



PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E BIODIVERSIDADE

**O PAPEL DAS ÁREAS ALAGÁVEIS NOS PADRÕES DE DIVERSIDADE DE
ESPÉCIES ARBÓREAS NA AMAZÔNIA**

BRUNO GARCIA LUIZE

Orientadora: Profa. Dra. Clarisse Palma da Silva

Coorientador: Prof. Dr. Thiago Sanna Freire Silva

DEZEMBRO – 2019

PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E BIODIVERSIDADE

**O PAPEL DAS ÁREAS ALAGÁVEIS NOS PADRÕES DE DIVERSIDADE DE
ESPÉCIES ARBÓREAS NA AMAZÔNIA**

BRUNO GARCIA LUIZE

Orientadora: Profa. Dra. Clarisse Palma da Silva

Coorientador: Prof. Dr. Thiago Sanna Freire Silva

Tese apresentada ao Instituto de Biociências do Campus de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Ecologia e Biodiversidade.

DEZEMBRO – 2019

L953p

Luize, Bruno Garcia

O papel das áreas alagáveis nos padrões de diversidade de espécies arbóreas na Amazônia / Bruno Garcia Luize. -- Rio Claro, 2019
159 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp),
Instituto de Biociências, Rio Claro

Orientadora: Clarisse Palma da Silva

Coorientador: Thiago Sanna Freire Silva

1. Ecologia. 2. Biogeografia. 3. Comunidades vegetais. 4. Solos
Inundação. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de
Biociências, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Rio Claro



CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: **O papel das planícies de inundação na origem e manutenção da diversidade de espécies arbóreas na Amazônia**

AUTOR: BRUNO GARCIA LUIZE

ORIENTADORA: CLARISSE PALMA DA SILVA

COORIENTADOR: THIAGO SANNA FREIRE SILVA

Aprovado como parte das exigências para obtenção do Título de Doutor em ECOLOGIA E BIODIVERSIDADE, área: Biodiversidade pela Comissão Examinadora:

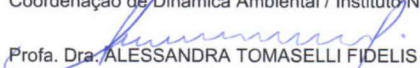


Profa. Dra. CLARISSE PALMA DA SILVA

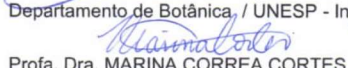
Departamento de Biologia Vegetal - Instituto de Biociências / UNICAMP - Universidade Estadual de Campinas / SP


Profa. Dra. MARIA TERESA FERNANDEZ PIEDADE

Coordenação de Dinâmica Ambiental / Instituto Nacional de Pesquisas da Amazônia - Manaus/AM


Profa. Dra. ALESSANDRA TOMASELLI FIDELIS

Departamento de Botânica / UNESP - Instituto de Biociências de Rio Claro - SP


Profa. Dra. MARINA CORREA CORTES

Departamento de Ecologia / UNESP - Instituto de Biociências de Rio Claro - SP


Profa. Dra. THAISE EMILIO LOPES DE SOUSA

Departamento de Biologia Vegetal - Instituto de Biociências / UNICAMP - Universidade Estadual de Campinas / SP

Rio Claro, 13 de dezembro de 2019

Título alterado para:

O papel das áreas alagáveis nos padrões de diversidade de espécies arbóreas na Amazônia

*Dedico às crianças com
seu olhar encantado e
curioso que nos ensina a
beleza da ciência.*

Agradecimentos

Agradeço aos meus orientadores Dra. Clarisse Palma da Silva e Dr. Thiago Sanna Freire Silva.

Agradeço aos meus supervisores de estágio em pesquisa no exterior Dr. Chris Dick e Dr. Simon Ferrier. E a UMich e o CSIRO, instituições que me abrigaram nesta experiência fantástica que pude vivenciar.

Agradeço a Fundação de Amparo à pesquisa do Estado de São Paulo pela bolsa de doutorado FAPESP #2015/24554-0 e pelas Bolsas de Estágio em Pesquisa no Exterior FAPESP # 2017/22233-8 e FAPESP # 2018/23532-1. Agradeço também aos revisores de meus projetos.

Agradeço ao apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (MCT/CNPq/CT-INFRA/GEOMA Projeto # 382728/2010) que possibilitou o início desta pesquisa. E as pesquisadoras Dr. Evelyn M. M. L. Novo e Dra. Claudia de Deus.

Agradeço a banca examinadora, desde meu plano de atividades, qualificação e defesa. Muito obrigado pelas críticas que fizeram este trabalho ser mais do que imaginei.

Agradeço aos muitos pesquisadores que me auxiliaram na construção desta tese. Meus primeiros orientadores, Dra. Patrícia Morellato, Dr. José Luís Camargo, Dr. Eduardo M. Venticinque. Aos colegas do EcoDyn, aos colegas do Dick Lab na UMich e aos pesquisadores do Macroecology Lab no CSIRO.

Agradeço aos diversos amigos que fiz e que me acompanham por esta vida. Não vou citar nomes aqui, vocês sabem que sou muito grato por ter Amigos como vocês.

Agradeço ao leitor, espero que encontre passagens interessantes e que este texto ajude você a ter novas ideias. Temos muito para observar e aprender por aí.

Agradeço a minha família. A companhia e paciência que me ajudaram a chegar até aqui.

E por fim agradeço por ter tido a oportunidade de viver tudo isso e muito mais.

*“E talvez, desesperado com tanta devastação
Pegou a primeira estrada, sem rumo, sem direção
Com os olhos cheios de água, sumiu levando essa mágoa
Dentro do seu coração”*

Vital Farias – Saga da Amazônia

*“Vou vendo o que o rio faz
Quando o rio não faz nada
Vejo os rastros que ele traz
Numa sequência arrastada
Do que ficou para trás”*

Dorival Caymmi – Na ribeira desse rio

Resumo

Áreas úmidas são ambientes na interface terrestre e aquática, onde sazonalmente a disponibilidade de água pode estar em excesso ou em escassez. A história geológica da bacia amazônica está intimamente relacionada com a presença de áreas úmidas em grandes extensões espaciais e temporais e em variadas tipologias. Dentre as tipologias de áreas úmidas presentes na Amazônia as áreas alagáveis ao longo das planícies de inundação dos grandes rios são possivelmente as que possuem maior extensão territorial. Esta tese aborda o papel das áreas úmidas para a diversidade de árvores na Amazônia. As florestas que crescem em áreas úmidas possuem menor diversidade de espécies arbóreas em relação às florestas em ambientes terrestres (i.e., florestas de terra-firme); possivelmente devido às limitações ecológicas e fisiológicas relacionadas a saturação hídrica do solo e as inundações periódicas. Entretanto, nas áreas úmidas da Amazônia já foram registradas 3,515 espécies de árvores (**Capítulo 2**), uma quantidade comparável à da diversidade na Floresta Atlântica. Em relação às florestas de terra-firme da Amazônia, as espécies de árvores que ocorrem em áreas úmidas tendem a apresentar maiores áreas de distribuição e amplitudes de tolerâncias de nicho ao longo da região Neotropical (**Capítulo 3**). A composição florística e a distância filogenética entre espécies arbóreas nas florestas de várzea da Amazônia central mudam amplamente entre localidades (**Capítulo 4**). O gradiente ambiental contido entre as manchas de floresta de várzea foi o fator preponderante para explicar a distância filogenética entre florestas, enquanto que a composição das espécies é influenciada principalmente pelas distâncias geográficas entre as localidades. Além disso, geralmente as espécies relativamente mais abundantes são as que apresentam maiores quantidades de associações de co-ocorrência (**Capítulo 5**). A estruturação destas co-ocorrências pode ser influenciada por interações bióticas de facilitação e de competição entre as espécies, mas também por características similares do nicho das espécies, indicada pela proximidade evolutiva entre elas. E ainda, por dispersão limitada, indicada por maiores ou menores sobreposições na distribuição geográfica das espécies. Uma das limitações do estudo da diversidade biológica está relacionada com o uso das espécies como unidade básica de análise. Para árvores em florestas tropicais nem sempre é possível distinguir espécies como unidades discretas. A carência de estruturas diagnósticas nas coletas botânicas, a baixa representatividade das variações fenotípicas entre os indivíduos de uma espécie nas coleções botânicas e as variações fenotípicas crípticas existentes entre espécies proximamente aparentadas são fatores que dificultam a descrição da diversidade em nível de espécie. O código de barras de DNA é apresentado como uma técnica promissora para diminuir a lacuna causada pela dificuldade na identificação das espécies com divergência recente como na família Lecythidaceae em florestas de várzea e terra firme na Amazônia central (**Capítulo 6**). Nesta tese nós estudamos a influência das áreas úmidas na a origem e manutenção da diversidade de árvores nas florestas da Amazônia e concluímos que estes ambientes promovem diferenças nas características ecológicas das espécies arbóreas e heterogeneidade biótica na região.

Palavras-chave: conjunto de espécies; área de distribuição geográfica; amplitude de nicho; tolerância ambiental; diversidade beta; diversidade filogenética; co-ocorrência de espécies; montagem de comunidades; código de barras de DNA; áreas úmidas; dispersão; ecologia de comunidades; biogeografia; biodiversidade; Amazônia.

Abstract

Wetlands are in the interface of terrestrial and aquatic environments, where seasonally water availability may be in excess or scarcity. Geological history of Amazon basin is closely linked with a huge temporal and spatial extents of wetlands. Nowadays, floodplains (i.e., Várzea and Igapó) are the wetlands with greatest coverage in Amazon. The present thesis is focused on the role of wetlands to tree species diversity in Amazon. Wetland forests have lower tree species diversity than upland forests (i.e., Terra-Firme); most likely due to ecological and physiological limitations. Notwithstanding, in Amazonian wetland forests 3,515 tree species already were recorded, (**Chapter 2**), which is comparable to tree species diversity in the Atlantic Forest. Wetland tree species show greater ranges sizes and niche breadth compared to tree species do not occur in wetlands (**Chapter 3**). Floristic compositional turnover and phylogenetic distances between floodplain forests in Central Amazon is high (**Chapter 4**). The most influential driver of floristic compositional turnover was the geographic distances between localities, whereas phylogenetic distances is driven mainly by the environmental gradients between forests. Furthermore, in general, the most abundant species are those that shows greater co-occurrence associations (**Chapter 5**). Co-occurrence structure is influenced by biotic interactions like facilitation and competition among species, but also by niche similarities indicated in the evolutionary distance among them. Moreover, dispersal limitation suggested by species range overlap is another factor structuring species pairs co-occurrences. One of the limitations of biological diversity studies is the use of species as basic unities of evaluation. However, sometimes it is impossible distinguish tree species in tropical forests as discrete unities. The lack of diagnosis characters in botanical collections, low representativity of phenotypic variation, and cryptic character variation between close related species were factors that beset species delimitation. DNA barcode is presented as a promising technique to reduce the identification shortfall of recent diverging Lecythidaceae species occurring in várzea and terra-firme forests of Central Amazon (**Chapter 6**). At this thesis we studied the role of wetlands on the origin and maintenance of tree species diversity in Amazonian forests, and we conclude that wetlands promote changes in ecological features of tree species and generates biotic heterogeneity in the region.

Keywords: species pool; range size; species distribution; niche breadth; environmental tolerances; beta-diversity; phylogenetic beta-diversity; species co-occurrences; community assembly; DNA barcode; wetlands; dispersal; community ecology; biogeography; biodiversity; Amazonia.

Lista de Ilustrações

Figure 2-1. Location of published species lists and herbaria records reporting tree species on Amazonian wetlands forests. The red dots are the location of tree species lists (TSL) from botanical inventories on Amazonian wetlands, blue dots are the voucher specimens from botanical collections (BC). The Amazonia sensu-latissimo region is defined in (EVA et al., 2005), wetland areas were obtained from (HESS et al., 2015b), and the classification of major Amazonian river types is given by (VENTICINQUE et al., 2016).	34
Figure 2-2. Cumulative collector's curve and sample-based species accumulation curve for tree species in Amazonian wetlands. (A) Tree species lists (TSL) ordered from 1950 to 2017 (see Table 2-2 for a list of reviewed studies). (B) Botanical collections (BC) from 1857 to 2016 (see Table 2-3 for a list of herbaria where records are available). The dots represent the cumulative number of species, the solid red line is the result of random interpolation of these points, and the dashed red line is the predicted number of recorded species with increased effort. The gray area denotes the 95% confidence interval of the estimated curves.....	37
Figure 2-3. Per-genus proportion of Amazonian tree species occurring and not occurring in wetlands. Proportions are calculated for the 803 genera listed the Amazon tree species checklist (CARDOSO et al., 2017) and ranked from higher to lower proportion of species on wetlands.....	42
Figure 3-1. Map of the Americas bound by latitudes 35° N and S: (a) geolocated records of Amazonian tree species that do occur in wetlands (black dots); (b) geolocated records of Amazonian tree species that do not occur in wetlands. In (a) and (b) the red line shows the boundaries of Amazonia sensu stricto. (c) Total annual precipitation (mm), (d) precipitation seasonality (CV), (e) actual evapotranspiration (mm/yr), and (f) water table depth (m below land surface).	51
Figure 3-2. Relationship between niche breadth and range size (logarithmic scale) for 5 150 Amazonian tree species in four different niche axes: (a) total annual precipitation; (b) precipitation seasonality; (c) actual evapotranspiration; and (d) water table depth. All models show a positive relation between niche breadth and range size. Dashed horizontal lines show the mean, median and the first and third quartiles for the logarithmic distribution of range size and the linear distribution of niche breadth.....	57
Figure 3-3. Relationship between niche breadth and range size (logarithmic scale) for Amazonian tree species that occur in Amazonian wetlands (n = 2 838 species) versus species that do not occur in Amazonian wetlands (n = 2 312 species). Linear models are shown for four different niche breadth axes: (a) total annual precipitation; (b) precipitation seasonality; (c) annual actual evapotranspiration; and (d) water table depth. The fitted curves for each group follow the model in Eq. 3.....	60
Figure 3-4. Relative importance of the four different hydrological niche breadth variables used to estimate log-transformed range sizes of Amazonian tree species. The goodness-of-fit partitioning was computed following the method by Lindeman, Merenda, and Gold (1980) to decompose total explained variance in multiple linear models.	61

Figure 3-5. Figure 3.S1. Histogram for the slopes estimated by the models $\log_{10}RS = \alpha + \beta_{1NB} \pm \varepsilon Eq. 1$ relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.	66
Figure 3-6. Figure 3.S2. Histogram for the standardized slopes estimated by the models $\log_{10}RS = \alpha + \beta_{1NB} \pm \varepsilon Eq. 1$ relating independent measurements of species niche breadth and range size in comparison with the observed standardized slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.	66
Figure 3-7. Figure 3.S3. Histogram for the slopes estimated by the models $\log_{10}RS = \alpha + \beta_{1NB} + \beta_{2W} \vee NW \pm \varepsilon Eq. 2$ relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.	67
Figure 3-8. Figure 3.S4. Histogram for the slopes estimated by the models ($\log_{10}RS = \alpha + \beta_{1NB} + \beta_{2NW} + \beta_{3NBW} \vee NW \pm \varepsilon Eq. 3$ relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.	67
Figure 4-1. Wetlands in the Central Amazon region and surveyed forest sites within the three floodplain landscapes considered in this study. Map colors show a simple reclassification of the dual-season wetland mask (HESS et al., 2015b).	72
Figure 4-2. Density plots for the components of a) compositional β -diversity, and b) phylogenetic β -diversity. Note the difference in axis scales.	78
Figure 4-3. The relationship between compositional and phylogenetic β -diversity components for all site pairs studied, where (a) is the relationship for Sørensen dissimilarity index without performing the additive partitioning decomposition; (b) is the relationship for the turnover component of β -diversity; and (c) is the relationship for the nestedness component of β -diversity.	79
Figure 4-4. Percentage of deviance explained, partitioned by geographic distance, environmental distance and shared among geographic distance and environmental distance, for GDMs fitted to compositional and phylogenetic β -diversity components. The GDMs were fitted for all the 903 pairwise comparisons between 43 plots established in central Amazonia, with the exception of the GDM fitted to the β_{sne} matrix, which was fitted only to site pairs with $\beta_{sne} \geq 0.06$ (i.e. β_{sne} values above the 3rd quartile) which involved 278 pairwise comparisons.	80
Figure 4-5. Generalized dissimilarity model for the compositional turnover β_{sim} through floodplain forests in Central Amazonia region. The three upper plots show the fitted I-spline functions in relation	

to the respective predictors (a) geographic distance, (b) wet season NDWI, and (c) ALOS 3D. In (d) the predicted ecological distances in relationship with the observed compositional dissimilarities. The colors in (e) shows the first three PCA axes of the transformed ecological distances given the GDM prediction, with each PCA axes assigned respectively to red, green and blue. Similar colors indicate locations predicted to support a similar composition of tree species while dissimilar colors depict locations diverging in composition. 83

Figure 4-6. Generalized dissimilarity model for phylogenetic turnover $\text{PhyloSor}_{\text{Turn}}$ through floodplain forests in Central Amazonia region. The three upper plots show the fitted I-spline functions in relationship with measured distance for the respective predictors (a) geographic distance, (b) wet season NDWI, and (c) mean VH backscatter from Sentinel-1. In (d) the predicted ecological distances in relationship with the observed phylogenetic turnover. The colors in (e) shows the first three PCA of the transformed ecological distances given the GDM prediction. The colors of the first three PCA was assigned respectively to red, green and blue. Similar colors indicate locations with comparable phylogenetic turnover while dissimilar colors depict locations with distinct phylogenetic turnover... 84

Figure 4-7. Percent of explained deviance for the generalized dissimilarity models taking account all the 28 environmental predictors plus geographic distances and all 28 environmental predictors but without the inclusion of geographic distances as predictor. The GDMs were fitted for all the 903 pairwise comparisons between the 43 floodplain forest surveys in central Amazonia, 592 pairs of sites between the three floodplain landscapes, 311 pairs of sites sampled within each of the three landscapes. With exception for the GDMs for the β_{sne} matrix that only was fitted for $\beta_{\text{sne}} \geq 0.06$ (i.e. β_{sne} above 3rd quartile) which means 278 pairwise comparisons, 171 pairs of sites between the three sampled floodplain landscapes and 107 pairs of sites within each of the three landscapes. Each panel shows results for the respective dissimilarity matrices the GDMs was fitted, the upper row shows results for the CBD and de lower row for the PBD matrices. 96

Figure 4-8. Percentage of variable importance for the set of predictors included in the generalized dissimilarity models explaining compositional and phylogenetic β -diversity matrices. The bars show the percentage of variable importance after backward selection with 999 permutations. Backward selection was applied to select the set of predictors to fit the simplest GDM and posteriorly used to map β -diversity through floodplain forests of Central Amazon. In (a) are highlighted the most important predictors for the β_{sor} and PhyloSor matrices, (b) gives the same information for β_{sim} and $\text{PhyloSor}_{\text{Turn}}$ matrices and (c) for β_{sne} and PhyloSorPD matrices. Note that for the β_{sne} only the bars for the model that do not include geographic distances is depicted, since the geographic distances do not have any importance for β_{sne} matrix explanation. 97

Figure 4-9. Pearson correlation between the set of 31 environmental predictors extracted for the 43 site locations sampled in floodplain forests of Central Amazonia. For a list of predictors names and descriptions please refer to table 4-1..... 98

Figure 4-10. Scatter plot for the predicted ecological distances in relationship with observed Sørensen dissimilarity index β_{sor} of compositional β -diversity. The GDM was fitted using geographic distances, ALOS_3D and ALOS_TopoDiver_TPI as model predictors (please see Table 4-1 for a description of predictors).	99
Figure 4-11. Hierarchical clustering build by unweighted pair group method with arithmetic mean (UPGMA clustering) for a) turnover of compositional and b) phylogenetic β -diversity between pairs of floodplain forests surveys.	100
Figure 5-1. Location of the 513 sampled plots in várzea forests alongside the Amazonas river in Brazil and the species richness distribution among localities.....	106
Figure 5-2. Rank for the number of times a species is detected co-occurring in positive and negative associations highlighting the species abundance classification into intermediate-rare and dominant species. Dominant species are those 36 species that together account for half of the 18 782 trees ≥ 10 cm d.b.h. sampled, while intermediate-rare are the 632 tree species with abundance accounting for the other 50% of the trees sampled. There are 291 species within 1861 positive, and 185 species within the 1095 negative co-occurrence species pairs detected. Note that species rank order may differs between negative and positive co-occurrences.	112
Figure 5-3. The relationship between probability of co-occurrence and range overlap for species pairs co-occurring more than expected by chance, highlighting the position of the species pairs with standardized residuals greater than 1 deviation. (a) for the positive co-occurrences probabilities (P_{gt}), and (b) for the negative co-occurrences probabilities (P_{lt}).....	117
Figure 5-4. Adjusted linear model for the relationship between probability of co-occurrence and divergence time for species pairs co-occurring more than expected by chance and highlighting the position of the species pairs with standardized residuals greater than 1 deviation. (a) positive co-occurrences probabilities (P_{gt}), and (b) negative co-occurrences probabilities (P_{lt}).	118
Figure 5-5. Proportional distribution of co-occurrences classified as exceeding/weak range overlap (R.O.) and close/distant divergence time (D.T.) based on the standardized residuals (i.e. $-1 \geq \text{ResStd} \geq +1$) of the linear models relating probability of co-occurrence, range overlap and divergence time. (a) Proportion of species pairs determined as negative co-occurrences, and (b) species pairs determined as positive co-occurrences, in each one of the classes assigned. (c) and (d) shows number of species pairs classified by both models as non-average residuals, (c) number of species pairs negative co-occurrences, and (d) number of species pairs positive co-occurrences.....	120
Figure 5-6. Figure 5.S1. Boxplot for the co-occurrence probability (P_{gt} and P_{lt}) computed for positive and negative species pairs following the probabilistic pairwise approach to detect species co-occurrences (Veech 2014). The lower the probabilities the greater are the chances of species pair co-occurrence.	123

Figure 6-1. Boxplots for the percentage of pairwise genetic similarities for comparisons within and among Lecythidaceae genus. Sequences evaluated are <i>matK</i> , <i>rbcL</i> , <i>ycf1</i> (1), <i>ycf1</i> (2) and concatenations for <i>matK+rbcL</i> , <i>ycf1</i> (1)+ <i>ycf1</i> (2), <i>matK+rbcL+ycf1</i> (1)+ <i>ycf1</i> (2).	134
Figure 6-2. Maximum clade credibility (MCC) tree showing the evolutionary relationship recovered for Lecythidaceae species from sequences of a) <i>matK+rbcL</i> (42 sequences, 1.606 bp, and 20 species), b) <i>ycf1</i> (1)+ <i>ycf1</i> (2) (42 sequences, 2.055 bp, and 19 species), and c) <i>matK+rbcL+ycf1</i> (1)+ <i>ycf1</i> (2) (38 sequences, 3.644 bp, and 18 species).	136
Figure 6-3. The evolutionary relationship between 63 <i>matK</i> (838 bp) sequences of Lecythidaceae species, 43 samples collected in Central Amazon (the present study) and 20 sequences available in GenBank for collections coming from French Guiana (Gonzalez et al. 2009 – sequences ending with label FJ5).	140
Figure 6-4. The evolutionary relationship between 70 <i>rbcL</i> (807 bp) sequences of Lecythidaceae species, 50 samples collected in Central Amazon (the present study) and 20 sequences available in GenBank for collections coming from French Guiana (Gonzalez et al. 2009 – sequences ending with label FJ03).	141

Lista de Tabelas

Table 2-1. Tree species richness for the ten richest botanical families found in Amazonian wetlands compared with their richness ranking according to the Amazon tree flora.	41
Table 2-2. Reference list for reviewed tree species lists. https://doi.org/10.1371/journal.pone.0198130.s001 (XLSX)	45
Table 2-3. List of herbaria available on SpeciesLink that contributed with records. https://doi.org/10.1371/journal.pone.0198130.s002 (XLSX)	45
Table 2-4. Checklist of the Amazonian wetlands tree species pool. https://doi.org/10.1371/journal.pone.0198130.s003 (XLSX)	45
Table 3-1. Estimated parameters for the generalized linear models relating niche breadth and the base-10 logarithm of range sizes for Amazonian tree species. The intercept and slopes for each model (Eqs. 1-3), standard deviance, <i>F</i> -statistic, AIC, and Δ AIC values are for comparison among the models. <i>p</i> -values were < 0.001 for all models.	59
Table 3-2. Table S1. Range sizes and niche breadths estimated for each one of the 5.150 Amazonian tree species with more than three occurrence records.	66
Table 4-1. Grid layers included as predictor variables (i.e., environmental gradients) for building the generalized dissimilarity models of compositional and phylogenetic β -diversity in floodplain forests of central Amazonia.	91
Table 5-1. Fitted linear models relating probability of co-occurrence and species pairs range overlap/divergence time, the applied classification based on the residual position of the species pair	

along with the evidence that the residuals indicate as likely drivers for the observed co-occurrence pattern.	110
Table 5-2. Matrix of expected cross combinations of species pairs classification according the standardized residuals of the fitted linear models and the likely drivers for the combination.....	111
Table 5-3. Rank of the dominant tree species in Amazonian seasonal flooded forest (<i>Várzea</i>) and number of times the species appear in a positive or negative co-occurrence pair along with their rank for the number of positive and negative co-occurrences.	114
Table 5-4. Significant co-occurrences and associated parameters: species names, species incidence, observed co-occurrences, probability of co-occurrence, expected co-occurrences, P_{lt} , P_{gt} , Patristic distance, and Range overlap.	123
Table 6-1. Lecythidaceae species vouchered from two Amazonian forest types (<i>várzea</i> – white water floodplain forests at PP-SDR and M-SDR; and <i>terra-firme</i> – upland forests at BDFFP) for DNA barcode analyses. Their accepted species names, species code and the number of DNA samples successfully sequenced for the three chloroplastidial DNA regions evaluated.....	129
Table 6-2. Regions sequenced, and primers utilized in the present study.....	132

Sumário

1. Introdução geral	17
2. The tree species pool of Amazonian wetland forests: Which species can assemble in periodically waterlogged habitats?.....	29
3. Consistently larger geographic range sizes and hydrological niche breadths suggest higher environmental tolerances for wetland-adapted Amazonian trees.....	46
4. Modelling compositional and phylogenetic beta-diversity in Central Amazonian floodplain forests	68
5. The identification of drivers for tree species pairs associations in Amazonian seasonal flooded forests.	101
6. Applying DNA barcoding for delimitation of species: a test with Lecythidaceae from central Amazonia	124
7. Conclusões.....	142
8. Referências	147

1. Introdução geral

Biodiversidade e diversificação na biota amazônica

Como a biodiversidade se origina e se mantém constituindo arranjos intrincados onde seres vivos interagem, entre si e com o meio-ambiente ao redor, e transformam, desde aquilo que está mais próximo até todo o planeta Terra? São questionamentos recorrentes na biogeografia e na ecologia que nesta tese serão abordados no contexto da flora arbórea que compõe a floresta tropical amazônica. Especificamente, a questão norteadora para a presente tese: *Como os habitats florestais das áreas úmidas da Amazônia têm contribuído para a diversificação de espécies arbóreas e na manutenção da diversidade regional destas comunidades?* Nada mais é do que uma variante da questão central – como se origina e se mantém a diversidade de espécies? A busca pela resposta, por vários pesquisadores e em diferentes especialidades, é uma das maneiras de compreendermos o mundo que habitamos. Se pensarmos que a Terra está vivendo um evento de extinção em massa (IPBES et al., 2019), com causas profundas no nosso modo de vida, podemos perceber a facilidade com que substituímos a biodiversidade e seus complexos arranjos; o que torna ainda mais necessária a busca pela compreensão sobre sua origem e manutenção.

Diferentes regiões da Terra possuem conjuntos distintos de espécies, partindo dos polos para a região equatorial o número de espécies coexistindo em uma dada área tende a aumentar (MACARTHUR, 1965) e a composição das assembleias de espécies tende a se diferenciar mais de um lugar para o outro (KRAFT et al., 2011). Além dos gradientes latitudinais de diversidade há descontinuidades na composição e na diversidade de espécies entre os continentes as quais definem regiões biogeográficas (ANTONELLI, 2017; FICETOLA; MAZEL; THUILLER, 2017). Os padrões de distribuição de diversidade biológica em escala global envolvem a atuação de eventos em escala tempo geológico que contam a história da formação de grandes feições geomorfológicas, da tectônica de placas e da dinâmica climática na origem de novas espécies e na distribuição geográfica que estas espécies assumem (FICETOLA; MAZEL; THUILLER, 2017).

Dentre todas as regiões biogeográficas, o Neotrópico é onde se pode encontrar a maior diversidade de espécies de plantas com flores em todo o globo (ANTONELLI; SANMARTÍN, 2011; GOVAERTS, 2001). A flora vascular conhecida para o continente Sul Americano, abrangendo a região Neotropical e Andina, é de aproximadamente 82,052 espécies, sendo 90% (73,552 espécies) delas endêmicas (ULLOA ULLOA et al., 2017). As florestas tropicais da Amazônia é possivelmente o maior centro de diversidade de plantas com flores da região Neotropical (ANTONELLI; SANMARTÍN, 2011; GENTRY, 1982). Por exemplo, estimativas indicam que possam existir na Amazônia até 16.000 espécies de árvores, no entanto, cerca de 70% dessas espécies ainda não foram documentadas (TER STEEGE et al., 2013). Levando a afirmativa de que nosso conhecimento sobre a distribuição geográfica e sobre as preferências bióticas e abióticas das espécies de árvores da Amazônia ainda é bastante limitado. Isso devido em parte pelo efeito combinado de baixas densidades de coletas botânicas para boa parte da região (FEELEY, 2015) e de um padrão de raridade para a maioria das espécies conhecidas (HUBBELL, 2013).

Hipóteses de cunho biogeográfico sobre a origem da alta diversidade biológica encontrada nas florestas tropicais amazônicas envolvem cenários onde longos períodos de estabilidade climática e de isolamento geográfico do continente Sul Americano possibilitaram a especiação e a acumulação de espécies (ANTONELLI; SANMARTÍN, 2011). Também a partir da observação realizada por biogeógrafos levantou-se a hipótese de os grandes rios amazônicos agirem como barreiras geográficas para a dispersão das espécies (WALLACE, 1854), influenciando a fragmentação das populações e consequente especiação por vicariância. As florestas tropicais cobrem grande extensão do norte da América do Sul desde o Paleoceno há aproximadamente 58 Milhões de anos (WING et al., 2009). Apesar de haver linhagens de espécies arbóreas que radiaram no Mioceno e permanecem nas florestas atuais (HOORN et al., 2011), também há evidência para a origem e diversificação de linhagens durante o Pleistoceno (ANTONELLI;

SANMARTÍN, 2011; LEAL; DA SILVA; PINHEIRO, 2016; RULL, 2011a, 2011b; TURCHETTO-ZOLET et al., 2013). O que sugere que a origem da diversidade biológica na região possa ter se dado tanto pela acumulação de espécies ao longo do tempo geológico profundo quanto em períodos mais curtos de tempo.

A distribuição da diversidade é influenciada por processos que permeiam diferentes escalas taxonômicas, espaciais e temporais (CAVENDER-BARES; KEEN; MILES, 2006; LEVIN, 1992; MAGURRAN, 2004; MAGURRAN; MCGILL, 2011; MCGILL, 2010), requisitando a conciliação de conhecimentos provenientes de diferentes disciplinas para compreensão da atuação destes processos (BAKER et al., 2014; MOUQUET et al., 2012). Pesquisas sobre a distribuição da diversidade em uma escala geográfica menor, como por exemplo em ilhas oceânicas (MACARTHUR; WILSON, 1967) ou em cadeias de montanhas (VON HUMBOLDT; BONPLAND, 1807; WHITTAKER, 1960), possibilitaram a realização de experimentos e de observações minuciosas da distribuição de abundância entre as espécies (PRESTON, 1948; WHITTAKER, 1965). O que possibilitou incluir a similaridade na preferência e no uso do espaço ecológico e as interações biológicas que as espécies realizam como fatores influenciadores da origem e manutenção da biodiversidade (MAGURRAN, 2004).

De modo geral, os processos que atuam na origem e manutenção da diversidade em um determinado conjunto de espécies (i.e., desde uma região biogeográfica até uma assembleia local) podem ser sintetizados como processos de entrada, de permanência e de saída de espécies (RICKLEFS; SCHLUTER, 1993; ZOBEL, 1997). A entrada das espécies ocorre com a especiação, originando uma nova espécie, ou com a dispersão entre regiões ou entre manchas de habitat. A permanência das espécies é modulada pelas taxas de dispersão e estabelecimento de populações em habitat favoráveis, onde o crescimento populacional é positivo e as interações biológicas não impossibilitam a coexistência entre as espécies (CHESSON, 2000). E a saída pode acontecer quando há exclusão local de uma

espécie ocasionada por exemplo pela predação ou pela competição, quando há deriva populacional devido à baixa abundância e/ou dispersão da espécie no local, ou quando há extinção da espécie em toda uma região. Deste modo, os processos centrais para a origem e manutenção da diversidade biológica seriam o processo de especiação, a dinâmica das populações dentro das diferentes manchas de habitat, o processo de dispersão (i.e., migração/imigração) entre estas manchas de habitat e o processo de extinção.

A extensão geográfica pela qual uma espécie se distribui é resultado da dispersão e da especiação onde tanto o espaço geográfico como o espaço ecológico são ocupados pelas espécies (COLWELL; RANGEL, 2009; SEXTON et al., 2009, 2017). Em geral, quanto maior a amplitude de nicho que uma espécie tolera maior extensão geográfica esta espécie tende a se distribuir (SLATYER; HIRST; SEXTON, 2013). Em uma região desprovida de barreiras geográficas e com condições climáticas relativamente homogêneas é de se esperar que as espécies ocupem grandes extensões geográficas. Por outro lado, espécies podem ter sua distribuição limitada quando encontram barreiras geográficas ou quando há descontinuidades espaciais nas manchas de habitats favoráveis. Durante a existência de uma espécie, a amplitude de nicho e a extensão geográfica ocupada pelas espécies se expandem e se retraem em uma dinâmica que influencia a evolução daquela espécie (DONOGHUE, 2008; SEXTON et al., 2009; WIENS, 2004). No entanto, por considerar uma grande escala espacial, a distribuição geográfica não demonstra como as diferentes manchas de habitat em uma paisagem estão ocupadas pelas espécies.

As condições ambientais e as interações biológicas variam gradativamente ao longo de paisagens e a ocupação das assembleias pelas espécies é conduzida pela dinâmica de dispersão, seleção de habitats (i.e., filtragem ambiental) e pelas interações biológicas realizadas (ver por exemplo, Figura 1A (RAPACCIUOLO; BLOIS, 2019). O padrão de troca de espécies na composição das assembleias de diferentes localidades, conceituado como diversidade β (WHITTAKER, 1960), demonstra como as espécies ocupam as

manchas de habitats. Sendo que a dissimilaridade na composição das espécies entre localidades é influenciada tanto por diferenças nas condições ambientais quanto distanciamento geográfico entre manchas de habitat (DUIVENVOORDEN; SVENNING; WRIGHT, 2002). Sugerindo, então, a atuação de processos de “filtragem” das espécies que possuem nichos compatíveis com as condições ambientais (i.e., seleção ambiental), e de processos neutros, melhor entendidos como a baixa probabilidade de as espécies ocuparem as localidades mesmo as condições sendo compatíveis.

Apesar de a diversidade β ser útil para elucidar o efeito relativo dos gradientes ambientais e do distanciamento geográfico na seleção de espécies em uma assembleia, ainda assim é complicado vislumbrar como as diferentes espécies interagem. O simples fato de duas espécies ocorrerem juntas não garante que uma interação biótica seja estruturada, isso porque este encontro pode se dar simplesmente ao acaso. No entanto, há espécies que ocorrem sistematicamente juntas e outras que estão segregadas e ocupam diferentes localidades, o que sugere a existência de pares de espécies onde há um padrão de co-ocorrência estruturado (GOTELLI, 2000; MORUETA-HOLME et al., 2016; ULRICH; GOTELLI, 2010; VEECH, 2013). Então, o padrão de co-ocorrência possibilita que sejam feitas inferências sobre as possíveis interações bióticas que se formam entre as espécies em uma assembleia (CONNOR; SIMBERLOFF, 1979; KOHLI; TERRY; ROWE, 2018; SFENTHOURAKIS; TZANATOS; GIOKAS, 2005; ULRICH, 2004; WEIHER; KEDDY, 2001).

A biodiversidade pode ser definida em diferentes níveis, desde os variados tipos de moléculas orgânicas até a diversidade de ecossistemas em uma região (WILSON, 1988). A diversidade biológica, por sua vez, operacionalmente é definida em termos de número de espécies e suas abundâncias relativas (MAGURRAN, 2004). Sendo assim, o reconhecimento de espécies, sua delimitação e sua identificação, são cruciais para o estudo da diversidade biológica (HORTAL et al., 2015). No entanto, a delimitação de espécies

em unidades discretas pode ser bastante complicada. No caso das espécies arbóreas em florestas tropicais a logística envolvida para acessar os locais e a altura das árvores são limitantes para obtenção de amostras botânicas. Geralmente, o trabalho especializado de escalada é essencial para a obtenção de um ramo de árvore com estruturas férteis que possibilitem a identificação da espécie. Na Amazônia, mesmo em locais de fácil acesso a densidade de coletas botânicas de árvores pode ser considerada baixa (FEELEY, 2015; HOPKINS, 2007, 2019). Somado a esta característica, as árvores na Amazônia tendem a florescer e frutificar em ciclos supra anuais (NEWSTROM et al., 1994) e em poucas semanas. A descrição e a distinção entre espécies de Angiospermas estão embasadas principalmente em características morfológicas das estruturas férteis (i.e., flores e frutos); então a falta destas características diagnósticas pode impossibilitar que uma espécie seja identificada. Vencido o desafio de realizar coletas botânicas com os caracteres diagnósticos, a próxima complicação é a própria diversidade presente em cada grupo e espécies proximamente aparentadas muito semelhantes. Em alguns casos, as espécies possuem diferenças crípticas ou a gradiente de variação em suas características diagnóstico não é completamente compreendido. Por exemplo, a família Lecythidaceae que tem centro de diversidade na Amazônia, onde são conhecidas cerca de 160 espécies de árvores, incluindo espécies de grande importância para estrutura da floresta devido as suas altas abundâncias, e que apesar de ser uma família facilmente reconhecida apenas pelos caracteres vegetativos é extremamente difícil ter suas espécies identificadas sem as características presentes nas flores e nos frutos. Isso pode ocorrer devido a presença de espécies em complexo (e.g., *Eschweilera coriacea*), possivelmente influenciado pela recente radiação no grupo, que ocorreu no Mioceno superior (Oscar Vargas, Comunicação pessoal) e pela introgressão genética entre linhagens. Dada a importância da delimitação de espécies para estudar a diversidade biológica diferentes técnicas têm sido testadas para avaliar a possibilidade de separar espécies, como por exemplo, assinaturas espectrais e comparação de sequências de DNA (LANG et al., 2015).

As áreas úmidas, a evolução da paisagem amazônica e a diversidade biológica na floresta

Áreas úmidas são ambientes sazonais onde os níveis do lençol freático ficam acima do solo permanentemente ou temporariamente oscilando entre fase terrestre e outra fase aquática (i.e., pântanos, planícies de inundação) (JUNK et al., 2011; KEDDY, 2010). Um dos tipos de áreas úmidas mais comumente encontrado na Amazônia são as planícies de inundação, que são áreas alagáveis moduladas pelo pulso de inundação experienciado pelos grandes rios tropicais (*sensu* (JUNK et al., 1989)). O pulso de inundação nos rios da região amazônica ocorre desde o Paleoceno (c.: 66 Ma) mesmo durante períodos climáticos mais secos (WITTMANN; SCHÖNGART; JUNK, 2010). Reconstruções paleológicas indicam que terras baixas da região oeste da bacia Amazônica estiveram sobre influência de um grande sistema inundável durante boa parte do Mioceno (c.: 30-23 Ma) (HOORN et al., 2010b; LATRUBESSE et al., 2010). Possivelmente sendo o sistema de áreas úmidas de maior extensão e duração na história do planeta Terra (HOORN et al., 2010a). O soerguimento da cordilheira dos Andes influenciou o regime atmosférico da América do Sul e possibilitou que mais chuvas se precipitassem sobre a bacia amazônica (INSEL; POULSEN; EHLERS, 2010), implicando em variação na precipitação acumulada ao longo das bacias de captação dos rios. Além disso, variações no nível dos oceanos, durante as sucessivas glaciações do Pleistoceno também influenciaram os níveis dos rios da Amazônia (BERTASSOLI-JR et al., 2019; IRION et al., 2010; PUPIM et al., 2019). Como exemplo, durante o período interglacial Sangamoniano, há 125.000 anos BP, o nível do mar atingiu cerca de 20 metros acima do nível atual fazendo com que os maiores rios do baixo Amazonas ficassem represados (IRION et al., 2010).

As áreas úmidas são interpretadas como habitats extremos para espécies de árvores (GENTRY, 1988; HONORIO CORONADO et al., 2015), podendo atuar como barreira para a dispersão de espécies vegetais e favorecer a especiação por alopatria (ANTONELLI

et al., 2009). A inundação e a seca periódica ocasionam oscilações temporais na disponibilidade de recursos e impõem restrições ecológicas e fisiológicas que influenciam as taxas de recrutamento e mortalidade das espécies lenhosas (JUNK et al., 2010; PAROLIN et al., 2004). Considerando que tanto o excesso quanto a escassez de água atuam na seleção e evolução do nicho e das distribuições geográficas das espécies (SILVERTOWN; ARAYA; GOWING, 2015) podemos esperar que o habitat de áreas úmidas deva ter influência na geração e manutenção da diversidade na região Amazônica (WITTMANN; SCHÖNGART; JUNK, 2010). Por exemplo, a diversidade de árvores é maior ao oeste e noroeste da bacia Amazônica, onde ao longo do tempo geológico os níveis de precipitação se mantiveram mais estáveis e atualmente são registradas os maiores volumes de precipitação anual, assim como menor sazonalidade climática (HOORN; WESSELINGH, 2010; TER STEEGE et al., 2003). Além disso, as florestas inundáveis na Amazônia podem ser consideradas hiperdiversas, com até nove vezes mais espécies do que a quantidade encontrada em tipos florestais semelhantes de outras regiões do globo (WITTMANN et al., 2006).

A estabilidade e recorrência do pulso de inundação, ao longo do tempo geológico, pode ter influenciado o acúmulo de espécies com características adaptativas para sobrevivência em ambientes inundáveis. Isto porque o pulso de inundação configura uma zona gradual de transição entre ambientes terrestres e aquáticos, na forma de um gradiente ambiental complexo (e.g. duração de inundação, teores de Oxigênio no solo) que implica em padrões característicos de distribuição das tipologias de vegetação e das espécies (LUIZE et al., 2015a; WITTMANN; JUNK; PIEDADE, 2004). Por outro lado, as áreas úmidas podem atuar como sumidouros para as espécies de árvores das florestas de terra firme (WITTMANN; SCHÖNGART; JUNK, 2010). Por exemplo, nas florestas de terra firme, inundações eventuais e de curta duração aumentam em até 5% a mortalidade de árvores em relação a locais não afetados por estas inundações (MORI; BECKER, 1991). Sugerindo que a maioria das espécies de terra firme não consigam sobreviver sob as

condições ocasionadas pelas inundações ou uma baixa probabilidade de imigração das terras-firme para áreas úmidas.

A oscilação na extensão ocupada pelas áreas úmidas da Amazônia e na sazonalidade do pulso de inundação implica as espécies e as florestas de planícies de inundação tenham se expandido e contraído em múltiplas ocasiões. Atualmente, estima-se que as áreas úmidas ocupem cerca de 14% - 17% da bacia amazônica (HESS et al., 2015a), mas ao considerar não apenas as áreas alagáveis ao longo dos maiores rios, mas também as áreas úmidas interfluviais ao longo de menores rios a proporção coberta pelas áreas úmidas da Amazônia pode chegar a 30% de toda a bacia (JUNK et al., 2011). As áreas úmidas são representadas por uma grande heterogeneidade de habitats (i.e., várzeas, igapós, brejos, campinaranas, savanas alagáveis e margens de rios de menor porte, JUNK et al., 2011). Os igapós, localizados nas planícies de inundação dos rios de águas claras ou negras que drenam os escudos cristalinos ao sul e ao norte da bacia (JUNK et al., 2011), ocupam substrato de menor fertilidade na Amazônia (JUNK et al., 2011) e provavelmente foram o primeiro tipo de planície de inundação na região (PUPIM et al., 2019). Florestas de várzea, localizadas nas planícies de inundação de formação recente no Quaternário (DE FATIMA ROSSETTI et al., 2005; IRION et al., 2010) dos rios que drenam os sopés da cordilheira dos Andes (i.e., rios de águas brancas) e que depositam uma grande carga de sedimentos ricos em nutrientes, formando solos de alta fertilidade (IRION, 1978; IRION et al., 2010). Várzeas estendem-se por aproximadamente 400.000 km² e representam o tipo de área úmida de maior extensão da bacia amazônica (JUNK et al., 2011). Várzeas e Igapós são áreas alagáveis moduladas pelo pulso de inundação dos grandes rios da bacia. Enquanto que pântanos, brejos, campinaranas alagadas e florestas de baixio margeando pequenos rios as quais são áreas úmidas moduladas pelo regime de precipitação local e pela posição que ocupam no relevo (JUNK et al., 2012, 2011, 2014).

Florestas que permanecem pouco tempo alagadas nas planícies de inundação tendem a ter composição de espécies similar àquela encontrada nas florestas de terra firme adjacentes (TERBORGH; ANDRESEN, 1998). A chegada de espécies nas áreas úmidas parece se dar a partir das florestas de terra-firme que é a matriz florestal dominante nas terras-baixas da Amazônia (WITTMANN; SCHÖNGART; JUNK, 2010). Estimativas indicam que até 90% das espécies de árvores que ocupam as florestas inundáveis podem ser encontradas também em florestas de terra firme (WITTMANN et al., 2013) exibindo, contudo, variações ecotípicas e genéticas de acordo com o ambiente que ocupam (FERREIRA et al., 2007, 2010). Apesar de as florestas inundáveis apresentarem uma menor diversidade local (i.e., diversidade α) e regional (i.e., diversidade γ) de espécies de árvores em relação às florestas de terra firme (WITTMANN et al., 2006), ambas possuem variações comparáveis na composição de espécies ao longo do espaço geográfico (i.e., diversidade β) (ALBERNAZ et al., 2012; DEXTER; TERBORGH; CUNNINGHAM, 2012; DUQUE et al., 2009).

Objetivos

Para avaliar a questão central apresentada anteriormente esta tese teve como objetivos:

1. Definir qual é o conjunto de espécies arbóreas que ocorrem em florestas de áreas úmidas da Amazônia e avaliar a proporção de espécies em áreas úmidas com relação a extensão ocupada pelo habitat na Amazônia.
2. Avaliar como a amplitude de nicho das espécies arbóreas amazônicas influencia a extensão da distribuição geográfica destas espécies; e como as distribuições ecológicas e geográficas das espécies variam entre aquelas que ocorrem em áreas úmidas e aquelas que não ocorrem em áreas úmidas.
3. Avaliar como a variação na composição de espécies e nas suas relações filogenética estão relacionados com os gradientes ambientais e geográficos entre as florestas de várzea da Amazônia.

4. Definir quais espécies tendem a co-ocorrer com maior ou menor frequência e avaliar quais os possíveis processos estruturando estas associações nas florestas de várzea da Amazônia.
5. Avaliar o uso de marcadores moleculares de DNA chloroplastidial para auxiliar na delimitação de espécies da família Lecythidaceae em florestas de várzea e terra-firme da Amazônia.

Estrutura da tese

A tese é apresentada em 7 capítulos, o primeiro é esta introdução geral. No segundo capítulo revisamos o conjunto de espécies de árvores que tem ocorrência registrada nas áreas úmidas das florestas tropicais da Amazônia. O conceito de conjunto de espécies é aplicado, mostrando que 3,615 (33% das 6,727 em CARDOSO et al., 2017) espécies de árvores possuem registros de ocorrência em florestas de áreas úmidas. O que pode ser considerado uma alta proporção, principalmente se considerarmos que a extensão das áreas úmidas é de aproximadamente 17% da área de $5.06 \times 10^6 \text{ km}^2$ ocupada pelas florestas tropicais nas terras baixas da bacia Amazônica (HESS et al., 2015a). No terceiro capítulo, foram avaliadas as diferenças na extensão da distribuição geográfica e na amplitude de nicho das espécies que ocorrem e que não ocorrem em áreas úmidas. O padrão ecológico de que maiores amplitudes de tolerância de nicho é relacionado a maiores áreas de distribuição geográfica foi verificado para 5,150 espécies arbóreas da Amazônia. Em comparação com as espécies que não ocorrem em áreas úmidas, as espécies de árvores que ocorrem em áreas úmidas tendem a ser mais amplamente distribuídas tanto em relação a tolerância climática quanto para extensão geográfica que ocupam. No quarto capítulo, estudamos a variação da diversidade de espécies entre florestas e mostra que os gradientes ambientais e a distância entre localidades influenciam as espécies e as linhagens que compõem as florestas de várzea na Amazônia central. As linhagens tendem a ser selecionadas pelas diferenças nas condições ambientais entre manchas de florestas,

enquanto a composição de espécies tende a ser selecionada pelo acaso envolvido na dispersão e colonização de locais distantes geograficamente. No quinto capítulo, analisamos o padrão de co-ocorrência entre pares de espécies ao longo das florestas de várzea as margens do rio Amazonas. Também foi estimada a sobreposição na distribuição geográfica e o distanciamento filogenético entre as co-ocorrências. O uso conjunto da informação sobre os pares de espécies co-ocorrendo permitiu fazer inferências sobre a influência de interações interespecíficas e similaridade de nicho na estruturação destas associações. No sexto capítulo, abordamos as dificuldades de identificação das espécies arbóreas em florestas tropicais e avaliamos a possibilidade de utilização de técnicas moleculares para a identificação de espécies da família da castanha do Brasil – Lecythidaceae cuja diversificação é muito recente. No sétimo capítulo, apresento sinteticamente as conclusões obtidas.

2. The tree species pool of Amazonian wetland forests: Which species can assemble in periodically waterlogged habitats?

Short running title: The tree species pool of Amazonian wetlands

Luize BG, Magalhães JLL, Queiroz H, Lopes MA, Venticinquê EM, Novo EMLM, and Silva TSF (2018) The tree species pool of Amazonian wetland forests: Which species can assemble in periodically waterlogged habitats? *PLOS ONE* 13(5): e0198130. <https://doi.org/10.1371/journal.pone.0198130>

Abstract

We determined the filtered tree species pool of Amazonian wetland forests, based on confirmed occurrence records, to better understand how tree diversity in wetland environments compares to tree diversity in the entire Amazon region. The tree species pool was determined using data from two main sources: 1) a compilation of published tree species lists plus one unpublished list of our own, derived from tree plot inventories and floristic surveys; 2) queries on botanical collections that include Amazonian flora, curated by herbaria and available through the SpeciesLink digital biodiversity database. We applied taxonomic name resolution and determined sample-based species accumulation curves for both datasets, to estimate sampling effort and predict the expected species richness using Chao's analytical estimators. We report a total of 3 615 valid tree species occurring in Amazonian wetland forests. After surveying almost 70 years of research efforts to inventory the diversity of Amazonian wetland trees, we found that 74% these records were registered in published species lists (2 688 tree species). Tree species richness estimates predicted from either single dataset underestimated the total pooled species richness recorded as occurring in Amazonian wetlands, with only 41% of the species shared by both datasets. The filtered tree species pool of Amazonian wetland forests comprises 53% of the 6 727 tree species taxonomically confirmed for the Amazonian tree flora to date. This large proportion is likely to be the result of significant species interchange among forest habitats within the Amazon region, as well as *in situ* speciation processes due to strong ecological filtering. The provided tree species pool raises the number of tree species previously reported as occurring in Amazonian wetlands by a factor of 3.2.

Introduction

Knowledge about the biodiversity expected for larger regions, known as the regional species pool (CORNELL; HARRISON, 2014), is important for inferring evolutionary processes in community assembly (CARSTENSEN et al., 2013). Empirical studies determining the species pool of large regions are central for disentangling the cross-scale processes that shape biodiversity patterns (RICKLEFS; HE, 2016) but identifying the species pool of a region is not a trivial task. It requires the accumulation of several biodiversity surveys, well-spaced across the region and covering all possible habitat types. The very definition of species pool as “the set of species able to assemble within a local community” (CORNELL; HARRISON, 2014; SRIVASTAVA, 1999; ZOBEL, 2016) must be considered before attempting its determination, as the species pool may be defined in terms of a delimited geographic region (i.e. unfiltered pool), or regarding a specific habitat type (i.e. filtered pool) (CORNELL; HARRISON, 2014; ZOBEL, 2016).

The Amazon encompasses more than one third of all Neotropical plant diversity (ANTONELLI; SANMARTÍN, 2011; GENTRY, 1982), distributed among several habitats with high levels of heterogeneity (OLSON et al., 2001). Two recently published checklists of the Amazonian flora report overall tree species richness between 6 727 (CARDOSO et al., 2017) and 11 676 (TER STEEGE et al., 2016) valid species recorded in herbaria, biodiversity repositories and/or inventories, with a predicted richness of c.a. 16 000 tree species (SLIK et al., 2015; TER STEEGE et al., 2013, 2016) based on inventory observations. The stark difference between checklists comes from a more thorough taxonomic review performed by (CARDOSO et al., 2017), but regardless of source, both lists can be considered as approximations of the regional unfiltered tree species pool of the Amazon region, in its broadest sense (DEXTER et al., 2017).

However, the Amazon region covers more than 7 million square kilometers, spanning 40° of longitude, 25° of latitude, and an elevational gradient of c.a. 6 000 m, and most of the several Amazonian habitats remain poorly sampled (FEELEY, 2015; TER STEEGE et al., 2013), strongly limiting our knowledge of the true regional species pool. It is unreasonable to expect that all Amazonian tree species are able to occupy every environment, and thus be part of the species pools of all habitats. Thus, to truly understand the processes controlling the assembly and maintenance of Amazon diversity, we must improve our knowledge regarding the filtered species pools (CORNELL; HARRISON, 2014) of the diverse habitats comprising the Amazon region.

Wetlands have been extensively present in the Amazon since at least the Miocene (30-23 Ma) (HOORN et al., 2010b; LATRUBESSE et al., 2010), and Pleistocene ocean level oscillations (2.5 Ma) may have strongly influenced their extent and distribution over time (IRION et al., 2010). Wetlands currently cover 8.4×10^5 km² of the Amazon lowlands (c.a. 17% (HESS et al., 2015a)), of which approximately 70% are covered by forests (MELACK; HESS, 2010). Total extent may be even higher, comprising up to 30% of the entire Amazon basin, if we consider hydromorphic soils along smaller streams (JUNK et al., 2011; PITMAN et al., 2014; WITTMANN et al., 2017). Most Amazonian wetlands show monomodal seasonal fluctuations in water stage and/or water table heights, known as the *flood pulse* (JUNK et al., 2011), which has been inferred to occur at least since the Paleocene (66 Ma) (WITTMANN; SCHÖNGART; JUNK, 2010).

Hydrological seasonality influences edaphic conditions, leading to hydrological segregation of species niches (SILVERTOWN, 2004; SILVERTOWN et al., 1999) as plants develop the physiological and ecological adaptations necessary to survive several floods and droughts during their lifespan (FERREIRA et al., 2009, 2007; PAROLIN et al., 2004; PAROLIN; AL., 2002; PAROLIN; FERREIRA; JUNK, 2003; WITTMANN; JUNK, 2003). The hydrological regime experienced by each individual tree occurring in

the Amazonian wetlands depends on local interactions between basin hydrology and local geomorphology (FERREIRA-FERREIRA et al., 2015), which create strong gradients of flood height and duration, shaping tree species diversification and geographical distribution across scales (ALBERNAZ et al., 2012; LUIZE et al., 2015a; MONTERO; PIEDADE; WITTMANN, 2012; WITTMANN et al., 2006, 2013, 2017) We can thus consider wetlands habitats as environmental filters, selecting individuals and species which can tolerate recurrent inundation and drought during their lifespan (*e.g.*: *Hymatanthus* (FERREIRA et al., 2007); *Inga* (DEXTER; TERBORGH; CUNNINGHAM, 2012)), and it is very likely that Amazonian wetland species have evolved into a particularly filtered species pool.

While most tree diversity studies in the Amazon still focus on upland forests, there has been growing interest in understanding the influence of water-saturated environments on questions related to tree richness (LUIZE et al., 2015a; PITMAN et al., 2014; SCHIETTI et al., 2014), compositional patterns (ALBERNAZ et al., 2012; SCHIETTI et al., 2014; TERBORGH; ANDRESEN, 1998), and phylogenetic diversity (ALDANA et al., 2017; DEXTER; TERBORGH; CUNNINGHAM, 2012; FINE; ZAPATA; DALY, 2014). Available tree species lists for Amazonian wetlands place the eutrophic floodplain (*várzea*) forests as the richest wetland forests in the world, with 918 confirmed tree species (WITTMANN et al., 2006), and a recent survey of Brazilian Amazonian wetlands raises this number to 1 119 tree species (WITTMANN et al., 2017), comprising 16% of the 6 727 tree species reported for overall Amazon lowland forests (CARDOSO et al., 2017). Furthermore, based on 542 taxa (species and morpho-species), three main biogeographic regions are supported by tree species compositional changes along the Brazilian Amazon river mainstem (ALBERNAZ et al., 2012). It is thus clear that we need a more comprehensive knowledge of the filtered species pool able to colonize these habitats, to better understand the hydrological dimension of niches occupied by Amazonian tree

species (SILVERTOWN, 2004; SILVERTOWN et al., 1999) and its role in the assembly and evolution of Amazon rainforests.

Here, we provide the most comprehensive estimate to date of the filtered tree species pool able to assemble in Amazonian wetlands, combining tree species records from herbaria databases and published and unpublished tree species surveys from different types of Amazonian wetland forests. We also discuss the possible role of wetlands in maintaining Amazon tree diversity, and offer a prediction to the expected number of species comprising the total filtered tree species pool that can survive in wetland environments, assessing how it compares to the known Amazon tree flora and predicted basin wide diversity. Finally, we discuss current limitations and best practices for increasing our biogeographical knowledge of the most tree species rich and diverse wetland forests in the world.

Materials and methods

Datasets

Our first dataset comprises a review of published tree species lists (TSL) from tree plot inventories and/or floristic surveys conducted in Amazonian wetland forests (Figure 2-1) complemented by one previously unpublished primary inventory of our own (Table 2-2). To construct TSL, we only considered studies that reported complete species lists, for any Amazonian wetland type (JUNK et al., 2012).

Our second dataset was built by querying botanical collections (BC) made in Amazonian wetland forests, curated by herbaria (Figure 2-1, Table 2-3) and included in the SpeciesLink digital biodiversity database (<http://www.splink.org.br>). We queried digitized voucher labels using the following keywords: “*Alagada*”; “*Alagado*”; “*Alagável*”; “*Aluvial*”; “*Alluvial*”; “*Área Úmida*”; “*Brejo*”; “*Chavascal*”; “*Flooded*”; “*Flood*”; “*Floodplain*”; “*Hidromórfico*”; “*Hydromorphic*”; “*Igapó*”; “*Inundada*”; “*Inundável*”; “*Restinga*”; “*Tahuampa*”; “*Várzea*”. We then merged all botanical records

returned for each keyword and filtered these records to include only Angiosperm species and only specimens collected in the Amazonia *sensu-latissimo* region, as defined by (EVA et al., 2005) (Figure 2-1).

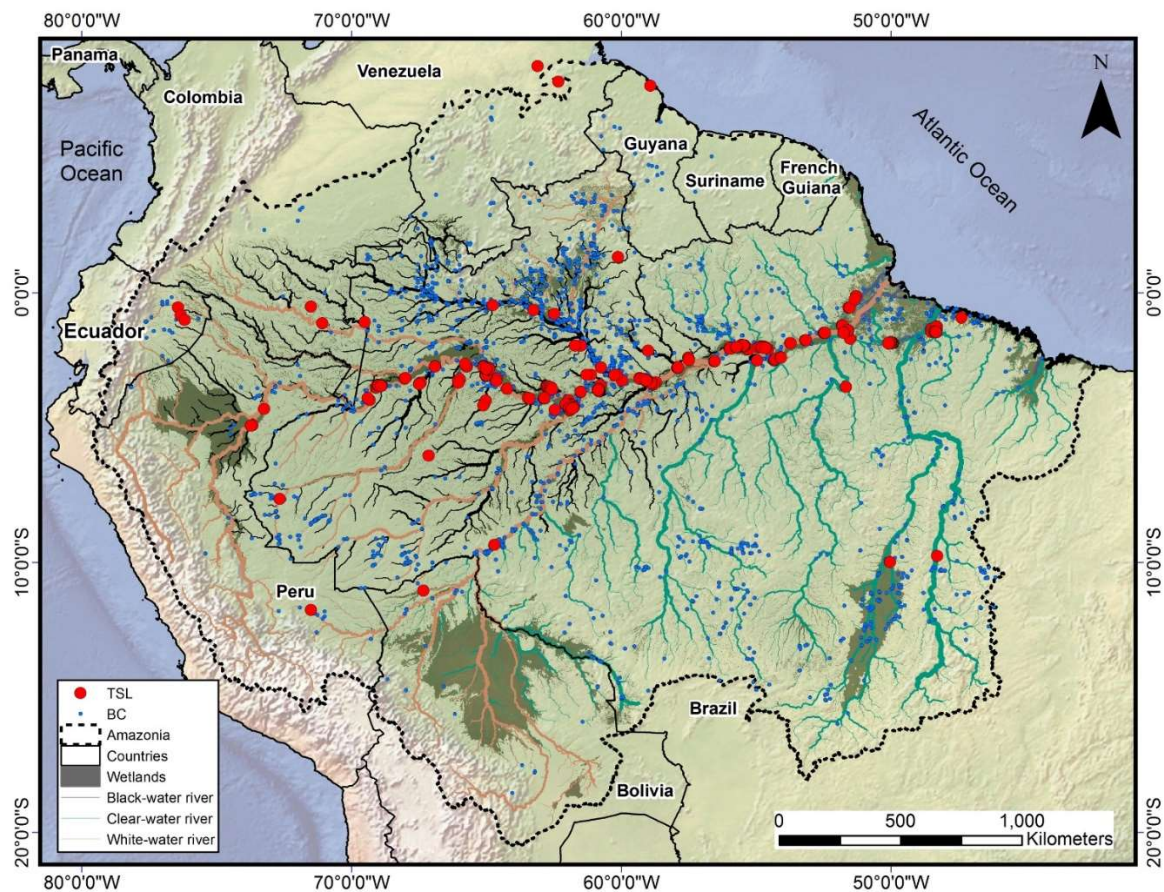


Figure 2-1. Location of published species lists and herbaria records reporting tree species on Amazonian wetlands forests. The red dots are the location of tree species lists (TSL) from botanical inventories on Amazonian wetlands, blue dots are the voucher specimens from botanical collections (BC). The Amazonia *sensu-latissimo* region is defined in (EVA et al., 2005), wetland areas were obtained from (HESS et al., 2015b), and the classification of major Amazonian river types is given by (VENTICINQUE et al., 2016).

Taxonomic standardization

Valid canonical names for species were achieved by performing taxonomic name resolution for both species datasets, using the Taxonomic Name Resolution Service - TNRS V. 4.0 online platform (BOYLE et al., 2013). We set TNRS to perform name resolution without allowing partial matches, and with a minimum match threshold > 0.85 . The authority sources consulted were, in order of relevance, TROPICOS (<http://www.tropicos.org>) and THE PLANT LIST (<http://theplantlist.org>), last updated on August 2015 (for details see: <http://tnrs.iplantcollaborative.org>). For the TSL dataset, after performing taxonomic name resolution, we filtered the resulting records to remove families known to comprise only non-tree life forms, and we assumed all remaining records after filtering corresponded to tree species. The filtered records from BC dataset were matched to the most recent Amazon tree flora checklist (CARDOSO et al., 2017), retaining only species names confirmed by taxonomic specialists as valid species names and having a tree life form (i.e. ligneous trunk reaching 10 cm DBH).

Richness estimation

We used the TSL and BC datasets to build separate species-by-sampling-unit incidence matrices, aggregating incidence by study for TSL, and by year of collection for BC. We used the resulting matrices to assess the chronological order of incidence of each recorded species, building a cumulative species collector's curve using the 'vegan' package (OKSANEN et al., 2013) and to obtaining the respective sample-based species accumulation curves for each dataset (CHAO et al., 2014). We then used the sample-based curves to predict the expected species richness if collection efforts were doubled. The inferred and estimated sample-based accumulation curves and predictions of species richness were calculated using rarefaction and extrapolation functions for incidence data provided by (CHAO et al., 2014), using the 'iNEXT' package (HSIEH; MA; CHAO, 2016). All analyses were performed in R 3.3.2. (R CORE TEAM, 2018).

Results and discussion

Determining the filtered species pool of Amazonian wetlands

In total, we reviewed 69 studies reporting tree species lists for inventories conducted on Amazonian wetland forests (Table 2-2), of which 16 (~ 20%) did not include a complete list of species and could not be added to the TSL dataset. From the 53 studies included in TSL, we recovered 21 446 records comprising 2 688 valid tree species names (Table 2-4). From these, we estimate that 3 380 (lower 95% = 3 305, upper 95% = 3 455) tree species would be recorded for Amazon wetland forests if sampling effort was doubled (Figure 2-2A). Neither the collector's curve, nor the estimated sample-based species accumulation curve showed signs of reaching an asymptote (Figure 2-2A), even after almost 70 years of inventories being conducted in Amazonian wetland forests.

We retrieved 231 119 plant occurrence records from the *SpeciesLink* database. After filtering for Angiosperms in the Amazon region, performing taxonomic name resolution and matching against the reference tree species lists, we retained 20 902 records for 2 408 valid tree species names (BC dataset – Table 2-3 and Table 2-4), lower than the observed or expected number of tree species obtained from the TSL dataset. For the BC dataset, we predicted an expected richness of 2 938 tree species (lower 95% = 2 867, upper 95% = 3 009) to be recorded for Amazonian wetland forests if collection efforts were doubled (Figure 2-2B).

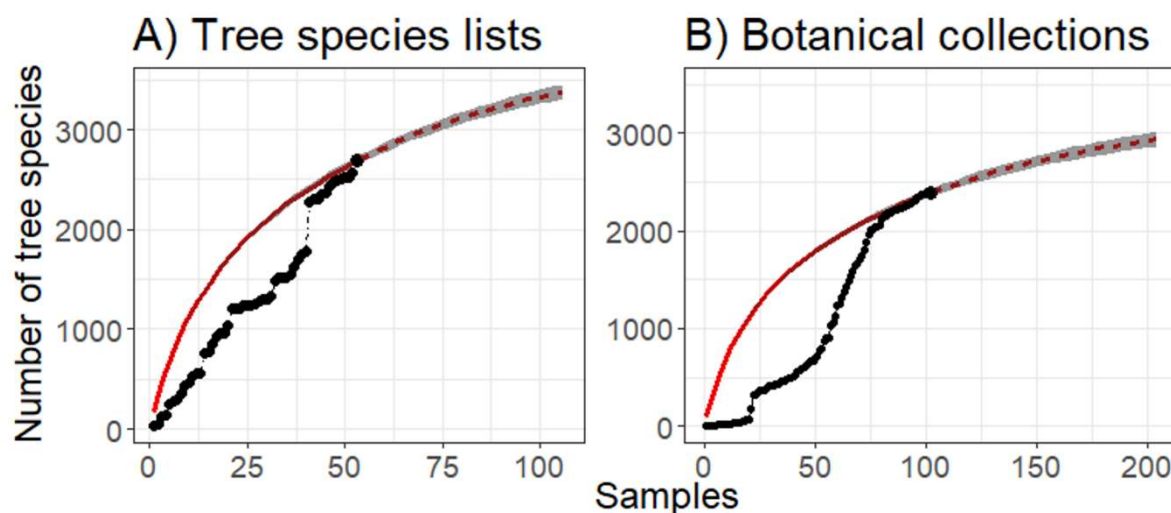


Figure 2-2. Cumulative collector's curve and sample-based species accumulation curve for tree species in Amazonian wetlands. (A) Tree species lists (TSL) ordered from 1950 to 2017 (see Table 2-2 for a list of reviewed studies). (B) Botanical collections (BC) from 1857 to 2016 (see Table 2-3 for a list of herbaria where records are available). The dots represent the cumulative number of species, the solid red line is the result of random interpolation of these points, and the dashed red line is the predicted number of recorded species with increased effort. The gray area denotes the 95% confidence interval of the estimated curves.

Pooling together the TSL and BC datasets confirmed a total of 3 615 valid tree species, comprising 42 348 records of trees occurring in Amazonian wetland forests (Table 2-3), a higher richness than the expected doubling-effort predictions from either isolated dataset. The two datasets shared 1 481 (c.a. 41%) tree species, with 1 207 (c.a. 33%) only recorded by TSL and 927 (c.a. 26%) tree species only recorded by BC.

Scope and limitations of the determined tree species pool

The determined tree species pool of Amazonian wetland forests comprises 3 615 valid species, encompassing environmental conditions found between diverse wetland types (JUNK et al., 2011). This is the most comprehensive estimate to date of the Amazonian tree species pool that can survive under extreme hydrological conditions. Although the sampling effort devoted to Amazonian upland forests is currently four times higher than to wetland forests (TER STEEGE et al., 2013; WITTMANN; SCHÖNGART; JUNK, 2010), our tree species list represents 53% of all the 6 727 tree species confirmed for the entire Amazon region (CARDOSO et al., 2017). Assuming this to be an accurate estimate

of the true proportion, Amazonian wetlands could harbor c.a. 8 500 of the 16 000 tree species expected to comprise the total Amazonian tree flora (TER STEEGE et al., 2013).

Most likely, other tree species reported for the Amazon may also occur in hydromorphic environments but have not yet been recorded in Amazonian wetlands. For instance, the average collection density recovered by us (TSL+BC) is 0.020 records per 100 km² of Amazonian wetlands, when considering the 2.1 million km² estimate of (JUNK et al., 2011), or 0.050 records per 100 km² if considering the more restrictive 840 000 km² mapped by (HESS et al., 2015a). These sampling densities are three orders of magnitude lower than the observed density of 10 records per 100 km² for Amazonian forests in general (CARDOSO et al., 2017; FEELEY; SILMAN, 2011). For this reason, we also expect that an important portion of tree species occurring in Amazonian wetlands may not be yet known to science. For example, from the 173 tree species discovered in the Amazon during the first decade of the 21st century (WWF, 2009), only 21 (12%) were identified in our estimated species pool, and of these, only six holotype specimens seem to come from vouchers collected in Amazonian wetland habitats. We thus emphasize the dire need for more intensive and comprehensive sampling of the Amazonian wetland environments.

A second limitation of the present list is introduced by the bias towards specific wetland types within the Amazon. Biodiversity assessments in the Amazon and elsewhere are generally biased towards major urban centers and along major rivers or roadways (CARDOSO et al., 2017; OLIVEIRA et al., 2016), and this bias is shown towards inventories of certain types of floodplain forests. The coverage of wetland habitat types and species occurrences recorded in our TSL and BC datasets show, as previously recognized by (WITTMANN; SCHÖNGART; JUNK, 2010), that eutrophic floodplain forests (*várzeas*) along large “white-water” rivers are the most sampled wetland forest type across the Amazon. Most of the Amazonian human population and major urban centers are adjacent to these areas, and we found the largest densities of botanical records along

the Amazonas and Negro river mainstems, near major urban centers with well-established research institutions (*e.g.*: Belém, Manaus, Tefé, Iquitos). A much lower record density was observed along the floodplain wetlands of other major Amazon tributaries (*e.g.*: Putumayo-Içá; Juruá; Purus and Madeira), or in riparian forests along interfluvial areas of the Amazon lowlands.

A third limitation is that we could not use one in every four (25%) published tree surveys conducted in Amazonian wetland forests, as the authors did not include explicit and complete species lists in the publications. Although the 21st century has seen the rise of collaborative networks, and comprehensive checklists for Neotropical forests provide large amounts of valuable information, we still need a deeper cultural shift among researchers, favoring data sharing and transparency, if we are to improve our combined knowledge of tropical tree biodiversity (BAKER et al., 2017). It is surprising that the two datasets we investigated shared less than half of the total number of valid tree species recorded, as we would expect complete overlap under an ideal scenario where at least one voucher specimen was deposited for each species recorded in each reviewed inventory (with vouchers properly digitized and made available online in herbaria databases). However, although most published inventory studies claimed to have deposited voucher specimens for their sampled plots, we were unable to find nearly a third of the species reported for inventory plots in the digitized herbaria sources. Very often, easily recognizable species and specimens without fertile structures are not included in voucher collections, creating a “data void” in the herbaria records (FEELEY; SILMAN, 2011). Thus, in practice, inventories and isolated botanical collections provide complementary floristic information for assessing tree species diversity. This reinforces the need for including the complete species lists in published inventories and shows that scientists need to keep performing both types of studies if we are to increase our knowledge of the Amazon wetland tree diversity.

Finally, a more comprehensive knowledge of the Amazon wetlands tree species pool can be achieved through efforts in reducing other biological shortfalls (*sensu* (HORTAL et al., 2015)). For instance, the uncertainty regarding actual life-form (i.e.: tree) of the recorded plant species (“Raunkiaeran shortfall”), and the lack of voucher determinations and taxonomic reviews for most *herbaria* records (“Linnean shortfall”), resulted in the removal of c.a. 25 000 records and 6 000 species names originally present in the BC dataset after taxonomic standardization and matching to the tree species list of (CARDOSO et al., 2017). Furthermore, many samples did not include information on habitat conditions, precluding a detailed assessment of species occurrence by wetland type (*e.g.*: *várzea*, *igapó*, *campinas*, *tidal várzeas*). More efforts should be made to ensure forthcoming botanical collections and inventories explicitly include life form and specific habitat conditions, as well as other ecologically relevant information.

How does the Amazonian wetland species pool compare to the basin-wide species pool?

The tree species pool of Amazon wetlands comprised 104 botanical families distributed into 689 genera, with eleven families having more than 100 tree species each. Leguminosae (578 tree species), Rubiaceae (220 tree species), Annonaceae (182 tree species), Lauraceae (175 tree species), and Myrtaceae (155 tree species) were the most diverse tree families in Amazonian wetland forests, comprising together 36% of the Amazonian wetlands tree species pool. The ten richest families in Amazonian wetlands accounted for 53% of the entire species pool (Table 2-1).

Table 2-1. Tree species richness for the ten richest botanical families found in Amazonian wetlands compared with their richness ranking according to the Amazon tree flora.

Family	Richness ranking for Amazonian wetlands tree species pool	¹Richness ranking for Amazonian tree flora	Number of valid tree species in Amazon Wetlands	Number of valid tree species in entire Amazon flora¹	Percent of species occurring in wetlands (%)
Leguminosae	1	1	578	1 042	55
Rubiaceae	2	5	220	338	65
Annonaceae	3	4	182	388	46
Lauraceae	4	2	175	400	43
Myrtaceae	5	3	155	393	39
Melastomataceae	6	6	136	263	51
Chrysobalanaceae	7	7	132	256	51
Sapotaceae	8	8	128	244	52
Euphorbiaceae	9	11	114	160	71
Moraceae	10	13	112	147	76

¹Following (CARDOSO et al., 2017).

Although 69 tree families had half or more of their Amazonian taxa occurring in Amazon wetlands, including some of the richest wetland families (Leguminosae, Euphorbiaceae and Moraceae, Table 2-1), we did not find any wetland records for 15 families with known occurrence in Amazon forests. Overall, c.a. 51% of the Amazonian tree species within each family occurred in wetland habitats, but there were noticeable differences in rank order and percentage of shared species between the ten richest wetland-occurring families and their respective richness ranking within the overall Amazon flora, as given by (CARDOSO et al., 2017) (Table 2-1).

At the genus level, 221 genera in the Amazon tree checklist [9] had all its known species recorded in the Amazon wetlands tree species pool (Figure 2-3). However, many of these genera (124) had only a single accepted species occurring in the Amazon, with only eight genera having 10 or more known species (max. 26 species). Conversely, 201 genera listed

on the Amazon tree checklist (CARDOSO et al., 2017) had no species recorded in Amazonian wetlands (Figure 2-3). The richest genus in Amazon wetlands is *Inga* (85 tree species), followed by *Licania* (69 species), *Miconia* (69 species), *Pouteria* (69 species), and *Eugenia* (59 species).

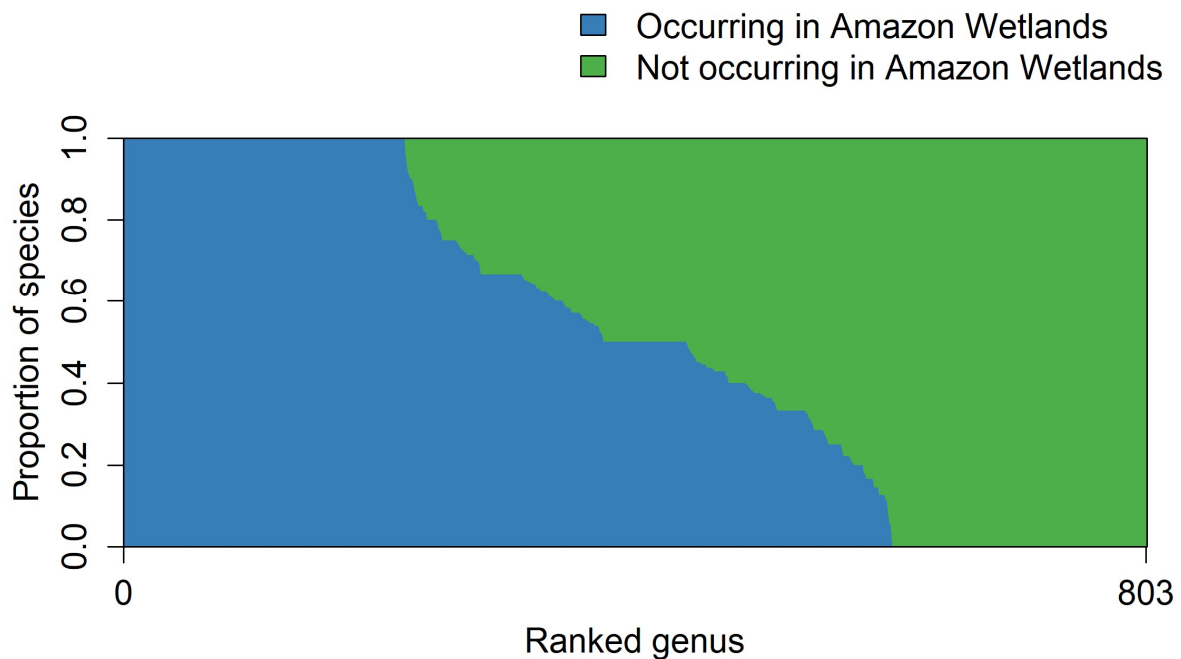


Figure 2-3. Per-genus proportion of Amazonian tree species occurring and not occurring in wetlands. Proportions are calculated for the 803 genera listed the Amazon tree species checklist (CARDOSO et al., 2017) and ranked from higher to lower proportion of species on wetlands.

The ecological and evolutionary role of Amazonian wetlands

The filtered tree species pool for Amazonian wetland forests includes almost all botanical families known to occur in Amazon forests. It is comparable to the 3 389 tree species acknowledged for the entire Brazilian Atlantic Forest (JARDIM BOTÂNICO DO RIO DE JANEIRO, 2020), one of the most biodiverse Neotropical biomes. One possible explanation for this richness is that, as Amazonian upland and wetland areas are contiguous habitats known to have an interchangeable flora (TERBORGH; ANDRESEN, 1998; WITTMANN et al., 2013), we can expect a high degree of lateral migrations among these habitats, with a large proportion of tree species in each lineage reaching and eventually adapting to both flooded and non-flooded forested habitats. Still, different patterns might also be plausible. For instance, the contribution of tree species occurring in Amazonian wetlands to the total diversity of the Amazon-centered genus suggest some taxa have evolved a high degree of *in situ* specialization on wetlands, only then colonizing upland habitats. Despite the high likelihood that a tree species will reach wetland habitats when migrating across the Amazon landscape, many Amazonian tree species do not show preference for flooded habitats; c.a. 64% of the 4 963 tree species recorded in ATDN database (TER STEEGE et al., 2013), with only 68 of the 600 most common tree species occurring in white-water Amazon floodplain forest seeming to be habitat endemics (WITTMANN et al., 2013). Assessing phylogenetic history and the relative contribution of each direction of migration to diversification could give us important insight on the origin and evolutionary history of several important taxa in the Amazon tree flora, and the role of strong environmental filtering and hydrological niche specialization in this process, as has been shown for Brazilian *Cerrado* species in relation to fire disturbance (SIMON et al., 2009).

Growing evidence suggests that it is reasonable to think of a tree species pool comprised by the entire Amazon region (DEXTER et al., 2017), but the role of ecological filtering in

the assembly of local communities cannot be excluded (MISIEWICZ; FINE, 2014). The continental dimensions of the Amazon biome and the virtual lack of geographic barriers for plant species across the lowlands implies few dispersal limitations for tree species (DEXTER et al., 2017). New environmental conditions are reached when species expand their distributions, and this floristic interchange between wetland and upland habitats might modulate source-sink population dynamics across marginal habitats. At ecological timescales, source-sink dynamics will affect population regulation and species coexistence (CHESSON, 2000; SHMIDA; WILSON, 1985); over evolutionary timescales, it will select ecotypes more prone to colonize certain habitats, leading to genetic and morphological differentiation among populations (FERREIRA et al., 2007, 2010; MISIEWICZ; FINE, 2014). In this context, although the Amazonian hydrological gradients are more idiosyncratic than the conspicuous and widely discussed temperature gradients along Andean mountain slopes, there is ample evidence for selective pressures acting on the hydrological niche dimension of Amazonian tree species, strongly affecting vegetation development and the distribution of species diversity across the region (SCHIETTI et al., 2014; SILVERTOWN; ARAYA; GOWING, 2015; TER STEEGE et al., 2013). Therefore, these lowland hydrological gradients are very likely to have had a strong historical role on tree species diversification, range expansion (ALDANA et al., 2017; DEXTER; TERBORGH; CUNNINGHAM, 2012; HOUSEHOLDER et al., 2016; WITTMANN et al., 2013), and local community assembly (LUIZE et al., 2015a; SCHIETTI et al., 2014).

Conclusions

We show that the tree species pool of Amazonian wetlands comprises 53% (3 615) of the confirmed tree species occurring in the overall Amazon, raising previous richness estimates by a factor of 3.2. It is very likely that many of these species will also occur in other forested habitats, or even other Neotropical regions. A large portion of the

Neotropical plant diversity is encompassed by Amazon-centered taxa and understanding their evolutionary and ecological histories can improve our knowledge of the development of this hyperdiverse biogeographic realm. Geographical barriers for plant dispersal are mostly absent in the Amazon region, which is instead characterized by a mosaic of habitat types and environmental gradients, including wetland habitats that have been pervasively present since before the Andean uplift. Further studies that can disassemble and then contrast the Amazon tree flora into the filtered species pools associated with each habitat type are necessary to open new avenues for exploring the ecological and geographic distribution of Amazonian tree species, functional types, and lineages, and unveil the relative role of dispersal and environmental filtering on community assembly and on the origins and maintenance of species diversity over time.

Acknowledgments

We thank Lucia G. Lohmann and Ana Carolina C. Carnaval for helpful discussions when planning the scope of this manuscript.

Supporting information

Table 2-2. Reference list for reviewed tree species lists.
<https://doi.org/10.1371/journal.pone.0198130.s001> (XLSX)

Table 2-3. List of herbaria available on SpeciesLink that contributed with records.
<https://doi.org/10.1371/journal.pone.0198130.s002> (XLSX)

Table 2-4. Checklist of the Amazonian wetlands tree species pool.
<https://doi.org/10.1371/journal.pone.0198130.s003> (XLSX)

3. Consistently larger geographic range sizes and hydrological niche breadths suggest higher environmental tolerances for wetland-adapted

Amazonian trees

Short running title: Amazon wetland trees have larger ranges sizes

Luize BG, Siqueira T, Silva TSF (2019) Consistently larger geographic range sizes and hydrological niche breadths suggest higher environmental tolerances for wetland-adapted Amazonian trees. Submetido em *Global Ecology and Biogeography*.

Abstract

Aim: Investigate the relationship between hydrological niche breadth and geographic range size for Amazonian tree species and explore the role of Amazonian wetland and upland habitats on the current distribution of tree species.

Location: Neotropics

Time period: Contemporaneous

Major taxa studied: Angiosperms

Methods: We obtained species occurrence records from GBIF and SpeciesLink. Hydrological niche breadth was measured on different unidimensional axes defined by 1) total annual precipitation; 2) precipitation seasonality; 3) actual evapotranspiration and 4) water table depth. Geographic range sizes were estimated using alpha-hull adjustments. We estimated range size and niche breadth for 76% of the valid Amazonian tree species (5 150 tree species), using 571 092 valid occurrence points. General linear models were used to relate niche breadth to range sizes while contrasting tree species occurring and not occurring in wetland habitats.

Results: The hydrological niche breadth of Amazonian tree species varied mostly along the water table depth axis. The average range size for an Amazonian tree species was 750 919 km² (median of 153 990 km², and standard deviation 1 547 540 km²). Niche breadth-range size relationships for Amazonian tree species were positive for all models, and the explanatory power of the models improved when including whether or not a species occurs in wetlands. Wetland occurrence resulted in steeper positive slopes for the niche breadth – range size relationship, and consistently larger range sizes for a given niche breadth.

Main conclusions: Amazonian tree species varied strongly in hydrological niche breadth and range size, but most species had narrow niche breadths and range sizes. The positive relationship between hydrological niche breadth and range size was stronger for tree species that occur in wetland habitats. Wetland species also had comparatively larger range sizes than non-wetland species, suggesting that the South American riverscape may have been acting as a corridor for species dispersal in the Neotropical lowlands.

Introduction

Species with larger niche breadths tend to have larger geographic distributions, and this positive relationship is well accepted as a general ecological pattern (SLATYER; HIRST; SEXTON, 2013). Plant lineages with distributions centered on Amazonian lowlands are likely to be generalist species with widespread ranges and broader niches (DEXTER et al., 2017). The relatively stable equatorial location through deep time, coupled with moderate topographic and climatic gradients along large expanses of the Amazonian region may have promoted the copious dispersal of species through the landscape. Indeed, Amazonian forests are considered a source of species for other neotropical biomes (ANTONELLI et al., 2018), suggesting that these species are more prone to range expansion, thus evolving broader niches to survive and colonize new environments. Furthermore, most of the Amazonian lowlands may be considered as lacking in effective geographic barriers for plant dispersal (DEXTER et al., 2017; NAZARENO; DICK; LOHMANN, 2017), and thus lowland plant species are more likely to be segregated by long distances among population patches than by geographic vicariance (DEXTER et al., 2017). Together, the climatic and physiographic features found in Amazonian lowlands may have favored plant species to colonize very large geographic areas and to evolve as environmentally generalist species.

Conversely, species with distributions skewed towards Andean regions are likely to have relatively narrow niche breadths and geographic ranges (HOORN et al., 2013). Tropical mountains have been suggested as evolutionary cradles (HOORN et al., 2013), promoting rapid speciation during the Quaternary through specialization within the short-range limits along the conspicuous environmental gradients that follow elevation (GENTRY, 1982; JANZEN, 1967). Mountain uplifting also had a direct influence on species distributions by creating effective geographic barriers and triggering allopatric speciation. The orogeny of the Andean mountain range has long been recognized for contributing directly and indirectly to Neotropical plant species diversity (GENTRY, 1982; HOORN et al., 2010b); including changes in lowland hydrography (HOORN;

WESSELINGH, 2010), climate (mainly precipitation patterns) (INSEL; POULSEN; EHLERS, 2010), and the exportation of migrants for Amazon lowlands (GENTRY, 1982; HOUSEHOLDER et al., 2016). Still, compared to montane habitats, Amazonian lowlands have been extensively present across space and time and are likely to harbor 33% of Neotropical plant species diversity (GENTRY, 1982), with almost 14 000 seed plant species, of which 6 727 are trees (CARDOSO et al., 2017). But apart from Andean orogeny, few other historical factors are raised to explain the origin of the enormous diversity of plants across the Amazonian lowlands.

One remarkable environmental feature with historical implications for the biota, which has been present throughout geological time in the Amazonian lowlands, is the large extent of wetland habitats, covering areas much larger than the spatially restricted montane habitats (above 500m altitude). The vast South American river network and associated wetland extent has influenced continental physiography since before the Superior Miocene (HOORN; WESSELINGH, 2010). Throughout the Miocene (c. 30-23 Ma), wetlands covered an area larger than 1.5×10^6 km² of the former Amazon basin, possibly one of the largest and longest-lived wetland systems in Earth's geological history and comprising much of the present-day western Amazonian lowlands (HOORN et al., 2010a; LATRUBESSE et al., 2010). To this day, South America still has some of the largest extents of wetlands worldwide (GUMBRICHT et al., 2017), and over half of the valid Amazonian tree species (3 615) are known to also occur in wetland habitats (LUIZE et al., 2018). Therefore, considering the prevalence and extent of wetlands in the Amazonian lowland through time and the large proportion of the world's richest tree flora that is adapted to waterlogged habitats, it is reasonable to expect a considerable influence of Amazonian wetlands on tree species diversification and dispersal (i.e. niche breadth and range size) through deep time.

Wetlands are regarded as harsh environments for species survival, demanding ecophysiological adaptations (PAROLIN et al., 2004) and likely promoting habitat

specialization (WITTMANN et al., 2013). However, environmental harshness can also contribute to larger geographic ranges by demanding a higher degree of tolerance from the adapted species, as shown for tree species in North America (MORIN; LECHOWICZ, 2013). Therefore, it is possible that once Amazonian species develop the necessary traits to colonize wetlands, they could also be able to increase their geographic ranges through much of the wet climates found in the Americas. Furthermore, the several meters of inundation during the flood season prevent tree root respiration and reduce water uptake, and therefore for some species flooding stress may be functionally similar to droughts (PAROLIN et al., 2010), perhaps enabling wetland species to also colonize dryer Neotropical regions.

In this study we investigate whether the distribution of niche breadth and geographic range size for Amazonian tree species may be a result of the pervasive presence of wetland habitats during the evolution of tree biota in Amazonian lowlands. As waterlogged habitats are regarded as more extreme habitats for tree species to colonize, we postulate two opposing hypotheses: H1) To be able to occur in wetlands, tree species develop wider niches, meaning that adaptation to wetland environments leads to more generalist species and results in larger geographic ranges; H2) the occurrence in wetland habitats requires niche specialization, reflecting adaptive constraints and resulting in narrower niches and smaller range sizes. To test these hypotheses, we derived continental range size and hydrological niche breadth estimates for all taxonomically recognized Amazonian tree species and estimated the relationship between these variables, accounting for the ability or not of each species to colonize wetland habitats.

Methods

Data acquisition

We queried the GBIF and SpeciesLink databases for occurrence records of all preserved vascular plants specimens recorded between 1970 and 2017 across the Americas on January 25, 2018. The search returned 6 528 962 records from GBIF and 3 877 675

from SpeciesLink. We merged the two sets of records and filtered the results by matching species names to the taxonomic vetted list of 6 727 tree species that have been confirmed for the Amazonian lowlands (CARDOSO et al., 2017). The remaining records included only specimens already logged with their most up-to-date scientific names in GBIF and SpeciesLink, not including occurrence records for specimens labeled as synonyms or misspelled, to ensure taxonomic correctness. We then used the workflow implemented in the “speciesgeocodeR” package (TÖPEL et al., 2017) of the R language (R CORE TEAM, 2018) to remove occurrence records with geographic issues such as records outside terrestrial limits, missing values in the coordinates, non-valid coordinates, coordinates that are equal zero, latitudes equals longitudes, and records located up to 0.5 degrees from country capitals. We also excluded species with less than three occurrence records. Our final dataset comprised 571 092 occurrence records in the Americas, for 5 150 Amazonian tree species. Finally, we classified these 5 150 tree species into occurring or not occurring in Amazonian wetlands, based on our previously published Amazonian wetlands tree species list (LUIZE et al., 2018). Species occurring in wetlands yielded a total of 461 666 occurrence records and 2 838 species (Figure 3-1a); species not occurring in wetlands yielded a total of 109 426 occurrence records and 2 312 species (Figure 3-1b). Note that species occurring in wetlands can also have occurrence records in upland areas (i.e. we are not able to identify wetland-exclusive species due to the lack of consistent habitat metadata on species records, see Luize et al. 2018).

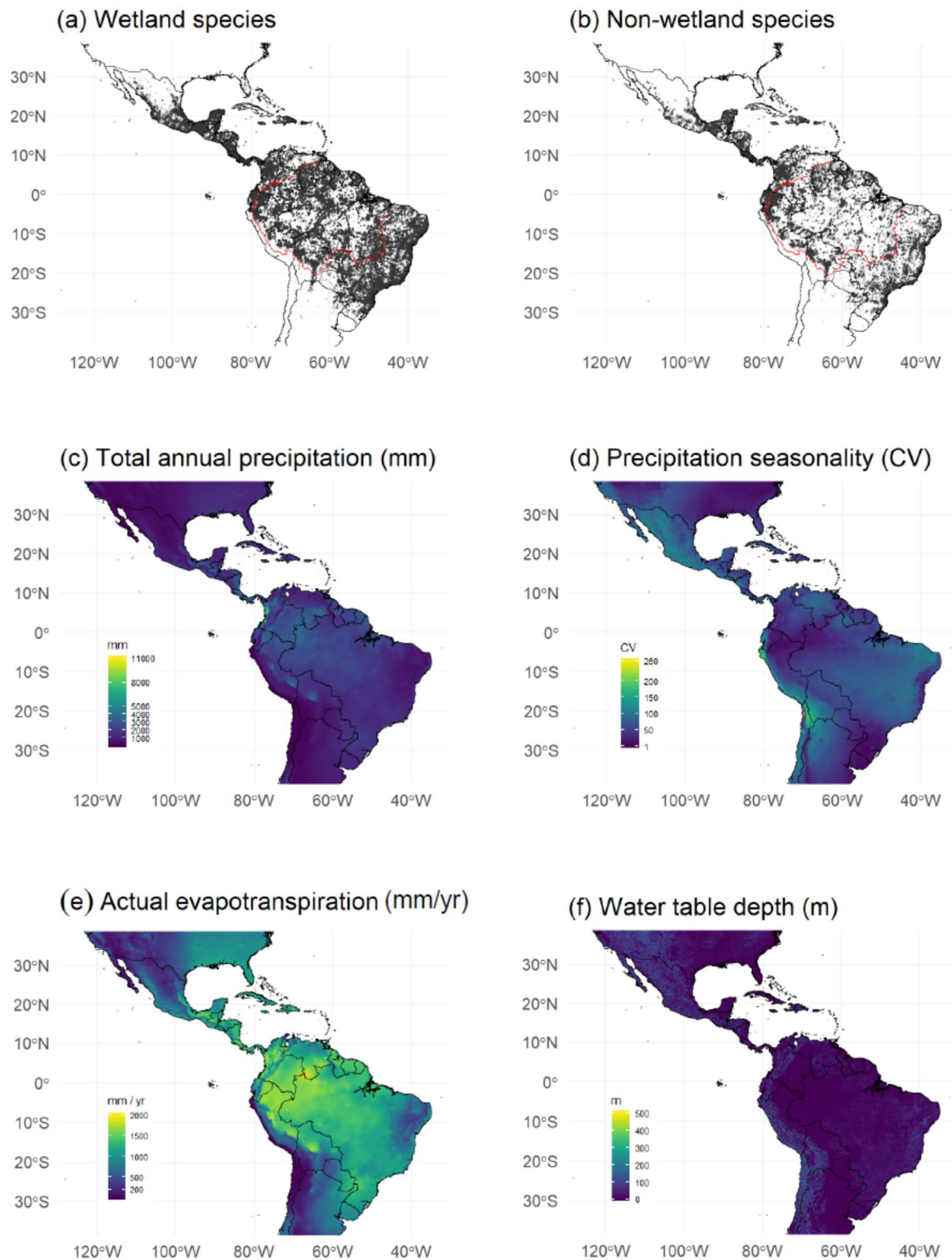


Figure 3-1. Map of the Americas bound by latitudes 35° N and S: (a) geolocated records of Amazonian tree species that do occur in wetlands (black dots); (b) geolocated records of Amazonian tree species that do not occur in wetlands. In (a) and (b) the red line shows the boundaries of Amazonia sensu stricto. (c) Total annual precipitation (mm), (d) precipitation seasonality (CV), (e) actual evapotranspiration (mm/yr), and (f) water table depth (m below land surface).

To describe the hydrological conditions found across the Americas, we chose four environmental variables among several possible candidates usually applied for modeling species niches, comprising three climatic and one edaphic variable. The first two climatic variables were determined using historical averages of total annual precipitation (mm) (Figure 3-1c) and precipitation seasonality (coefficient of variation) (Figure 3-1d) from 1960 to 1990, from the WorldClim v.1.4 database (i.e. bio12 and bio15 (HIJMANS et al., 2005)). Total annual precipitation estimates the expected amount of water reaching the soil per year, and precipitation seasonality indicates the likelihood of seasonal water deficit. The third climatic environmental condition was average actual evapotranspiration (mm/yr., Figure 3-1e), between 1950 and 2000, from the CGIAR-CSI soil-water balance model (TRABUCCO; ZOMER, 2010). Actual evapotranspiration characterizes a climatic condition of water-energy balance that is closer to real water availability to plants than precipitation amounts (STEPHENSON, 1990). Together, these three climatic environmental conditions approximate the amount and variability of plant water supply in time and space. As an edaphic environmental condition, we used the results from a global model for water table depth (m below land surface) (Figure 3-1f) that is constrained by ground observations and calibrated by climate, terrain and sea level (FAN; LI; MIGUEZ-MACHO, 2013). All environmental information was obtained as grids layers with cell spatial resolution equal to 30 arc-seconds (c.a. 1 km² at Equatorial latitudes).

Species niche breadths

To measure niche breadth, we extracted the raw and standardized z -values of the cells intersecting species occurrences for each selected environmental grid layer, using the function “extract” from the R package “raster” (HIJMANS, 2017). To compute the standardized z -values, we cropped each grid layer for the extended tropical American region (bound by 35° N and S), and then applied a standard z -normalization ($z = x_i - \mu / \sigma$, where: x_i is the grid cell value, and μ and σ are the sample mean and standard deviation of all grid cells across the region). We repeated this procedure for each selected

environmental layer. Then we computed species niche breadths as the univariate interquantile range between the 10th and 90th quantiles for each of the four environmental conditions. We chose interquantile distance rather than minimum and maximum values to reduce the influence of extreme environmental conditions that could arise from remaining geolocation errors or wrongful environmental estimates.

Species geographic range sizes

Geographic range sizes were calculated by fitting an α -convex hull to the occurrence records of each species across the entire tropical range of the Neotropics, as this method has been successfully used to measure and compare species ranges (GALLAGHER, 2016). The α -convex hull algorithm is based on the Voronoi diagram and Delaunay triangulation of spatial coordinates points and is suitable only when there are more than three georeferenced occurrence points (PATEIRO-LÓPEZ; RODRÍGUEZ-CASAL, 2010). To estimate the α -convex hull for each tree species, we used the function “EOO.computing” of the “ConR” R package (DAUBY et al., 2017), which also imports functions from the package “alphahull” (PATEIRO-LÓPEZ; RODRÍGUEZ-CASAL, 2010). The method implemented in “ConR” produces estimates biased towards wider distributions, but it is the standard method used to assess the conservation status following IUCN red-list standards (DAUBY et al., 2017). We first fit α -hulls using five different values of the α parameter ($\alpha = \{0.5, 1, 3, 5, 10\}$) and the default package value for the α buffer (0.1°) for all tree species. A value of α close to 0 will simply correspond to a distance buffer around each occurrence record, while α tending to infinity will lead to a generalized convex hull encompassing all point coordinates (GALLAGHER, 2016). For tree species with occurrence points following a straight line, range size was computed building a buffer polygon of 0.1° width around the line segment (DAUBY et al., 2017). After visually inspecting the resulting estimated ranges and plotting the overall range size distribution from each parameter combination, we selected $\alpha=3$ as the parameter with the best

compromise between locality and generality for estimating range size for all species. Areas were computed geodetically in reference to the WGS 84 ellipsoid.

Niche breadth and range size relationships

We evaluated separately the overall relationship between each species niche breadth dimension and range size, and the effect of wetland adaptations, using a series of general linear models.

$$\log_{10}(RS) = \alpha + \beta_1 NB \pm \varepsilon (Eq. 1)$$

$$\log_{10}(RS) = \alpha + \beta_1 NB + \beta_2 (W \vee NW) \pm \varepsilon (Eq. 2)$$

$$\log_{10}(RS) = \alpha + \beta_1 NB + \beta_2 (NW) + \beta_3 NB (W \vee NW) \pm \varepsilon (Eq. 3)$$

Where: RS is range size; NB is the measured niche breadth for each of the four studied niche dimensions; W is the set of tree species classified as occurring in wetlands and NW is the set of tree species that does not occur in wetlands; and α , β and ε are the estimated parameters of the models, respectively the intercept, slope and remaining deviance. To reduce non-linearity and heteroscedasticity we applied a base-10 logarithmic transformation to species range sizes. As the expected relationship between niche breadth and range size is positive, we compared the effect size of contrasting models using standardized slope coefficients. We used F-tests and AIC to compare the competing models.

As species niche breadths and range sizes were calculated using the same occurrence records, the estimates of the niche breadth – range size relationship were not completely independent. Although each estimate is based on a separate set of records (niche breadth uses all records while range size is based on the outermost records within the occurrence range), a remaining bias may still affect the estimations. To quantify this possible circularity bias we applied a randomization procedure with 500 iterations to produce independent estimates of the relationship between range size and niche breadth measurements. For each iteration, species occurrence records were split randomly into two

independent sets of equal size; the first half was used to determine species range size and the second half to measure species niche breadth, using the same procedure described previously. To refrain from obtaining spurious estimates, this randomization was applied only for tree species with more than 10 occurrence records (4 239 species). We then computed 500 coefficient estimates for all general linear models (Eqs. 1-3) using the independently derived variables. The difference between model slopes computed for all available occurrence records (β_{all}) and the model slopes computed from the randomized sets (β_{rand}) allowed the evaluation of the bias arising from lack of independence.

To analyze which niche breadth dimension had the greatest explanatory power when modeling Amazonian tree species range size, we fitted a multiple linear model with all niche variables.

$$\log_{10}(RS) = \alpha + \beta_1 NB_{TAP} + \beta_2 NB_{PS} + \beta_3 NB_{AET} + \beta_4 NB_{WTD} \pm \varepsilon (Eq. 4)$$

Where *TAP* is total annual precipitation, *PS* is precipitation seasonality, *AET* is actual evapotranspiration and *WTD* is water table depth. We then partitioned the total explained variance of the model following the approach by (LINDEMAN; MERENDA; GOLD, 1980), as implemented on the “relaimpo” R package (GRÖMPING, 2006). This approach decomposes the R^2 into non-negative contributions to the multiple linear model. The order of the explanatory variables is permuted, and the average of each variable’s contribution is computed over the different sets of models, without weighting the explanatory variables among those different models (GRÖMPING, 2006). Before applying the variance partitioning approach, we assessed the correlation between niche breadth measurements, and found a moderate correlation for most of the variable comparisons, with maximum Pearson's *r* coefficient of 0.65 between niche breadth for precipitation seasonality and actual evapotranspiration.

Results

Distribution of hydrological niche breadths for Amazonian tree species

The niche breadth of Amazonian tree species varied the most along the axis of water table depth (range = 0 – 410 m; z-values = 0 – 9.9; coefficient of variation = 93.8%). The niche breadth axis for total annual precipitation ranged between 0 and 6 252 mm (z-values = 0 – 7.2; cv = 54.2%), followed by actual evapotranspiration (range = 0–1 670 mm/yr.; z-values = 0 – 3.9; cv = 50.8%), and precipitation seasonality (range = 0 – 170; z-values = 0 – 7.2; cv = 43.05%). A total of 65 species had the measured niche breadth for at least one dimension equal to 0, and there were 23, 53, 20, and 31 species with niche breadth equal 0 for total annual precipitation, precipitation seasonality, actual evapotranspiration, and water table depth, respectively. In general, species with zero niche breadth had less than 5 occurrence records, but there was one species (*Guarea zepivae* T.D. Penn.) with a total of 11 occurrence records that still yielded a niche breadth of zero for precipitation seasonality. Niche breadth was weakly related to the number of available occurrence records used for its estimation ($r = 0.23 \text{ NB}_{\text{TAP}}$; $r = 0.32 \text{ NB}_{\text{PS}}$; $r = 0.30 \text{ NB}_{\text{AET}}$; $r = 0.09 \text{ NB}_{\text{WTD}}$).

Distribution of range sizes for Amazonian tree species

The number of occurrences records available to estimate range sizes was highly variable, with 51 tree species having the minimum necessary three points of occurrence, while 437 tree species had ≥ 300 records. The maximum number of records (4 059) was observed for *Myrcia splendens* (Sw.) DC., native from South and North America. The range size estimated for each species was strongly related to the number of occurrence records used for estimation ($r = 0.85$), and the distribution of range sizes for Amazonian tree species varied from 370 km² to 16 560 000 km², with a skew towards small ranges but clearly showing species with very large ranges.

Niche breadth – range size relationships for Amazonian tree species

All niche breadth variables were positively related to species range size (Figure 3-2). The steepest model slope was observed when niche breadth was characterized using precipitation seasonality ($\beta_{\text{std}} = 0.48$, Figure 3-2b), followed by actual evapotranspiration

($\beta_{\text{std}} = 0.42$, Figure 3-2c), total annual precipitation ($\beta_{\text{std}} = 0.39$, Figure 3-2a), and water table depth ($\beta_{\text{std}} = 0.06$, Figure 3-2d).

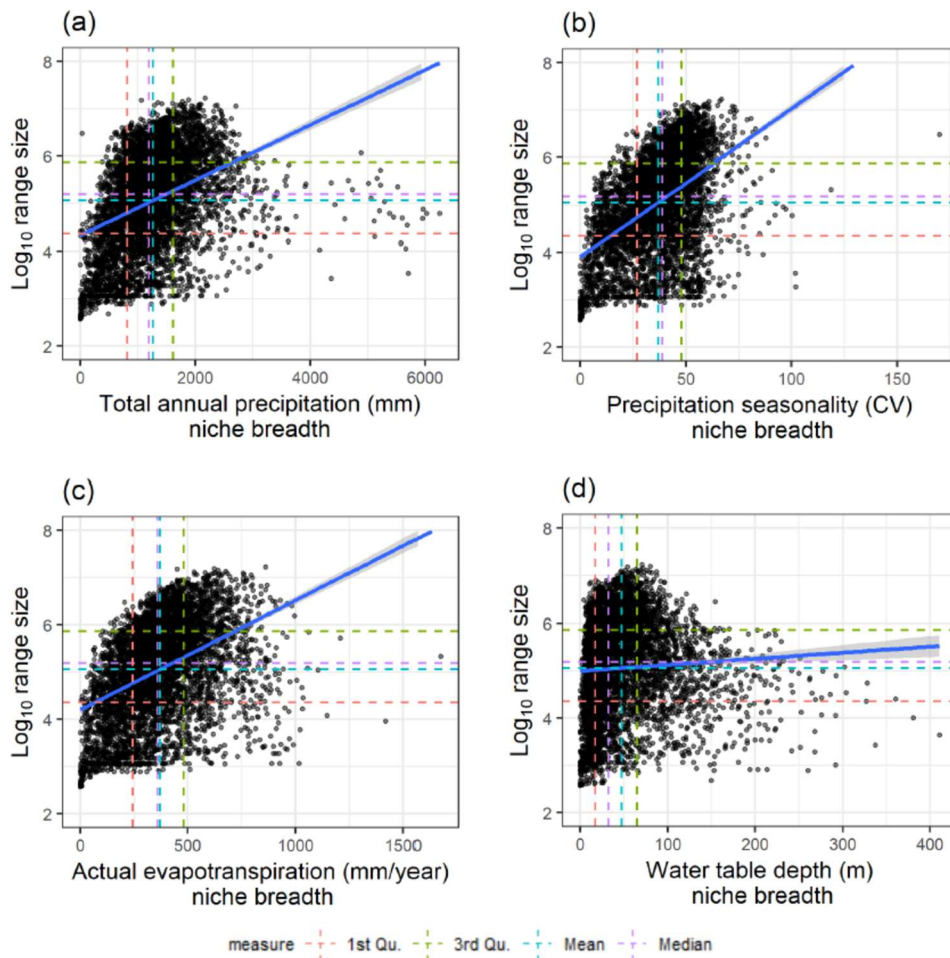


Figure 3-2. Relationship between niche breadth and range size (logarithmic scale) for 5 150 Amazonian tree species in four different niche axes: (a) total annual precipitation; (b) precipitation seasonality; (c) actual evapotranspiration; and (d) water table depth. All models show a positive relation between niche breadth and range size. Dashed horizontal lines show the mean, median and the first and third quartiles for the logarithmic distribution of range size and the linear distribution of niche breadth.

The lack of independence between estimations of niche breadth and range size did not change the observed positive relationship between niche breadth and range size (Figure 3-5 –3-8), but slopes estimated using randomization showed that lack of independence may lead to both slope underestimation and overestimation (Figure 3-5, Figure 3-6). The maximum difference between slopes estimated with the randomization procedure and the observed β_{std} slopes was small (β_{std} difference = - 0.31), with most β_{std} differences = $|0.03|$.

In general, slopes obtained by randomization were higher than slopes obtained using all available occurrence records (Figure 3-6). The only exception was the model for NB_{TAP} where slopes estimated using independent sets were lower than observed slopes (Figure 3-5).

The role of wetland adaptations on the niche breadth – range size relationships of Amazonian species

All general linear models yielded p-values lower than 0.001 as expected from the large sample sizes (Table 3-1). The lowest support for an estimated niche breadth – range size relationship was found for water table depth ($p = 0.00006$, Table 3-1). The models that included an interaction term between niche breadth and wetland occurrence (Eq. 3) had the lowest AIC values (Table 3-1), despite the larger number of estimated parameters.

Table 3-1. Estimated parameters for the generalized linear models relating niche breadth and the base-10 logarithm of range sizes for Amazonian tree species. The intercept and slopes for each model (Eqs. 1-3), standard deviance, F -statistic, AIC, and Δ AIC values are for comparison among the models. p -values were < 0.001 for all models.

Niche breadth	α (\pm sd)	β_1 (\pm sd)	β_2 (\pm sd)	β_3 (\pm sd)	R^2 (adj R^2)	F-statistic	AIC	Δ AIC
Annual precipitation		0.0006						
	4.321 (0.027)	(1.9e-05)	—	—	0.15 (0.15)	919.3	14096	1416
	3.843 (0.027)	(1.6e-05)	0.92 (2.3e-02)	—	0.34 (0.34)	1384	12727	47.66
Precipitation seasonality		0.0004	0.61 (4.9e-02)	0.00024 (3.4e-05)	0.35 (0.35)	948.3	12679	0
	3.909 (0.031)	(2.1e-05)	—	—	0.23 (0.23)	1543	13592	828.14
	3.758 (0.029)	(7.6e-04)	0.72 (2.4e-02)	—	0.34 (0.34)	1333	12795	31.38
Actual evapotranspiration		0.0208	0.38 (6.3e-02)	0.00902 (1.5e-03)	0.34 (0.34)	905.2	12763	0
	4.199 (0.028)	(6.9e-05)	—	—	0.18 (0.18)	1133	13918	1195.87
	3.844 (0.027)	(6.2e-05)	0.84 (2.3e-02)	—	0.34 (0.34)	1354	12767	45.04
Water table depth		0.0016	0.52 (5.2e-02)	0.00086 (1.2e-04)	0.35 (0.35)	926.5	12722	0
	4.998 (0.020)	(3.2e-04)	—	—	0.0031 (0.0029)	16.13	14926	1394.95
	4.300 (0.025)	(2.8e-04)	1.03 (2.5e-02)	—	0.23 (0.23)	808	13538	6.83
	4.332 (0.027)	(3.3e-04)	0.95 (3.8e-02)	0.00192 (6.4e-04)	0.24 (0.23)	542.4	13531	0

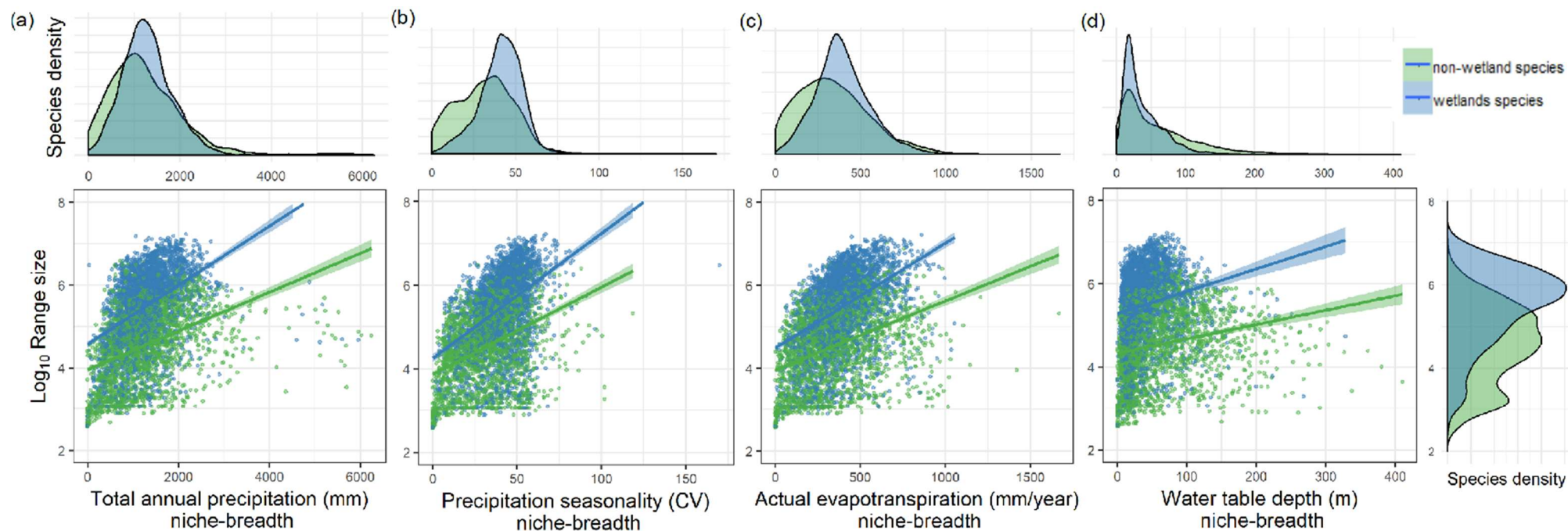


Figure 3-3. Relationship between niche breadth and range size (logarithmic scale) for Amazonian tree species that occur in Amazonian wetlands (n = 2,838 species) versus species that do not occur in Amazonian wetlands (n = 2,312 species). Linear models are shown for four different niche breadth axes: (a) total annual precipitation; (b) precipitation seasonality; (c) annual actual evapotranspiration; and (d) water table depth. The fitted curves for each group follow the model in Eq. 3.

The inclusion of wetland/non-wetland species as an explanatory variable yielded higher intercepts and slightly steeper slopes for wetland Amazonian tree species than for non-wetland species (Figure 3-3 and Table 3-1). The strongest difference in intercept and slope between species groups was observed when using water table depth as explanatory variable, followed by actual evapotranspiration (Table 3-1). All models showed a relatively steeper increase in the logarithm of range size along the gradient of niche breadth (i.e. steeper slope) for wetland tree species when compared with non-wetland species.

Variance partitioning among all four niche variables showed that precipitation seasonality had the highest explanatory power (12%), followed by actual evapotranspiration (7.5%) and total annual precipitation (7.2%), while water table depth had very low explanatory power (0.5%) in estimating the logarithm of species range size (Figure 3-4).

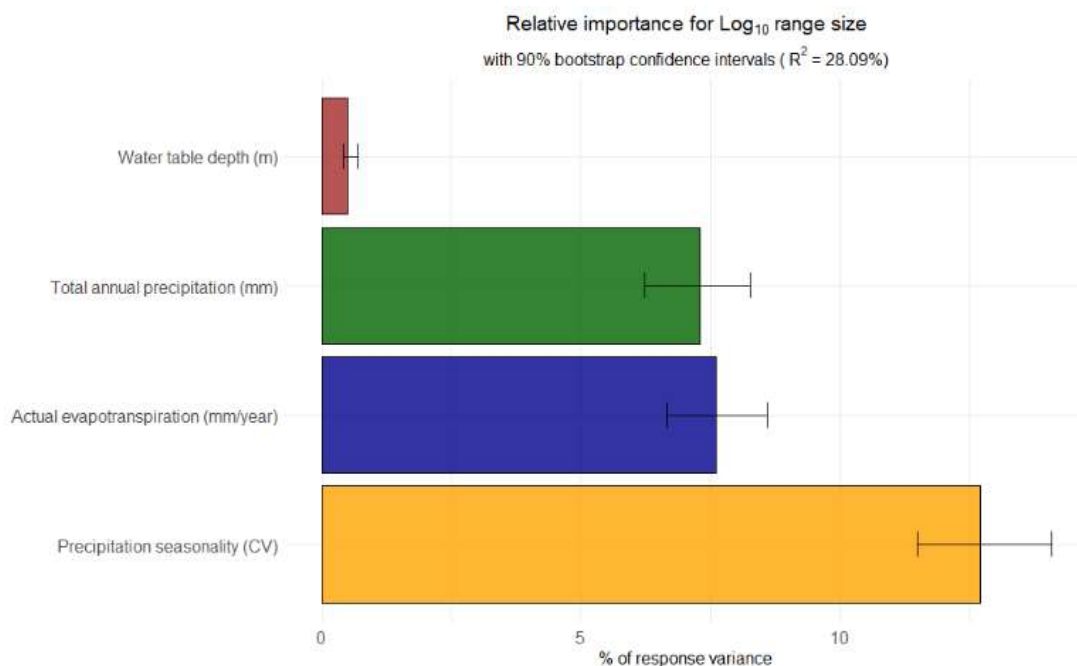


Figure 3-4. Relative importance of the four different hydrological niche breadth variables used to estimate log-transformed range sizes of Amazonian tree species. The goodness-of-fit partitioning was computed following the method by Lindeman, Merenda, and Gold (1980) to decompose total explained variance in multiple linear models.

Discussion

We provide empirical support to the general ecological pattern of positive relationships between niche breadth and range size (SLATYER; HIRST; SEXTON, 2013). Furthermore, our results show that tree species occurring in Amazonian wetland forests have a steeper increment in range size for each increment in niche breadth, compared to tree species only occurring in upland forests. The niche breadths and range sizes of Amazonian tree species varied greatly, showing a pattern where most species tolerate moderate climatic and edaphic variation in hydrological conditions, and fewer species tolerate broader variations in hydrological conditions (Figure 3-2a – 3-2c). For climatic niche breadths (*TP*, *PS* and *AET*) we observed a higher density of non-wetland species with narrower climatic hydrological niches, in comparison with wetland-adapted species, which had their peak density at wider niche breadths. Consequently, the geographic range size of most Amazonian tree species is relatively narrow (median range size of only 154 000 km²), and do not cover more than c.a. 2% of the extent of the Amazon basin. Conversely, at least 25% of Amazonian tree species are extremely widespread across Neotropical forests, with range sizes twice as large as the Amazon basin. Taken together, these observations suggest that wetland-adapted Amazonian tree species have a higher potential for acclimation to both wetter and dryer places, supporting our initial H1 hypothesis (to be able to occur in wetlands, tree species develop wider niches, meaning that adaptation to wetland environments leads to more generalist species and results in larger geographic ranges) in detriment of H2 (the occurrence in wetland habitats requires niche specialization, reflecting adaptive constraints and resulting in narrower niches and smaller range sizes).

The positive niche breadth – range size relationship together with the observation of a more left-skewed distribution of hydrological niche breadth indicates that few species can survive both extremes of the hydrological gradients, as both water surplus and deficit play a prominent role in plant physiology and species distribution (ESQUIVEL-

MUELBERT et al., 2017; KREFT; JETZ, 2007; MOULATLET et al., 2014; SCHIETTI et al., 2014; SILVERTOWN; ARAYA; GOWING, 2015; TER STEEGE et al., 2003). Hydrological niche segregation between species is expected to act at a very local scale (SILVERTOWN; ARAYA; GOWING, 2015), and findings showing water supply (i.e. water table depth) as the most important predictor of local scale distribution of Amazonian tree species support this expectation (MOULATLET et al., 2014; SCHIETTI et al., 2014). However, hydrological niches are also fundamental to define plant species distribution at very large scales. For instance, dry season length has been shown to be the strongest climatic predictor for the east-west gradient of tree α -diversity in Amazonia (TER STEEGE et al., 2003); and the global distribution of vascular plant diversity is driven by water-energy balance (KREFT; JETZ, 2007), emphasizing the coupled effect of seasonal energy input and water supply on plant establishment and coexistence.

Most Amazonian wetlands are forested floodplains, where soils are waterlogged annually from a few days to half of the year, but where seasonal droughts also take place (WITTMANN et al., 2013). This hyper-seasonality implies that tree species colonizing wetlands need to survive both flooding and droughts during their lifespan (PAROLIN et al., 2004), and our results demonstrate that this increased tolerance to contrasting hydrological conditions found in wetlands promotes wider geographic distributions. Moreover, wetland habitats may act as corridors for tree species dispersal, as wetlands cover large extents of the Neotropics, and particularly South America, creating a network of suitable habitat connections both within and among Neotropical biomes. Tolerance to wetlands may thus contribute to the explanation of why most species occurring in distinct Neotropical biomes have an Amazonian origin (ANTONELLI et al., 2018). One implication of the larger range sizes found for Amazonian tree species occurring in wetland is that those tree species are likely to comprise far-apart isolated populations, and experience genetic divergence through isolation-by-distance. Although we did not test for that hypothesis, isolation-by-distance has been demonstrated as an influential process

acting on *Inga* diversification (DEXTER et al., 2017), which is the most species-rich genus occurring in Amazonian wetland forests (LUIZE et al., 2018).

The range size of Amazonian tree species has been estimated before (FEELEY; SILMAN, 2009; GOMES et al., 2018; HUBBELL et al., 2008; TER STEEGE et al., 2015, 2016), but the total extent and methodological approach applied to define species distribution differs among studies, precluding further comparisons. To date, the most comprehensive estimate of range sizes for Amazonian plant species was produced with the aim to estimate species extinction risk (FEELEY; SILMAN, 2009). Subsequently, such estimates were applied to investigate broad scale macroecological patterns (DEXTER; CHAVE, 2016). The range size measures provided in our study are in accordance with IUCN standards, allowing a better evaluation of their conservation status by offering the extent of occurrences for Amazonian tree species throughout the entire Neotropics, and providing a fast and easily updatable characterization of tree species distribution. Considering our finding that most Amazonian tree species have small range sizes, their conservation status may be worse than previously evaluated (FEELEY; SILMAN, 2009; TER STEEGE et al., 2015). Furthermore, the α -convex hull approach has already been used to estimate global range sizes of the Australian seed flora (GALLAGHER, 2016) and to estimate species richness of the tribe Bignononieae in South America (MEYER; DINIZ-FILHO; LOHMANN, 2017), thus offering a common basis for future comparisons.

Our estimates of range size support that at least 25% of the recognized Amazonian tree flora has geographic ranges that extrapolate the limits of the Amazon basin, highlighting the need of biodiversity studies that go beyond Amazonian boundaries if we are to better understand species distributions, abundances, and niche breadths. There is a synergy between species niche breadth and range sizes (SEXTON et al., 2017), where the increase in niche breadth translates to an increase in range size, at the same time that the

occupation of dissimilar environments during range expansion is likely to broaden the species niches. Additionally, we conclude that tree species in wetlands can “go further” when compared with tree species that only occur in upland forests, both being relatively more generalists to hydrological conditions and colonizing widespread geographic area. If we are to move forward our understanding of the origins and maintenance of Neotropical tree diversity, one important functional aspect to be comprehensively considered and described is thus a species ability to cope with the dry and waterlogged conditions that dominate much of Neotropical forests and savannas.

Data Accessibility

The datasets used for this analysis are freely available and can be accessed online. Species occurrence records are available from the Global Biodiversity Information Facility – GBIF (<https://www.gbif.org/occurrence/search>) and SpeciesLink (<http://inct.splink.org.br/>) databases. Climate data are available from WorldClim (<https://www.worldclim.org/bioclim>). The global dataset of actual evapotranspiration is available from the CGIAR-CSI GeoPortal (<https://cgiarcsi.community/data/global-high-resolution-soil-water-balance/>). Water table depth model available upon request from the Authors (Fan et al., 2013). R scripts to compute and export the species range size area and the associated shapefile and to extract and compute the species niche breadth were available upon request to the correspondence author.

Supporting information

Table 3-2. Table S1. Range sizes and niche breadths estimated for each one of the 5.150 Amazonian tree species with more than three occurrence records.

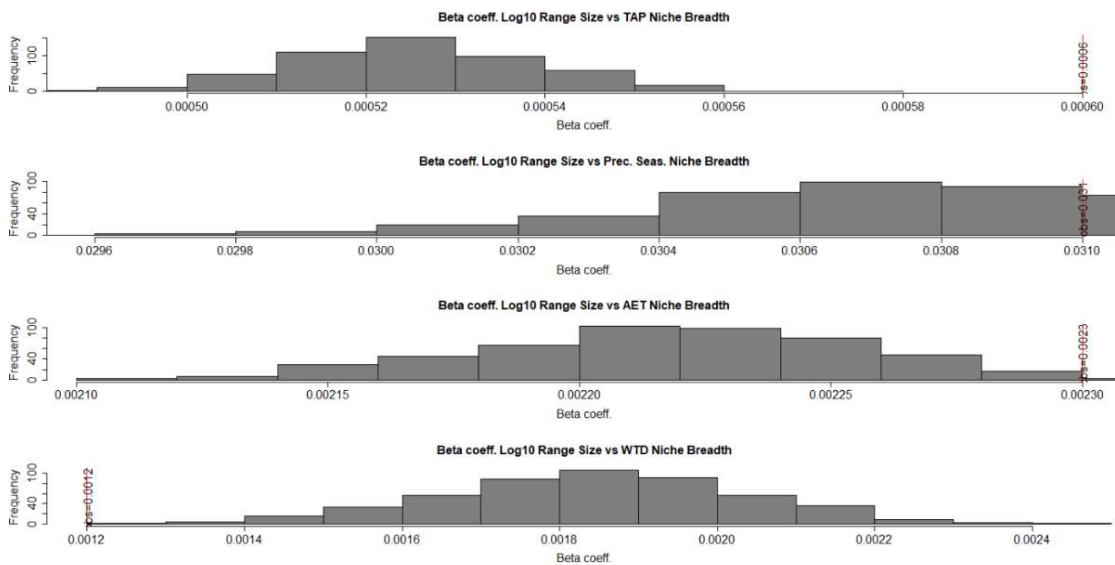


Figure 3-5. Figure 3.S1. Histogram for the slopes estimated by the models $\log_{10}(RS) = \alpha + \beta_1 NB \pm \varepsilon$ (Eq. 1) relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.

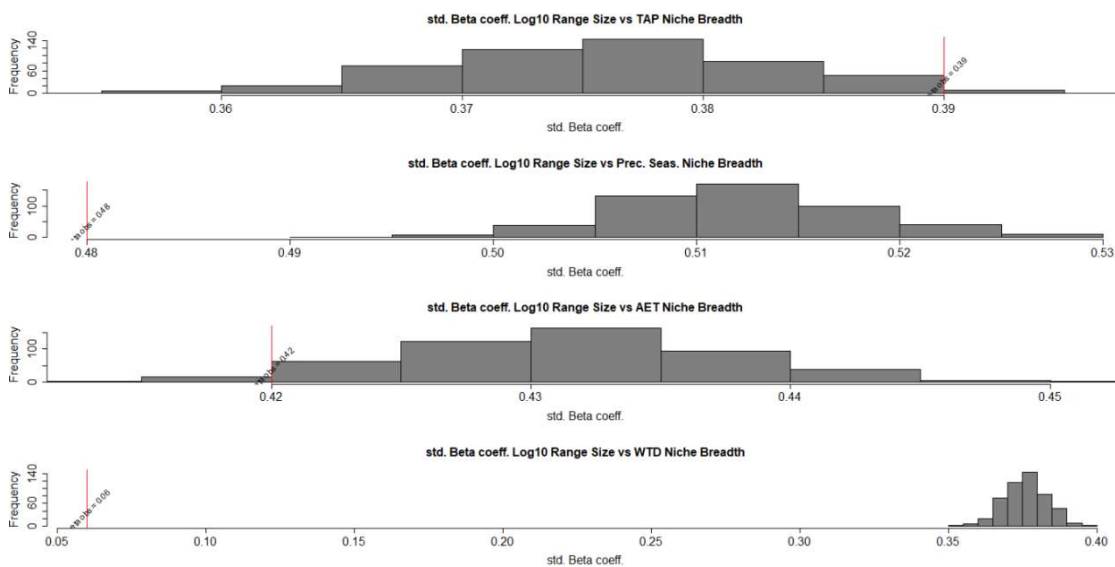


Figure 3-6. Figure 3.S2. Histogram for the standardized slopes estimated by the models $\log_{10}(RS) = \alpha + \beta_1 NB \pm \varepsilon$ (Eq. 1) relating independent measurements of species niche breadth and range size in comparison with the observed standardized slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.

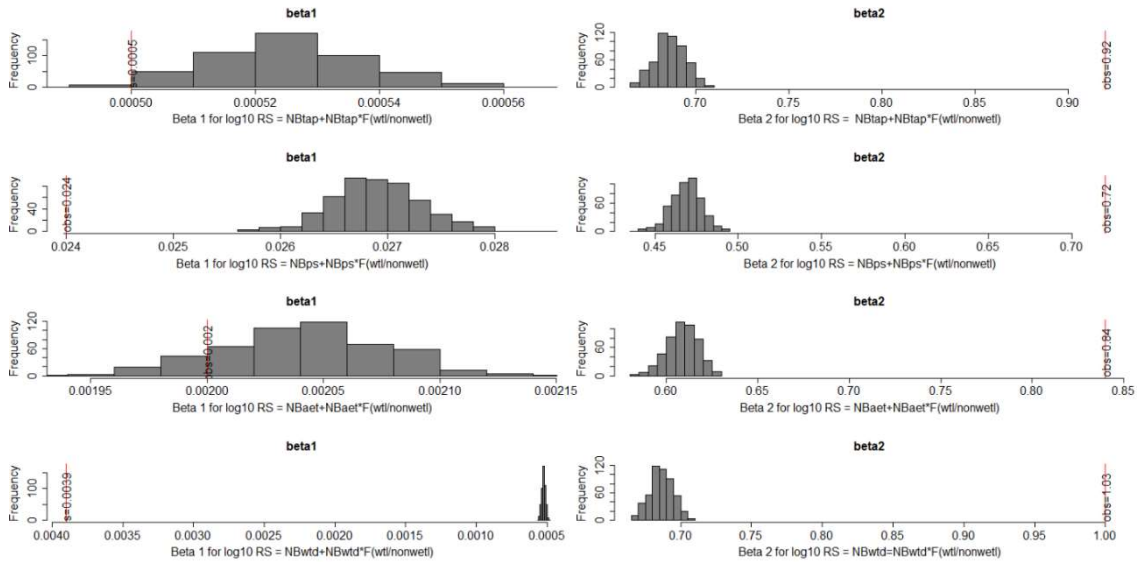


Figure 3-7. Figure 3.S3. Histogram for the slopes estimated by the models $\log_{10}(RS) = \alpha + \beta_1 NB + \beta_2 (W \vee NW) \pm \varepsilon (Eq. 2)$ relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.

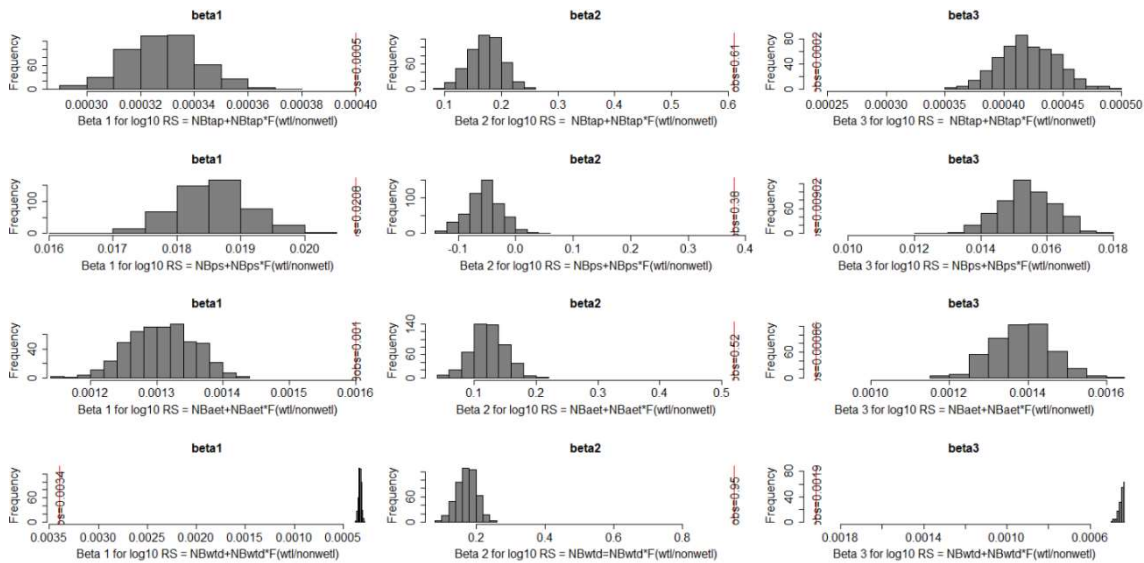


Figure 3-8. Figure 3.S4. Histogram for the slopes estimated by the models $(\log_{10}(RS) = \alpha + \beta_1 NB + \beta_2 (NW) + \beta_3 NB(W \vee NW) \pm \varepsilon (Eq. 3)$ relating independent measurements of species niche breadth and range size in comparison with the observed slopes for model relating measurements of species niche breadth and range size computed with all available occurrence records (red line). For details please see main text.

4. Modelling compositional and phylogenetic beta-diversity in

Central Amazonian floodplain forests

Luize BG; Silva TSF; Venticinque EM; Novo EMLM; Silva CP; Harwood TD, Ware C, Mokany K, Rosauer DF, Ferrier S (2019) Modelling compositional and phylogenetic beta-diversity in Central Amazonian floodplain forests. Em preparação para *Ecography*.

Abstract

Amazonian floodplains show high number of tree species able to survive recurrent long-lasting inundations. Local tree diversity is related to flood duration, an indirect environmental gradient, tight coupled to fluvial dynamics, environmental heterogeneity, soil fertility, and forest succession. The strong environmental gradients sorting tree species in distinct communities may explain high β -diversity in floodplains, which can be as high as in upland forests. However, arrival of species from adjacent upland forests, and neutral processes also may play an influence β -diversity. Our aim is to improve our understanding of the distributional patterns of compositional and phylogenetic β -diversity in floodplain forests of central Amazonia, and explore the relative role of environmental gradients, neutral, and historical processes influencing these patterns. We applied generalized dissimilarity modelling to evaluate and predict compositional and phylogenetic β -diversity. Model deviance partitioning was applied to evaluate the relative role of geographic separation and environmental gradients in explaining the turnover and nestedness components of compositional and phylogenetic β -diversity (Sørensen index) between all possible pairs of 43 floodplain forest sites separated from 1 to 500 kilometers apart each other. We spatially projected the GDMs to map expected patterns in species composition and phylogenetic turnover. Results provide evidence for high compositional and phylogenetic turnover throughout central Amazon floodplain forests. The compositional and the phylogenetic components of turnover exhibit markedly different relationship with environmental and geographic gradients. While compositional turnover is mostly explained by the geographic separation, phylogenetic turnover is mostly explained by environmental differences between sites. We conclude that neutral processes are a pronounced driver influencing species composition, while ecological sorting is the most influential process involved in assembling phylogenetic lineages in floodplain forests. Our analyses provide a foundation for future monitoring of ongoing compositional changes in Amazonian floodplain forest.

Keywords: Generalized Dissimilarity Model; Wetland forests; Tree diversity; Community Assembly; Beta diversity additive partitioning

Introduction

The study of biodiversity through space and time has shown that intensive land use and climatic change are modifying forest dynamics, floristic and functional composition of tree assemblages in Amazonia (ALEIXO et al., 2019; ESQUIVEL-MUELBERT et al., 2019; LAURANCE et al., 2006; RESENDE et al., 2019). This understanding of temporal dynamics is derived largely from data collected at long term monitoring sites. Such monitoring requires continued and considerable input of resources, which limits the number of sites which can be surveyed (ESTES et al., 2018) , and therefore the extent to which these sites representatively sample environmental and geographic gradients shaping spatial patterns in the distribution of biodiversity, thereby allowing inferences to be made across large regions. The understanding of, and the ability to predict, temporal variation in floristic composition may be enhanced by an improved knowledge of spatial variation, through space-for-time substitution, and vice versa (BLOIS et al., 2013). In the present study we evaluate tree species compositional and phylogenetic β -diversity, providing spatially explicit estimates for β -diversity through central Amazonia floodplain forests.

An important distinction exists between Amazonian forests which occur in uplands, and do not become flooded, and those in wetlands that are seasonally flooded. Compared to Amazonian upland non-flooded forests, local tree diversity (i.e. α -alpha diversity) is lower in wetland seasonal flooded forests (TER STEEGE et al., 2000). However, the tree species pool (i.e. γ -gamma diversity) recorded for Amazonian wetland forests is high and accounts for half of the total number of species recorded within the entire Amazonian region (LUIZE et al., 2018). Since Amazonian wetland forests have low α -diversity and high γ -diversity, we can expect high levels of compositional dissimilarity between stands of wetland forests (i.e. high β -beta diversity). Visual comparisons of the distance decay of similarity reported for Amazonian wetlands (ALBERNAZ et al., 2012; WITTMANN et al., 2006) with that reported for upland forests (e.g. CONDIT et al., 2002)

suggest analogous β -diversity levels for both forests. Indeed, when evaluated side by side, the distance decay of similarity for *Inga* species shows indistinguishable intercept and slopes for both upland and floodplain communities (DEXTER; TERBORGH; CUNNINGHAM, 2012), providing support for high levels of β -diversity within wetland forests. However, much of the explanatory value which geographic distance provides for the understanding of β -diversity is actually shared with underlying environmental predictors (e.g. DUIVENVOORDEN; SVENNING; WRIGHT, 2002). Therefore, we need to consider both ecological and evolutionary processes, species dispersal limitations (i.e. neutral processes), environmental niches of species (i.e. ecological determinism), and historical constraints (i.e. biogeography and evolution)(CAVENDER-BARES et al., 2009; GRAHAM; FINE, 2008; WEBB et al., 2002) to better understand spatial variation of β -diversity in these forests.

In addition, compositional changes within Amazonian floodplain forests are influenced by various processes, including forest disturbance due to fluvial dynamics (SALO et al., 1986), forest succession (WITTMANN; JUNK; PIEDADE, 2004), flooding gradients (ALBERNAZ et al., 2012; MONTERO; PIEDADE; WITTMANN, 2012; WITTMANN; ANHUF; FUNK, 2002), heterogeneity in soil proprieties (DE ASSIS et al., 2017), and species lateral migration from adjacent upland forests (TERBORGH; ANDRESEN, 1998). Indeed, Amazonian floodplains show a mosaic of habitats patches that vary greatly across the landscape and through short environmental gradients (FERREIRA-FERREIRA et al., 2015), and a strong negative gradient for α -diversity tightly correlated with small differences in flood height between forests (LUIZE et al., 2015a).

Despite this accumulated knowledge of an array of factors explaining compositional dissimilarity between floodplain forests, most previous studies have been based on the application of multivariate ordination techniques (e.g. DE ASSIS et al., 2017;

MONTERO; PIEDADE; WITTMANN, 2014; TERBORGH; ANDRESEN, 1998). A few studies have also applied mantel correlations between geographic distances and compositional dissimilarities (ALBERNAZ et al., 2012; WITTMANN et al., 2006), both to determine distinct biogeographic areas/ecoregions (ALBERNAZ et al., 2012), and to compare the decay in similarity between discrete communities within floodplain forests (WITTMANN et al., 2006). Such methods, although performing well in describing the complexity of the system and to order the samples in relation to field-observed environmental variables, they offer limited capacity for inference and extrapolation of results beyond surveyed locations. Recent advances in remote sensing provide geographically-complete mapping of environmental gradients, thereby allowing the extrapolation of β -diversity across whole regions (ASNER et al., 2017; HIGGINS et al., 2014). However, successful application of remote sensing datasets still depends on ground-based measurements of species composition (FERRIER et al., 2007; FERRIER; GUIBAN, 2006; ROCCHINI et al., 2018).

In this study, we evaluate how tree species compositional and phylogenetic β -diversity of central Amazonian floodplain forests is related to environmental gradients derived from remote sensed datasets, and with geographic distance between sites, thereby allowing spatially explicit prediction of expected β -diversity patterns across the region. We hypothesized that the patterns of compositional and phylogenetic β -diversity would be highly concordant, shaped in a similar way by geographic distance and environmental gradients. By providing spatially explicit predictions of compositional and phylogenetic β -diversity at fine resolution we establish a baseline for future surveys and comparisons that can assist ongoing monitoring of temporal and spatial changes in floodplain forest biodiversity.

Methods

Floodplain forest inventory plots

We established 43 inventory plots in whitewater floodplain forests located in three landscapes of the Central Amazon region: 1) Badajos Lake (LG_BAD), n=11 plots; 2) Sustainable Development Reserve Piagaçu–Purus (SDR_PP), n=20 plots; and 3) Madeira–Amazonas Confluence (MAD_AM)), n=12 plots (Figure 1). Plot area was 0.375 ha (150 x 25 m), keeping homogeneous flood variation in each sample. In each landscape, plots were located at least 1 km apart each other and comprising available flood height gradient on that floodplains. A total of 8,975 trees ≥ 10 cm DBH (i.e.: diameter breast height or 1.3m from the tree base) were marked with numerated aluminum tags and 432 taxonomic units identified (i.e. 376 species – 87% valid species, and 56 morphotypes – 13% identified to genus level). Morphotypes were standardized and reference vouchers were included in INPA and IFAM herbarium.

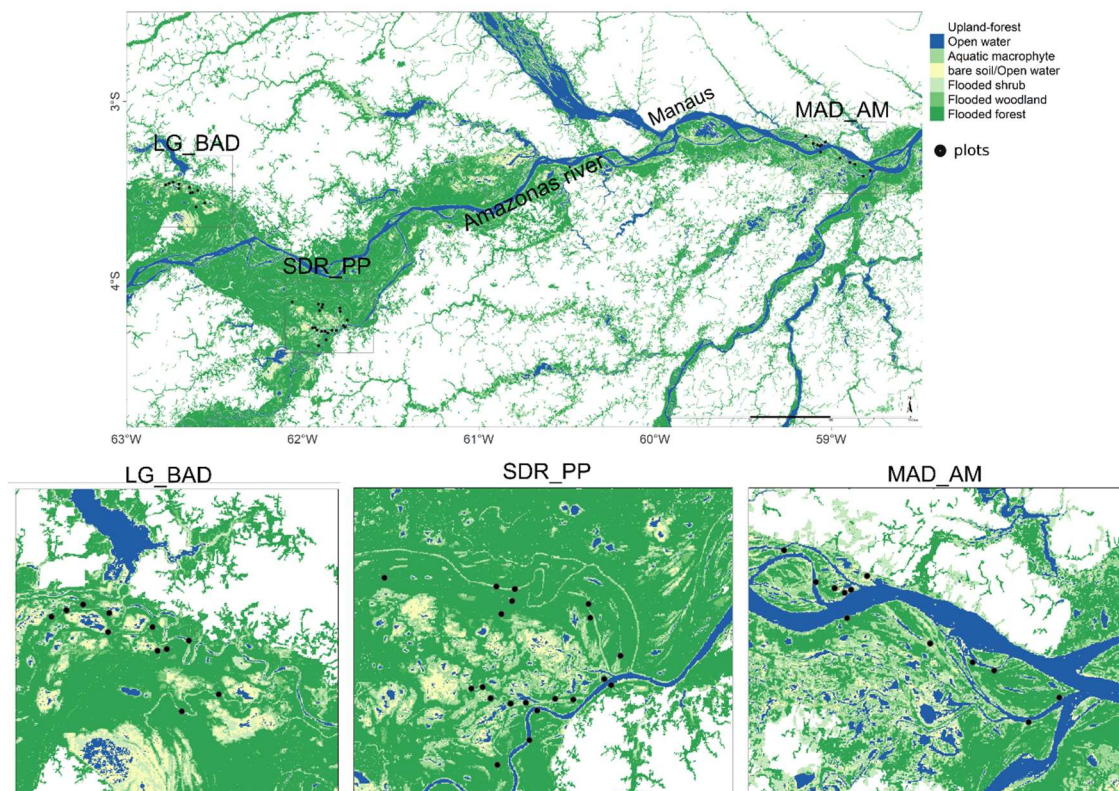


Figure 4-1. Wetlands in the Central Amazon region and surveyed forest sites within the three floodplain landscapes considered in this study. Map colors show a simple reclassification of the dual-season wetland mask (HESS et al., 2015b).

Floristic and phylogenetic dissimilarity measurements

Compositional and phylogenetic β -diversity were measured with the Sørensen index (β_{sor} and PhyloSor respectively) and decomposed into turnover and nestedness components (BASELGA, 2010, 2012; BRYANT et al., 2008; LEPRIEUR et al., 2012). The Sørensen index is a monotonic transformation of β -diversity that measures the proportion of unique species for a given pair of sites. For compositional β -diversity the dissimilarity arising from species replacement/turnover was measured with the Simpson dissimilarity index (β_{sim}), while dissimilarity due to the species loss/gain from place to place was measured with the nestedness index (β_{sne}) (BASELGA, 2010, 2012). Sørensen phylogenetic dissimilarity (PhyloSor), instead of using the number of shared species between communities, considers the sum of the branch lengths shared between pairs of communities (BRYANT et al., 2008; LEPRIEUR et al., 2012). PhyloSor ranges from 0 to 1. Two locations with distinct species sharing very small amounts of evolutionary history will yield PhyloSor value close to 1, whereas if they share exactly the same species PhyloSor will be 0 (BRYANT et al., 2008; LEPRIEUR et al., 2012). PhyloSor was decomposed into ‘true’ phylogenetic turnover (PhyloSor_{TURN}) and phylogenetic diversity gradient (PhyloSor_{PD}) (LEPRIEUR et al., 2012).

A comprehensive synthesis phylogeny for vascular plants (ALLOTB) including 353,187 taxa (SMITH; BROWN, 2018) was employed to compute phylogenetic β -diversity metrics. ALLOTB is a dated phylogeny that comprises the largest number of species and the best resolution available for the placement of taxa without molecular data. Morphotypes and species names not matching ALLOTB names were removed (56 morphotypes and 51 species), eliminating 25% of the taxonomic unities. From the 51 species removed from the original matrix, most occur at low frequency across the plots (3^o quartile: 19% of the plots). The species richness gradient between sampled plots with and without filtering species remains unchanged (spearman’s $\rho = 0.98$). Compositional β -diversity matrices (β_{sor} , β_{sim} and β_{sne}) and phylogenetic β -diversity matrices (PhyloSor,

PhyloSor_{TURN}, PhyloSor_{PD}) were computed with the R package ‘betapart’ (BASELGA et al., 2018; BASELGA; ORME, 2012).

Remote sensing data

A total of 28 grid layers in a 30 meters spatial resolution (Table 4-1) for the entire region of interest (bounded in and Longitude 63, 58.5 W, Latitude 2.5 and 4.8 S, Figure 4-1) were compiled using Google Earth Engine platform (GORELICK et al., 2017). Before modelling, the respective remote-sensing values of the cells in the 28 grid layers were extracted using the geographic coordinates of the surveyed sites, thereby producing a table for those environmental attributes.

Topography and terrain heterogeneity were obtained from the digital surface model produced from Advanced World 3D (AW3D version 1.0, TAKAKU et al., 2016) as 1) elevational position, 2) multi-scale topographic position index (MTPI) and 3) topographic heterogeneity (TopoDiver) (THEOBALD; HARRISON-ATLAS; MONAHAN, 2015).

Vegetation structure and productivity (e.g., MORTON et al., 2014) were characterized by the Normalized Difference Vegetation Index (NDVI) complemented by the Normalized Difference Water Index (NDWI; GAO, 1996). NDVI and NDWI were computed with information gathered from the Landsat 5 ETM imagery collection (1997 to 2012). These indices differ only in the channel reflectance applied to normalize the vegetation index, since by using SWIR wavelengths NDWI gives information on leaf water content providing a better discrimination on vegetation liquid water status (GAO, 1996). NDVI and NDWI are sensitive to background soil reflectance, meaning canopy openness and/or vegetation coverage affects index computation, producing lower values where forest canopy is less dense, with NDWI showing even lower values under flooded forest conditions. NDVI and NDWI are sensitive to background soil reflectance, meaning canopy openness and/or vegetation coverage affects index computation, producing lower

values where forest canopy is less dense, with NDWI showing even lower values under flooded forest conditions. For each normalized vegetation index, four layers assessing 1) the total annual forest productivity, 2) forest productivity seasonality, 3) vegetation productivity in the flooded (wet), and 4) vegetation productivity in the non-flooded (dry) season in central Amazonian floodplains were produced using 90th quantile NDVI and NDWI annual values (Table 4-1), to avoid outliers, and capture ‘pure’ reflectance signals.

For modelling purposes, a set of 17 statistical descriptors for C-band backscatter signal captured by the Synthetic Aperture Radar onboard on the satellite Sentinel-1 was computed (Table 4-1). The data collection of the Sentinel-1 were obtained for a 5 years’ time-series (2014 to 2019) as level-1 ground range detected scenes for the interferometric wide swath mode, both in vertical and horizontal polarization. Sentinel-1 is designed to provide of Earth surface regardless of weather, and the radar imagery is designed to contribute to the mapping of standing water (floods), land cover, soil moisture and vegetation biomass (MALENOVSKÝ et al., 2012).

β-diversity modelling

Generalized Dissimilarity Modelling (GDM) was used to analyze the relationship between compositional and phylogenetic β -diversity components to the set of environmental attributes and the geographic distance between floodplain forests sites. GDM is a statistical method for analyzing and predicting spatial patterns of dissimilarity in species composition over large regions (FERRIER et al., 2007), and has been extended to accommodate phylogenetic and genetic distances (FERRIER et al., 2007; FITZPATRICK; KELLER, 2015; ROSAUER et al., 2014). The method relates biotic and abiotic variables, while accounting for non-linearity in the relationship between biological dissimilarity and ecological distance, and for the non-stationarity in the rate of compositional turnover along any given environmental gradient (FERRIER et al., 2007). To fit each GDM model, the standard three I-spline basis functions (i.e. defined by knots

at 0 (minimum), 50 (median), and 100 (maximum) quantiles) were calculated as implemented in the R package “gdm” (MANION et al., 2017).

The total deviance explained by each GDM including all the 28 environmental attributes plus geographic distances between sites, was partitioned into three components: 1) deviance explained purely by geographic distances; 2) deviance explained purely by environmental distances; and 3) explained deviance shared between environmental and geographic distances. The relative role of the environmental and geographic gradients in shaping compositional and phylogenetic β -diversity was then assessed by comparing the amounts of deviance explained by each of these components.

As the use of large numbers of predictors may result in model overfitting, and because some of the predictors may have no explanatory value, a backward selection procedure was undertaken to progressively removing less important predictors to deviance explanation of a given model. Variable importance is quantified as the percent change in deviance explained between a model fitted with and without that predictor (Manion et al. 2018). This was achieved by permutate 999 times the GDMs fitting but leaving one predictor aside on each permutation; at the of each permutation only those predictors with higher importance to deviance explained is kept. The model permutation and the dropping of unimportant predictors continues until all remaining predictors do not reduce total deviance explained. Backward selection was performed as implemented by the function ‘gdm.varImp’ from the R package “gdm” (MANION et al., 2017). After reaching a solution, we have adjusted a simple GDM that was then applied to transform the relevant environmental and geographic grid layers into a multi-dimensional grid with predicted ecological distances. The reduced GDM resulting from this backward selection procedure was then used to transform the relevant environmental and geographic grid layers into a multi-dimensional grid according to the fitted GDM functions. To reduce redundancy, the resulting multi-dimensional grid was submitted to a Principal Component Analysis (PCA),

and the first three principal components were mapped using a scaled (i.e. 0-255) RGB color space to depict patterns of β -diversity across the region of interest.

Results

Floodplain forests compositional and phylogenetic β -diversity

Compositional β -diversity varied from a minimum β_{sor} of 0.39 to a maximum of 1.00 (β_{sor} : 1st quartile = 0.70, 3rd quartile 0.88, Figure 4-2). This component of β -diversity arises mainly from turnover, with β_{sim} values ranging from a minimum of 0.21 to complete replacement of species between forest stands, $\beta_{\text{sim}} = 1.00$ (β_{sim} : 1st quartile = 0.65, 3rd quartile 0.83, Figure 4-2). Maximum β_{sne} was 0.45, but most of nestedness index values were lower than 0.06 (β_{sne} : 1st quartile = 0.01, 3rd quartile 0.06), including pairs of plots which do not show any nestedness (i.e. $\beta_{\text{sne}} = 0$).

Phylogenetic β -diversity varied from minimum PhyloSor of 0.11 to 0.75 (PhyloSor: 1st quartile = 0.40, 3rd quartile 0.55, Figure 4-2). The higher influence on phylogenetic β -diversity comes from phylogenetic turnover, where the maximum PhyloSor_{TURN} was 0.63 indicating a pairs of forests plots assembled with distantly related species (i.e., sharing little evolutionary history), but minimum PhyloSor_{TURN} values were very close to 0 (minimum PhyloSor_{TURN} 0.0015), indicating pairs of forest plots with all tree species very closely related (PhyloSor_{TURN} 1st quartile = 0.30, 3rd quartile 0.43). In contrast to what was observed for nestedness component of compositional β -diversity, the phylogenetic diversity component of phylogenetic β -diversity achieved high values (i.e. maximum PhyloSor_{PD}=0.54), suggesting a strong influence of the phylogenetic diversity gradient on observed phylogenetic β -diversity (i.e. PhyloSor). However, minimum PhyloSor_{PD} was 0.00001 (1st quartile = 0.03, 3rd quartile 0.16), suggesting that the phylogenetic diversity gradient influenced only a few pairs of forest plots.

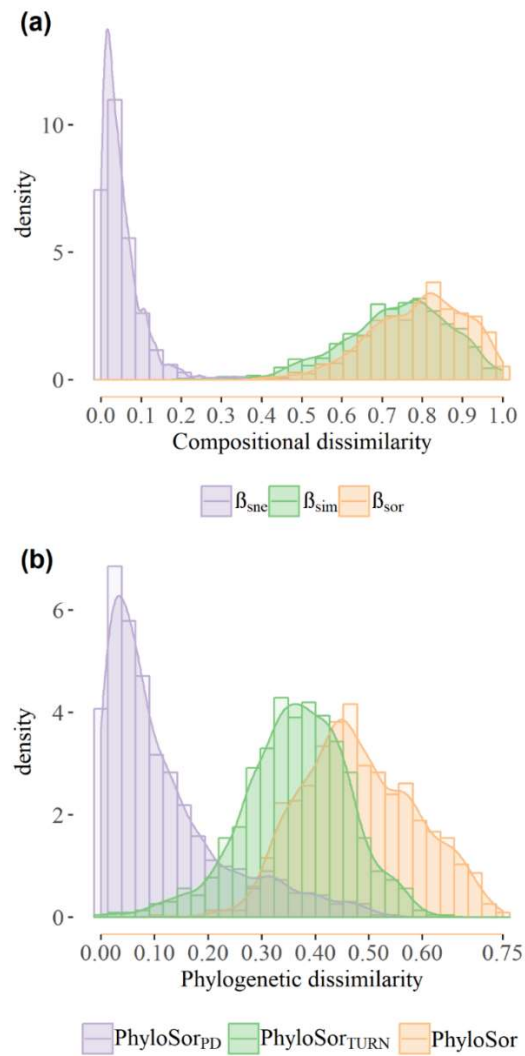


Figure 4-2. Density plots for the components of a) compositional β -diversity, and b) phylogenetic β -diversity. Note the difference in axis scales.

Observed compositional and phylogenetic β -diversity was direct correlated (β_{sor} and PhyloSor: $r_{mantel}=0.87$, $P=0.001$, Fig 3a). The turnover (β_{sim} and PhyloSor_{TURN} $r_{mantel}=0.59$, $P=0.001$, Fig 3b) and nestedness (β_{sne} and PhyloSor_{PD}: $r_{mantel}=0.60$, $P=0.001$, Fig 3c) components of compositional and phylogenetic β -diversity were also positively correlated, but with lower correlation coefficients.

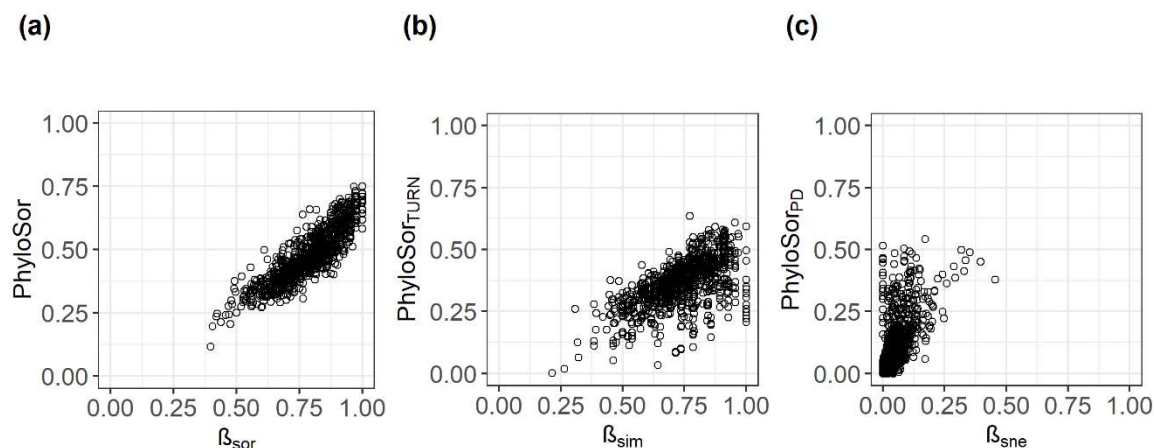


Figure 4-3. The relationship between compositional and phylogenetic β -diversity components for all site pairs studied, where (a) is the relationship for Sørensen dissimilarity index without performing the additive partitioning decomposition; (b) is the relationship for the turnover component of β -diversity; and (c) is the relationship for the nestedness component of β -diversity.

Geographic distances or environmental distances: which set of predictors explain most amounts of deviances in compositional and phylogenetic β -diversity?

The full GDM explains 38% of the β_{sor} , and 41% of the PhyloSor adjusted deviance (Figure 4-4). Geographic distance between sites was the greatest contributor to deviance explained for both β_{sor} and PhyloSor (Figure 4-4). However, important differences emerged for the GDMs fitted to the turnover and nestedness components of β -diversity. For the turnover component of compositional β -diversity (β_{sim}) the environmental predictors were comparatively of lower importance in explaining model deviance than for the turnover component of phylogenetic β -diversity (PhyloSor_{TURN}) (Figure 4-4). Geographic distances were the predictor accounting for the higher amounts of model deviance of β_{sim} , while the environmental predictors accounts for most of PhyloSor_{TURN} model explanation (Figure 4-4). Furthermore, the amount of deviance in β_{sim} explained by geographic distance exceeded the total amount of deviance explained by both geographic and environmental distances for the PhyloSor_{TURN} model. The same does not hold for the β_{sne} and PhyloSor_{PD} comparison, where environmental distances explained most of the deviance of these models (Figure 4-4). It is worth noting, that the GDM achieving the

highest percentage of deviance explained was for PhyloSor, while the GDM with the lowest percentage of deviance explained was for the PhyloSor_{TURN} (Figure 4-4).

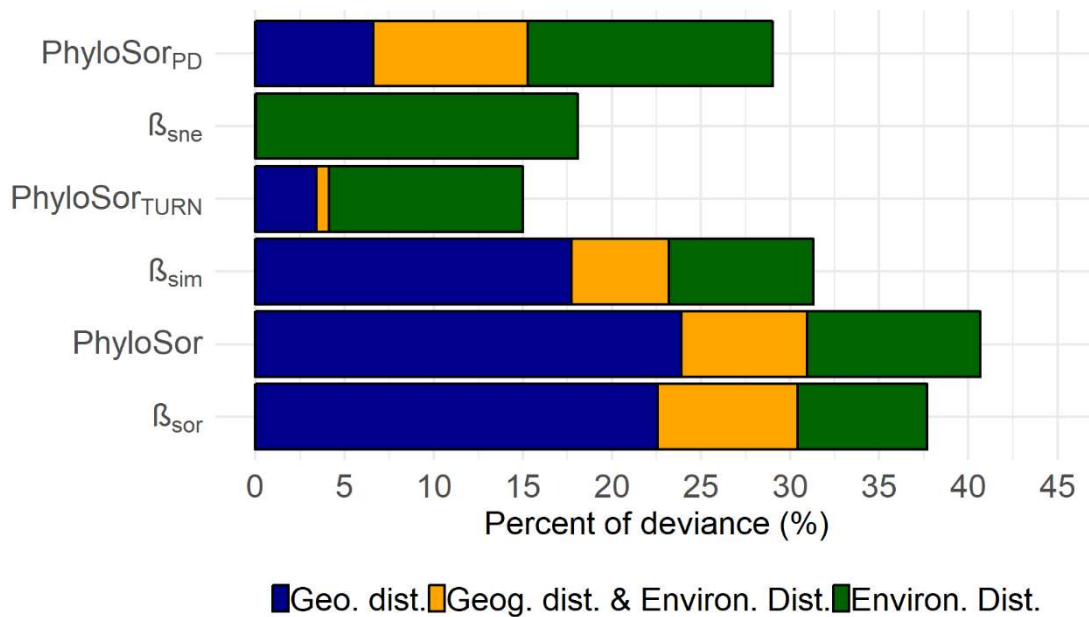


Figure 4-4. Percentage of deviance explained, partitioned by geographic distance, environmental distance and shared among geographic distance and environmental distance, for GDMs fitted to compositional and phylogenetic β -diversity components. The GDMs were fitted for all the 903 pairwise comparisons between 43 plots established in central Amazonia, with the exception of the GDM fitted to the β_{sne} matrix, which was fitted only to site pairs with $\beta_{sne} \geq 0.06$ (i.e. β_{sne} values above the 3rd quartile) which involved 278 pairwise comparisons.

Distributional pattern for the compositional and phylogenetic β -diversity

After the backward selection, the set of environmental predictors remaining as the most important in the model depends on which β -diversity index is included as the dependent variable (Figure 4-7). This may result, at least in part, from the presence of high correlations between some of the environmental predictors employed (Figure 4-8). The most consistently selected environmental predictor was the wet season NDWI index, which was selected by the GDM models for β_{sor} , PhyloSor, β_{sim} and PhyloSortURN (Figure 4-7). Furthermore, the consistent selection of geographic distance as an important predictor agrees with the results obtained by the partitioning of explained deviance.

The mapping of the predicted values for the PCA axes illustrates the expected spatial pattern of β -diversity across the region of interest, both for compositional and phylogenetic turnover (Figure 4-5 and Figure 4-6). In short, the compositional turnover, as measured by β_{sim} , is greater and more complex in the LG_BAD and SDR_PP floodplain landscapes, where the map shows a mosaic of very different colours – reds, greens and blues (Figure 4-5). Compositional β -diversity is less pronounced in the MAD_AM landscape showing almost pure blue colors related to the third PCA axis, furthermore this landscape depicts a pattern of β -diversity strikingly different to that observed in the western part of the region. Compositional β -diversity changes gradually in the section of the Amazonas river that goes from the Purus river confluence to the Negro river confluence, and then shows an abrupt compositional change close to Manaus (Figure 4-5). Also, the floodplains in the Madeira-Purus inter-fluvial area have a composition that is a mixture of the expected turnover of both river floodplains (Figure 4-5).

The phylogenetic turnover map (PhyloSortURN, Figure 4-6) shows a different pattern when compared with the compositional turnover map (β_{sim}). Overall, the phylogenetic composition of the eastern part of the region is linked most strongly with the first PCA axis, depicting reddish colors. On the other side of the region, the LG_BAD and

the SDR_PP landscapes are more linked with the second PCA axis (i.e. green colors, Figure 4-6). However, LG_BAD and SDR_PP landscapes show scattered patches of orange indicating heterogeneity on phylogenetic composition when compared with the observed turnover on the MAD_AM landscape. The region closer to Manaus shows a striking break for the phylogenetic β -diversity as was the case for the compositional β -diversity. Furthermore, the Madeira-Purus inter-fluvial floodplains seems to have a phylogenetic composition more closely linked with that observed in the Madeira river floodplains.

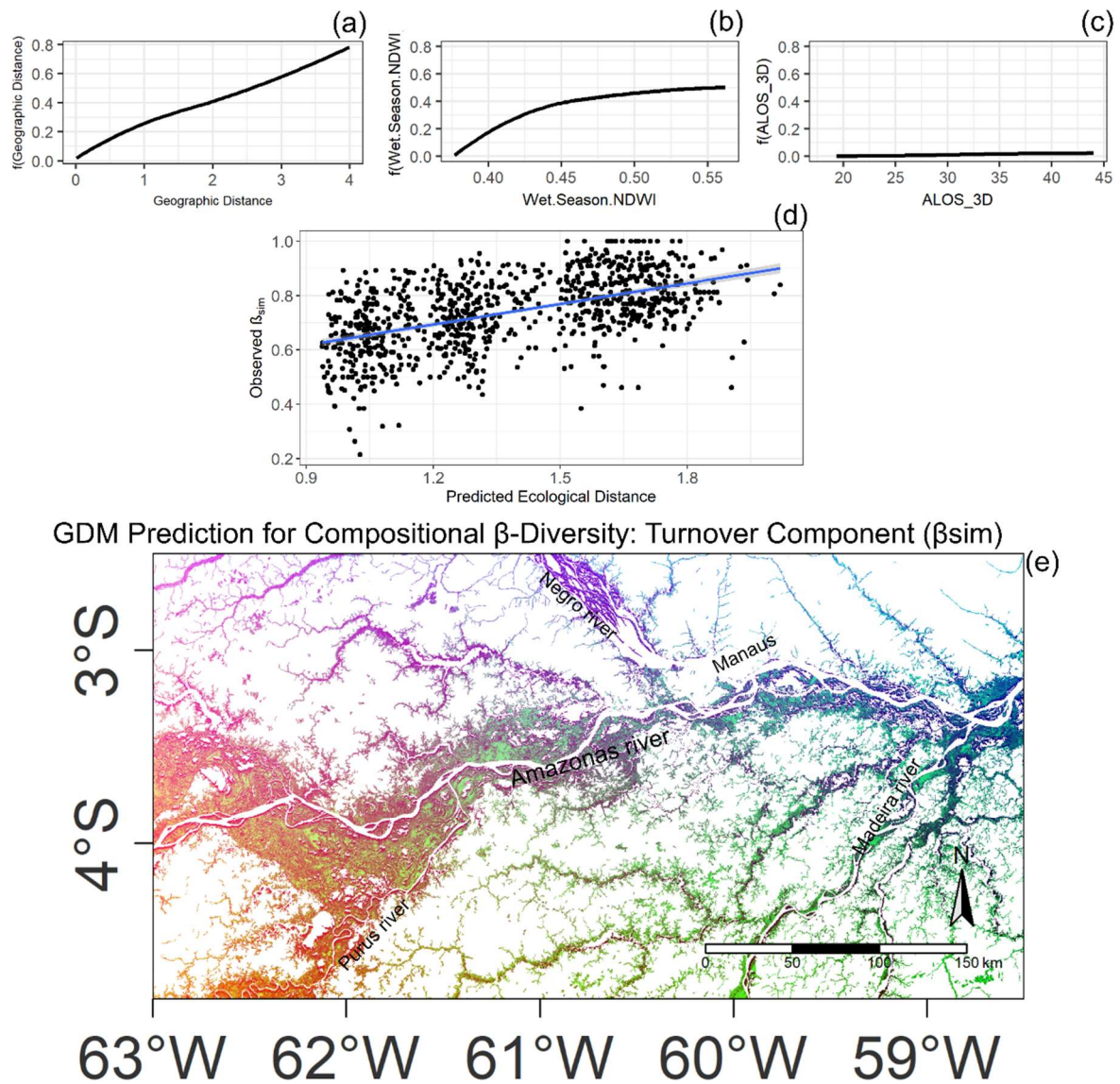


Figure 4-5. Generalized dissimilarity model for the compositional turnover β_{sim} through floodplain forests in Central Amazonia region. The three upper plots show the fitted I-spline functions in relation to the respective predictors (a) geographic distance, (b) wet season NDWI, and (c) ALOS 3D. In (d) the predicted ecological distances in relationship with the observed compositional dissimilarities. The colors in (e) shows the first three PCA axes of the transformed ecological distances given the GDM prediction, with each PCA axes assigned respectively to red, green and blue. Similar colors indicate locations predicted to support a similar composition of tree species while dissimilar colors depict locations diverging in composition.

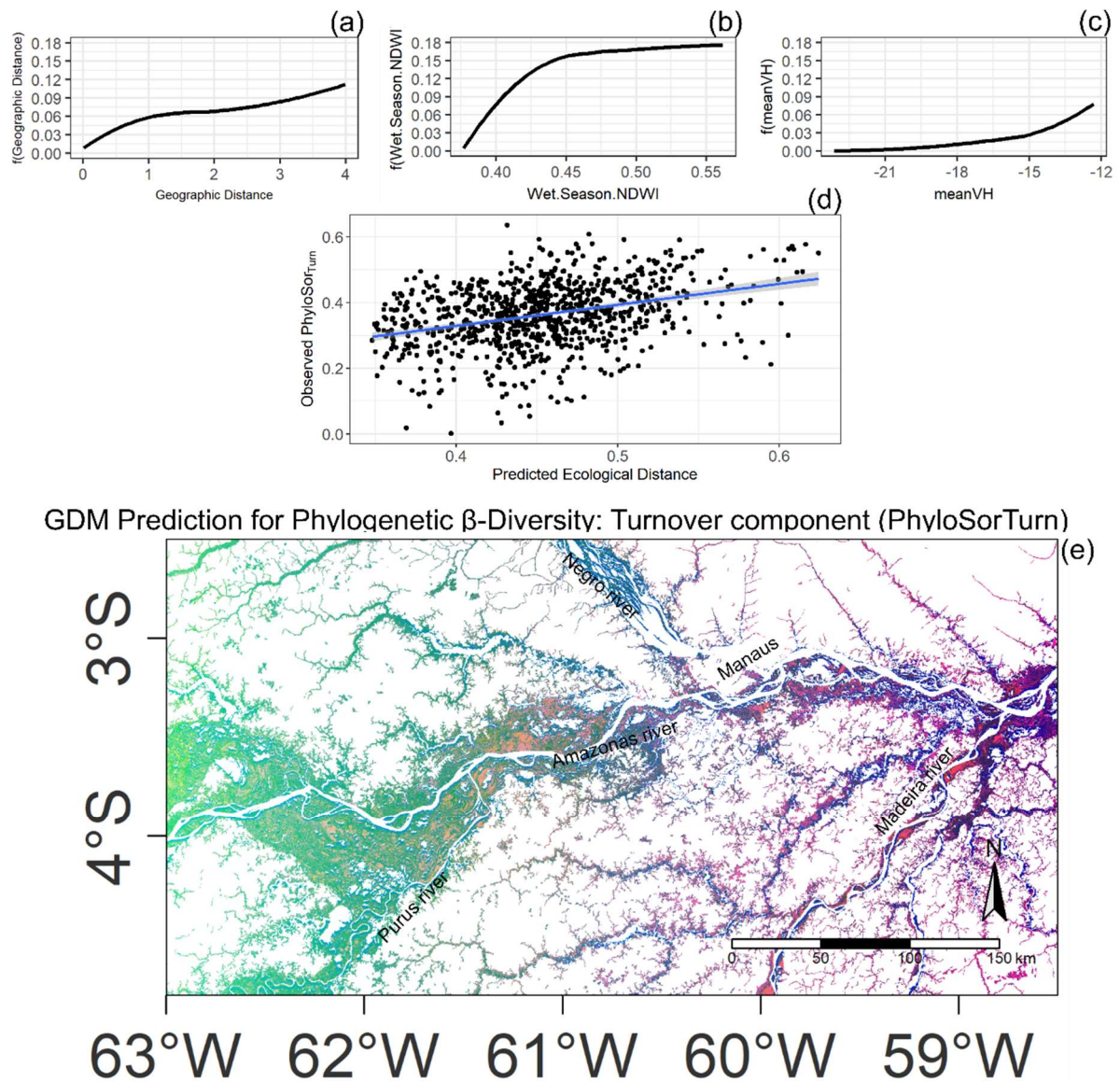


Figure 4-6. Generalized dissimilarity model for phylogenetic turnover $\text{PhyloSor}_{\text{Turn}}$ through floodplain forests in Central Amazonia region. The three upper plots show the fitted I-spline functions in relationship with measured distance for the respective predictors (a) geographic distance, (b) wet season NDWI, and (c) mean VH backscatter from Sentinel-1. In (d) the predicted ecological distances in relationship with the observed phylogenetic turnover. The colors in (e) shows the first three PCA of the transformed ecological distances given the GDM prediction. The colors of the first three PCA was assigned respectively to red, green and blue. Similar colors indicate locations with comparable phylogenetic turnover while dissimilar colors depict locations with distinct phylogenetic turnover.

Discussion

Compositional and phylogenetic β -diversity in Central Amazon floodplains

Pairwise measurements of compositional and phylogenetic β -diversity include a turnover and a nestedness component (BASELGA, 2010, 2012; LEPRIEUR et al., 2012). For central Amazon floodplain forests, the nestedness component of compositional β -diversity is almost imperceptible, indicating that the high levels of β -diversity observed in this system are not simply a function of variation in α -diversity. This suggests that, even though a α -diversity is known to vary strongly along the flood duration gradient (LUIZE et al., 2015a), the assemblage of tree species occurring in areas of low α -diversity is not simply a subset of the assemblage occurring in high-diversity areas. Despite the lower values for β_{sne} , our results show that there is phylogenetic diversity gradient in floodplain forests related to underlying environmental variation. Importantly, the turnover component of compositional and phylogenetic β -diversity has the greatest influence on the measured Sørensen index, indicating high levels of species and lineages replacement between pairs of communities.

The high level of tree species turnover in Amazonian floodplain forests ($\beta_{\text{sor}} = 0.3$ to 0.9) is comparable with that reported for a much extensive altitudinal and climatic gradient in Panamanian tropical forests (JONES et al., 2013). Moreover, reported levels of β_{sor} are in line with the high compositional dissimilarity reported elsewhere for Amazonian floodplain forests (e.g. ALBERNAZ et al., 2012; DE ASSIS et al., 2017; WITTMANN et al., 2006). We are not aware of other studies reporting phylogenetic β -diversity for Amazonian floodplain forests, however, compared with the phylogenetic β -diversity reported between white-sand forests distributed all over the Amazon (GUEVARA et al., 2016), we can affirm phylogenetic β -diversity in central Amazon floodplain forests is relatively high.

What we can learn from the differences in predictors explaining compositional and phylogenetic β -diversity?

Compositional and phylogenetic β -diversity indexes are correlated and the set of predictors of most importance in explaining deviance in the β_{sor} and PhyloSor models are quite similar, as is the amount of deviance explained by the models (Figure 4-4). Moreover, relatively short distances between surveyed sites (i.e. 1 – 400km) have an influence on the observed compositional and phylogenetic β -diversity, reinforcing and providing evidence that geographic distances between 250 to 500 km are enough to cause a decay in tree species composition similarity by half (ALBERNAZ et al., 2012; DEXTER; TERBORGH; CUNNINGHAM, 2012; WITTMANN, 2012). Compositional and phylogenetic β -diversity components have previously been shown to be correlated by means of simulation (LEPRIEUR et al., 2012), leading to the expectation that equivalent predictors will explain each β -diversity component. Indeed, both β_{sor} and PhyloSor shows a strong relationship with geographic distances suggesting dispersal limitation and ecological drift as influential processes driving floodplain forests assemblage. On the other hand, the moderate correlation observed for the compositional and phylogenetic turnover components (Figure 4-2, LEPRIEUR et al., 2012) may have a direct influence on the contrasting results for the most important set of predictors explaining model deviances for these components. For instance, the GDM for β_{sim} has most of the deviance explained by the geographic distances between sites, while the GDM for PhyloSorTurn has most of the deviance explained by environmental gradients (Figure 4-4).

These contrasting results between the amounts of deviance in turnover component of compositional and phylogenetic β -diversity explained by geographic distance versus environment suggest distinct processes driving community assembly in floodplain forests. For compositional turnover, the stronger effect of geographic distance suggests a prominent role of dispersal limitation and/or ecological drift on the selection of species

forming each community. Conversely, the stronger influence of environmental gradients explaining phylogenetic β -diversity turnover suggests a role for environmental sorting on the selection of tree species lineages that were able to colonize and assemble floodplain forest communities. The distinct response of compositional and phylogenetic β -diversity turnover to environmental gradients means that while distant sites with similar environmental conditions may not share the same species, they are nevertheless likely to share closely related species. This might occur if the isolation between sites is strong and old enough to allow speciation, but the species niche has been retained (i.e. niche conservatism) (e.g. GRAHAM; FINE, 2008). Another possible process explaining this pattern might be a selection of lineages that are able to colonize floodplain forest via lateral migration from the core inter-fluvial areas surrounding the floodplain landscapes.

Could the effect of geographic distance be an indication of lateral migration from upland forests?

Lateral migration of species from upland to floodplain forests is regarded as an influential historical factor driving tree species compositional changes in Amazonian forests (TERBORGH; ANDRESEN, 1998). The three surveyed landscapes are likely to receive tree propagules from distinct inter-fluvial areas of upland forests. For instance, the LG_BAD landscape most likely will receive species from the Negro–Japura interfluvial area, while the SDR_PP landscape encompasses the Purus–Jurua and the Madeira–Purus inter-fluvial areas, and the MAD_AM may include arrivals from the Madeira–Purus and the Trombetas–Negro. Since large rivers may act as barriers for some upland tree species, each of these inter-fluvial areas, are likely to have different biogeographic species pools, and only a selected group of upland tree species in the surroundings may spillover to floodplains forests.

The arrival of species from upland to floodplain forests may occur by chance or may be governed by density dependent processes similar to a mass effect (SHMIDA;

WILSON, 1985). Therefore, it is expected that tree species with large populations in adjacent upland forests are more likely to reach and persist in floodplain forests. Conversely, lateral migration also may happen if some rare trees specialize in flooded riparian habitats alongside small streams in the upland forests, subsequently colonize long-lasting inundated habitats of major floodplains. Our study was not designed to disentangle the influence of lateral migration on floodplain forests, keeping open an idea proposed 20 years ago (TERBORGH; ANDRESEN, 1998). Future studies surveying both floodplain forests and surrounding upland forests could help elucidate the historical process involved with habitat lateral migration within Amazonian forests. Notwithstanding, the adjacency with distinct interfluvial areas have influence on the high levels of compositional and phylogenetic β -diversity in Central Amazonian floodplain forests as most of the surveyed sites β -diversity clustering within landscapes (Figure 4-11).

Spatial patterns in compositional and phylogenetic β -diversity

Compositional and phylogenetic turnover maps illustrate expected phytogeographic change in tree species composition of Amazonian floodplain forests (ALBERNAZ et al., 2012; PRANCE, 1979). For instance, floodplains alongside the Amazonas river are characterized by eutrophic fine-grained soils and are flooded by white-water rivers with high sediments load (i.e. *Várzea* forests). *Várzea* forests show a distinct species composition to the floodplains alongside the Negro river, which are characterized by oligotrophic coarse-grained soils and are flooded by black-water rivers with low sediment load and extremely high dissolved organic matter concentration (i.e. *Igapó* forest) (PRANCE, 1979). However, some tree species that occur in *Várzea* forests are also found in *Igapó* forests (LUIZE et al., 2018; WITTMANN et al., 2017), therefore, to show how much compositionally and phylogenetically dissimilar the two floras are, both floodplain habitats should be surveyed.

The compositional and phylogenetic β -diversity maps illustrate a biogeographic discontinuity previously indicated for tree species composition in várzea forests alongside the Amazonas river floodplains (ALBERNAZ et al., 2012). Both β_{sor} and PhyloSortURN maps support such biogeographic discontinuity along the Amazonas river floodplains between the Purus and Negro river mouths (Figure 4-5 and Figure 4-6). However, the lower compositional turnover observed close to MAD_AM landscape may be the result of anthropogenic driven biotic homogenization instead of indicating a distinct biogeographic region.

An important observation coming from the compositional and phylogenetic β -diversity maps is the influence the Madeira–Purus inter-fluvial floodplains may have as source of species to rescue the species composition at the Madeira–Amazonas confluence. This observation has practical implication for the selection and establishment of future protected areas in that inter-fluvial area. Moreover, due to the high heterogeneity in the compositional and phylogenetic β -diversity, the LG_BAD landscape is suggested as an important conservation area, thereby including under protection a white-water floodplain landscape north of the Amazonas river. Speaking more generally, our results suggest that for the design of protected areas in floodplain forests, if one wants to protect the highest number of tree species, the creation of geographically dispersed protected areas will serve as a good approach. However, if the aim is to protect tree species evolutionary history, and to keep the ongoing ecological and evolutionary processes in place, it is desirable to design protected areas based on environmental conditions and adjacency to upland forest inter-fluvial blocks.

One of the key findings of the recent released Global Assessment from the Intergovernmental Platform for Biodiversity and Ecosystems Services is that biological communities are becoming more similar to each other (IPBES et al., 2019). This could be happening in the MAD_AM floodplain forests, where intensive anthropogenic disturbance

caused by selective logging and deforestation for cattle ranching is relatively more pronounced. The historical intensive use in the floodplains located on the Madeira and Amazonas river confluence may have had an imprint in the tree species composition, already perceptible as a relatively lower compositional species turnover observed in that forests. On the other hand, our results suggest that the relationship between the phylogenetic β -diversity and the environmental gradient is a process playing out long evolutionary time, and that a few tree species from the selected lineages will be able to survive and disperse through long distances tracking favorable conditions. Therefore, an unnoticed biotic homogenization due to intensive human disturbances may result in a rapid loss of a unique flora found within the Amazonian forests. In conclusion, spatial pattern for tree species compositional and phylogenetic β -diversity distribution in Central Amazonian floodplain forests are not completely concordant; and while environmental gradients are relatively more important to define phylogenetic β -diversity, geographic separation is the most influential driver of species compositional β -diversity between those floodplain forests.

Data Accessibility Statement

Environmental grid layers can be assessed and downloaded in Google Earth Engine. A comprehensive phylogeny for vascular seed plants can be downloaded in https://github.com/FePhyFoFum/big_seed_plant_trees/releases. Species occurrence matrix, plot locations and R codes for the analysis is available upon request to the corresponding author.

Supplementary Material

Table 4-1. Grid layers included as predictor variables (i.e., environmental gradients) for building the generalized dissimilarity models of compositional and phylogenetic β -diversity in floodplain forests of central Amazonia.

layer name	satellite/sensor		data metric	data information	ecological variable – description	temporal coverage	citation	source
ALOS_3D (AW3D30_V1_1)	ALOS/PRISM		meters	digital surface model	topography	–	Takaku et al. 2016	https://developers.google.com/earth-engine/datasets/catalog/JAXA_ALOS_AW3D30_V1_1
ALOS_MTPi	ALOS/PRISM		dimensionless	multi-scale topographic position index	landform distinction (ranging from negative (valleys) to positive (ridges) values)	–	Theobald et al. 2015	https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_mTPI
ALOS_TopoDiver_TP I	ALOS/PRISM		dimensionless	Shannon's equitability index	topography and climate heterogeneity	–	Theobald et al. 2015	https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_topoDiversity
rVVVH	Sentinel band SAR	1/C–	dimensionless	interferometric wide swath mode – wave polarization ratio (backscatter ratio index)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
cvVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode – coefficient variation vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
cvVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode – coefficient variation vertical–	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1

layer name	satellite/sensor		data metric	data information	ecological variable – description	temporal coverage	citation	source
				vertical polarization (backscatter statistic)				
maxVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –maximum vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
maxVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –maximum vertical–vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
meanVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –mean vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
meanVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –mean vertical–vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
medianVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	median vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1

layer name	satellite/sensor		data metric	data information	ecological variable – description	temporal coverage	citation	source
medianVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –median vertical–vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
minVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –minimum vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
minVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –minimum vertical–vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
rangeVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –range vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
rangeVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –range vertical–vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
sdVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –standard	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1

layer name	satellite/sensor	data metric		data information	ecological variable – description	temporal coverage	citation	source
				deviation vertical–horizontal polarization (backscatter statistic)	terrain. Water content.	Ground Range Detected (GRD) scenes)	missions/sentinel-1	
sdVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –standard deviation Vertical Vertical (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
skewVH	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –skewness vertical–horizontal polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
skewVV	Sentinel band SAR	1/C–	backscatter coefficient (σ^0) in decibels (dB)	interferometric wide swath mode –skewness Vertical Vertical polarization (backscatter statistic)	Biophysical characteristics of the terrain. Water content.	From 2014–12–19 to 2019–05–17 (1570 Ground Range Detected (GRD) scenes)	https://earth.esa.int/web/sentinel/missions/sentinel-1	https://developers.google.com/earth-engine/sentinel1
Full.year.NDVI	Landsat 5/ETM		dimensionless	Normalized Difference Vegetation Index (reflectance – band ratio index)	vegetation productivity – average for the 90th percentile NDVI value in the time series	From 1997–01–01 to 2012–12–31	USGS	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
NDVI.variability	Landsat 5/ETM		dimensionless	Normalized Difference Vegetation Index (reflectance – band ratio index)	vegetation productivity seasonality – standard deviation for	From 1997–01–01 to 2012–12–31	USGS	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR

layer name	satellite/sensor	data metric	data information	ecological variable – description	temporal coverage	citation	source
				the NDVI value in the time series			
Wet.Season.NDVI	Landsat 5/ETM	dimensionless	Normalized Difference Vegetation Index (reflectance – band ratio index)	vegetation productivity – average for the 90th percentile NDVI value in the time series along the wet season	From 1997–01–01 to 2012–12–31; Scenes from May–15th to Jul–15th	USGS	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
Dry.Season.NDVI	Landsat 5/ETM	dimensionless	Normalized Difference Vegetation Index (reflectance – band ratio index)	vegetation productivity – average for the 90th percentile NDVI value in the time series along the dry season	From 1997–01–01 to 2012–12–31; Scenes from Sep–14th to Nov–13th	USGS	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
Full.year.NDWI	Landsat 5/ETM	dimensionless	Normalized Difference Water Index (reflectance – band ratio index)	vegetation moisture – average for the 90th percentile NDVI value in the time series	From 1997–01–01 to 2012–12–31	Gao 1996	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
NDWI.variability	Landsat 5/ETM	dimensionless	Normalized Difference Water Index (reflectance – band ratio index)	vegetation moisture seasonality – standard deviation for the NDVI value in the time series	From 1997–01–01 to 2012–12–31	Gao 1996	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
Wet.Season.NDWI	Landsat 5/ETM	dimensionless	Normalized Difference Water Index (reflectance – band ratio index)	vegetation moisture – average for the 90th percentile NDVI value in the time series along the wet season	From 1997–01–01 to 2012–12–31; Scenes from May–15th to Jul–15th	Gao 1996	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
Dry.Season.NDWI	Landsat 5/ETM	dimensionless	Normalized Difference Water Index (reflectance – band ratio index)	vegetation moisture – average for the 90th percentile NDVI value in the time series along the dry season	From 1997–01–01 to 2012–12–31; Scenes from Sep–14th to Nov–13th	Gao 1996	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR

Supplementary text for figure S1

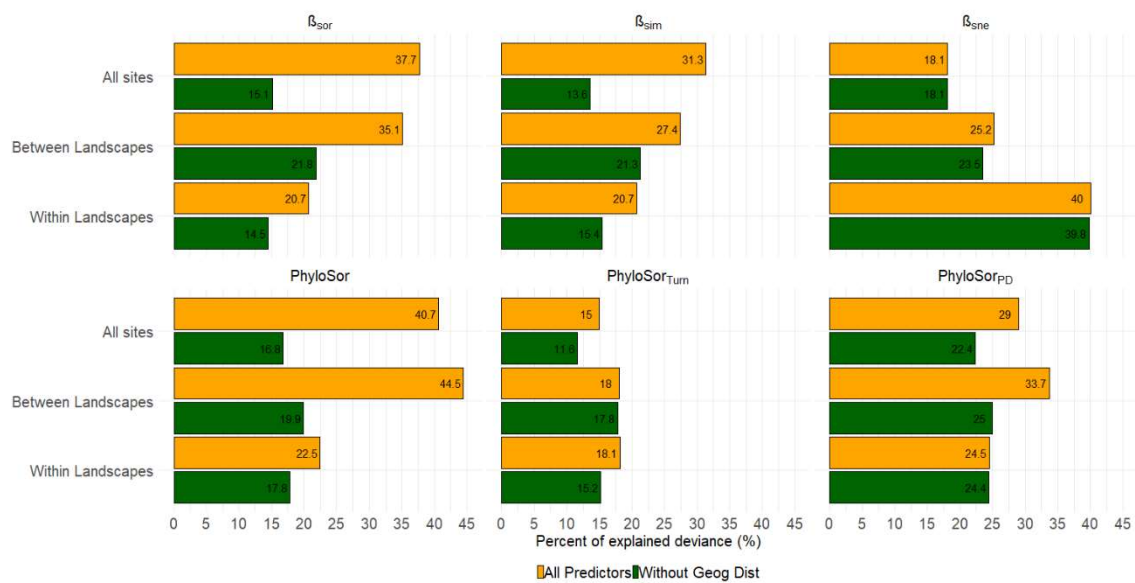


Figure 4-7. Percent of explained deviance for the generalized dissimilarity models taking account all the 28 environmental predictors plus geographic distances and all 28 environmental predictors but without the inclusion of geographic distances as predictor. The GDMs were fitted for all the 903 pairwise comparisons between the 43 floodplain forest surveys in central Amazonia, 592 pairs of sites between the three floodplain landscapes, 311 pairs of sites sampled within each of the three landscapes. With exception for the GDMs for the β_{sne} matrix that only was fitted for $\beta_{sne} \geq 0.06$ (i.e. β_{sne} above 3rd quartile) which means 278 pairwise comparisons, 171 pairs of sites between the three sampled floodplain landscapes and 107 pairs of sites within each of the three landscapes. Each panel shows results for the respective dissimilarity matrices the GDMs was fitted, the upper row shows results for the CBD and the lower row for the PBD matrices

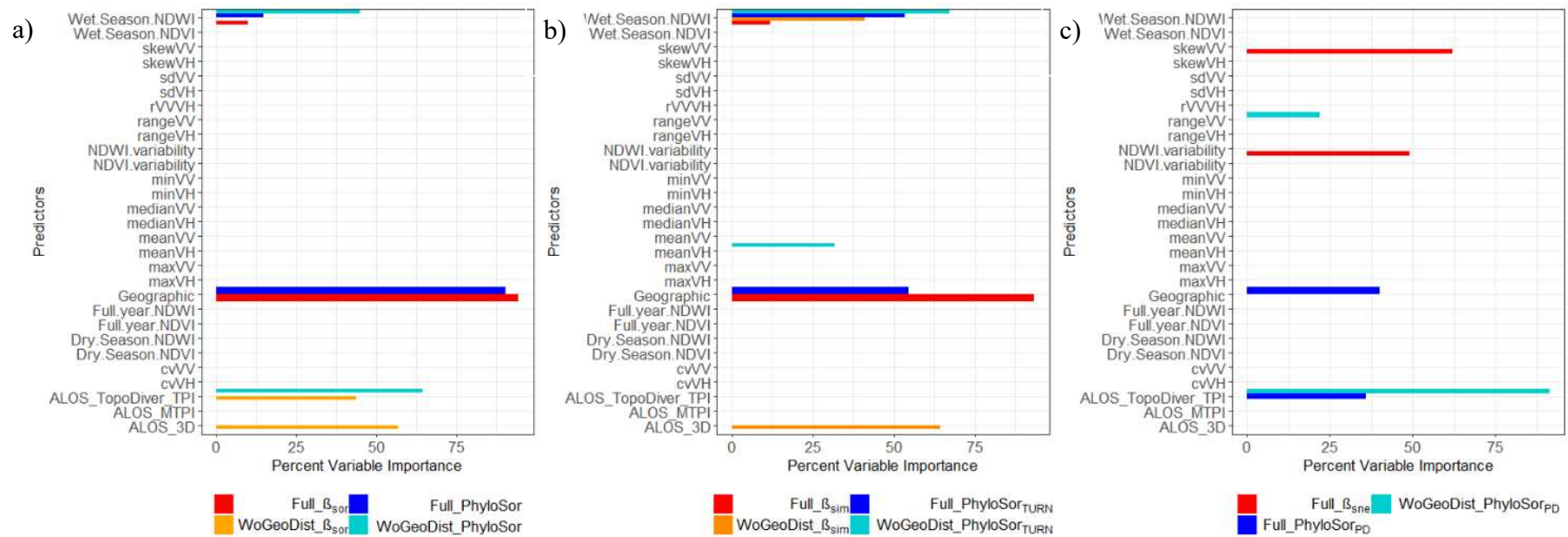


Figure 4-8. Percentage of variable importance for the set of predictors included in the generalized dissimilarity models explaining compositional and phylogenetic β -diversity matrices. The bars show the percentage of variable importance after backward selection with 999 permutations. Backward selection was applied to select the set of predictors to fit the simplest GDM and posteriorly used to map β -diversity through floodplain forests of Central Amazon. In (a) are highlighted the most important predictors for the β_{sor} and PhyloSor matrices, (b) gives the same information for β_{sim} and PhyloSor_{Turn} matrices and (c) for β_{sne} and PhyloSor_{PD} matrices. Note that for the β_{sne} only the bars for the model that do not include geographic distances is depicted, since the geographic distances do not have any importance for β_{sne} matrix explanation.

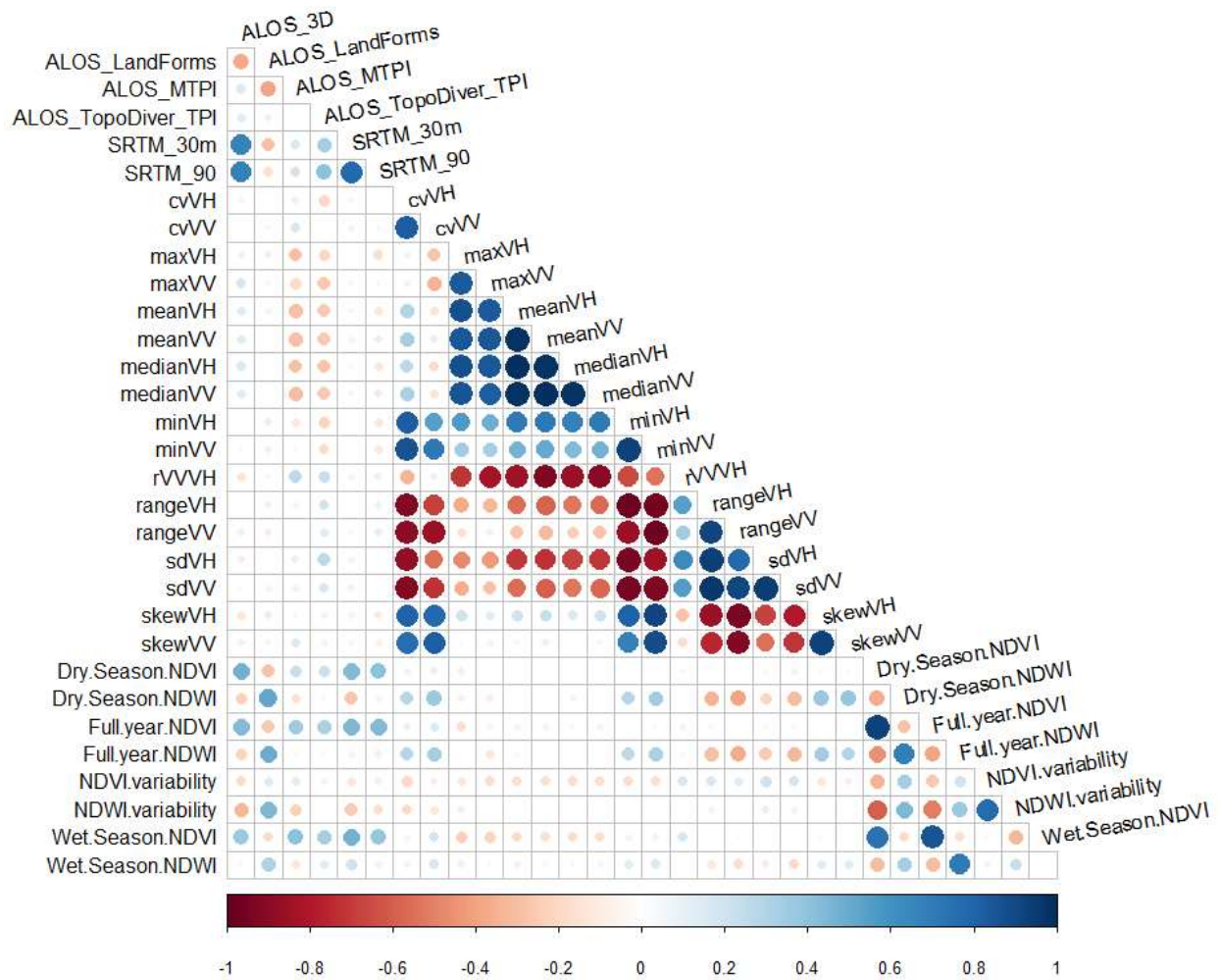


Figure 4-9. Pearson correlation between the set of 31 environmental predictors extracted for the 43 site locations sampled in floodplain forests of Central Amazonia. For a list of predictors names and descriptions please refer to table 4-1.

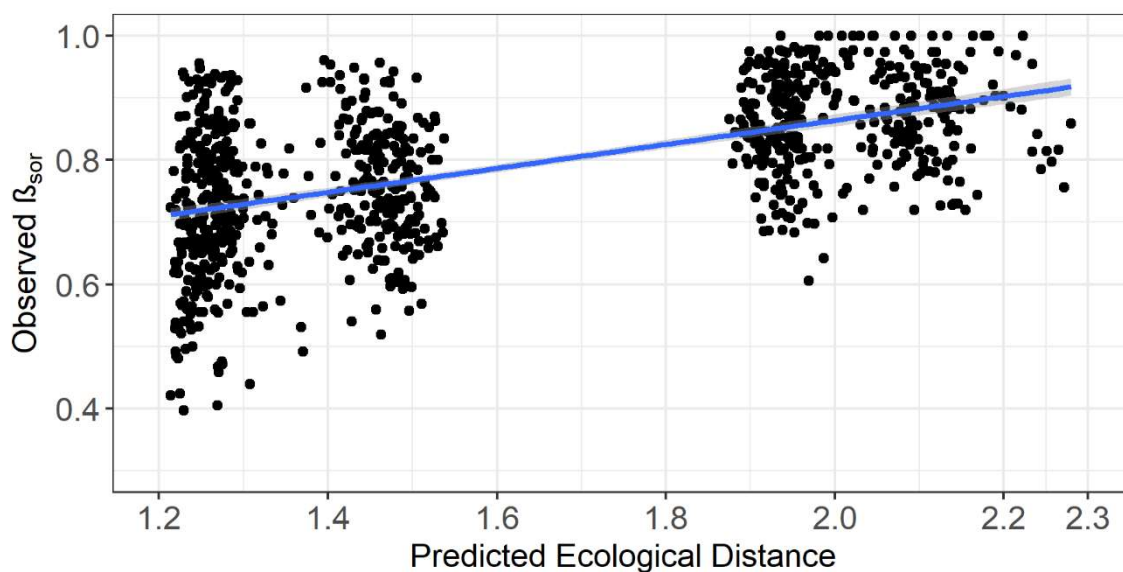


Figure 4-10. Scatter plot for the predicted ecological distances in relationship with observed Sørensen dissimilarity index β_{sor} of compositional β -diversity. The GDM was fitted using geographic distances, ALOS_3D and ALOS_TopoDiver_TPI as model predictors (please see Table 4-1 for a description of predictors).

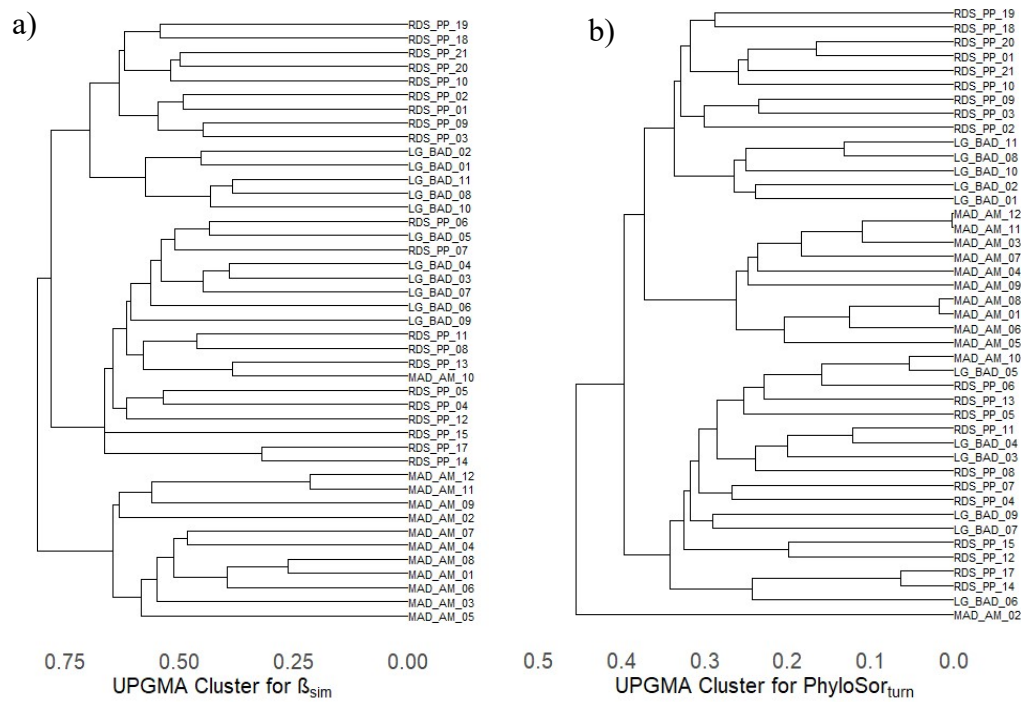


Figure 4-11. Hierarchical clustering build by unweighted pair group method with arithmetic mean (UPGMA clustering) for a) turnover of compositional and b) phylogenetic β -diversity between pairs of floodplain forests surveys.

5. The identification of drivers for tree species pairs associations in Amazonian seasonal flooded forests.

Luize BG, Magalhães JLL, Queiroz H, Lopes MA, Venticinque EM, Novo EMLM, and Silva TSF (2020) The identification of drivers for tree species pairs associations in Amazonian seasonal flooded forests. Em preparação para *Journal of Vegetation Science*.

Abstract

Questions: An ongoing goal of community ecology is to define interspecific associations, further applying to understand community assemblage processes. We investigate which is the contribution of the most abundant species on tree species positive and negative co-occurrences, and how the probability of species pair co-occurrence relates with species geographic range overlap and species divergence time.

Location: Floodplain forests (*Várzea*) alongside the Amazonas river mainstem.

Methods: We evaluated a forest inventory plot network containing 513 (0.25x0.25) sites and 667 tree species. Pairwise probabilistic approach defined species co-occurrences that were more or less frequent than those expected by chance. Overall, relative abundance across all sites defined the most abundant tree species. Geographic range overlap was defined based on alpha-hulls adjusted to the records of species occurrence, and species divergence time was computed using a synthesis phylogeny. We classify the likely driver of species pairs co-occurrences based on distribution of linear model residuals. Weak species range overlap or above average range overlap and species divergence time distant or closer than that expected allowed the identification of drivers structuring co-occurrences.

Results: From a total of 222 111 possible species pairs, only 2956 (1.3%) species pairs clearly assemble positive (1861) or negative (1095) co-occurrences. At all 291 tree species were detected in at least one co-occurrence, most forming positive (291 species) than negative co-occurrences (185 species). The 36 dominant tree species are included in 1314 (70%) positive and 1046 (95%) negative co-occurrences. The relationship between the probability of co-occurrence with range overlap, and with divergence time, indicate dispersal limitation, environmental filtering and biotic interactions structuring the co-occurrences associations.

Conclusions: Despite accounting for only 5% of the 667 sampled tree species 79% of the 2956 co-occurrences were assembled with at least one dominant tree species. Dominant tree species produce more positive co-occurrences than negative co-occurrences. Environmental filtering, dispersal limitation and biotic interaction are determinant of a few co-occurrences.

Introduction

Amazonian tropical forests may harbor close to 3.9×10^{11} individual trees distributed in an expected 16 000 species (TER STEEGE et al., 2013), meaning that 127 992 000 species pairwise combinations were possible to co-occur in those forests. However, many of those tree species barely will meet each other during an evolutionary time scale, while many others will assemble neutral, positive, or negative associations (MORUETA-HOLME et al., 2016; VEECH, 2014). The large extent of forest cover coupled with low abundance (TER STEEGE et al., 2013), niche specificity (GOMES et al., 2018) and restricted geographic distribution (FEELEY; SILMAN, 2009) of several Amazonian tree species imply in low encounter rates between most species, precluding their association or imprinting a seemingly random aspect on them. On the other hand, positive co-occurrences are expected to be favored either when tree species share large portions of their ranges, show niche similarities, or coevolve positive biotic interactions such as species facilitation for habitat occupancy (MORUETA-HOLME et al., 2016). Also, negative species associations may be structured when tree species evolve niche dissimilarities, avoiding each other, or even if species niche specificities are quite similar but species compete by local resources (SFENTHOURAKIS; TZANATOS; GIOKAS, 2005). Detecting species associations and the drivers that structure such co-occurrences are key to understand how the large number of tree species coexisting in tropical forests is achieved and maintained.

A single Amazon forest hectare may be assembled by over to 300 tree species (PITMAN et al., 2001), implying a very unlikely chance that a neighbor tree will belong to the same species. Indeed, average conspecific neighborhood in Amazonian forests is two percent, but at the same time a few tree species will account for a high proportion of the stems in each forest hectare (PITMAN et al., 2001). For instance, 5% of the 2031 tree species observed in an Amazonian forest plot network at Loreto in Peru accounted for 50% of the 60 000 individual trees sampled (DRAPER et al., 2019). Despite, the most dominant species in general show high abundances in few forest types and locations (DRAPER et al., 2019; TER STEEGE et al., 2013), suggesting species habitat selection. It is worth to mentioning, that species turnover pattern for dominant tree species mirror the turnover emerging from the analyses of all the available regional tree species pool which indicates that regional turnover is maintained by dominant tree species (DRAPER

et al., 2019; PITMAN et al., 2001). Therefore, it is expected that species having intermediate to small population sizes have high likelihood to frequently meet with one of the most dominant species regionally. The recurrent meeting between species may shape pairwise interactions and forest community assembly, and tree species achieving highest relative abundances are likely to coordinate which species could be part of the community.

The analysis of species co-occurrence is a contentious statistical topic in community ecology and biogeography. Different approaches for the detection of species co-occurrences include the construction of null models for a species vs. sites presence-absence matrix (GOTELLI, 2000), analytical probabilistic pairwise approach (VEECH, 2013, 2014), network approach (MORUETA-HOLME et al., 2016), among others. Despite differences, the main aim of each approach is to detect species pairs or hubs (i.e. triplets - networks) that co-occur at random, or that co-occur more often than at random (positive/aggregated) or less often than at random (negative/segregated), and to reveal underlying factors influencing the observed co-occurrence patterns (GOTELLI, 2000; MORUETA-HOLME et al., 2016; VEECH, 2014). The patterns of species co-occurrences could be linked to the patterns of species turnover, and it is reasonable to assume the same factors driving both community assembling and species co-occurrences. For instance, dispersal limitation may constrain range overlap between co-occurring species, reducing species encounter frequency, but also may keep one species restricted to the geographic extent of another species when there is a large range overlap. Environmental filtering also is expected to influence co-occurrence when both species track similar environmental conditions or when each species have dissimilar environmental preferences. Finally, biotic interactions may be inferred from co-occurrence patterns (e.g. KOHLI; TERRY; ROWE, 2018), indicating the relative role of positive-positive (i.e. facilitation) and positive-negative (i.e. competition) interactions on the assembling of species co-occurrences.

The high local tree species richness, low species densities and high compositional turnover characteristic of Amazonian *terra-firme* forests (DUIVENVOORDEN; SVENNING; WRIGHT, 2002; TER STEEGE et al., 2003; TERBORGH; ANDRESEN, 1998) do not completely characterize other major Amazonian forest types. For instance, while in a *terra-firme* hectare 126 – 217 tree species coexist (PITMAN et al., 2001), the same area will average 190, 105, and 60 tree species coexisting in *várzea*, white-sand and *igapó* forests, respectively (MONTERO; PIEDADE; WITTMANN, 2012; PITMAN et al., 2014; STROPP et al., 2011). As expected, the compositional turnover is highest between

major Amazonian forest types (TERBORGH; ANDRESEN, 1998), but when evaluated separately the distance decay in similarity for seasonally flooded forest types (i.e. *várzea* forests) is quite similar to that observed for *terra-firme* forests (DRAPER et al., 2019). The tree diversity gradient among major Amazonian forest types and similar compositional similarity decay suggest the existence of species co-occurrence gradient, implying that environmental filtering, biotic interactions and dispersal limitation may vary according to forest type. However, to our knowledge, until now no study has focused on the detection of associations of tree species pairwise associations and to provide clues for the underlying factors determining those co-occurrences in Amazonian forests.

In this study we sought for the positive and negative tree species co-occurrence in seasonally flooded forest along the floodplains of the Amazonas River to infer drivers structuring such associations. First, from the identified positive and negative co-occurrences, we evaluate the role of dominant species on those associations. Since high relative abundances may indicate a higher ability on resource uptake, we expect that dominant tree species produce more negative associations compared to species pairs assembled by intermediate and rare species. Second, we evaluate the relationship between the co-occurrence probability and proportion of range overlap. Species range overlap assist as a proxy for a degree of sympatry between species, thus we hypothesize that species pairs positively associated will show larger size of range overlaps than species pairs negatively associated. A third line of evidence we evaluated was the relationship between the probability of negative and positive co-occurrence and the divergence time between species pairs. Close related species tend to have relatively more similar niches than distant related species, thus we hypothesize the greater the divergence time between species the more likely is of a positive association formation. Furthermore, we hypothesize that divergence time directly correlates with positive species associations and indirectly correlates with negative species associations. Finally, to infer the role of environmental filtering, dispersal limitation, competition, and facilitation structuring negative/positive co-occurrences we classified the species pairs regarding the deviation from the expected average range size overlap and divergence time they assume.

Methods

Tree inventory

We sampled 18 782 trees ≥ 10 cm dbh over the 513 forest inventory plots of 0.065m² (25x25m) established in várzea forests between 2009 and 2014 (Figure 5-1). Our plot network is nested grouped in ten small regions (c.a. 30 km²) distributed over 2000 km in the floodplains of the middle and lower Amazonas river sections in the Brazilian Amazon. In each region the plots were established at different flood levels comprising the available flood gradient. Representative vouchers were prepared and included in INPA and MPEG herbariums. Besides tree species tagging and identification, we measure tree dbh, recorded plot location, and took measurements for the height of the flood marks left in the tree trunks during the previous flooded phase. Tree species identification was standardized across sampled plots and matched to the Tropicos database, resolving species synonyms and excluding non-valid species names and morpho-types from further analysis. The final assembled presence absence matrix contains 667 tree species and 513 plots, allowing the possibility 222 111 species pairwise combinations.

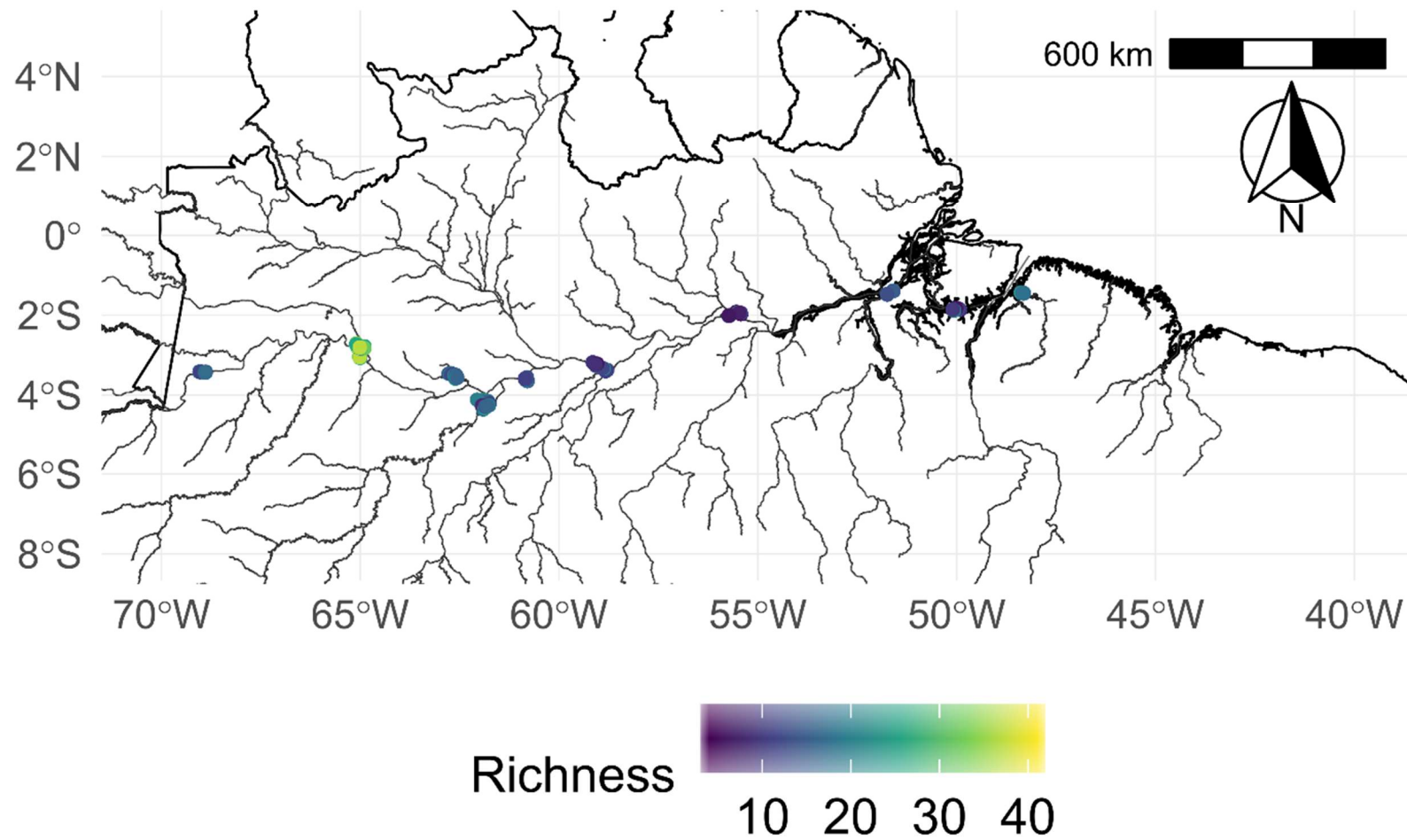


Figure 5-1. Location of the 513 sampled plots in várzea forests alongside the Amazonas river in Brazil and the species richness distribution among localities.

Co-occurrence determination

To determine positive and negative co-occurrences we applied the probabilistic pairwise approach (VEECH, 2013, 2014) to the presence – absence community matrix, as implemented in the R package ‘cooccur’ (GRIFFITH; VEECH; MARSH, 2016). The probabilistic approach analytically determines the probability that two species co-occur at an observed frequency greater than (P_{gt}) or smaller than (P_{lt}) the expected frequency if the two species were distributed randomly in relation to each other (VEECH, 2013, 2014). The computation of exact probabilities for two species co-occur either more or less frequently than they actually do (VEECH, 2013) allowed the construction of two datasets, one for the determined positive and other for the negative co-occurrences. We used $P \leq 0.05$ as cutoff to classify species pairs as positive ($P_{gt} \leq 0.05$) or negative ($P_{lt} \leq 0.05$) co-occurrences, and we filtered from further analysis, random associations (i.e. $P > 0.05$) and species pairs with insufficient information to analytically compute probability of co-occurrence (i.e. pairs that do not occur at all or pairs with expected co-occurrence lower than 1). The probabilistic pairwise approach is another way of formulating the Fisher’s exact test for statistical association between discrete variables and is identical to matrix-level null models for species co-occurrence that randomize the distribution of species among equiprobable sites (i.e. F-E algorithm) (ARITA, 2016).

Species pairs classification: dominance, range overlap, and divergence time

We determined the dominant tree species in *várzea* forests by computing overall relative abundance of each species while accounting the individual trees assigned to morpho-species. We designate as dominants, those species that, together, comprised 50% of the total number of trees sampled, using the same cutoff criteria previous applied for Amazonian tree dominance classification elsewhere (DRAPER et al., 2019; TER STEEGE et al., 2013). To evaluate the role of dominant tree species on species co-occurrence pattern, we compared the frequency of positive and negative species associations between species pairs containing a dominant tree species and pairs that do not include a dominant species. We expected a higher number of co-occurrence pairs assembled by a dominant/non-dominant tree species. Moreover, if tree species dominance suggested any competitive advantage, we expected the dominant/non-dominant pairs would be more preponderant on negative co-occurrences pairs while non-dominant/non-dominant pairs showing relatively more positive co-occurrences.

To determine species geographic range overlap, we computed extent of occurrence polygons by applying alpha-hull algorithm to the curated georeferenced occurrences of the species included in our analysis. Details regarding the occurrence records search and cleaning as well the estimation of the species geographic distribution can be found in Luize, Siqueira, and Silva (2019). Species range polygon was rasterized on a 1-degree grid cell raster produced for South America, where to be considered for presence, species range polygon should cover 10% of a grid cell. We then computed a presence – absence matrix by applying standard parameters of the function ‘lets.presab’ on the R package letsR and the “Chesser&Zink” method to compute the proportion of the smaller species range that overlaps within the larger species range (VILELA; VILLALOBOS, 2015). To estimate divergence time between species pairs, we used the most inclusive and up-to-date dated phylogeny for vascular plants (i.e. ALLOTB, SMITH; BROWN, 2018), and computed patristic distances that are the sum of the branches length linking two tips of a phylogeny (i.e. species). Patristic distances were computed by applying the function ‘fastDist’ from the R package phytools (REVELL, 2012). As the branch length in the vascular plants phylogeny (SMITH; BROWN, 2018) depict time values the resulting phylogenetic distance matrix values denote divergence time between species pairs. Due to a lack of species occurrences available in the biodiversity databases, and in the synthesis phylogeny, neither the range overlap measurements nor the divergence time could be computed for all species resulting in a decrease in the number of positive and negative co-occurrences available for fitting the models.

Range overlap and divergence time as a function of the positive and negative co-occurrences

The range overlap and the divergence time of co-occurring species were correlated with the positive and negative probability of co-occurrence. The positive and negative probability of co-occurrence was included as explanatory variable to fit four simple linear models, two models relating P_{gt} and P_{lt} with species range overlap and two models relating P_{gt} and P_{lt} with species divergence time. The standardized residuals of the fitted models were used to categorize the species pairs and to infer the expected drivers compelling co-occurrences (Table 5-1). The species pairs from the positive co-occurrence dataset ($P_{gt} \leq 0.05$) and from the negative co-occurrence dataset ($P_{lt} \leq 0.05$) were classified as follow: 1) exceeding range overlap (Ero), those co-occurrences with standardized residuals greater

than or equal to +1 deviation from the modelled range overlap, 2) weak range overlap (Wro) co-occurrences with negative standardized residuals lower than or equal to -1 deviation, 3) distant divergence (DD) for standardized residuals $\geq +1$, and 4) close divergence (CD) for standardized residuals ≤ -1 for the model relating the time of divergence between species pairs and their probability of co-occurrence. We cross compared the likely outputs of the models (Table 5-1 and Table 5-2) and compared the number of species pairs in each category to infer the existence of outstanding drivers of the co-occurrence pattern based on the preponderance of one category over the others.

Table 5-1. Fitted linear models relating probability of co-occurrence and species pairs range overlap/divergence time, the applied classification based on the residual position of the species pair along with the evidence that the residuals indicate as likely drivers for the observed co-occurrence pattern.

Models		Standardized residuals	Category	Evidence	Likely drivers
I	Range overlap	Positive ≥ 1	Exceeding range overlap	Species distributed over similar geography	Biotic interaction: Facilitation
		Negative ≤ -1	Weak range overlap	Species distributed over distinct geography	Dispersal limitation
II	Divergence time	Positive ≥ 1	Distant divergence	Species are distantly related	Environmental filtering; Biotic interaction: Facilitation
		Negative ≤ -1	Close divergence	Species are closely related	Environmental filtering
II I	Range overlap	Positive ≥ 1	Exceeding range overlap	Species distributed over similar geography	Environmental filtering Biotic interaction: Competition
		Negative ≤ -1	Weak range overlap	Species distributed over distinct geography	Dispersal limitation Biotic interaction: Competition
I V	Divergence time	Positive ≥ 1	Distant divergence	Species are distantly related	Environmental filtering Biotic interaction: Competition
		Negative ≤ -1	Close divergence	Species are closely related	Biotic interaction: Competition

Table 5-2. Matrix of expected cross combinations of species pairs classification according the standardized residuals of the fitted linear models and the likely drivers for the combination.

Positive co-occurrence P_{gt}		Range overlap	
		Exceeding range overlap	Weak range overlap
Divergence time	Distant divergence	Biotic interaction: Facilitation; Environmental filtering	Dispersal limitation; Environmental filtering; Biotic interaction: Facilitation
	Close divergence	Biotic interaction: Facilitation; Environmental filtering	Dispersal limitation; Environmental filtering
Negative co-occurrence P_{lt}		Range overlap	
		Exceeding range overlap	Weak range overlap
Divergence time	Distant divergence	Environmental filtering; Biotic interaction: Competition	Dispersal limitation; Biotic interaction: Competition
	Close divergence	Environmental filtering; Biotic interaction: Competition	Dispersal limitation; Biotic interaction: Competition

Results

Tree species co-occurrences and the influence of dominance

A total of 12042 species pairs (5.4% of 222 111 the possible pairs) were classified as either random (9086 specie pairs), positive (1861 species pairs) or negative (1095 species pairs) co-occurrences. Most of the species pairs (210 069, 94.6%), was removed from the analysis including those which expected co-occurrence probabilities was lower than 1, as well as those 9086 (4.1%) species pairs classified as random associations (i.e. $P_{lt} < 0.05 > P_{gt}$). The positive co-occurrences were assembled by 291 species and the negative co-occurrence by 185 species (Figure 5-2, Figure 5-3), all species that assembled negative co-occurrences were present in the positive co-occurrences. Also, the species assembling the highest number of positive and negative co-occurrences show similar frequencies (Figure 5-2, Table 5-3).

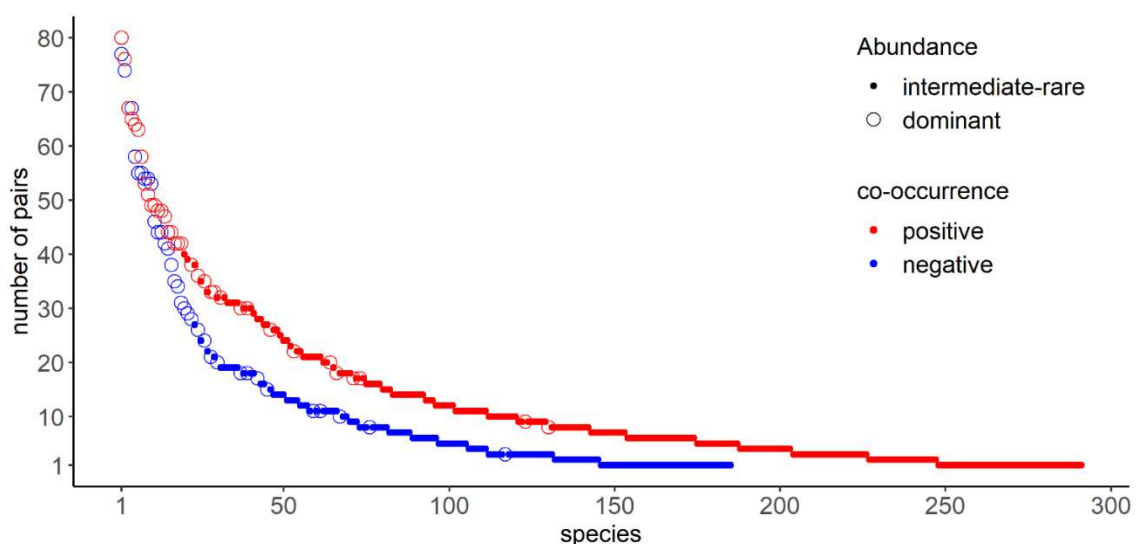


Figure 5-2. Rank for the number of times a species is detected co-occurring in positive and negative associations highlighting the species abundance classification into intermediate-rare and dominant species. Dominant species are those 36 species that together account for half of the 18 782 trees ≥ 10 cm d.b.h. sampled, while intermediate-rare are the 632 tree species with abundance accounting for the other 50% of the trees sampled. There are 291 species within 1861 positive, and 185 species within the 1095 negative co-occurrence species pairs detected. Note that species rank order may differs between negative and positive co-occurrences.

Overall, 36 tree species were classified as dominant, together summed 50.3 percent of the 18 782 trees sampled and five percent of the 667 species identified (Table 5-3). The dominant tree species in *várzea* forests figure amongst the foremost positions on the

ranking of the number of times a species is present on a co-occurrence association (Figure 5-2, Table 5-3), providing evidence that species abundance has an influence on tree species-pairs encounter and avoidance. For instance, *Triplaris weigeltiana*, the most abundant tree species sampled in *várzea* forests is present respectively in 80 and 77 positive and negative co-occurrence associations, being the first ranked species based on number of associations (Figure 5-3, Table 5-3). For the 1861 positive co-occurrences, 1314 (70%) species pairs were assembled by a dominant species, including 149 species pairs assembled by two dominant tree species. From the 1095 negative co-occurrences, 1046 (95%) species pairs were assembled by a dominant species, including 226 species pairs assembled by two of the dominant tree species.

Table 5-3. Rank of the dominant tree species in Amazonian seasonal flooded forest (*Várzea*) and number of times the species appear in a positive or negative co-occurrence pair along with their rank for the number of positive and negative co-occurrences.

Rank relative abundance	Species	Total abundance	Relative abundance	Cumulative relative abundance	Number of co - occurrences positive	Number of co - occurrences negative	Rank number of co - occurrences positive	Rank number of co - occurrences negative
1	<i>Triplaris weigeltiana</i>	702	4.10	4.1	80	77	1	1
2	<i>Euterpe oleracea</i>	539	3.15	7.3	76	74	2	2
3	<i>Pseudobombax munguba</i>	495	2.89	10.1	67	67	3	3
4	<i>Vitex cymosa</i>	425	2.48	12.6	65	67	4	4
5	<i>Viola surinamensis</i>	362	2.12	14.7	64	58	5	5
6	<i>Pterocarpus amazonum</i>	357	2.09	16.8	63	55	6	6
7	<i>Handroanthus barbatus</i>	342	2.00	18.8	58	55	7	7
8	<i>Pterocarpus santalinoides</i>	330	1.93	20.8	53	54	8	8
9	<i>Mabea subsessilis</i>	313	1.83	22.6	51	54	9	9
10	<i>Astrocaryum murumuru</i>	290	1.70	24.3	49	53	10	10
11	<i>Pentaclethra macroloba</i>	285	1.67	26	49	46	11	11
12	<i>Astrocaryum jauari</i>	257	1.50	27.5	48	44	12	12
13	<i>Spondias mombin</i>	248	1.45	28.9	48	44	13	13
14	<i>Luehea cymulosa</i>	245	1.43	30.3	47	42	14	14
15	<i>Himatanthus articulatus</i>	236	1.38	31.7	44	41	15	15
16	<i>Laetia corymbulosa</i>	226	1.32	33	44	38	16	16
17	<i>Pouteria elegans</i>	223	1.30	34.3	42	35	17	17
18	<i>Oxandra riedeliana</i>	222	1.30	35.6	42	34	18	18
19	<i>Crateva tapia</i>	178	1.04	36.7	42	31	19	19
20	<i>Eschweilera albiflora</i>	174	1.02	37.7	38	30	20	22
21	<i>Hura crepitans</i>	165	0.96	38.7	36	29	21	24
22	<i>Hevea brasiliensis</i>	160	0.94	39.6	35	28	22	26
23	<i>Maclobium bifolium</i>	151	0.88	40.5	33	26	24	28

Rank relative abundance	Species	Total abundance	Relative abundance	Cumulative relative abundance	Number of co - occurrences positive	Number of co - occurrences negative	Rank number of co - occurrences positive	Rank number of co - occurrences negative
24	<i>Vatairea guianensis</i>	151	0.88	41.4	33	24	26	29
25	<i>Leonia glycyarpa</i>	150	0.88	42.2	32	21	28	31
26	<i>Carapa guianensis</i>	145	0.85	43.1	30	20	30	37
27	<i>Cecropia membranacea</i>	138	0.81	43.9	30	18	37	39
28	<i>Guazuma ulmifolia</i>	134	0.78	44.7	26	18	39	46
29	<i>Eschweilera parviflora</i>	132	0.77	45.4	22	17	42	53
30	<i>Euterpe precatoria</i>	129	0.75	46.2	20	15	45	64
31	<i>Gustavia augusta</i>	122	0.71	46.9	18	11	59	66
32	<i>Cecropia latiloba</i>	119	0.70	47.6	17	11	61	71
33	<i>Attalea phalerata</i>	117	0.68	48.3	17	10	67	73
34	<i>Pirahea trifoliata</i>	116	0.68	49	9	8	76	123
35	<i>Unonopsis guatterioides</i>	115	0.67	49.6	8	3	117	130
36	<i>Licania heteromorpha</i>	110	0.6	50.3	8	3	117	130

Pairwise co-occurrences and species range overlap

Range overlap measurements were obtained for 1387 (74%) species pairs of the positive co-occurrences, and for 864 (78%) species pairs of the negative co-occurrences. Range overlap was not correlated with the probability of co-occurrence neither for the positive ($r = 0.010$) nor for the negative co-occurrences ($r = 0.034$). Average proportion of range overlap between the positive co-occurrences was 0.62, the same value for the intercept computed by the model relating the probability of co-occurrence and range overlap (Figure 5-3a). The negative co-occurrence dataset estimates the intercept for range overlap equal to 0.53, very close to average range overlap of 0.54 (Figure 5-3b). The standardized residuals for model relating positive co-occurrences and range overlap vary from -2.3 and 1.4 (Figure 5-3a), whereas for the negative co-occurrence they vary -1.9 – 1.6 (Figure 5-3b).

Pairwise co-occurrences and species divergence time

Divergence time between species pairs varies between 0.92 – 271.82 Ma for the two co-occurrence datasets pooled together. The positive co-occurrence dataset includes 1589 species pairs (85% of the positive co-occurrences) and the negative co-occurrence dataset 921 species pairs (84% of the negative co-occurrences) with divergence time available. Average divergence time for the positive co-occurrences is 238.58 Ma, the same value estimated by the intercept of 238.85 Ma for the model relating positive co-occurrence and divergence time (Figure 5-4a). The adjusted model for the negative co-occurrences estimates an intercept of 241.30 Ma for divergence time between species pairs, that is quite similar to the average divergence time of 240.68 Ma computed for the negative co-occurrence species pairs dataset (Figure 5-4b). Standardized residuals for the model relating divergence time and positive co-occurrences varies between -5.5 – 0.8 (Figure 5-4a). For the positive co-occurrences any species pair showed standardized residuals equal or greater than 1 deviance, while 130 species pairs (8.1%) have standardized residuals equal or lower than -1 deviance, and 1459 species pairs (92%) have average standardized residuals between -1 and 1 deviance (Figure 5-4a). For the model relating negative co-occurrences probabilities with species pairs divergence time, there are 73 (8%) species pairs with standardized residuals lower than -1, and 848 (92%) species pairs with average standardized residuals (Figure 5-4b).

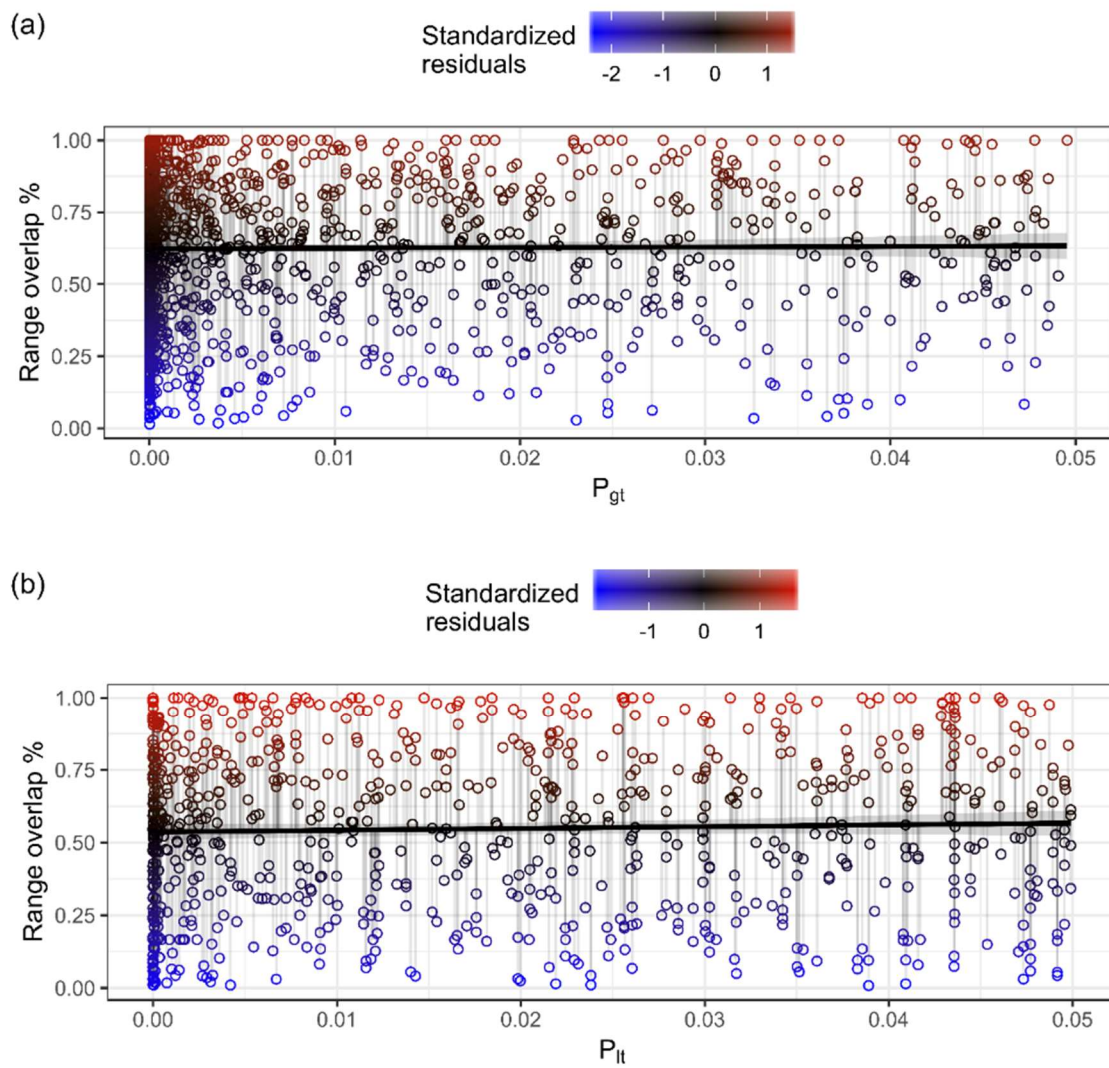


Figure 5-3. The relationship between probability of co-occurrence and range overlap for species pairs co-occurring more than expected by chance, highlighting the position of the species pairs with standardized residuals greater than 1 deviation. (a) for the positive co-occurrences probabilities (P_{gt}), and (b) for the negative co-occurrences probabilities (P_{lt}).

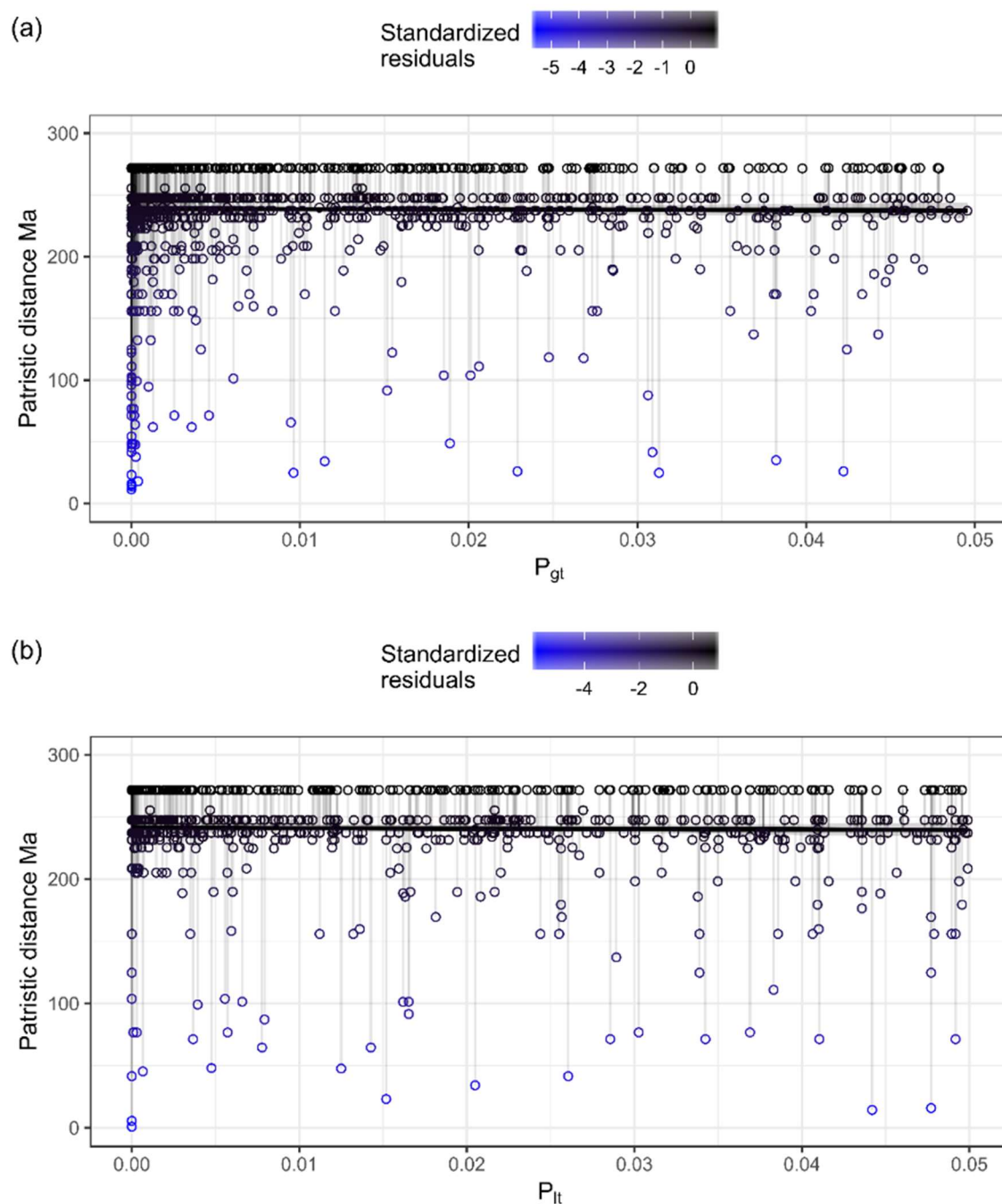


Figure 5-4. Adjusted linear model for the relationship between probability of co-occurrence and divergence time for species pairs co-occurring more than expected by chance and highlighting the position of the species pairs with standardized residuals greater than 1 deviation. (a) positive co-occurrences probabilities (P_{gt}), and (b) negative co-occurrences probabilities (P_{lt}).

Range overlap and divergence time classification

Overall, 1947 (65% positive and negative co-occurrences) species pairs also had measurements for range overlap and divergence time, allowing the classification of likely drivers structuring co-occurrences. More species pairs from positive co-occurrences (1200 – 64% of positive co-occurrences) than species pairs from negative co-occurrences (747 – 68% negative co-occurrences) had measurements for range overlap and divergence time. For the negative co-occurrences most of the species pairs show large range overlap and relatively short divergence time (Figure 5-5 a, c). While for the positive co-occurrences the number of pairs showing large range overlap and short divergence is almost the same of the number of pairs showing weak ranges overlap and short divergence time (Figure 5-5 b, d).

Discussion

For the first time 2956 co-occurrences clearly deviating from expected co-occurrence frequencies were detected for trees in Amazonian seasonally flooded forests, but not all the 667 tree species sampled in várzea forest participate in a co-occurrence. Additionally, there were recurrent species both in positive and negative inter-species pairs, being those species generally the dominant tree species in várzea forests (Figure 5-2). Differently from the hypothesized, neither the range overlap nor the divergence time between species pairs were correlated with their probability of co-occurrence. Notwithstanding, underlying drivers of species co-occurrences can be inferred from the position assumed by species pairs on the linear models relating the positive and negative co-occurrences probabilities with species range overlap and with species divergence time. The higher number of positive co-occurrences compared to that of negative co-occurrences (Figure 5-2) suggests a definite role of environmental filtering, and biotic facilitation on community assembly of várzea forests (Table 5-2; Figure 5-5). Although, those inferences may not be attained as definitive, they rather provide work hypothesis for refinements on our understanding on the causes of species associations in tropical flooded forests.

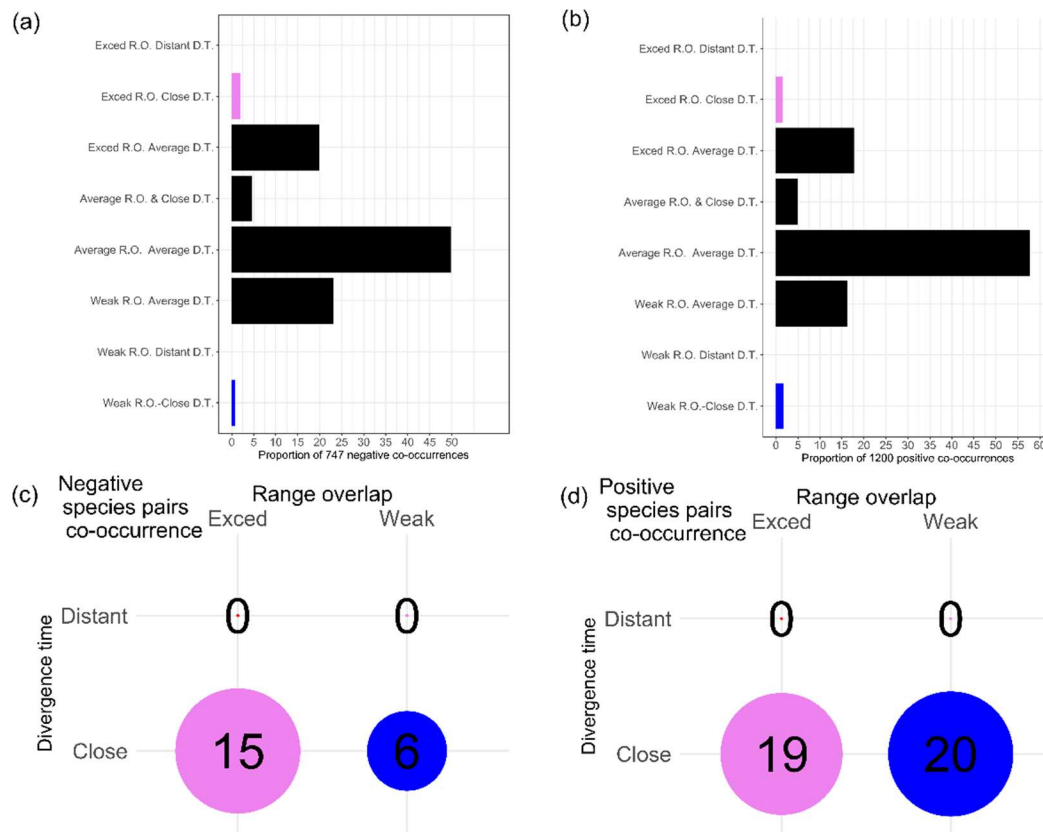


Figure 5-5. Proportional distribution of co-occurrences classified as exceeding/weak range overlap (R.O.) and close/distant divergence time (D.T.) based on the standardized residuals (i.e. $-1 \geq \text{ResStd} \geq +1$) of the linear models relating probability of co-occurrence, range overlap and divergence time. (a) Proportion of species pairs determined as negative co-occurrences, and (b) species pairs determined as positive co-occurrences, in each one of the classes assigned. (c) and (d) shows number of species pairs classified by both models as non-average residuals, (c) number of species pairs negative co-occurrences, and (d) number of species pairs positive co-occurrences.

Co-occurrence patterns may change over time or in the presence of a third species (MORUETA-HOLME et al., 2016; VEECH, 2014), allowing a positive co-occurrence to become a negative co-occurrence or vice-versa. In general, the majority of possible species pairs could not be evaluated, or did not co-occur at all in the sampled plots. Among the classified species pairs, most co-occurred at random, suggesting that a stochastic meeting between species play a prominent role on community assembly. Previous studies on species co-occurrences seems to detect more positive associations than negative associations, independently of the method applied (null model, probabilistic and network) (VEECH, 2013). Positive co-occurrences are linked to species clumping or aggregation suggesting species similar environmental preferences and/or positive-positive biotic interaction such as facilitation (KOHLI; TERRY; ROWE, 2018). The prevalence of positive over negative co-occurrences may be a strong indication of environmental filtering and species facilitation acting on community assembly of *várzea* forests. However, it is worth noting, nearly 10% of the species pairs were classified as negative co-occurrences meaning that the role of negative biotic interactions, such as competition and spatial/environmental avoidance between species pairs, also play a role on community assembly. Moreover, 75% of co-occurrences are random, indeed the determination of the underlying process structuring species association is not as straightforward for the positive and negative co-occurrences as it is for the random co-occurrences (KOHLI; TERRY; ROWE, 2018).

Although our approach differs from the proposed trait-based framework (KOHLI; TERRY; ROWE, 2018), both approaches can be applied to infer mechanisms structuring pairwise co-occurrence patterns of species across heterogeneous sites. For instance, our analysis allowed to classify of 57 species pairs indicating the effect of environmental filtering, dispersal limitation, facilitation, and competition on the structure of co-occurrences (Figure 5-5). Facilitation compared with competition seems to be a relatively more prominent driver of species co-occurrences in *várzea* forests. Facilitation within species may arise from mycorrhiza interactions contributing on species soil nutrients exploitation (FOUGNIES et al., 2007). We found that tree species from Fabaceae family are present in most of the 39 positive co-occurrences classified as weak/exceeding range overlap and close/distant divergence time. Fabaceae are acknowledged as the most important family assembling Neotropical forests and is well known by producing

arbuscular mycorrhiza (MOREIRA et al., 1992). Species from the genus *Pterocarpus*, accounted for most of the 39 positive co-occurrences highlighted, and is known to have their growth favored by the development of arbuscular mycorrhiza even under flooded conditions (FOUGNIES et al., 2007).

Regarding the negative co-occurrences, most of the species pairs was assembled by at least one dominant species. As the computation of the co-occurrence probabilities does not consider species abundance but only species incidence, we infer that the high number of negative co-occurrences assembled by a dominant species is an indication of competing pairs of species. For instance, all the 21 negative co-occurrences showing exceeding/weak range overlap and close divergence time are assembled by a dominant species, including species pairs assembled by two dominant species and showing weak range overlap and close divergence time (*Astrocaryum jauari* and *Astrocaryum murumuru*; *Pterocarpus amazonum* and *Pentaclethra macroloba*; *Pterocarpus amazonum* and *Pterocarpus santalinoides*; *Pterocarpus amazonum* and *Swartzia racemosa*). Surprisingly the same species involved in positive co-occurrences are also present in negative co-occurrences. Furthermore their dominance position illustrates that there is an idiosyncratic response of the species accordingly the neighbor species identity. Species of the genus *Astrocaryum*, *Pterocarpus*, *Pentaclethra*, *Swartzia* and *Ormosia* are among those tree species assembling both negative and positive co-occurrences and show the highest residuals for both range overlap and divergence time, highlighting their importance on *várzea* forest structure.

In this study we defined tree species pairs occurring less or more often than random, and inferred the role of environmental filter, dispersal limitation, and biotic interactions structuring those co-occurrences. The approach adopted in this study is a step towards our understanding of relative role for different drivers acting upon the assembling of *várzea* forests. The identified co-occurrences may promote the design of experimental studies aiming a better understanding of biotic and environmental interactions among tree species in *várzea* forests. Moreover, identified co-occurrences are useful for planning restoration programs and sustainable use of wetlands forest, indicating tree species that are more prone to be cultivated together, for example. Dominant tree species throughout Amazonian forests show a pronounced role on the amount of forest biomass and on species diversity patterns (FAUSET et al., 2015; TER STEEGE et al., 2013). We conclude that

dominant tree species also are central to structure species co-occurrences influencing floodplain forest assemblage.

Supplementary Material

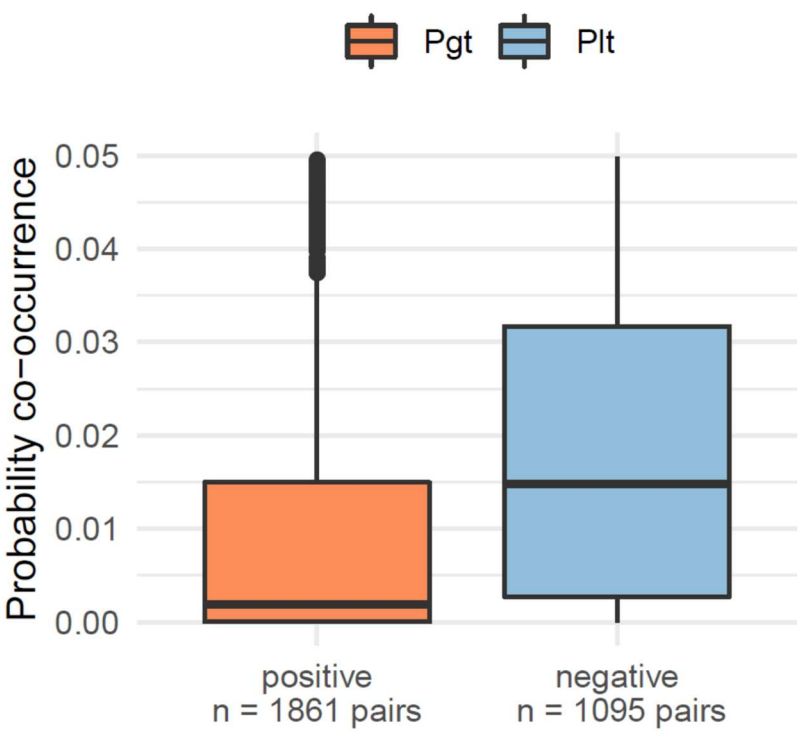


Figure 5-6. Figure 5.S1. Boxplot for the co-occurrence probability (Pgt and Plt) computed for positive and negative species pairs following the probabilistic pairwise approach to detect species co-occurrences (Veech 2014). The lower the probabilities the greater are the chances of species pair co-occurrence.

Table 5-4. Significant co-occurrences and associated parameters: species names, species incidence, observed co-occurrences, probability of co-occurrence, expected co-occurrences, P_{lt} , P_{gt} , Patristic distance, and Range overlap.

6. Applying DNA barcoding for delimitation of species: a test with

Lecythidaceae from central Amazonia

Bruno Garcia Luize; Oscar Mauricio Vargas Hernandez; Drew Larson; Diego Alvarado-Serrano; Thiago Sanna Freire Silva; Alberto Vicentini; Clarisse Palma da Silva and; Christopher William Dick. Resultados preliminares.

Abstract

Species identification set the basis of ecological and evolutionary studies. Appropriate species determination, however, is specially challenging in the tropics because of limited collections with informative taxonomic characters and the existence of cryptic species. DNA barcoding is a promising technique to species identification, but recent studies have shown that traditional universal barcode regions (*rbcL* and *matK*) have a limited power to identify species in hyperdiverse regions like Amazon forest, suggesting taxon specific development of markers. We tested the application of recent developed taxon-specific barcodes, *ycf1*, in the ecological dominant and species rich Brazil Nut family of trees (Lecythidaceae). We compared the efficiency of *ycf1* against *rbcL* and *matK* of 52 samples comprising 22 Lecythidaceae species collected from forests in Central Amazonia. We produced 184 new sequences to three *cpDNA* regions (*matK*; *rbcL*; *ycf1*). Traditional DNA barcode markers (*matK* and *rbcL*) underperformed when compared with *ycf1*; *matK* and *rbcL* are unable to differentiate species and presented low resolution when included in a phylogenetic analysis. New developed marker (*ycf1*) presented more polymorphism between sequences, improving the Lecythidaceae species delimitation, but still infer polyphyletic relationship among species. Concatenation of the sequences for the three regions improved species delimitation and phylogenetic inference. Shortcomings of the markers tested in Lecythidaceae include failing in the amplification and sequencing, a sparse representation of Lecythidaceae sequences in GenBank, a very close resemblance between sequences of different species, likely arising from recent speciation in the family.

Introduction

The Neotropical realm comprises one third of all known vascular plant species worldwide (ULLOA ULLOA et al., 2017). Despite a nearly constant rate of new plant species descriptions over the last 25 years (ULLOA ULLOA et al., 2017), most likely humanity will never acknowledge all the plant species living on the Americas before the industrial age. Furthermore, current alarming rates of habitat loss and species extinctions (CURTIS et al., 2018) combined with the reduced number of trained taxonomists and inadequate funding to collect, describe, and catalogue vascular plant diversity in the Americas also contribute to that unknowledge scenario (DICK; KRESS, 2009). The taxonomic shortfall (Linnaean shortfall) threatens scientific development in other biodiversity research areas (HORTAL et al., 2015).

Taxonomy is an area of intense scientific debate that gradually is incorporating innovative techniques to describe species (e.g. PINHEIRO; DANTAS-QUEIROZ; PALMA-SILVA, 2018; PRATA et al., 2018). For instance, molecular and chemical evidences are being used to circumscribe plant species orders and families (CHASE et al., 2016). Nevertheless, the foundation of flowering plants taxonomy has traditionally relied on the morphology of apomorphic characters in flower and fruits. Only recently, following advances in microscopy, chemical and molecular techniques and in computational tools, taxonomists are including molecular and vegetative characters to delineate species. Even after taxonomists formally describe a species, researchers are confronted with the hard task to find a feasible identification for samples (LANG et al., 2015), especially when informative morphological structures are absent on collections, or when samples belong to species complexes that have similar morphological characters. Furthermore, as new information about species character variation become available, species determination constantly needs to be reviewed.

DNA barcoding is an innovative technique that can inform both taxonomists and field biologists cataloging biodiversity (DICK; KRESS, 2009). In general, most of the DNA molecules is quite similar among species; however, small regions of the DNA are unique and highly variable among close related species. DNA barcoding is a technique that aims to find standardized short DNA regions as an internal species tag for taxa identification (HEBERT; GREGORY, 2005). The very beginning of DNA barcode application shows mitochondrial DNA (*mtDNA*) is useful to delimitate animal species. For land plants, the desired support for species identification based on *mtDNA* is not provided due the low alleles substitution rates of *mtDNA* (CBOL PLANT WORKING GROUP et al., 2009). Instead of the single CO1 *mtDNA* broadly accepted for animals, for land plants barcode standards is a combination of two or more gene regions from chloroplast and/or nuclear DNA (CBOL PLANT WORKING GROUP et al., 2009). The CBOL Plant Working Group et al. (2009) proposed the plastome regions *matK* + *rbcL* as universal barcodes for land plants, but recent testing of such markers in tropical system have shown low performance of those markers. Such observation promoted the search and identification for a vast array of possible candidate regions for DNA barcode of land plants.

Further assessments for possible candidates as a universal DNA-barcode for land plants provide evidence that the most accurate DNA-barcode should be the entire *cpDNA* genome (i.e. super-barcode) (LI et al., 2015). However, Next Generation Sequencing (NGS) methods still are costly, time demanding and unavailable for most researchers. As a compromise between the use of a less accessible super-barcode and the most widely accessible single region DNA-barcode sequences Li and colleagues (2015) argued for a definition of specific DNA-barcodes for different plant lineages. Therefore, a short

fragment of DNA with sufficiently high mutational rates to provide accurate species identification should be developed for a given taxonomic group (Li et al., 2015), but few complete plastome sequences are available for most plant groups and the definition of specific DNA barcode proceed gradually.

Recently, 24 plastomes for Lecythidaceae were assembled and annotated (THOMSON; VARGAS; DICK, 2018); using the alignment of those 24 plastomes sequenced, authors identified multiple regions with sequence variability and concluded that *ycf1* and *rpl16-rps3* are promising for barcoding in the family (THOMSON; VARGAS; DICK, 2018). The present study aims to evaluate to what extent the standard (*matK* + *rbcL*) and newly indicated region (*ycf1*) are useful for identification of Lecythidaceae species. Neotropical Lecythidaceae (subfamily Lecythidoideae) currently comprise 217 named species and 10 accepted genera (MORI et al., 2010). The center of diversity in Lecythidaceae is within the Amazon region (S = 157 species) (MORI; LEPSCH-CUNHA, 1995), where trees and treelets colonize the canopy and understory of tropical forests (MORI; LEPSCH-CUNHA, 1995; MORI; PRANCE; DE ZEEUW, 1990; PRANCE; MORI, 1979). Lecythidaceae flowers and fruits (i.e., pixidium) morphology are very characteristic and suggest adaptation to buzz pollination and endocoric dispersal (MORI et al., 2015). Furthermore, in Amazonian floodplain forests (i.e., *Várzea* and *Igapó* forests) propagules show adaptation to be dispersed by water and fish (MORI; PRANCE; DE ZEEUW, 1990; PRANCE; MORI, 1979). However, the high similarity among species vegetative characters, and the difficulty to find and collect individuals with fertile structures, hinder the achievement of proper species identifications. Furthermore, Lecythidaceae undergo a recent radiation, with most of the speciations during Miocene (c.a., 11 Ma) (VARGAS et al., in prep.), therefore difculting the species delimitation

(PINHEIRO; DANTAS-QUEIROZ; PALMA-SILVA, 2018). For instance, the most abundant trees in Amazonian forests (e.g., *Eschweilera coriacea* (DC.) S.A.Mori ; TER STEEGE et al., 2013) may be a species complex (Souza et al., unpubl.). Nevertheless, Lecythidaceae is an ecologically important family among Amazonian forests, also being dominant on the forest carbon cycle (FAUSET et al., 2015).

Methods

Focal species and collection locations

We focused our sampling in the three most species rich genera of Neotropical Lecythidaceae: 1) *Gustavia* comprise 47 accepted species names, 2) *Lecythis* (34 accepted species), and 3) *Eschweilera* (~100 accepted species) (MORI et al., 2010). Given our sampling these three genera are segregated into two major clades: 1) *Gustavia* within the actinomorphic-flowered Lecythidoideae grade, and 2) *Lecythis* and *Eschweilera* within the zygomorphic-flowered Lecythidoideae grade (MORI et al., 2007; MORI; PRANCE; DE ZEEUW, 1990; PRANCE; MORI, 1979). *Gustavia* species comprise a clade that is sister of the Lecythidoideae zygomorphic-flowered clade. The genera *Lecythis* and *Eschweilera*, however, are not monophyletic (MORI et al., 2007). We sampled 52 individuals distributed in 22 species (*Gustavia*: 8 individuals, 4 species; *Lecythis*: 11 individuals, 6 species; *Eschweilera*: 33 individuals, 12 species – Table 6-1). Overall, 184 new DNA sequences were produced, we are unable to produce cpDNA sequences for all the four markers in each sample but produced sequences for at least for two different markers in each one of the 22 focal species (Table 6-1).

Table 6-1. Lecythidaceae species vouchered from two Amazonian forest types (*várzea* – white water floodplain forests at PP-SDR and M-SDR; and *terra-firme* – upland forests at BDFFP) for DNA barcode analyses. Their accepted species names, species code and the number of DNA samples successfully sequenced for the three chloroplastial DNA regions evaluated.

Forest type	species code	Species	# samples	# samples sequenced <i>matK+rbcL+ycf1(1)+ycf1(2)</i>	Markers sequenced
Várzea	GUAU	<i>Gustavia augusta</i> L.	4	4	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	GUEL	<i>Gustavia elliptica</i> S.A.Mori	1	0	<i>rbcL</i> , <i>ycf1a</i>
Várzea	GEHE	<i>Gustavia hexapetala</i> (Aubl.) Sm.	2	2	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Várzea	GUPO	<i>Gustavia poeppigiana</i> O.Berg	1	1	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Várzea	ESAL	<i>Eschweilera albiflora</i> (DC.) Miers	11	7	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	ESAT	<i>Eschweilera atropetiolata</i> S.A.Mori	3	2	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	ESBR	<i>Eschweilera bracteosa</i> (Poepp. ex O.Berg) Miers	1	1	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	ESCR	<i>Eschweilera coriacea</i> (DC.) S.A.Mori	5	4	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	ESGR	<i>Eschweilera grandiflora</i> (Aubl.) Sandwith	2	1	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>
Terra-firme	ESMI	<i>Eschweilera micrantha</i> (O.Berg) Miers	2	2	<i>matK</i> , <i>rbcL</i> , <i>ycf1(1)</i> , <i>ycf1(2)</i>

Forest type	species code	Species	# samples	# samples sequenced <i>matK+rbcL+ycf1(1)+ycf1(2)</i>	Markers sequenced
Várzea	ESOV	<i>Eschweilera ovalifolia</i> (DC.) Nied.	1	1	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Várzea	ESPA	<i>Eschweilera parvifolia</i> Mart. ex DC.	2	2	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	ESPE	<i>Eschweilera pedicellata</i> (Rich.) S.A.Mori	2	1	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	ESRC	<i>Eschweilera romeu-cardosoi</i> S.A.Mori	2	2	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	ESTE	<i>Eschweilera tessmannii</i> R.Knuth	1	1	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	ESWA	<i>Eschweilera wachenheimii</i> (Benoist) Sandwith	1	0	<i>matK, rbcL, ycf1(1)</i>
Terra-firme	LEBA	<i>Lecythis barnebyi</i> S.A.Mori	4	3	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	LEGR	<i>Lecythis gracieana</i> S.A.Mori	1	1	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	LEPA	<i>Lecythis parvifructa</i> S.A.Mori	1	0	<i>rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	LEPR	<i>Lecythis prancei</i> S.A.Mori	3	2	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
Terra-firme	LE05	<i>Lecythis</i> sp.05.LECY	1	0	<i>matK, ycf1(1)</i>
Terra-firme	LESP	<i>Lecythis</i> spp.	1	1	<i>matK, rbcL, ycf1(1), ycf1(2)</i>
			52 samples, 22 species	38 samples, 18 species	

Botanical collections were done on three areas of the Brazilian Amazonian forests. The first area we focused our collections is a 100-ha upland forest (*terra-firme*) plot established in 1987, within the Biological Dynamics of Forest Fragments Project – BDFFP (LAURANCE et al., 2011), specifically to surveying Lecythidaceae, and were 7 791 trees in 38 species in 8 genera of Lecythidaceae (≥ 10 cm DBH) was tagged and mapped (MORI; LEPSCH-CUNHA, 1995). Over the years, the taxonomy of the species on Lecythidaceae plot was extensively studied by the expert in the family, and over 90% of the tagged trees have been identified (MORI; LEPSCH-CUNHA, 1995). But despite 30 year of research, still there are Lecythidaceae trees waiting fertile collections to be described by the first time. The other two focal areas are located in white-water floodplain forests (*várzea*) approximately 380 km apart each other, the northern portion of the Piagaçu-Purus Sustainable Development Reserve (PP-SDR) (LUIZE et al., 2015a), and the Mamirauá Sustainable Development Reserve (M-SDR) – Jarauá sector (AYRES, 1993). White-water floodplain forests have a reduced number of Lecythidaceae species (c.a. 10 species in 3 genera; LUIZE et al., 2015b) relatively to the upland forests of Central Amazonia (MORI; LEPSCH-CUNHA, 1995).

Leaves of each voucher specimen were dried in silica-gel until DNA extraction. Vouchers were deposited on INPA and PDBFF herbarium. The species identifications for the samples of the Lecythidaceae plot were done by Scott Mori and others, and samples from the floodplain forests sites were done by Bruno Garcia Luize.

DNA extraction, primers utilized, DNA amplification and sequencing

Total DNA extraction were carried out combining pre-washing steps using β -mercaptoethanol (BME) solution and CTAB buffer and then proceeded using NucleoSpin™ Plant II Kit (Macherey- Nagel, Bethlehem, Pennsylvania, USA) following the manufacturer's protocol. PCR (Polymerase chain reaction) amplification of *matK*, *rbcL*, *ycf1*(1), *ycf1*(2) were

performed in an Eppendorf EP gradient S thermocycler (Table 6-2; Thermocycler settings: 1) 94°C – 2min, 2) 50°C – 1min, 3) 72°C – 1min, 4) 94°C – 30 sec, 5) 54°C – 30 sec, 6) 72°C – 1min, 7) 72°C – 7 min. Repeating 40 times steps 4, 5 and 6). Amplifications were performed with a final volume of 25 µL, PCR reactions were done using 10 ng DNA template, 0.25 µL of 1U Taq polymerase (Promega), and 0.25 µL of both the forward and reverse primers (Table 6-2). PCR amplification success were checked in electrophoresis using agarose gel 1.5% stained with GelRed.

Table 6-2. Regions sequenced, and primers utilized in the present study.

<i>cpDNA</i> region	DNA strand	primer sequence	Reference
<i>matK</i>	F	5'AATTTACGATCAATTCATTCAAY3'	THOMSON 2015 <i>unpubl.</i>
<i>matK</i>	R	5' WCTTTATTCGATACAAACTCAT3'	THOMSON 2015 <i>unpubl.</i>
<i>rbcL</i>	F	5'TGGATTCAAAGCTGGTGTTA3'	THOMSON 2015 <i>unpubl.</i>
<i>rbcL</i>	R	5'GATGTGAAGAAGTAGGCCAT3'	THOMSON 2015 <i>unpubl.</i>
<i>ycf1</i> (1)	F	5'AGAACCTTTGATTATGTCTCGACG3'	THOMSON ; VARGAS; DICK, 2018
<i>ycf1</i> (1)	R	5'AGAGACATGCTATAAAAATAGCCCA 3'	THOMSON ; VARGAS; DICK, 2018
<i>ycf1</i> (2)	F	5'TGATTCTGAATCTTTTAGCATTAKAAC T3'	THOMSON ; VARGAS; DICK, 2018
<i>ycf1</i> (2)	R	5'KCGTCGAGACATAATCAAAGGT3'	THOMSON ; VARGAS; DICK, 2018

DNA was sequenced in an Applied Biosystems 3730xl DNA Analyzer platform at the Biomedical Research Core Facility of the University of Michigan. The forward and reverse sequences were visually inspected using Geneious™ 11.1.5 software. First, we assembled the two DNA strands using de-novo assemble, and then we manually resolve the peaks ambiguities

among the strands. We aligned the DNA consensus sequences using the MAFFT multiple alignment V. 1.3.7 algorithm implemented in Geneious™.

Data analysis, and phylogenetic inferences

We inferred the evolutionary relationship among species building a consensus phylogeny using a strict molecular clock that assume all branches with same evolutionary rate, 10 00000 steps for a Markov Chain Monte Carlo (MCMC), and the Yule process model for speciation (GERNHARD, 2008; YULE, 1925) with standard priors as implemented on BEAST (DRUMMOND et al., 2012; DRUMMOND; RAMBAUT, 2007). One consensus phylogeny was inferred for each one of the markers separately: 1) *matK*; 2) *rbcL*; 3) *ycf1*(1); 4) *ycf1*(2); and for the markers concatenated: 5) *matK* + *rbcL*; 6) *ycf1*(1) + *ycf1*(2); 7) *matK* + *rbcL* + *ycf1*(1) + *ycf1*(2). To achieve a consensus phylogeny, we applied the Maximum clade credibility (MCC) on Tree Annotator specifying 10 000 states as burn-in. Finally, a Mega BLAST query in the nucleotide collection of GenBank was performed for each sequence. The query setting was: 1) word size = 28; 2) linear gap cost; 3) scoring 1-2. The resulting e-values lower than 1 e^{-1} was tabulated and species identity with higher bit-scores was compared with species identifications done beforehand.

Results

Sequence similarities

The *matK* average pairwise sequence similarity varies from 96.3% between *Eschweilera* and *Gustavia* to 100% for same species comparisons as well for two different species in the same genus (Figure 6-1). When compared with *matK*, *rbcL* shows higher genetic similarity between species, *rbcL* sequence similarity varies from 98.6% for two different genera (*Eschweilera* vs. *Gustavia*) to 100% between samples of the same genus (Figure 6-1). When *matK* and *rbcL* are evaluated in concatenation for the 42 sequences produced to both

markers, the lower genetic similarity between samples was 97.5%, while for different *Eschweilera* species similarity was 100% (Figure 6-1). The *ycf1*(1) sequences similarities range from a minimum of 93.7% for sequences comparisons of *Eschweilera* vs. *Gustavia* and of *Lecythis* vs. *Gustavia* (Figure 6-1), to 100% within genus comparisons (Figure 6-1). For *ycf1*(2) sequences the lower genetic similarity between genera was 93.3% for *L. gracieana* vs. *G. hexapetala* / *G. poeppigiana* and 100% between congeneric species. The concatenation of *ycf1*(1)+ *ycf1*(2) produced longer sequences (i.e. 2.055 bp), with a sequence similarity of 93.4% between genera and 100 % for comparisons between two different *Eschweilera* species.

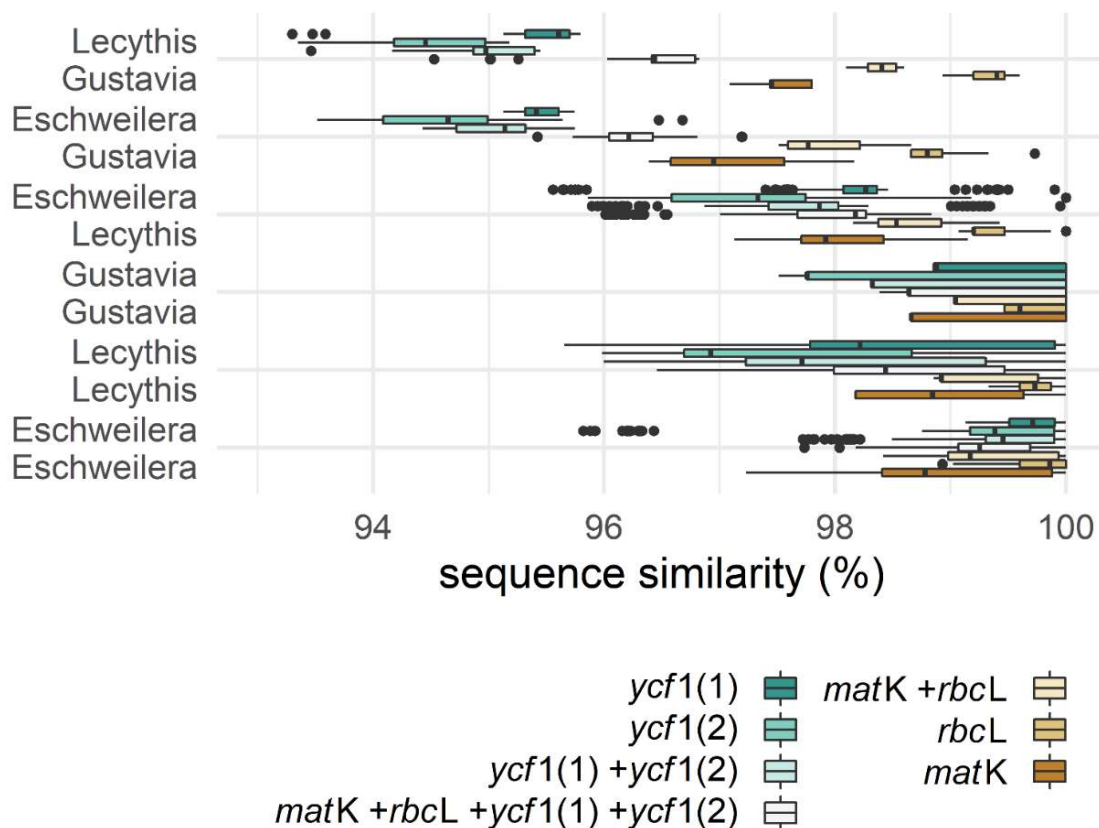


Figure 6-1. Boxplots for the percentage of pairwise genetic similarities for comparisons within and among Lecythidaceae genus. Sequences evaluated are *matK*, *rbcL*, *ycf1*(1), *ycf1*(2) and concatenations for *matK*+*rbcL*, *ycf1*(1)+*ycf1*(2), *matK*+*rbcL*+*ycf1*(1)+*ycf1*(2).

Phylogenetic inferences

Maximum clade credibility phylogeny inferred for the concatenation of the *matk+rbcL* was not able to recover the monophyly of *Gustavia* in relation to *Eschweilera* and *Lecythis* and does not tease apart *Eschweilera* from *Lecythis* species (Figure 6-2a). The same is observed for the phylogeny inferred for the concatenation of *ycf1(1)+ycf1(2)*, furthermore the *ycf1(1)+ycf1(2)* mixture *Lecythis* and *Eschweilera* species, and show a polytomy for *E. coriacea*, *E. albiflora*, and *E. romeu-cardosoi* (Figure 6-2b). The phylogeny inferred with concatenated sequences (*matK+rbcL+ycf1(1)+ycf1(2)*) recovered *Gustavia* species as sister of *Eschweilera* and *Lecythis* species (Figure 6-2c).

The BLAST search, using the 43 *matK* sequences, in the GenBank nucleotide database, correctly find 7 species from 20 species parsed, and only *E. wachenheimi* had the better hit ranked first place. For the *rbcL* sequences the BLAST correctly hits 15 similar sequences identifying 7 species from the 22 species included in the query, the better ranked species with a correct identification was *E. wachenheimi* and *G. hexapetala*. The BLAST search for the *ycf1(1)* sequences only distinguished 7 sequences and three species from the 47 sequences and 22 species submitted to the query. While, for the *ycf1(2)* the BLAST correctly returns the identification for 6 sequences in two species from the 44 sequences in 19 species passed in the query.

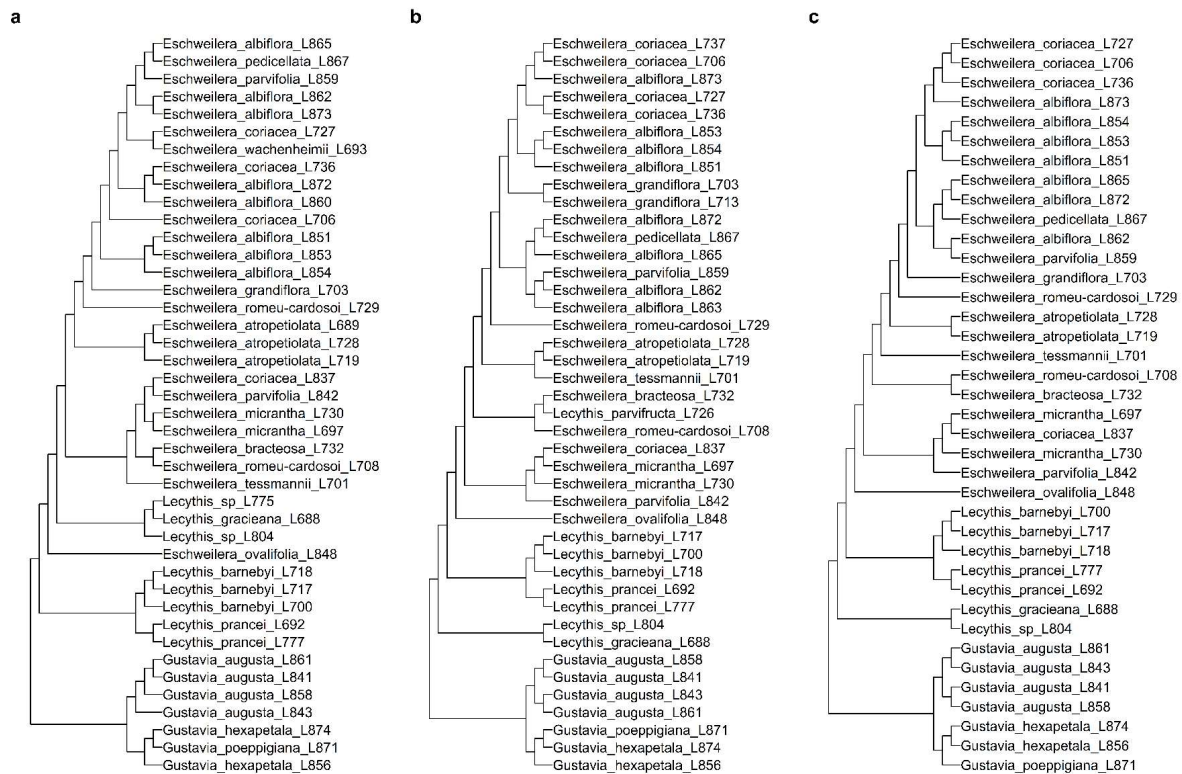


Figure 6-2. Maximum clade credibility (MCC) tree showing the evolutionary relationship recovered for Lecythidaceae species from sequences of a) *matK+rbcL* (42 sequences, 1.606 bp, and 20 species), b) *ycf1(1)+ycf1(2)* (42 sequences, 2.055 bp, and 19 species), and c) *matK+rbcL+ycf1(1)+ycf1(2)* (38 sequences, 3.644 bp, and 18 species).

Discussion

Following CBOL plant work group (2009) there are three main criteria for region selection to barcoding species, the first criteria is the primer “*Universality: Which loci can be routinely sequenced across the land plants?*”. One of the main claims for barcoding application is the readily amplification and sequencing of highly variable and short length sequences (CBOL PLANT WORKING GROUP et al., 2009). Among the regions we had evaluated, the highest amplification and sequencing success achieved was for the *rbcL* (96% of the 52 DNA samples) followed by the *ycf1(1)* (90% of the 52 DNA samples), suggesting these two regions are easily to be sequenced. Successfully amplification and sequencing were not possible for all samples, *ycf1(2)* and *matK* showed greater number of fails in amplification/sequencing (c.a.

20% of the 52 samples was not sequenced). We successfully generated sequences in all of the four markers for 73% of the samples, implying an amplification/sequencing fail rate from 10% of the samples for each marker to an overall 27% of fail for assemble the longer concatenated sequences.

The second criteria suggested by CBOL is the “*Sequence quality and coverage: Which loci are most amenable to the production of bidirectional sequences with few or no ambiguous base calls?*” It is acknowledged that for Sanger sequencing the quality of the sequences decreases with their length (dropping drastically for sequences with more than > 800 bp). We find that the quality of the forward and reverse sequences generated for the regions with longer sequences lengths (~ 1,000 bp, *ycf1*) falls considerably while for the shorter *matK* and *rbcL* (~700bp), the sequences showed good resolutions. The *rbcL* produced high quality sequences and we found very few ambiguities among strands, the same is valid for the *matK*, however for the *matK* we were able to sequence less samples than for the *rbcL*.

The third criteria suggested by the CBOL is “*Discrimination: Which loci enable most species to be distinguished?*” The *matK* and the *rbcL* do not show good resolution for Lecythidaceae species identifications as suggested by the high genetic similarities found between sequences and on the low resolution found in the phylogenies inferred with those regions. Despite *ycf1* shows greater polymorphism and improving species discrimination, that region still produce polytomies. Furthermore, none of the regions we evaluated were able to recover expected monophyly between the clade formed by *Gustavia* from the clade formed by *Lecythis* and *Eschweilera*, with exception of the three regions concatenated. Thus, suggesting that for infer the topology of the sampled Lecythidaceae species the entire plastome should be applied as previously suggested by THOMSON; VARGAS; DICK, 2018.

Previous barcode research includes 20 *matK* and 20 *rbcL* sequences of Lecythidaceae in GenBank collected in French Guiana (GONZALEZ et al., 2009). Combining those sequences and inferring a new phylogeny produced low node resolution (51-100%), and a paraphyly among *Lecythis* and *Eschweilera* species (Figure 6-4 and Figure 6-3). The *matK* phylogeny recovered *Gustavia* as monophyletic (Figure 6-4), however *matK* phylogeny include less *Gustavia* species compared with those included on *rbcL* phylogeny. The *rbcL* phylogeny do not recover the monophyly of *Gustavia* in relation to *Lecythis* and *Eschweilera*, and produces polytomies grouping *Lecythis*, *Eschweilera* and *Couratari* species (Figure 6-3).

A reduced number of sequences for the three regions we evaluated were available in GenBank (*matK* 105 records, *rbcL* 159, and only 29 records considering *ycf1*(1) and *ycf1*(2)). Therefore, low representation of sequences available in GenBank may have influenced the hit of BLAST algorithm to match species identification. A likely influence for such low sequences representation in GenBank may be the early choice of regions applied for Lecythidaceae phylogenetic inferences. For instance, early phylogeny for Lecythidaceae has been inferred from a combination of nuclear ITS, and chloroplastidial *ndhF* (312 sequences for Lecythidaceae in GenBank), *trnL-F* (473 sequences for Lecythidaceae in GenBank), and *trnH-psbA* (325 sequences for Lecythidaceae in GenBank) (HUANG; MORI; KELLY, 2015; MORI et al., 2015). Moreover, despite the *ycf1* be useful for Lecythidaceae barcoding (THOMSON; VARGAS; DICK, 2018) due to be a relatively short, easily sequenced and high variable region of the Lecythidaceae *cpDNA*, the *ycf1* still have poor sample and species representation in sequence repositories hindering their full barcoding application.

Finally, incomplete lineage sorting and introgression are factors that may also limit the use of the regions evaluated for broadly access species identification on Lecythidaceae. In conclusion, the standard (*matK* + *rbcL*) and the newly indicated regions (*ycf1*) do not perform

as desired for barcoding identification of recently diverging Lecythidaceae species. A better taxonomic resolution from Lecythidaceae barcoding could be achieved by combining nuclear and plastome DNA. Furthermore, a greater representation of DNA sequences in repository databases may improve the performance of searching algorithms such BLAST, making the barcoding technique straightforward.

Supplementary material

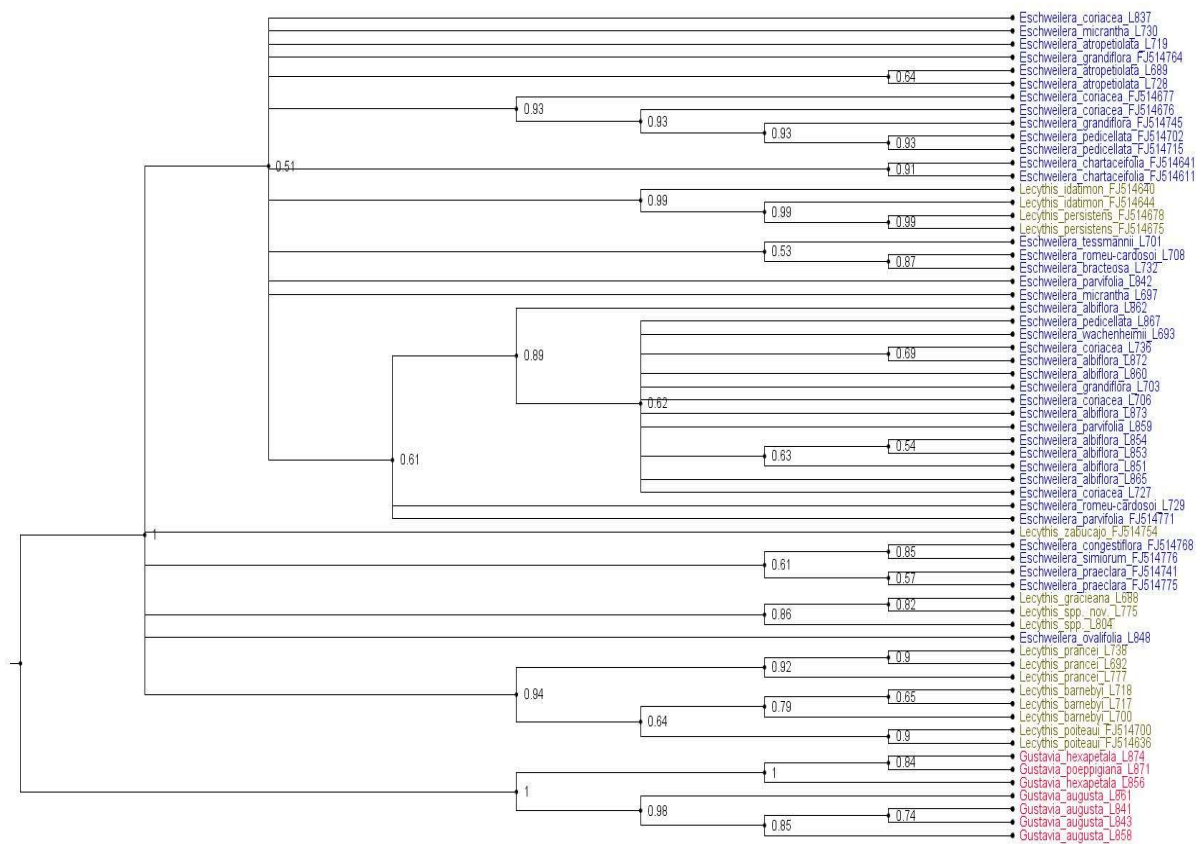


Figure 6-3. The evolutionary relationship between 63 *matK* (838 bp) sequences of Lecythidaceae species, 43 samples collected in Central Amazon (the present study) and 20 sequences available in GenBank for collections coming from French Guiana (Gonzalez et al. 2009 – sequences ending with label FJ5).

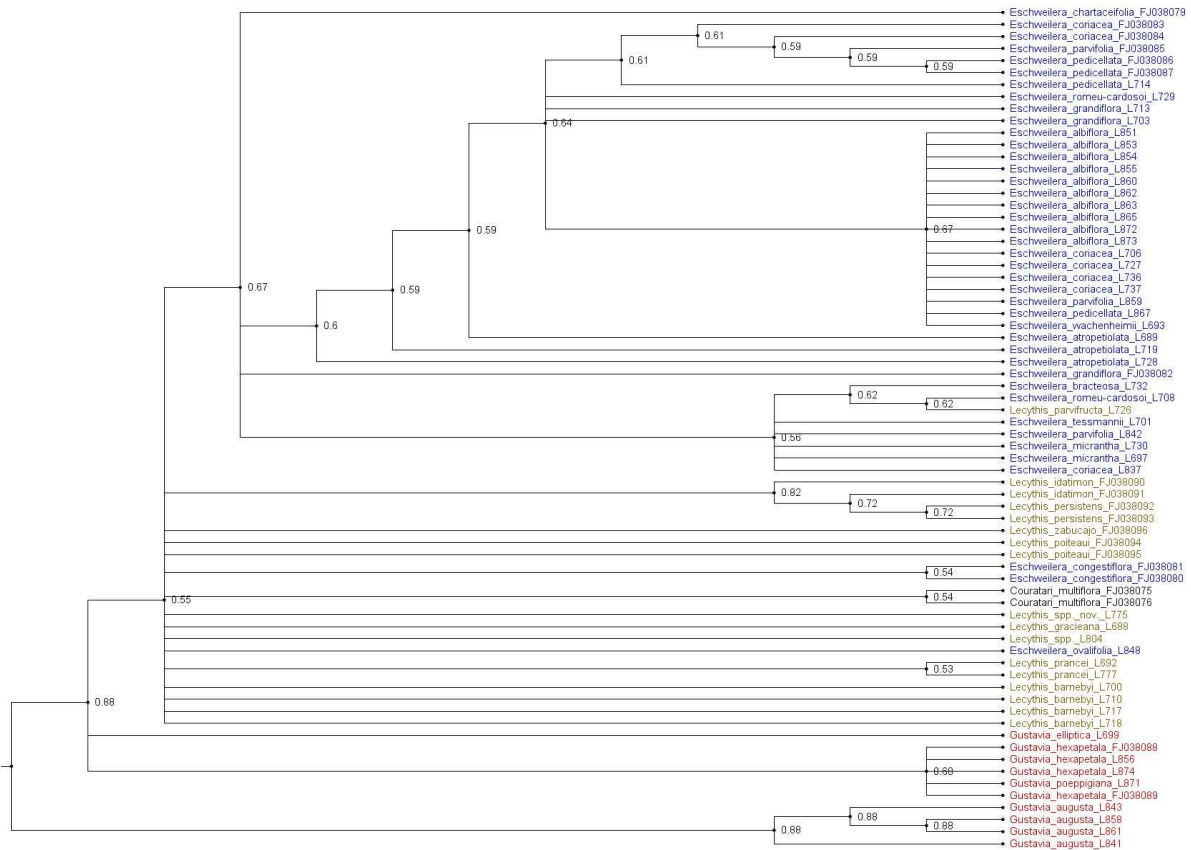


Figure 6-4. The evolutionary relationship between 70 *rbcL* (807 bp) sequences of Lecythidaceae species, 50 samples collected in Central Amazon (the present study) and 20 sequences available in GenBank for collections coming from French Guiana (Gonzalez et al. 2009 – sequences ending with label FJ03).

7. Conclusões

A pergunta que esta tese busca responder *Como os habitats florestais das áreas úmidas da Amazônia tem contribuído para a diversificação de espécies arbóreas e na manutenção da diversidade regional destas comunidades?* Foi abordada por diferentes perspectivas da biogeografia e da ecologia e possibilitou que novos questionamentos surgissem. Alguns destes questionamentos foram abordados nesta tese e nos trazem evidências para o papel das áreas úmidas na diversidade das florestas Amazônicas.

No segundo capítulo da tese, a partir de uma revisão em trabalhos de inventários florestais assim como em coletas botânicas depositadas em museus, mostramos que a diversidade de espécies arbóreas ocorrendo em áreas úmidas da Amazônia é 3.2 vezes maior do que previamente considerada. Dois motivos para a maior quantidade de espécies encontrada foram a utilização de coletas botânicas depositadas em herbários com anotações para o tipo de habitat onde a coleta foi realizada e a utilização de uma definição ampla para áreas úmidas. Enquanto as pesquisas anteriores focaram principalmente em inventários florestais e em tipos específicos de áreas úmidas da Amazônia (MONTERO; PIEDADE; WITTMANN, 2012; WITTMANN, 2012). A definição do conjunto de espécies arbóreas com ocorrência registrada em florestas de áreas úmidas da Amazônia mostra que pouco mais da metade das espécies conhecidas para as florestas da Amazônia também ocorrem em áreas úmidas. Em geral, a proporção de espécies arbóreas que podem ocorrer em áreas úmidas da Amazônia é alta, uma vez que atualmente cerca de 14 - 17% da área bacia Amazônica é classificada como área úmida. O que fortalece a hipótese de que a migração de espécies das florestas de terra-firme para as florestas em áreas úmidas deva ser recorrente e mais pronunciada do que a contrário (i.e., espécies de áreas úmidas migrando para terra-firme). A dificuldade em separar quais espécies são exclusivas nos habitats de áreas úmidas daquelas espécies que ocorrem apenas esporadicamente nas áreas úmidas é

uma questão que permanece em aberto. A delimitação dos conjuntos de espécies para os habitats florestais da Amazônia (e.g. terra-firme, campinaranas) e o refinamento dos conjuntos de espécies registrados nos diferentes tipos de áreas-úmidas (e.g. Várzea, Igapó, Pântanos, Baixios) pode contribuir para um melhor entendimento das histórias evolutivas das espécies de árvores da Amazônia.

No terceiro capítulo da tese, foi realizada uma busca por registros de ocorrência para cada uma das espécies de árvores registradas para as florestas da Amazônia e listadas em CARDOSO e colaboradores (2017). Os registros de ocorrência com coordenadas geográficas foram utilizados para estimar o tamanho da área de distribuição das espécies ao longo da região Neotropical e a amplitude de tolerância climática que as espécies ocorrem (i.e., amplitude do nicho). Nesta pesquisa mostramos que quanto maior a amplitude de nicho que uma espécie apresenta maior é o tamanho da área que a espécie se distribui. Além desta relação direta entre amplitude de nicho e tamanho da área de distribuição, nossos resultados indicaram que a tolerância para a variação sazonal na quantidade de precipitação e a tolerância para a demanda de evapotranspiração são os descritores ambientais que estão mais relacionados com a amplitude da distribuição geográfica que as espécies assumem. Utilizando o levantamento das espécies que podem ocorrer em áreas úmidas, mostramos que estas espécies tendem a ter uma área de distribuição e uma amplitude de tolerância climática maior do que as espécies que foram registradas apenas em florestas de terra firme. Esta pesquisa conclui que uma em cada quatro espécies de árvores que ocorre na Amazônia possuem distribuição geográfica que vai além dos limites do bioma e sugere que as espécies que conseguem ocupar as florestas em áreas úmidas sejam relativamente mais generalistas do que as espécies que ocupam as florestas de terra firme.

No quarto capítulo, foi avaliada a diversidade β em florestas de várzea na Amazônia central. Esta pesquisa foi realizada em uma escala espacial menor que as apresentadas anteriormente e mostra que a diversidade β nas florestas de várzea é alta tanto para a variação na composição de espécies como para a variação nas histórias evolutivas que estão contidas entre as manchas de floresta. A distância geográfica entre manchas de floresta é o descritor com maior influência na variação da composição de espécies, enquanto a distância ecológica, medidas em termos de condições ambientais e estruturais das florestas, é o descritor com maior influência na variação das histórias evolutivas. Em outras palavras, as linhagens que irão compor as diferentes manchas de florestas são selecionadas pela similaridade nas condições ambientais e as espécies são selecionadas devido à proximidade geográfica entre os locais. Este resultado tem implicações teóricas uma vez que sugere que o efeito de seleção ambiental se dá nas linhagens evolutivas enquanto distanciamento geográfico terá influência em quais as espécies de cada linhagem irão ocupar as manchas de floresta. As predições espacialmente explícitas apresentadas sugerem locais que podem estar sofrendo homogeneização biótica devido ao uso intensivo das florestas de várzea. Levando em consideração os processos envolvidos na estruturação da β -diversidade e a maneira que se distribuí ao longo da região é apresentada uma estratégia para a delimitação de áreas de proteção para as florestas de várzea. Em resumo, para proteger diferentes linhagens e manter o processo evolutivo acontecendo o mais adequado seria delimitar áreas de proteção com base na diferença ambiental entre os locais - assim mais diferenças de histórias evolutivas tendem a ser mantidas. No entanto, se proteger as espécies sem preocupação com o processo evolutivo a delimitação de áreas de proteção poderia ser feita com base nas distâncias das áreas de proteção que já existem. No mais, o estudo indica a região do interflúvio Purus – Madeira e as planícies de inundação da calha norte do rio Amazonas (i.e., Lago Badajós) como paisagens importantes para a proteção da variação na composição e nas histórias evolutivas das árvores que compõem as florestas de várzea na Amazônia central.

No quinto capítulo, foram definidos pares de espécies que tendem a co-ocorrer positiva e negativamente em florestas de várzea. A grande maioria das combinações possíveis entre pares de espécies não chega a acontecer, sugerindo que o acaso seja um dos fatores responsáveis pela estruturação das comunidades. Entre os pares de espécies que possuem co-ocorrências estruturadas, as espécies com maiores abundâncias relativas são as que apresentam maiores quantidades de associações. Nas co-ocorrências estruturadas positivamente as espécies tendem a co-ocorrer com maior frequência do que o acaso e essa agregação possivelmente é influenciada por interações de facilitação ou por similaridades na ocupação do espaço de geográfico e ecológico entre as espécies. Para as co-ocorrências estruturadas negativamente há uma tendência de segregação entre as espécies, possivelmente indicando interações de competição ou distribuições geográficas e ecológicas não coincidentes. Esta pesquisa indica que processos estocásticos sejam preponderantes na definição das co-ocorrências das espécies de árvores em florestas de várzea.

No sexto capítulo, utilizamos marcadores de DNA chloroplastidial para auxiliar na identificação de espécies em Lecythidaceae. Esta família está entre as de maior importância para a estrutura das florestas da Amazônia, sendo um grupo de difícil delimitação e que radiou recentemente. As três regiões do DNA avaliadas (i.e., *matK*, *rbcL*, *ycf1*) não foram suficientes para a delimitação das espécies conforme previamente identificadas com base em suas características morfológicas diagnósticas, devido à alta similaridade entre as sequências de DNA. Os marcadores avaliados agrupam clados de Lecythidaceae que foram previamente demonstrados como sendo distintos, porém não há clara monofilia de espécies pertencentes a gêneros irmãos. A hibridização/introgressão entre espécies proximamente relacionadas e/ou a falta de fixação de alelos específicos nas regiões avaliadas do DNA são fatores que podem influenciar a dificuldade na delimitação de espécies nesta família de árvores. No entanto, a

aplicação da técnica de código de barras de DNA para identificação de espécies de Lecythidaceae poderá ser beneficiada pelo uso de um número maior de sequências de DNA. Por exemplo, técnicas de sequenciamento em larga escala têm permitido a obtenção de sequências completas do DNA das espécies, e em 2018 foi publicado 24 plastomas anotados para espécies de Lecythidaceae (Thomsom et al. 2018). A combinação de marcadores de DNA nuclear e chloroplastidial também pode beneficiar uma melhor resolução para a separação das espécies.

A combinação dos resultados apresentados nos mostra como as áreas úmidas promovem diferenças nas características ecológicas das espécies e heterogeneidade biótica na região. O conjunto de espécies de árvores das áreas úmidas da Amazônia indica que nem todas as famílias botânicas possuem espécies ocorrendo nestes habitats. As espécies de árvores que ocorrem nos habitats de área úmida tendem a ter maiores distribuições geográficas e maiores amplitudes de tolerância de nicho. Nas áreas úmidas das planícies de inundação da Amazônia central há alta variação na composição de espécies e nas histórias evolutivas entre manchas de floresta; as quais são moduladas pela grande área ocupada pelo habitat e pelos gradientes ambientais que se forma entre as localidades. As espécies mais comuns em florestas de várzea ao longo do rio Amazonas tendem a estruturar mais co-ocorrências em relação as espécies de abundância intermediária e raras. O entendimento da origem e manutenção da diversidade biológica requer que se reconheça a influência dos diferentes habitats nas histórias de vida das espécies. E como conclusão geral indicamos que para um melhor entendimento da origem e manutenção da diversidade nas florestas da Amazônia é necessário incluir e avaliar os diferentes habitats florestais como um todo.

8. Referências

- ALBERNAZ, A. L. et al. Tree species compositional change and conservation implications in the white-water flooded forests of the Brazilian Amazon. **Journal of Biogeography**, v. 39, p. 869–883, 2012.
- ALDANA, A. M. et al. Environmental filtering of eudicot lineages underlies phylogenetic clustering in tropical South American flooded forests. **Oecologia**, v. 183, n. 2, p. 327–335, 2017.
- ALEIXO, I. et al. Amazonian rainforest tree mortality driven by climate and functional traits. **Nature Climate Change**, 2019.
- ANTONELLI, A. et al. Tracing the impact of the Andean uplift on Neotropical plant evolution. **Proceedings of the National Academy of Sciences**, v. 106, n. 24, p. 9749–9754, 2009.
- ANTONELLI, A. Biogeography: Drivers of bioregionalization. **Nature Ecology & Evolution**, v. 1, n. 4, p. 0114, 6 mar. 2017.
- ANTONELLI, A. et al. Amazonia is the primary source of Neotropical biodiversity. **Proceedings of the National Academy of Sciences**, v. 115, n. 23, p. 6034–6039, 5 jun. 2018.
- ANTONELLI, A.; SANMARTÍN, I. Why are there so many plant species in the Neotropics? **Taxon**, v. 60, n. 2, p. 403–414, 2011.
- ARITA, H. T. Species co-occurrence analysis: pairwise versus matrix-level approaches. **Global Ecology and Biogeography**, v. 25, n. 11, p. 1397–1400, 2016.
- ASNER, G. P. et al. Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. **Science**, v. 355, n. 6323, p. 385–389, 2017.
- AYRES, J. M. **As matas de várzea do Mamirauá: médio rio Solimões**. 2. ed. ed. [s.l.] Conselho Nacional de Desenvolvimento Científico e Tecnológico e Sociedade Civil Mamirauá, 1993.
- BAKER, P. A. et al. The emerging field of geogenomics: Constraining geological problems with genetic data. **Earth-Science Reviews**, v. 135, p. 38–47, 2014.
- BAKER, T. R. et al. Maximising Synergy among Tropical Plant Systematists, Ecologists, and Evolutionary Biologists. **Trends in Ecology and Evolution**, v. 32, n. 4, p. 258–267, 2017.
- BASELGA, A. Partitioning the turnover and nestedness components of beta diversity. **Global Ecology and Biogeography**, v. 19, n. October, p. 134–143, 2010.
- BASELGA, A. The relationship between species replacement , dissimilarity derived from nestedness , and nestedness. **Global Ecology and Biogeography**, v. 21, p. 1223–1232, 2012.
- BASELGA, A. et al. **Partitioning beta diversity into turnover and nestedness components. Package betapart, Version**, 2018.
- BASELGA, A.; ORME, C. D. L. betapart: an R package for the study of beta diversity. **Methods in Ecology and Evolution**, v. 3, n. 5, p. 808–812, 2012.
- BERTASSOLI-JR, D. J. et al. Spatiotemporal variations of riverine discharge within the Amazon Basin during the late Holocene coincide with extratropical temperature anomalies. **Geophysical Research Letters**, v. 46, n. 15, p. 9013–9022, 2019.
- BLOIS, J. L. et al. Space can substitute for time in predicting climate-change effects on biodiversity. **Proceedings of the National Academy of Sciences**, v. 110, n. 23, p. 9374–9379, 2013.
- BOYLE, B. et al. The taxonomic name resolution service: an online tool for automated standardization of plant names. **BMC bioinformatics**, v. 14, n. 1, p. 16, 2013.

- BRYANT, J. A. et al. Microbes on mountainsides : Contrasting elevational patterns of bacterial and plant diversity. **Proceedings of the National Academy of Sciences**, v. 105, n. 1, p. 11505–11511, 2008.
- CARDOSO, D. et al. Amazon plant diversity revealed by a taxonomically verified species list. **Proceedings of the National Academy of Sciences**, p. 201706756, 18 set. 2017.
- CARSTENSEN, D. W. et al. Introducing the biogeographic species pool. **Ecography**, v. 36, n. 12, p. 1310–1318, 2013.
- CAVENDER-BARES, J. et al. The merging of community ecology and phylogenetic biology. **Ecology Letters**, v. 12, n. 7, p. 693–715, 2009.
- CAVENDER-BARES, J.; KEEN, A.; MILES, B. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. **Ecology**, v. 87, n. sp7, p. S109–S122, 2006.
- CBOL PLANT WORKING GROUP et al. A DNA barcode for land plants. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 31, p. 12794–7, 2009.
- CHAO, A. et al. Rarefaction and extrapolation with Hill numbers : A framework for sampling and estimation in species diversity studies Rarefaction and extrapolation with Hill numbers : a framework for sampling and estimation in species diversity studies. **Ecological Monographs**, v. 84, n. 1, p. 45–67, 2014.
- CHASE, M. W. et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. **Botanical Journal of the Linnean Society**, v. 181, n. 1, p. 1–20, maio 2016.
- CHESSON, P. Mechanisms of maintenance of species diversity. **Annual Review of Ecology and Systematics**, v. 31, p. 343–66, 2000.
- COLWELL, R. K.; RANGEL, T. F. Hutchinson’s duality: The once and future niche. **Proceedings of the National Academy of Sciences**, v. 106, n. Supplement 2, p. 19651–19658, 2009.
- CONDIT, R. et al. Beta-diversity in tropical forest trees. **Science**, v. 295, n. 5555, p. 666–669, 2002.
- CONNOR, E. F.; SIMBERLOFF, D. The Assembly of Species Communities: Chance or Competition? **Ecology**, v. 60, n. 6, p. 1132–1140, 1979.
- CORNELL, H. V.; HARRISON, S. P. What Are Species Pools and When Are They Important? **Annual Review of Ecology, Evolution, and Systematics**, v. 45, n. 1, p. 45–67, 2014.
- CORREA, S. B.; WINEMILLER, K. Terrestrial – aquatic trophic linkages support fish production in a tropical oligotrophic river. **Oecologia**, v. 186, n. 4, p. 1069–1078, 2018.
- CURTIS, P. G. et al. Classifying drivers of global forest loss. **Science**, v. 361, n. 6407, p. 1108–1111, 2018.
- DAUBY, G. et al. ConR : An R package to assist large-scale multispecies preliminary conservation assessments using distribution data. **Ecology and Evolution**, v. 7, n. 24, p. 11292–11303, dez. 2017.
- DE ASSIS, R. L. et al. Patterns of floristic diversity and composition in floodplain forests across four Southern Amazon river tributaries, Brazil. **Flora: Morphology, Distribution, Functional Ecology of Plants**, v. 229, p. 124–140, 2017.
- DE FATIMA ROSSETTI, D. et al. New geological framework for Western Amazonia (Brazil) and implications for biogeography and evolution. **Quaternary Research**, v. 63, n. 1, p. 78–89, 2005.
- DEXTER, K.; CHAVE, J. Evolutionary patterns of range size, abundance and species richness in Amazonian angiosperm trees. **PeerJ**, v. 4, p. e2402, 2016.

- DEXTER, K. G. et al. Dispersal assembly of rain forest tree communities across the Amazon basin. **Proceedings of the National Academy of Sciences**, v. 114, n. 10, p. 2645–2650, 2017.
- DEXTER, K. G.; TERBORGH, J. W.; CUNNINGHAM, C. W. Historical effects on beta diversity and community assembly in Amazonian trees. **Proceedings of the National Academy of Sciences**, v. 109, n. 20, p. 7787–7792, 2012.
- DICK, C. W.; KRESS, W. J. Dissecting Tropical Plant Diversity with Forest Plots and a Molecular Toolkit. **BioScience**, v. 59, n. 9, p. 745–755, 2009.
- DONOGHUE, M. J. A phylogenetic perspective on the distribution of plant diversity. **Proceedings of the National Academy of sciences**, v. 105, p. 11549–11555, 2008.
- DRAPER, F. C. et al. Dominant tree species drive beta diversity patterns in Western Amazonia. **Ecology**, v. 100, n. 4, p. e02636, 2019.
- DRUMMOND, A. J. et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7. **Molecular Biology and Evolution**, v. 29, n. 8, p. 1969–1973, 2012.
- DRUMMOND, A. J.; RAMBAUT, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC evolutionary biology**, v. 7, n. 1, p. 214, 2007.
- DUIVENVOORDEN, J. F.; SVENNING, J.-C.; WRIGHT, S. J. Beta diversity in tropical forests. **Science**, v. 295, n. 5555, p. 636–637, 2002.
- DUQUE, Á. et al. Distance decay of tree species similarity in protected areas on terra firme forests in Colombian Amazonia. **Biotropica**, v. 41, n. 5, p. 599–607, 2009.
- ESQUIVEL-MUELBERT, A. et al. Seasonal drought limits tree species across the Neotropics. **Ecography**, v. 40, n. 5, p. 618–629, maio 2017.
- ESQUIVEL-MUELBERT, A. et al. Compositional response of Amazon forests to climate change. **Global change biology**, v. 25, n. 1, p. 39–56, 2019.
- ESTES, L. et al. The spatial and temporal domains of modern ecology. **Nature Ecology and Evolution**, v. 2, n. 5, p. 819–826, 2018.
- EVA, H. et al. **A proposal for defining the geographical boundaries of Amazonia; synthesis of the results from an expert consultation workshop organized by the European Commission in collaboration with the Amazon Cooperation Treaty Organization-JRC Ispra, 7-8 June 2005**. 1. ed. [s.l.] European Commission, 2005.
- FAN, Y.; LI, H.; MIGUEZ-MACHO, G. Global patterns of groundwater table depth. **Science**, v. 339, n. 6122, p. 940–943, 2013.
- FAUSET, S. et al. Hyperdominance in Amazonian forest carbon cycling Hyperdominance in Amazonian forest carbon cycling. **Nature Communications**, p. 6:6857, 2015.
- FEELEY, K. Are we filling the data void? An assessment of the amount and extent of plant collection records and census data available for tropical South america. **PloS one**, v. 10, n. 4, p. e0125629, jan. 2015.
- FEELEY, K. J.; SILMAN, M. R. Extinction risks of Amazonian plant species. **Proceedings of the National Academy of Sciences**, v. 106, n. 30, p. 12382–12387, 2009.
- FEELEY, K. J.; SILMAN, M. R. The data void in modeling current and future distributions of tropical species. **Global Change Biology**, v. 17, n. 1, p. 626–630, 2011.
- FERREIRA-FERREIRA, J. et al. Combining ALOS/PALSAR derived vegetation structure and inundation patterns to characterize major vegetation types in the Mamirauá Sustainable Development

Reserve, Central Amazon floodplain, Brazil. **Wetlands Ecology and Management**, v. 23, n. 1, p. 41–59, 11 fev. 2015.

FERREIRA, C. D. S. et al. The role of carbohydrates in seed germination and seedling establishment of *Himatanthus sucuuba*, an Amazonian tree with populations adapted to flooded and non-flooded conditions. **Annals of Botany**, v. 104, n. 6, p. 1111–1119, 2009.

FERREIRA, C. S. et al. Floodplain and upland populations of Amazonian *Himatanthus sucuuba*: Effects of flooding on germination, seedling growth and mortality. **Environmental and Experimental Botany**, v. 60, n. 3, p. 477–483, 2007.

FERREIRA, C. S. et al. Genetic variability, divergence and speciation in trees of periodically flooded forests of the Amazon: a case study of *Himatanthus sucuuba* (SPRUCE) WOODSON. In: JUNK, W. J. et al. (Eds.). **Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management**. [s.l.] Springer, 2010. p. 301–312.

FERRIER, S. et al. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. **Diversity and Distributions**, v. 13, n. 3, p. 252–264, 2007.

FERRIER, S.; GUIBAN, A. Spatial modelling of biodiversity at the community level. **Journal of Applied Ecology**, v. 43, n. 3, p. 393–404, 2006.

FICETOLA, G. F.; MAZEL, F.; THUILLER, W. Global determinants of zoogeographical boundaries. **Nature Ecology & Evolution**, v. 1, n. 4, p. 0089, 6 mar. 2017.

FINE, P.; ZAPATA, F.; DALY, D. Investigating processes of Neotropical rain forest tree diversification by examining the evolution and historical biogeography of the *Protieae* (Burseraceae). **Evolution**, v. 68, n. 7, p. 1988–2004, 2014.

FITZPATRICK, M. C.; KELLER, S. R. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. **Ecology letters**, v. 18, n. 1, p. 1–16, 2015.

FOUGNIES, L. et al. Arbuscular mycorrhizal colonization and nodulation improve flooding tolerance in *Pterocarpus officinalis* Jacq. seedlings. **Mycorrhiza**, v. 17, n. 3, p. 159–166, 2007.

GALLAGHER, R. V. Correlates of range size variation in the Australian seed-plant flora. **Journal of Biogeography**, v. 43, n. 7, p. 1287–1298, 2016.

GAO, B. NDWI A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space. **Remote sensing of environment**, v. 58, n. 3, p. 257–266, 1996.

GENTRY, A. H. 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**, v. 69, n. 3, p. 557–593, 1982.

GENTRY, A. H. Changes in Plant Community Diversity and Floristic Composition on Environmental and Geographical Gradients. **Annals of the Missouri Botanical Garden**, v. 75, n. 1, p. 1–34, 1988.

GERNHARD, T. The conditioned reconstructed process. **Journal of theoretical biology**, v. 253, n. 4, p. 769–778, 2008.

GOMES, V. H. F. F. et al. Species Distribution Modelling: Contrasting presence-only models with plot abundance data. **Scientific Reports**, v. 8, n. 1, p. 1003, 17 dez. 2018.

GONZALEZ, M. A. et al. Identification of amazonian trees with DNA barcodes. **PLoS ONE**, v. 4, n. 10, 2009.

GORELICK, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote**

Sensing of Environment, v. 202, p. 18–27, 2017.

GOTELLI, J. N. Null model analysis of species co-occurrence patterns. **Ecology**, v. 81, n. 9, p. 2606–21, 2000.

GOVAERTS, R. How Many Species of Seed Plants Are There? **Taxon**, v. 50, n. 4, p. 1085–1090, 2001.

GRAHAM, C. H.; FINE, P. V. A. Phylogenetic beta diversity : linking ecological and evolutionary processes across space in time. **Ecology Letters**, v. 11, n. 12, p. 1265–1277, 2008.

GRIFFITH, D. M.; VEECH, J. A.; MARSH, C. J. **cooccur** : Probabilistic Species Co-Occurrence Analysis in R. **Journal of Statistical Software**, v. 69, n. Code Snippet 2, p. 1–17, 2016.

GRÖMPING, U. Relative Importance for Linear Regression in R : The Package **relaimpo**. **Journal of Statistical Software**, v. 17, n. 1, 2006.

GUEVARA, J. E. et al. Low phylogenetic beta diversity and geographic neo-endemism in Amazonian white-sand forests. **Biotropica**, v. 48, n. 1, p. 34–46, 2016.

GUMBRICHT, T. et al. An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. **Global Change Biology**, v. 23, n. 9, p. 3581–3599, set. 2017.

HAFFER, J. Speciation in Amazonian forest birds. **Science**, v. 165, n. 3889, p. 131–137, 1969.

HEBERT, P. D. N.; GREGORY, T. R. The promise of DNA barcoding for taxonomy. **Systematic Biology**, v. 54, n. 5, p. 852–859, 2005.

HESS, L. L. et al. Wetlands of the Lowland Amazon Basin: Extent, Vegetative Cover, and Dual-season Inundated Area as Mapped with JERS-1 Synthetic Aperture Radar. **Wetlands**, v. 35, n. 4, p. 745–756, 2015a.

HESS, L. L. et al. **LBA-ECO LC-07 wetland extent, vegetation, and inundation: lowland Amazon Basin** ORNL DAAC Oak Ridge, Tennessee, USA ORNL DAAC, , 2015b.

HIGGINS, M. A. et al. Linking imaging spectroscopy and LiDAR with floristic composition and forest structure in Panama. **Remote Sensing of Environment**, v. 154, n. C, p. 358–367, 2014.

HIJMAN, R. J. et al. Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology**, v. 25, n. 15, p. 1965–1978, 2005.

HIJMAN, R. J. **raster: Geographic Data Analysis and Modeling**. [s.l.: s.n.].

HONORIO CORONADO, E. N. et al. Phylogenetic diversity of Amazonian tree communities. **Diversity and Distributions**, v. 21, n. 11, p. 1295–1307, nov. 2015.

HOORN, C. et al. The Development of the Amazonian Mega-Wetland (Miocene; Brazil, Colombia, Peru, Bolivia). In: HOORN, C.; WESSELINGH, F. P. (Eds.). . **Amazonia, Landscape and Species Evolution: A Look into the Past**. 1. ed. [s.l.] Blackwell Publishing, 2010a. p. 123–142.

HOORN, C. et al. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. **Science**, v. 330, n. 6006, p. 927–931, 12 nov. 2010b.

HOORN, C. et al. Origins of biodiversity—response. **Science**, v. 331, n. 6016, p. 399–400, 2011.

HOORN, C. et al. Biodiversity from mountain building. **Nature Geoscience**, v. 6, n. 3, p. 154, 2013.

HOORN, C.; WESSELINGH, F. **Amazonia: landscape and species evolution: a look into the past**. 1. ed. [s.l.] Blackwell Publishing, 2010.

- HOPKINS, M. J. G. Modelling the known and unknown plant biodiversity of the Amazon Basin. **Journal of Biogeography**, v. 34, n. 8, p. 1400–1411, 2007.
- HOPKINS, M. J. G. Are we close to knowing the plant diversity of the Amazon ? **Anais da Academia Brasileira de Ciências**, v. 91, n. 3, p. e20190396, 2019.
- HORTAL, J. et al. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. **Annual Review of Ecology, Evolution, and Systematics**, v. 46, n. 1, p. 523–549, 4 dez. 2015.
- HOUSEHOLDER, E. et al. A Diversity of Biogeographies in an Extreme Amazonian Wetland Habitat. In: **Forest structure, function and dynamics in Western Amazonia**. Chichester, UK: John Wiley & Sons, Ltd, 2016. p. 145–157.
- HSIEH, T. C.; MA, K. H.; CHAO, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). **Methods in Ecology and Evolution**, v. 7, n. 12, p. 1451–1456, 2016.
- HUANG, Y.-Y.; MORI, S. A.; KELLY, L. M. Toward a phylogenetic-based Generic Classification of Neotropical Lecythidaceae—I. Status of Bertholletia, Corythophora, Eschweilera and Lecythis. **Phytotaxa**, v. 203, n. 2, p. 85, mar. 2015.
- HUBBELL, S. P. **The unified neutral theory of biodiversity and biogeography (MPB-32)**. [s.l.] Princeton University Press, 2001.
- HUBBELL, S. P. et al. How many tree species are there in the Amazon and how many of them will go extinct? **Proceedings of the National Academy of Sciences**, v. 105, n. Supplement 1, p. 11498–11504, 12 ago. 2008.
- HUBBELL, S. P. Tropical rain forest conservation and the twin challenges of diversity and rarity. **Ecology and Evolution**, v. 3, n. 10, p. 3263–3274, 2013.
- INSEL, N.; POULSEN, C. J.; EHLERS, T. A. Influence of the Andes Mountains on South American moisture transport, convection, and precipitation. **Climate Dynamics**, v. 35, n. 7–8, p. 1477–1492, 29 dez. 2010.
- IPBES et al. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. **IPBES**, 2019.
- IRION, G. Soil infertility in the Amazonian rain forest. **Naturwissenschaften**, v. 65, n. 10, p. 515–519, 1978.
- IRION, G. et al. Development of the Amazon Valley During the Middle to Late Quaternary: Sedimentological and Climatological Observations. In: JUNK, W. J. et al. (Eds.). **Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management**. 1. ed. Dordrecht Heidelberg London New York: Elsevier, 2010. p. 27–42.
- JANZEN, D. H. Why Mountain Passes are Higher in the Tropics. **The American Naturalist**, v. 101, n. 919, p. 233–249, 1967.
- JARDIM BOTÂNICO DO RIO DE JANEIRO. **Flora do Brasil 2020 [under construction]**. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/>>.
- JONES, M. M. et al. Strong congruence in tree and fern community turnover in response to soils and climate in central Panama. **Journal of Ecology**, v. 101, n. 2, p. 506–516, 2013.
- JUNK, W. et al. A classification of major natural habitats of Amazonian white-water river floodplains (várzeas). **Wetlands Ecology and Management**, v. 20, n. 6, p. 461–475, 2012.
- JUNK, W. J. et al. The flood pulse concept in river-floodplain systems. **Canadian special**

publication of fisheries and aquatic sciences, v. 106, n. 1, p. 110–127, 1989.

JUNK, W. J. et al. **Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management**. 1. ed. Dordrecht Heidelberg London New York: Springer, 2010. v. 210

JUNK, W. J. et al. A Classification of Major Naturally-Occurring Amazonian Lowland Wetlands. **Wetlands**, v. 31, n. 4, p. 623–640, 2011.

JUNK, W. J. et al. Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. **Aquatic Conservation: Marine and Freshwater Ecosystems**, v. 24, n. 1, p. 5–22, 2014.

KEDDY, P. A. **Wetland ecology: principles and conservation**. [s.l.] Cambridge University Press, 2010.

KOHLI, B. A.; TERRY, R. C.; ROWE, R. J. A trait-based framework for discerning drivers of species co-occurrence across heterogeneous landscapes. **Ecography**, v. 41, p. 1921–1933, 2018.

KRAFT, N. J. B. et al. Disentangling the Drivers of β Diversity Along Latitudinal and Elevational Gradients Nathan. **Science**, v. 333, n. 1755, 2011.

KREFT, H.; JETZ, W. Global patterns and determinants of vascular plant diversity. **Proceedings of the National Academy of sciences**, v. 104, n. 14, p. 5925–5930, 2007.

LANG, C. et al. Near infrared spectroscopy facilitates rapid identification of both young and mature Amazonian tree species. **PLOS ONE**, v. 10, n. 8, p. e0134521, 27 ago. 2015.

LATRUBESSE, E. M. et al. The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. **Earth Science Reviews**, v. 99, n. 3–4, p. 99–124, 27 abr. 2010.

LAURANCE, W. F. et al. Rapid decay of tree-community composition in Amazonian forest fragments. **Proceedings of the National Academy of sciences**, v. 103, n. 50, p. 19010–19014, 2006.

LAURANCE, W. F. et al. The fate of Amazonian forest fragments: A 32-year investigation. **Biological Conservation**, v. 144, n. 1, p. 56–67, 2011.

LEAL, B. S. S.; DA SILVA, C. P.; PINHEIRO, F. Phylogeographic Studies Depict the Role of Space and Time Scales of Plant Speciation in a Highly Diverse Neotropical Region. **Critical Reviews in Plant Sciences**, p. 1–16, 2016.

LEPRIEUR, F. et al. Quantifying Phylogenetic Beta Diversity : Distinguishing between ‘ True ’ Turnover of Lineages and Phylogenetic Diversity Gradients. **PloS one**, v. 7, n. 8, p. e42760, 2012.

LEVIN, S. A. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. **Ecology**, v. 73, n. 6, p. 1943–1967, 1992.

LI, X. et al. Plant DNA barcoding: from gene to genome. **Biological reviews of the Cambridge Philosophical Society**, v. 90, n. 1, p. 157–166, 2015.

LINDEMAN, R. H.; MERENDA, P. F.; GOLD, R. Z. **Introduction to bivariate and multivariate analysis** Glenview IL: Scott, Foresman. London Foresman and co., , 1980.

LUIZE, B. G. et al. Effects of the Flooding Gradient on Tree Community Diversity in Várzea Forests of the Purus River, Central Amazon, Brazil. **Biotropica**, v. 47, n. 2, 2015a.

LUIZE, B. G. et al. A floristic survey of angiosperm species occurring at three landscapes of the Central Amazon várzea, Brazil. **Check List**, v. 11, n. 6, 2015b.

LUIZE, B. G. B. G. et al. The tree species pool of Amazonian wetland forests: Which species can

- assemble in periodically waterlogged habitats? **PLOS ONE**, v. 13, n. 5, p. e0198130, 29 maio 2018.
- MACARTHUR, R. H. Patterns of species diversity. **Biological Reviews**, v. 40, p. 510–533, 1965.
- MACARTHUR, R. H.; WILSON, E. O. **The theory of island biogeography**. [s.l.] Princeton University Press, 1967.
- MAGURRAN, A. E. **Measuring biological diversity**. [s.l.] John Wiley & Sons, 2004.
- MAGURRAN, A. E.; MCGILL, B. J. **Biological diversity: frontiers in measurement and assessment**. [s.l.] Oxford University Press, 2011.
- MALENOVSKÝ, Z. et al. Sentinels for science : Potential of Sentinel-1 , -2 , and -3 missions for scientific observations of ocean , cryosphere , and land. **Remote Sensing of Environment**, v. 120, p. 91–101, 2012.
- MANION, G. et al. **gdm: Generalized Dissimilarity Modeling (R package version 1.3. 1)**, 2017.
- MCGILL, B. J. Matters of Scale. **Science**, v. 328, n. 5978, p. 575–576, 2010.
- MELACK, J. M.; HESS, L. L. Remote sensing of the distribution and extent of wetlands in the Amazon basin. In: JUNK, W. J. et al. (Eds.). . **Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management**. 1. ed. Dordrecht Heidelberg London New York: Springer, 2010. v. 210p. 43–59.
- MEYER, L.; DINIZ-FILHO, J. A. F.; LOHMANN, L. G. A comparison of hull methods for estimating species ranges and richness maps. **Plant Ecology & Diversity**, v. 10, n. 5–6, p. 389–401, 2 nov. 2017.
- MISIEWICZ, T. M.; FINE, P. V. A. Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subseratum* (Burseraceae). **Molecular Ecology**, v. 23, n. 10, p. 2543–2558, 2014.
- MONTERO, J. C.; PIEDADE, M. T. F.; WITTMANN, F. Floristic variation across 600 km of inundation forests (Igapó) along the Negro River, Central Amazonia. **Hydrobiologia**, v. 729, n. 1, p. 1–18, dez. 2012.
- MOREIRA, M. D. E. S. et al. Occurrence of nodulation in legume species in the Amazon region of Brazil. **New Phytologist**, v. 121, p. 563–570, 1992.
- MORI, S. A. et al. Evolution of Lecythidaceae with an emphasis on the circumscription of neotropical genera: Information from combined ndhF and trnL-F sequence data. **American Journal of Botany**, v. 94, n. 3, p. 289–301, 2007.
- MORI, S. A. et al. **The Lecythidaceae Pages**.
- MORI, S. A. et al. Toward a phylogenetic-based generic classification of neotropical Lecythidaceae—II. Status of *Allantoma*, *Cariniana*, *Couratari*, *Couroupita*, *Grias* and *Gustavia*. **Phytotaxa**, v. 203, n. 2, p. 122–137, 2015.
- MORI, S. A.; BECKER, P. Flooding Affects Survival of Lecythidaceae in Terra Firme Forest Near Manaus , Brazil. **Biotropica**, v. 23, n. 1, p. 87–90, 1991.
- MORI, S. A.; LEPSCH-CUNHA, N. **Lecythidaceae of a Central Amazonian moist forest**. 1. ed. New York: The New York Botanical Garden, 1995. v. 75
- MORI, S. A.; PRANCE, G. T.; DE ZEEUW, C. H. **Lecythidaceae: Part II. The Zygomorphic-flowered New World Lecythidaceae (Couroupita, Corythophora, Bertholletia, Couratari, Eschweilera, and Lecythis), With a Study of Secondary Xylem of Neotropical Lecythidaceae**. [s.l.] Flora Neotropica, 1990. v. 21

- MORIN, X.; LECHOWICZ, M. J. Niche breadth and range area in North American trees. **Ecography**, v. 36, n. 3, p. 300–312, 2013.
- MORTON, D. C. et al. Amazon forests maintain consistent canopy structure and greenness during the dry season. **Nature**, v. 506, n. 7487, p. 221–224, 2014.
- MORUETA-HOLME, N. et al. A network approach for inferring species associations from co-occurrence data. **Ecography**, v. 39, n. 12, p. 1139–1150, 2016.
- MOULATLET, G. et al. Local Hydrological Conditions Explain Floristic Composition in Lowland Amazonian Forests. **Biotropica**, v. 46, n. 4, p. 395–403, 2014.
- MOUQUET, N. et al. Ecophylogenetics: advances and perspectives. **Biological Reviews**, v. 87, n. 4, p. 769–785, 2012.
- NAZARENO, A. G.; DICK, C. W.; LOHMANN, L. G. Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. **Molecular Ecology**, v. 26, n. 14, p. 3636–3648, abr. 2017.
- NEWSTROM, L. E. et al. Diversity of long-term flowering patterns. In: MCDADE, L. A. et al. (Eds.). **La Selva: ecology and natural history of a neotropical rain forest**. Chicago: University of Chicago Press, 1994. p. 142–160.
- OKSANEN, J. et al. **Package ‘vegan’. R Packag. ver. 2.0--8 254**, 2013.
- OLIVEIRA, U. et al. The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. **Diversity and Distributions**, v. 22, n. 12, p. 1232–1244, 2016.
- OLSON, D. M. et al. Terrestrial Ecoregions of the World: A New Map of Life on Earth. **BioScience**, v. 51, n. 11, p. 933, 2001.
- PAROLIN, P. et al. Central Amazonian floodplain forests: tree adaptations in a pulsing system. **The Botanical Review**, v. 70, n. 3, p. 357–380, 2004.
- PAROLIN, P. et al. Drought responses of flood-tolerant trees in Amazonian floodplains. **Annals of Botany**, v. 105, n. 1, p. 129–139, 1 jan. 2010.
- PAROLIN, P.; AL., E. Review of tree phenology in central amazonian floodplains. **Pesquisas. Botanica**, v. 52, p. 195–222, 2002.
- PAROLIN, P.; FERREIRA, L. V.; JUNK, W. J. Germination characteristics and establishment of trees from central Amazonian flood plains. **Tropical Ecology**, v. 44, n. 2, p. 155–167, 2003.
- PATEIRO-LÓPEZ, B.; RODRÍGUEZ-CASAL, A. Generalizing the Convex Hull of a Sample : The R Package Alphahull. **Journal of Statistical software**, v. 34, n. 5, p. 1–28, 2010.
- PINHEIRO, F.; DANTAS-QUEIROZ, M. V.; PALMA-SILVA, C. Plant Species Complexes as Models to Understand Speciation and Evolution: A Review of South American Studies. **Critical Reviews in Plant Sciences**, v. 0, n. 0, p. 1–27, 2018.
- PITMAN, N. C. A. et al. Dominance and distribution of tree species in upper Amazonian terra firme forests. **Ecology**, v. 82, n. 8, p. 2101–2117, 2001.
- PITMAN, N. C. A. et al. Distribution and abundance of tree species in swamp forests of Amazonian Ecuador. **Ecography**, v. 37, n. 9, p. 902–915, 2014.
- PRANCE, G. T. Notes on the vegetation of Amazonia III. The terminology of Amazonian forest types subject to inundation. **Brittonia**, v. 31, n. 1, p. 26–38, 1979.
- PRANCE, G. T.; MORI, S. A. **Lecythidaceae: Part I: The Actinomorphic-Flowered New World**

Lecythidaceae (Asteranthos, Gustavia, Grias, Allantoma, & Cariniana). [s.l.] Flora Neotropica, 1979. v. 21

PRATA, E. M. B. et al. Towards integrative taxonomy in Neotropical botany : disentangling the *Pagamea guianensis* species complex (Rubiaceae). **Botanical Journal of the Linnean Society**, v. 188, p. 213–231, 2018.

PRESTON, F. W. The Commonness, And Rarity, of Species. **Ecology**, v. 29, n. 3, p. 254–283, 1948.

PUPIM, F. N. et al. Chronology of Terra Firme formation in Amazonian lowlands reveals a dynamic Quaternary landscape. **Quaternary Science Reviews**, v. 210, p. 154–163, 2019.

R CORE TEAM. **R: A Language and Environment for Statistical Computing** Vienna, Austria, 2018.

RAPACCIUOLO, G.; BLOIS, J. L. Understanding ecological change across large spatial, temporal, and taxonomic scales: integrating data and methods in light of theory. **Ecography**, 16 abr. 2019.

RESENDE, A. F. DE et al. Massive tree mortality from flood pulse disturbances in Amazonian floodplain forests: The collateral effects of hydropower production. **Science of The Total Environment**, v. 659, p. 587–598, abr. 2019.

REVELL, L. J. phytools: An R package for phylogenetic comparative biology (and other things). **Methods in Ecology and Evolution**, v. 3, n. 2, p. 217–223, 2012.

RICKLEFS, R. E.; HE, F. Region effects influence local tree species diversity. **Proceedings of the National Academy of Sciences**, v. 113, n. 3, p. 674–679, 2016.

RICKLEFS, R. E.; SCHLUTER, D. **Species diversity in ecological communities: historical and geographical perspectives**. 1. ed. Chicago: University of Chicago Press, 1993.

ROCCHINI, D. et al. Measuring β -diversity by remote sensing: a challenge for biodiversity monitoring. **Methods in Ecology and Evolution**, v. 9, p. 1787–1798, 2018.

ROCHA, D. G. DA; KAEFER, I. L. What has become of the refugia hypothesis to explain biological diversity in Amazonia? **Ecology and Evolution**, n. November 2018, p. ece3.5051, 2019.

ROSAUER, D. F. et al. Phylogenetic generalised dissimilarity modelling : a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. **Ecography**, v. 37, n. 1, p. 21–32, 2014.

ROSENZWEIG, M. L. **Species diversity in space and time**. 1. ed. Cambridge: Cambridge University Press, 1995.

RULL, V. Origins of biodiversity. **Science**, v. 331, n. 6016, p. 398–399, 2011a.

RULL, V. Neotropical biodiversity: timing and potential drivers. **Trends in Ecology and Evolution**, v. 26, n. 10, p. 508–513, 2011b.

SALO, J. et al. River dynamics and the diversity of Amazon lowland forest. **Nature**, v. 322, n. 6076, p. 254–258, 1986.

SCHIETTI, J. et al. Vertical distance from drainage drives floristic composition changes in an Amazonian rainforest. **Plant Ecology & Diversity**, v. 7, n. 1–2, p. 241–253, 3 abr. 2014.

SEXTON, J. P. et al. Evolution and Ecology of Species Range Limits. **Annual Review of Ecology, Evolution, and Systematics**, v. 40, n. 1, p. 415–436, 2009.

SEXTON, J. P. et al. Evolution of Ecological Niche Breadth. **Annual Review of Ecology, Evolution, and Systematics**, v. 48, n. 1, p. 183–206, 2 nov. 2017.

- SFENTHOURAKIS, S.; TZANATOS, E.; GIOKAS, S. Species co-occurrence : the case of congeneric species and a causal approach to patterns of species association. **Global Ecology and Biogeography**, v. 15, p. 39–49, 2005.
- SHMIDA, A.; WILSON, M. V. Biological determinants of species diversity. **Journal of Biogeography**, v. 12, n. 1, p. 1–20, 1985.
- SILVA, S. M. et al. A dynamic continental moisture gradient drove Amazonian bird diversification. **Science Advances**, v. 5, p. eaat5752, 2019.
- SILVERTOWN, J. et al. Hydrologically defined niches reveal a basis for species richness in plant communities. **Nature**, v. 400, n. July, p. 61–63, 1999.
- SILVERTOWN, J. Plant coexistence and the niche. **Trends in Ecology and Evolution**, v. 19, n. 11, p. 605–611, 2004.
- SILVERTOWN, J.; ARAYA, Y.; GOWING, D. Hydrological niches in terrestrial plant communities: A review. **Journal of Ecology**, v. 103, n. 1, p. 93–108, 2015.
- SIMON, M. F. et al. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. **Proceedings of the National Academy of Sciences**, v. 106, n. 48, p. 20359–20364, 2009.
- SLATYER, R. A.; HIRST, M.; SEXTON, J. P. Niche breadth predicts geographical range size: a general ecological pattern. **Ecology Letters**, v. 16, n. 8, p. 1104–1114, ago. 2013.
- SLIK, J. W. F. et al. An estimate of the number of tropical tree species. **Proceedings of the National Academy of Sciences**, p. 201423147, 1 jun. 2015.
- SMITH, S. A.; BROWN, J. W. Constructing a broadly inclusive seed plant phylogeny. **American Journal of Botany**, v. 105, n. 3, p. 302–314, 2018.
- SRIVASTAVA, D. S. Using Local-Regional Richness Plots to Test for Species Saturation- Pitfalls and Potentials. **Journal of Animal Ecology**, v. 68, p. 1–16, 1999.
- STEPHENSON, N. L. Climatic control of vegetation distribution: the role of the water balance. **The American Naturalist**, v. 135, n. 5, p. 649–670, 1990.
- STROPP, J. et al. Tree communities of white-sand and terra-firme forests of the upper Rio Negro. **Acta Amazonica**, v. 41, n. 4, p. 521–544, 2011.
- TAKAKU, J. et al. VALIDATION OF ‘ AW3D ’ GLOBAL DSM GENERATED FROM ALOS PRISM. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. III-4, n. XXIII ISPRS Congress, p. 25–31, 2016.
- TER STEEGE, H. et al. An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana Shield. **Journal of Tropical Ecology**, v. 16, n. 6, p. 801–828, 2000.
- TER STEEGE, H. et al. A spatial model of tree α -diversity and tree density for the Amazon. **Biodiversity and Conservation**, v. 12, n. 11, p. 2255–2277, 2003.
- TER STEEGE, H. et al. Hyperdominance in the Amazonian tree flora. **Science**, v. 342, n. 6156, 2013.
- TER STEEGE, H. et al. Estimating the global conservation status of more than 15,000 Amazonian tree species. **Science Advances**, v. 1, n. 10, 2015.
- TER STEEGE, H. et al. The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. **Scientific Reports**, v. 6, p. 29549, 2016.
- TERBORGH, J.; ANDRESEN, E. The composition of Amazonian forests: patterns at local and

regional scales. **Journal of Tropical Ecology**, v. 14, n. 5, p. 645–664, 1998.

THEOBALD, D. M.; HARRISON-ATLAS, D.; MONAHAN, W. B. Ecologically-Relevant Maps of Landforms and Physiographic Diversity for Climate Adaptation Planning. **PloS one**, v. 10, n. 12, p. e0143619, 2015.

THOMSON, A. M.; VARGAS, O. M.; DICK, C. W. Complete plastome sequences from *Bertholletia excelsa* and 23 related species yield informative markers for Lecythidaceae. **Applications in Plant Sciences**, v. 6, n. 5, p. e01151, maio 2018.

TÖPEL, M. et al. SpeciesGeoCoder: Fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. **Systematic Biology**, v. 66, n. 2, p. 145–151, 2 ago. 2017.

TRABUCCO, A.; ZOMER, R. J. Global soil water balance geospatial database. **CGIAR Consortium for Spatial Information. Published online, available from the CGIAR-CSI GeoPortal at: <http://www.cgiar-csi.org>**, 2010.

TURCHETTO-ZOLET, A. C. et al. Phylogeographical patterns shed light on evolutionary process in South America. **Molecular Ecology**, v. 22, n. 5, p. 1193–1213, 2013.

ULLOA ULLOA, C. et al. An integrated assessment of the vascular plant species of the Americas. **Science**, v. 358, n. 6370, p. 1614–1617, 22 dez. 2017.

ULRICH, W. Species co-occurrences and neutral models : reassessing J . M . Diamond ’ s assembly rules. **Oikos**, v. 107, p. 603–609, 2004.

ULRICH, W.; GOTELLI, N. J. Null model analysis of species associations using abundance data Null model analysis of species associations abundance using. **Ecology**, v. 91, n. 11, p. 3384–3397, 2010.

VEECH, J. A. A probabilistic model for analysing species co-occurrence. **Global Ecology and Biogeography**, v. 22, n. 2, p. 252–260, 2013.

VEECH, J. A. The pairwise approach to analysing species co-occurrence. **Journal of Biogeography**, v. 41, n. 6, p. 1029–1035, 2014.

VENTICINQUE, E. et al. An explicit GIS-based river basin framework for aquatic ecosystem conservation in the Amazon. **Earth System Science Data**, v. 8, n. 2, p. 651–661, 2016.

VILELA, B.; VILLALOBOS, F. letsR: a new R package for data handling and analysis in macroecology. **Methods in Ecology and Evolution**, v. 6, p. 1229–1234, 18 maio 2015.

VON HUMBOLDT, A.; BONPLAND, A. **Essay on the geography of plants—with a physical tableau of the equinoctial regions**. 1. ed. [s.l: s.n.].

WALLACE, A. R. On the Monkeys of the Amazon. **Journal of Natural History Series**, v. 2, n. 14:84, p. 451–454, 1854.

WEBB, C. O. et al. Phylogenies and Community Ecology. **Annual Review of Ecology and Systematics**, v. 33, n. 1, p. 475–505, nov. 2002.

WEIHER, E.; KEDDY, P. **Ecological assembly rules: perspectives, advances, retreats**. [s.l.] Cambridge University Press, 2001.

WHITTAKER, R. H. Vegetation of the Siskiyou Mountains, Oregon and California. **Ecological Monographs**, v. 30, n. 3, p. 279–338, 1960.

WHITTAKER, R. H. Dominance and Diversity in Land Plant Communities. **Science**, v. 147, n. 3655, p. 250–260, 1965.

WIENS, J. J. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. **Evolution**, v. 58, n. 1, p. 193–197, 2004.

WILSON, E. O. **Biodiversity**. Washington, DC: The National Academies Press, 1988.

WING, S. L. et al. Late Paleocene fossils from the Cerrejón Formation, Colombia, are the earliest record of Neotropical rainforest. **Proceedings of the National Academy of Sciences**, v. 106, n. 44, p. 18627–18632, 2009.

WITTMANN, F. et al. Tree species composition and diversity gradients in white-water forests across the Amazon Basin. **Journal of Biogeography**, v. 33, n. 8, p. 1334–1347, ago. 2006.

WITTMANN, F. Tree species composition and diversity in Brazilian freshwater floodplains. In: PAGANO, M. (Ed.). **Mycorrhiza: Occurrence in Natural and Restored Environments**. 1. ed. [s.l.] Nova Science Publishers, Inc., 2012. p. 223–264.

WITTMANN, F. et al. Habitat specificity, endemism and the neotropical distribution of Amazonian white-water floodplain trees. **Ecography**, v. 36, n. 6, p. 690–707, jun. 2013.

WITTMANN, F. et al. The Brazilian freshwater wetlandscape: Changes in tree community diversity and composition on climatic and geographic gradients. **PLoS ONE**, v. 12, n. 4, p. 1–18, 2017.

WITTMANN, F.; ANHUF, D.; FUNK, W. J. Tree species distribution and community structure of central Amazonian várzea forests by remote-sensing techniques. **Journal of Tropical Ecology**, v. 18, n. 6, p. 805–820, nov. 2002.

WITTMANN, F.; JUNK, W. J. Sapling communities in Amazonian white-water forests. **Journal of Biogeography**, v. 30, n. 10, p. 1533–1544, 2003.

WITTMANN, F.; JUNK, W. J.; PIEDADE, M. T. F. The várzea forests in Amazonia: Flooding and the highly dynamic geomorphology interact with natural forest succession. **Forest Ecology and Management**, v. 196, n. 2–3, p. 199–212, 2004.

WITTMANN, F.; SCHÖNGART, J.; JUNK, W. J. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplain forests. In: JUNK, W. J. et al. (Eds.). **Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management**. 1. ed. Dordrecht Heidelberg London New York: Springer, 2010. p. 61–102.

WWF. Amazon Alive ! **Documentos**, p. 60, 2009.

YULE, G. U. II.—A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FR S. **Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character**, v. 213, n. 402–410, p. 21–87, 1925.

ZOBEL, M. The relative role of species pools in determining plant species richness: An alternative explanation of species coexistence? **Trends in Ecology and Evolution**, v. 12, n. 7, p. 266–269, 1997.

ZOBEL, M. The species pool concept as a framework for studying patterns of plant diversity. **Journal of Vegetation Science**, v. 27, n. 1, p. 8–18, 2016.