Review of Ecology, Evolution, and Systematics, 49, 477-497.

Girardin, C. A. J., Malhi, Y., Aragao, L. E. O. C., Mamani, M., Huaraca Huasco, W., Durand, L., ... & Whittaker, R. J. (2010). Net primary productivity allocation and cycling of carbon along a tropical forest elevational transect in the Peruvian Andes. Global Change Biology, 16(12), 3176-3192.

Girardin, C. A., Farfan-Rios, W., Garcia, K., Feeley, K. J., Jørgensen, P. M., Murakami, A. A., ... & Malhi, Y. (2014). Spatial patterns of above-ground structure, biomass and composition in a network of six Andean elevation transects. Plant Ecology & Diversity, 7(1-2), 161-171.

González-Caro, S., Umaña, M. N., Álvarez, E., Stevenson, P. R., & 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(2), 145-153.

González-Caro, S., Duque, Á., Feeley, K. J., Cabrera, E., Phillips, J., Ramirez, S., & Yepes, A. (2020). The legacy of biogeographic history on the composition and structure of Andean forests. Ecology, 101(10), e03131.

González-Caro, S., Duivenvoorden, J. F., Balslev, H., Cavelier, J., Grandez, C., Macia, M. J., ... & Duque, A. (2020). Scale-dependent drivers of the phylogenetic structure and similarity of tree communities in northwestern Amazonia. Journal of Ecology.

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

Guevara, J. E., Damasco, G., Baraloto, C., Fine, P. V., Peñuela, M. C., Castilho, C., ... & ter Steege, H. (2016). Low phylogenetic beta diversity and geographic neo-endemism in Amazonian white-sand forests. Biotropica, 48(1), 34-46.

Heywood, V. H., Brummitt, R. K., Culham, A., & Seberg, O. (2007). Flowering plant families of the world (Vol. 88). Ontario: Firefly Books.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology: A Journal of the Royal Meteorological Society, 25(15), 1965-1978.

Honorio Coronado, E. N., Dexter, K. G., Pennington, R. T., Chave, J., Lewis, S. L., Alexiades, M. N., ... & Phillips, O. L. (2015). Phylogenetic diversity of Amazonian tree communities. Diversity and Distributions, 21(11), 1295-1307.

Hooghiemstra, H., & Van der Hammen, T. (2004). Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359(1442), 173-181.

Hooghiemstra, H., Wijninga, V. M., & Cleef, A. M. (2006). The paleobotanical record of Colombia: implications for biogeography and biodiversity. Annals of the Missouri Botanical Garden, 93(2), 297-325.

Hoorn, C. (2006). Mangrove forests and marine incursions in Neogene Amazonia (lower Apaporis River, Colombia). Palaios, 21(2), 197-209.

Hoorn, C., Wesselingh, F. P., Ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., ... & Antonelli, A. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. science, 330(6006), 927-931.

Hubbell, S. P. (2001). The unified neutral theory of biodiversity and biogeography (MPB-32) (Vol. 32). Princeton University Press.

Hughes, C., & Eastwood, R. (2006). Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. Proceedings of the National Academy of Sciences, 103(27), 10334-10339.

Hulshof, C. M., Swenson, N. G., & Weiser, M. D. (2015). Tree height–diameter allometry across the United States. Ecology and evolution, 5(6), 1193-1204.

Igea, J., & Tanentzap, A. J. (2020). Angiosperm speciation cools down in the tropics. Ecology letters, 23(4), 692-700.

Jablonski, D., Roy, K., & Valentine, J. W. (2006). Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. Science, 314(5796), 102-106.

Jaramillo, C., Rueda, M. J., & Mora, G. (2006). Cenozoic plant diversity in the Neotropics. Science, 311(5769), 1893-1896.

Jaramillo, C., & Cárdenas, A. (2013). Global warming and neotropical rainforests: a historical perspective. Annual Review of Earth and Planetary Sciences, 41, 741-766.

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., ... & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics, 26(11), 1463-1464.

Kerkhoff, A. J., Moriarty, P. E., & Weiser, M. D. (2014). The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. Proceedings of the National Academy of Sciences, 111(22), 8125-8130.

Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. Proceedings of the National Academy of Sciences, 104(14), 5925-5930.

Körner, C. (2007). The use of 'altitude'in ecological research. Trends in ecology & evolution, 22(11), 569-574.

Körner, C., & Spehn, E. M. (Eds.). (2019). Mountain biodiversity: a global assessment (Vol. 7). Routledge.

Kooyman, R. M., Wilf, P., Barreda, V. D., Carpenter, R. J., Jordan, G. J., Sniderman, J. K., ... & Weston, P. H. (2014). Paleo-Antarctic rainforest into the modern Old World tropics: The rich past and threatened future of the "southern wet forest survivors". American Journal of Botany, 101(12), 2121-2135.

Laiolo, P., Pato, J., & Obeso, J. R. (2018). Ecological and evolutionary drivers of the elevational gradient of diversity. Ecology letters, 21(7), 1022-1032.

Larjavaara, M. (2014). The world's tallest trees grow in thermally similar climates. New Phytologist, 202(2), 344-349.

Latham, R. E., & Ricklefs, R. E. (1993). Global patterns of tree species richness in moist forests: energy-diversity theory does not account for variation in species richness. Oikos, 325-333.

Leibold, M. A., & Chase, J. M. (2017). Metacommunity Ecology, Volume 59 (Vol. 59). Princeton University Press.

Legendre, P., & Legendre, L. (2012). Numerical ecology. Elsevier.

Leprieur, F., Albouy, C., De Bortoli, J., Cowman, P. F., Bellwood, D. R., & Mouillot, D. (2012). Quantifying phylogenetic beta diversity: distinguishing between 'true'turnover of lineages and phylogenetic diversity gradients. PloS one, 7(8), e42760.

Lessard, J. P., Belmaker, J., Myers, J. A., Chase, J. M., & Rahbek, C. (2012). Inferring local ecological processes amid species pool influences. Trends in ecology & evolution, 27(11), 600-607.

Linan, A. G., Myers, J. A., Edwards, C. E., Zanne, A. E., Smith, S. A., Arellano, G., ... & Tello, S. J. (2021). The evolutionary assembly of forest communities along environmental gradients: recent diversification or sorting of pre-adapted clades?. bioRxiv, 2020-12.

Lips, J. M., & Duivenvoorden, J. F. (1996). Regional patterns of well drained upland soil differentiation in the middle Caquetá basin of Colombian Amazonia. Geoderma, 72(3-4), 219-257.

Macía, M. J. (2011). Spatial distribution and floristic composition of trees and lianas in different forest types of an Amazonian rainforest. Plant Ecology, 212(7), 1159-1177.

Malizia, A., Blundo, C., Carilla, J., Osinaga Acosta, O., Cuesta, F., Duque, A., ... & Young, K. R. (2020). Elevation and latitude drives structure and tree species composition in Andean forests: Results from a large-scale plot network. PloS one, 15(4), e0231553.

Mayfield, M. M., & Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology letters, 13(9), 1085-1093.

Marshall, A. R., Willcock, S., Platts, P. J., Lovett, J. C., Balmford, A., Burgess, N. D., ... & Lewis, S. L. (2012). Measuring and modelling above-ground carbon and tree allometry along a tropical elevation gradient. Biological Conservation, 154, 20-33.

Misiewicz, T. M., & Fine, P. V. (2014). Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: Protium subserratum (Burseraceae). Molecular Ecology, 23(10), 2543-2558.

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... & Turelli, M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology letters, 10(4), 315-331.

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, V., ... & Niño, H. (2015). Middle Miocene closure of the Central American seaway. Science, 348(6231), 226-229.

Mori, A. S., Shiono, T., Koide, D., Kitagawa, R., Ota, A. T., & Mizumachi, E. (2013). Community assembly processes shape an altitudinal gradient of forest biodiversity. Global Ecology and Biogeography, 22(7), 878-888.

Morlon, H., Schwilk, D. W., Bryant, J. A., Marquet, P. A., Rebelo, A. G., Tauss, C., ... & Green, J. L. (2011). Spatial patterns of phylogenetic diversity. Ecology letters, 14(2), 141-149.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature, 403(6772), 853-858.

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. Journal of biogeography, 26(4), 867-878.

Olson, M. E., Soriano, D., Rosell, J. A., Anfodillo, T., Donoghue, M. J., Edwards, E. J., ... & Méndez-Alonzo, R. (2018). Plant height and hydraulic vulnerability to drought and cold. Proceedings of the National Academy of Sciences, 115(29), 7551-7556.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Wagner, H. (2019). Community ecology package. R package version, 2(0).

Parolin, P. D., De Simone, O., Haase, K., Waldhoff, D., Rottenberger, S., Kuhn, U., ... & Junk, W. J. (2004). Central Amazonian floodplain forests: tree adaptations in a pulsing system. The Botanical Review, 70(3), 357-380.

Peña, M. A., Feeley, K. J., & Duque, A. (2018). Effects of endogenous and exogenous processes on aboveground biomass stocks and dynamics in Andean forests. Plant Ecology, 219(12), 1481-1492.

Phillips, O. L., Vargas, P. N., Monteagudo, A. L., Cruz, A. P., Zans, M. E. C., Sánchez, W. G., ... & Rose, S. (2003). Habitat association among Amazonian tree species: a landscape-scale approach. Journal of ecology, 91(5), 757-775.

Pitman, N. C., Terborgh, J., Silman, M. R., & Nuñez V, P. (1999). Tree species distributions in an upper Amazonian forest. Ecology, 80(8), 2651-2661.

Pitman, N. C., Mogollón, H., Dávila, N., Ríos, M., García-Villacorta, R., Guevara, J., ... & Valderrama, E. (2008). Tree community change across 700 km of lowland Amazonian forest from the Andean foothills to Brazil. Biotropica, 40(5), 525-535.

Pitman, N. C., Andino, J. E. G., Aulestia, M., Cerón, C. E., Neill, D. A., Palacios, W., ... & Terborgh, J. W. (2014). Distribution and abundance of tree species in swamp forests of Amazonian Ecuador. Ecography, 37(9), 902-915.

Qian, H., & Ricklefs, R. E. (2016). Out of the tropical lowlands: latitude versus elevation. Trends in ecology & evolution, 31(10), 738-741.

Rabosky, D. L., Title, P. O., & Huang, H. (2015). Minimal effects of latitude on present-day speciation rates in New World birds. Proceedings of the Royal Society B: Biological Sciences, 282(1809), 20142889.

Ramírez, S., González-Caro, S., Phillips, J., Cabrera, E., Feeley, K. J., & Duque, Á. (2019). The influence of historical dispersal on the phylogenetic structure of tree communities in the tropical Andes. Biotropica, 51(4), 500-508.

Rangel, T. F., Colwell, R. K., Graves, G. R., Fučíková, K., Rahbek, C., & Diniz-Filho, J. A. F. (2015). Phylogenetic uncertainty revisited: Implications for ecological analyses. Evolution, 69(5), 1301-1312.

Rapp, J. M., Silman, M. R., Clark, J. S., Girardin, C. A., Galiano, D., & Tito, R. (2012). Intra-and interspecific tree growth across a long altitudinal gradient in the Peruvian Andes. Ecology, 93(9), 2061-2072.

Rezende, V. L., Dexter, K. G., Pennington, R. T., & Oliveira-Filho, A. T. (2017). Geographical variation in the evolutionary diversity of tree communities across southern South America. Journal of biogeography, 44(10), 2365-2375.

Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. Ecology letters, 7(1), 1-15.

Ricklefs, R. E. (2006). Evolutionary diversification and the origin of the diversity–environment relationship. Ecology, 87(sp7), S3-S13.

Ricklefs, R. E. (2007). Estimating diversification rates from phylogenetic information. Trends in ecology & evolution, 22(11), 601-610.

Ricklefs, R. E., & He, F. (2016). Region effects influence local tree species diversity. Proceedings of the National Academy of Sciences, 113(3), 674-679.

Ruokolainen, K., Moulatlet, G. M., Zuquim, G., Hoorn, C., & Tuomisto, H. (2019). Geologically recent rearrangements in central Amazonian river network and their importance for the riverine barrier hypothesis. Frontiers of Biogeography, 11(3).

Schepaschenko, D., Chave, J., Phillips, O. L., Lewis, S. L., Davies, S. J., Réjou-Méchain, M., ... & Zo-Bi, I. C. (2019). The Forest Observation System, building a global reference dataset for remote sensing of forest biomass. Scientific data, 6(1), 1-11.

Schluter, D., & Pennell, M. W. (2017). Speciation gradients and the distribution of biodiversity. Nature, 546(7656), 48-55.

Segovia, R. A., & Armesto, J. J. (2015). The Gondwanan legacy in South American biogeography. Journal of biogeography, 42(2), 209-217.

Segovia, R. A., Pennington, R. T., Baker, T. R., De Souza, F. C., Neves, D. M., Davis, C. C., ... & Dexter, K. G. (2020). Freezing and water availability structure the evolutionary diversity of trees across the Americas. Science Advances, 6(19), eaaz5373.

Shipley, B. (2016). Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference with R. Cambridge University Press.

Silvertown, J., Dodd, M., Gowing, D., Lawson, C., & McConway, K. (2006). Phylogeny and the hierarchical organization of plant diversity. Ecology, 87(sp7), S39-S49.

Slik, J. F., Arroyo-Rodríguez, V., Aiba, S. I., Alvarez-Loayza, P., Alves, L. F., Ashton, P., ... & Valencia, R. (2015). An estimate of the number of tropical tree species. Proceedings of the National Academy of Sciences, 112(24), 7472-7477.

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. American journal of botany, 105(3), 302-314.

Stevens, G. C. (1992). The elevational gradient in altitudinal range: an extension of Rapoport's latitudinal rule to altitude. The American Naturalist, 140(6), 893-911.

Stropp, J., Ter Steege, H., Malhi, Y., ATDN, & RAINFOR. (2009). Disentangling regional and local tree diversity in the Amazon. Ecography, 32(1), 46-54.

Swenson, N. G., Enquist, B. J., Thompson, J., & Zimmerman, J. K. (2007). The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. Ecology, 88(7), 1770-1780.

Tello, J. S., Myers, J. A., Macía, M. J., Fuentes, A. F., Cayola, L., Arellano, G., ... & Jørgensen, P. M. (2015). Elevational gradients in β-diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. PLoS One, 10(3), e0121458.

Terborgh, J., & Andresen, E. (1998). The composition of Amazonian forests: patterns at local and regional scales. Journal of Tropical Ecology, 645-664.

Tilman, D. (1987). The importance of the mechanisms of interspecific competition. The American Naturalist, 129(5), 769-774.

ter Steege, H., Pitman, N. C., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., ... & Vásquez, R. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. Nature, 443(7110), 444-447.

ter Steege, H., Pitman, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., ... & Silman, M. R. (2013). Hyperdominance in the Amazonian tree flora. Science, 342(6156).

ter Steege, H., Henkel, T. W., Helal, N., Marimon, B. S., Marimon-Junior, B. H., Huth, A., ... & Pinto, L. F. M. (2019). Rarity of monodominance in hyperdiverse Amazonian forests. Scientific reports, 9(1), 1-15.

Tsirogiannis, C., & Sandel, B. (2016). PhyloMeasures: a package for computing phylogenetic biodiversity measures and their statistical moments. Ecography, 39(7), 709-714.

Tuomisto, H., Ruokolainen, K., & Yli-Halla, M. (2003). Dispersal, environment, and floristic variation of western Amazonian forests. Science, 299(5604), 241-244.

Umaña, M. N., Norden, N., Cano, Á., & Stevenson, P. R. (2012). Determinants of plant community assembly in a mosaic of landscape units in Central Amazonia: ecological and phylogenetic perspectives. PloS one, 7(9), e45199.

Unger, M., Homeier, J., & Leuschner, C. (2012). Effects of soil chemistry on tropical forest biomass and productivity at different elevations in the equatorial Andes. Oecologia, 170(1), 263-274.

van der Hammen, T., & Cleef, A. M. (1983). Trigonobalanus and the tropical amphi-pacific element in the North Andean forest. Journal of Biogeography, 437-440.

Vellend, M. (2016). The Theory of Ecological Communities (MPB-57) (pp. 69-90). Princeton University Press.

Venter, M., Dwyer, J., Dieleman, W., Ramachandra, A., Gillieson, D., Laurance, S., ... & Bird, M. I. (2017). Optimal climate for large trees at high elevations drives patterns of biomass in remote forests of Papua New Guinea. Global change biology, 23(11), 4873-4883.

von Humboldt, A., & Bonpland, A. (1807). Essai sur la géographie des plantes.

Wang, S., Chen, A., Fang, J., & Pacala, S. W. (2013). Why abundant tropical tree species are phylogenetically old. Proceedings of the National Academy of Sciences, 110(40), 16039-16043.

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. The American Naturalist, 156(2), 145-155.

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. Annual review of ecology and systematics, 33(1), 475-505.

Webb, C. O., & Donoghue, M. J. (2005). Phylomatic: tree assembly for applied phylogenetics. Molecular ecology notes, 5(1), 181-183.

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. Trends in ecology & evolution, 19(12), 639-644.

Wiens, J. J., & Graham, C. H. (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. Annual review of ecology and systematics 36, 519-539.

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., ... & Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. Nature, 506(7486), 89-92.

Zeil, W. (1979). The Andes-a geological review).