







Biogeographic history and habitat specialization shape floristic and phylogenetic composition across Amazonian forests

CHRISTOPHER BARALOTO ^{1,2,12} JASON VLEMINCKX ¹ JULIEN ENGEL,³ PASCAL PETRONELLI,⁴ NÁLLARETT DÁVILA,⁵ MARCOS RÍOS,⁵ ELVIS HARRY VALDERRAMA SANDOVAL,⁶ ITALO MESONES,⁷ JUAN ERNESTO GUEVARA ANDINO ⁸, CLAIRE FORTUNEL ³, ELODIE ALLIE,² C. E. TIMOTHY PAINE ⁹, AURÉLIE DOURDAIN,⁴ JEAN-YVES GORET,² OSCAR J. VALVERDE-BARRANTES,¹ FREDDIE DRAPER ^{1,10,11} AND PAUL V. A. FINE⁷

¹*Institute of Environment, Department of Biological Sciences, Florida International University, 11200 Southwest 8th Street, Miami, Florida 33199 USA*

²*INRAE, UMR Ecologie des Forêts de Guyane, Université de Guyane, Université des Antilles, Campus agronomique, BP 316, Kourou Cedex 97379 France*

³*AMAP (botAnique et Modélisation de l'Architecture des Plantes et des végétations), Université de Montpellier, CIRAD, CNRS, INRAE, IRD, Boulevard de la Lironde, Montpellier Cedex 5 TA A-51/PS234398 France*

⁴*CIRAD, UMR Ecologie des Forêts de Guyane, Université de Guyane, Université des Antilles, Campus agronomique, BP 316, Kourou Cedex 97379 France*

⁵*Instituto de Investigaciones de la Amazonia Peruana, Iquitos, Peru, Avenida José A. Quiñones km 2.5, Iquitos, Loreto Perú*

⁶*Facultad de Ciencias Biológicas—Universidad Nacional de la Amazonia Peruana, Pevas 5ta Cdra, Iquitos, Loreto Perú*

⁷*Department of Integrative Biology and Jepson Herbaria, University of California, Berkeley, 3040 Valley Life Sciences Building 3140, Berkeley, California 94720-3140 USA*

⁸*Ciencias Ambientales, Universidad de las Americas, Quito, Ecuador*

⁹*Environmental and Rural Sciences, University of New England, Armidale, New South Wales 2351 Australia*

¹⁰*Center for Global Discovery and Conservation Science, Arizona State University, 1001 South McAllister Avenue Tempe, Tempe, Arizona 85287 USA*

¹¹*School of Geography, University of Leeds, Woodhouse, Leeds LS2 9JT UK*

Citation: Baraloto, C., J. Vleminckx, J. Engel, P. Petronelli, N. Dávila, M. Ríos, E. H. Valderrama Sandoval, I. Mesones, J. E. Guevara Andino, C. Fortunel, E. Allie, C. E. T. Paine, A. Dourdain, J.-Y. Goret, O. J. Valverde-Barrantes, F. Draper, and P. V. A. Fine. 2021. Biogeographic history and habitat specialization shape floristic and phylogenetic composition across Amazonian forests. *Ecological Monographs* 00(00):e01473. 10.1002/ecm.1473

Abstract. A major challenge remains to understand the relative contributions of history, dispersal, and environmental filtering to the assembly of hyperdiverse communities across spatial scales. Here, we examine the extent to which biogeographical history and habitat specialization have generated turnover among and within lineages of Amazonian trees across broad geographic and environmental gradients. We replicated standardized tree inventories in 102 0.1-ha plots located in two distant regions—the western Amazon and the eastern Guiana shield. Within each region, we used a nested design to replicate plots on contrasted habitats: white-sand, terra firme, and seasonally flooded forests. Our plot network encompassed 26,386 trees that together represented 2,745 distinct taxa, which we standardized across all plots and regions. We combined taxonomic and phylogenetic data with detailed soil measurements and climatic data to: (1) test whether patterns of taxonomic and phylogenetic composition are consistent with recent or historical processes, (2) disentangle the relative effects of habitat, environment, and geographic distance on taxonomic and phylogenetic turnover among plots, and (3) contrast the proportion of habitat specialists among species from each region. We found substantial species turnover between Peru and French Guiana, with only 8.8% of species shared across regions; genus composition remained differentiated across habitats and regions, whereas turnover at higher taxonomic levels (family, order) was much lower. Species turnover across plots was explained primarily by regions, but also substantially by habitat differences and to a lesser extent by spatial distance within regions. Conversely, the composition of higher taxonomic levels was better explained by habitats (especially comparing white-sand forests to other habitats) than spatial distance. White-sand forests harbored most of the habitat specialists in both regions, with stronger habitat specialization in Peru than in French Guiana. Our results suggest that recent diversification events have resulted in extremely high turnover in species and genus composition with relatively little change in the composition of higher lineages. Our results also emphasize the contributions of rare habitats, such as white-sand forests, to the

Manuscript received 26 February 2021; accepted 16 April 2021; final version received 23 June 2021. Corresponding Editor: Walter P. Carson.

¹²E-mail: cbaralot@fiu.edu

extraordinary diversity of the Amazon and underline their importance as conservation priorities.

Key words: beta diversity; climate; community assembly; flooded forest; forest structure; French Guiana; Peru; soil properties; tropical rainforest; white-sand forest.

INTRODUCTION

The Amazon basin is home to the highest concentration of tree diversity in the world, boasting at least 16,000 species in its lowland rain forests alone (ter Steege et al. 2020). Gentry (1981, 1986) hypothesized that Amazonia's extraordinarily high species totals were related in large part to the substantial environmental heterogeneity in the region causing selection for habitat specialization by plants, contributing to high beta diversity across space. Indeed, many studies since have demonstrated significant correlations between tree species composition and environmental variables including climate (especially seasonality), soil resource availability, and seasonal flooding (Wittman et al. 2004, Fine and Kembel 2011, Baldeck et al. 2013, Van Breugel et al. 2019). Nevertheless, it remains unclear which environmental gradients cause the strongest beta-diversity patterns and how dispersal limitation interacts with environmental variation to shape floristic composition across the region.

Amazonia includes two major large-scale environmental gradients involving precipitation and soil fertility. Northwestern Amazonia is covered by a seasonal ever-wet rainforest, whereas eastern and southern Amazonia experience dry seasons of 2–4 months (Terborgh and Andresen 1998, ter Steege et al. 2006). The uplift of the Andes and the erosion of newly deposited volcanic sediments created a large area of relatively fertile clay soils in the western Amazon, which contrasts with the ancient and highly eroded lateritic clays of the central and eastern Amazon basin and Guianan shields (Hoorn et al. 2010). Seasonal Andean snowmelt causes large fluctuation in the water levels of its large rivers, with more significant duration of flooding occurring towards the West, whereas the Guianan shield and eastern Amazonia have more predictable and shorter-duration seasonal flooding (Wittman et al. 2010).

On top of these environmental filters, the mode and average dispersal distance of tree taxa also influence the diversity and composition of forests (Hubbell 2001, Condit et al. 2013, Dexter et al. 2017). Some tree taxa are excellent dispersers and may be able to maintain gene flow across enormous geographic ranges, spanning the Amazon basin (Dick et al. 2003). Others may be poor dispersers, which could increase the probability that a species becomes a geographic or habitat endemic, and thereby increase beta diversity across regions.

Considering community assembly over larger time scales, there may be historical reasons why certain lineages have prospered in certain regions and habitats and not others. One simple contributing factor may be the

relative age of different taxa. Some of the most diverse lineages of Amazonian trees (e.g., *Inga* (Fabaceae)) are thought to have radiated within the last few million years (Richardson et al. 2001, Nichols et al. 2015), whereas others, like *Protium* (Burseraceae) and *Eschweilera*, have a longer history in the region (Fine et al. 2014, Vargas and Dick 2020). In addition, some lineages could have adapted traits that promote success to one habitat type, and that habitat may have increased or decreased in land area over the past millennia (Fine and Baraloto 2016). Finally, major biogeographical events, like Andean uplift and its influence on the hydrology of the basin, or Pleistocene climate cycles, had major effects on tree composition and diversity, and the signal of these geologically recent events can still be detected in contemporary forests (Antonelli et al. 2009).

One way to investigate the relative importance of historical and recent processes in assembling plant communities involves the comparison of taxonomic (species level) with phylogenetic community structure patterns across spatial scales (Graham and Fine 2008, Fine and Kembel 2011, Kraft et al. 2011, Baraloto et al. 2012, McFadden et al. 2019). For example, if ancient events such as the appearance and disappearance of major geographic barriers (e.g., Lake Pevas dividing western and central Amazonia [Hoorn et al. 2010]) have been important in assembling these communities, we would expect strong phylogenetic structure among regions, with coexisting species intermediately related within regions due to a limited number of older lineages, each with many young species (Gerhold et al. 2018). On the other hand, if trans-Amazon dispersal has been common throughout the past several million years, we would expect strong taxonomic turnover but little phylogenetic structure among regions (Fine and Kembel 2011, Dexter et al. 2017). At the same time, if local niche-related processes such as environmental filtering and negative density dependence are related to recent diversification events within lineages, then we would predict higher taxonomic but lower phylogenetic structure, due to convergent adaptations within and among lineages (Fine and Baraloto 2016).

Our understanding of the local and regional diversity patterns of Amazonian tree communities has greatly improved with the publication of studies linking hundreds of 1-hectare plots. ter Steege et al. (2013) reported that half of all individual trees represented by 1,200 plots across the basin were represented by 227 hyperdominant trees that were generally habitat specialists restricted to one or two Amazonian subregions. Yet, even though these plot network studies have given us an unprecedented picture of Amazonian tree

diversity patterns, they have some serious limitations. First, the plots were set up haphazardly across the landscape, usually in areas that were the most accessible for botanists (Hopkins 2007). Thus, reported patterns of habitat specialization become difficult to separate from patterns of geographic endemism because habitats in any given subregion are usually spatially autocorrelated. Second, taxonomic standardization has not been attempted in any of these cross-Amazonian studies. Many of the Amazon Tree Diversity Network (ATDN) plots do not have vouchered collections; and even when they do, no one has checked to make sure that species identifications conform to a common standard. Thus, even the common species (the hyperdominants) could be misidentified in some cases, because so many of them belong to species complexes that are often lumped together into one species name (Damasco et al. 2019, Baker et al 2017).

The broad scale differences in seasonality and soil fertility between western and eastern Amazonia are marked, yet within each region there exists substantial environmental heterogeneity, consisting of a mosaic of contrasted edaphic habitats such as white-sand deposits, clay-rich terra firme forests, and seasonally flooded forests adjacent to large rivers at local scales that are replicated across the landscape (Hoorn et al. 2010). Both seasonally flooded and white-sand forests have been shown to host a unique assemblage of endemic species with suites of functional traits (Fine and Baraloto 2016) adapted to the particular challenges of each habitat (e.g., nutrient availability, flooding, etc. [Wittman et al. 2004, Fortunel et al. 2014, 2016]).

Here we investigate the assembly of Amazonian tree communities from local to regional spatial scales, using a balanced sampling design with tree species inventories carried out in 102 0.1-ha plots across two regions located at extreme geographic ends of Amazonia, one in the Peruvian Amazon and the other in French Guiana, 2,500 km to the northeast. This sampling was designed to cover the most representative habitats in both regions (white-sand, terra firme, and seasonally flooded forests) and thereby to disentangle the relative influences of geographic distance and habitat heterogeneity on the taxonomic and phylogenetic composition of tree communities. All tree species were vouchered in each plot where they occurred, and all identifications were standardized across the entire data set by the same botanists using the largest reference collection of Neotropical tree specimens, the Missouri Botanical Garden (MOBOT). We address the following questions:

- 1) How different are eastern and western Amazonian forests; and to what extent do taxonomic and phylogenetic turnover differ across Amazonia?
- 2) What is the relative strength of environmental factors (climate, soil, and flooding) and spatial distance (both regional and local) in explaining floristic and phylogenetic differences among tree communities?

- 3) How strongly are species associated with different habitats (white-sand, terra firme, and seasonally flooded forests)? Is habitat specialization more frequent in less seasonal forests with a steeper gradient of soil resource availability (Peru vs. French Guiana)?

MATERIALS AND METHODS

Study areas

We established a nested experimental design with replicated plots in habitats displaying contrasting soil conditions characteristic of lowland Amazonian forests—white-sand (WS), terra firme (TF), and seasonally flooded forests (SF) (Baraloto et al. 2011, Fortunel et al. 2014) —at both regional (c.100 km) and basin-wide (2,500 km) distances. A total of 102 0.1-ha plots were inventoried between 2008 and 2018 in 10 subregions in French Guiana (hereafter FG; 64 plots) and between 2008 and 2011 in three subregions in Peru (38 plots) (Fig. 1). Each plot was inventoried once, with subregions visited during different field missions within the mentioned period. We tried to maintain at least 50 km between subregions, and at least 500 m between plots.

French Guianan forests stand on a Precambrian tableland, with old, highly weathered and nutrient-depleted soils (Gourlet-Fleury et al. 2004). Mean annual rainfall across inventory subregions ranges is between 2,160 and 3,130 mm,¹³ and is distributed seasonally throughout the year (Table 1). The wet season stretches from December to July, and it is usually interrupted in February or March by a short dry period; whereas the dry season occurs from August to November with monthly rainfall never exceeding 100 mm. Mean daily temperatures oscillate between 23.0°C and 26.6°C, with low seasonal variation (Gourlet-Fleury et al. 2004). Elevation among subregions ranged from 42 to 529 m.

Western Amazonian forests in Peru occur on a more heterogeneous series of substrates because of the Andean uplift and the concomitant erosion of volcanic sediments and marine incursions (Hoorn et al. 2010). Climate conditions are less variable during the year. Mean annual rainfall across inventory subregions ranges between 2,405 and 2,750 mm¹⁴ and is less seasonal than in French Guiana (Table 1). Mean temperature is more stable between 26.3°C and 26.7°C with low seasonal variation. Elevation among subregions was also much less variable, from 95 to 173 m. Further details on the climate and geology of the regions and subregions are provided in Appendix S1.

¹³<http://www.worldclim.com/>

¹⁴<http://www.worldclim.com/>

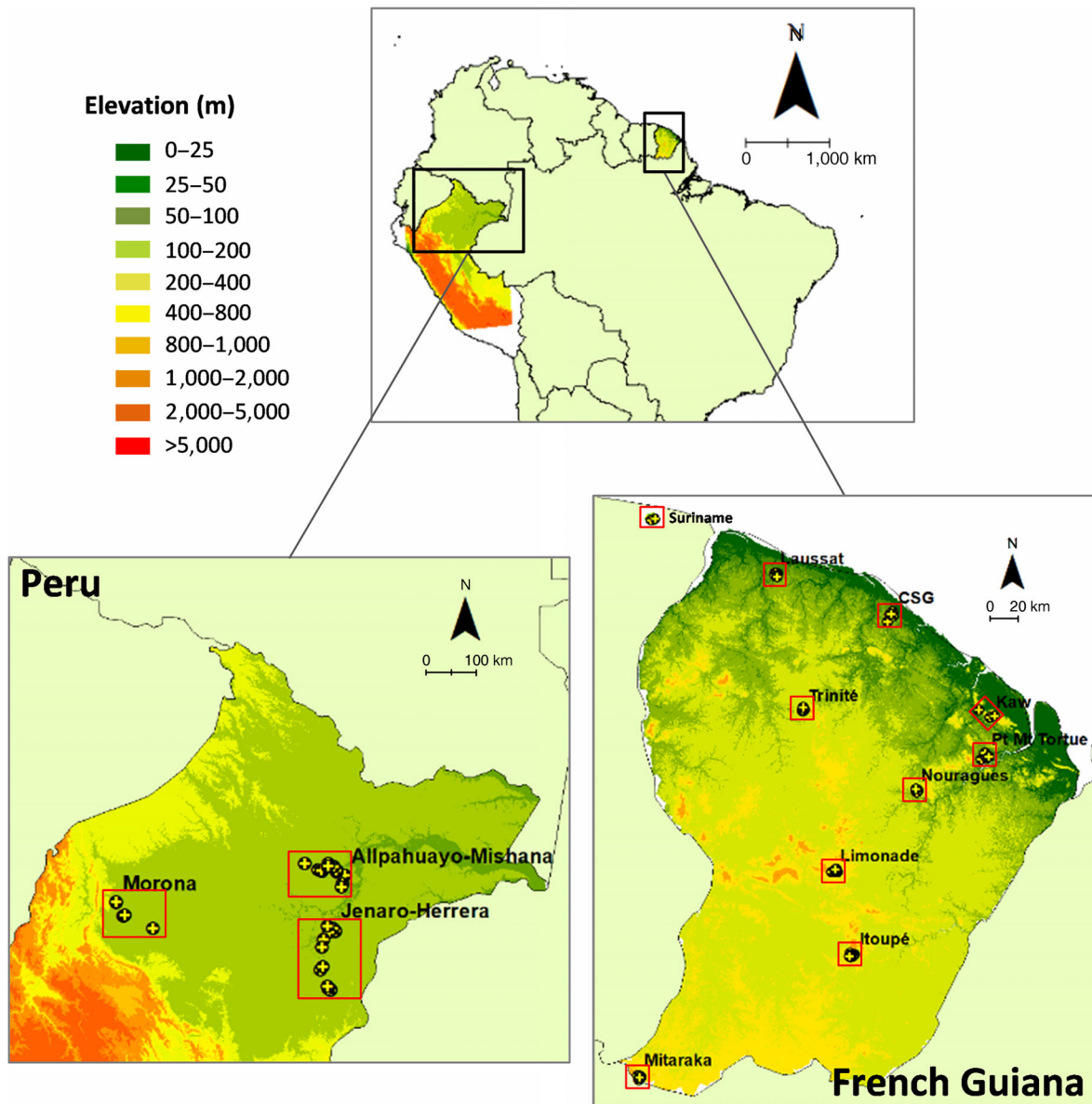


FIG. 1. Geographical distribution of the 13 subregions in the Amazon (red rectangles), showing the position of the plots (yellow cross symbols) within the two study regions (Peru and French Guiana).

Tree species inventories

Trees were inventoried following a modified version of the Gentry plots proposed by Phillips et al. (2003) and described in Baraloto et al. (2013). Each plot consisted of 10 parallel 50-m-long transects departing perpendicularly from a main 190-m-long central line, successively in alternate directions every 20 m along the line (a schematic illustration of a plot is provided in Appendix S2). All stems with a circumference ≥ 8 cm at 1.3 m above the ground (c. 2.5 cm DBH) were inventoried over a 2-m width along each transect. At least one individual of every putatively distinct taxon encountered was collected in the field to create plot-level herbarium vouchers. In

rare cases (0.2% of all stems sampled), no identification was made, nor could vouchers be collected, because of lack of leaves or obstructed canopies. Further sorting resulted in standardized project type collections for all distinct taxa, which were identified at regional herbaria for the Peru (AMAZ) and FG (CAY) collections. We then further standardized and resolved vouchers from both these collections during a 2-month period at the herbarium of the Missouri Botanical Garden (MOBOT), such that any unnamed, putative novel species could be compared to other congeners from the other region. At the end, we provide a full detail of all project vouchers describing our standardized inventories (see Baraloto et al. 2021 for complete voucher list).

TABLE 1. Sampling and geographical characteristics of each subregion.

	N	Alt	MAP	MAT	S	ENS ₂
Guianas	64	42–529	2,157–3,129	23.0–26.6	81.2 (3.5)	49.5 (5.2)
Saül-Limonade	12	196–253	2,421	24.6	66.1 (4.2)	24.4 (4.5)
Trinité	6	126–320	2,671	25.1	112.8 (12.0)	78.0 (14.3)
Itoupé	3	521–529	2,530	23.0	79.7 (6.9)	30.1 (7.0)
Mitaraka	9	317–347	2,157	24.9	70.9 (10.6)	66.1 (17.2)
Laussat	10	49–57	2,402	26.2	73.1 (3.4)	28.3 (3.6)
Nouragues	8	108–345	3,328	24.8	86.6 (14.0)	67.9 (20.4)
Petite Montagne Tortue	9	47–136	3,729	25.4	95.6 (10.6)	66.3 (17.0)
Center spatial Guyanais	4	43–63	2,932	25.8	75.8 (9.1)	25.3 (5.1)
Kaw	2	254–282	3,720	24.5	96.0 (6.0)	74.0 (18.3)
Suriname	2	196–229	2,241	26.6	35.5 (3.5)	9.4 (5.7)
<i>Terra firme</i>	20	45–347	2,775	25.2	96.6 (4.4)	69.7 (7.4)
<i>Seasonally flooded</i>	35	43–529	2,723	24.9	64.4 (5.2)	24.4 (4.2)
<i>White-sand</i>	10	39–345	2,908	25.7	64.6 (5.4)	23.8 (5.3)
Peru	38	95–173	2,405–2,750	26.3–26.7	101.6 (5.00)	61.4 (6.8)
Morona	6	143–173	2,405	26.7	108.3 (5.9)	89.6 (17.2)
North Loreto	18	105–149	2,750	26.3	101.6 (8.7)	60.0 (10.6)
South Loreto	14	95–139	2,499	26.8	98.6 (7.6)	51.0 (9.1)
<i>Terra firme</i>	11	95–158	2,597	26.6	129.1 (5.6)	102.0 (9.6)
<i>Seasonally flooded</i>	13	106–156	2,636	26.5	87.8 (6.1)	53.6 (9.2)
<i>White-sand</i>	14	106–173	2,625	26.6	86.8 (7.9)	29.8 (5.0)

Notes: N = number of plots; Alt = altitudinal range (m); MAP = mean annual precipitations (mm); MAT = mean annual temperature (°C). S = mean number of species per plot; ENS₂ = mean (calculated at the plot level) effective number of species expected from 1,000 random samplings (with replacement) of two individuals. The three last lines of each region represent the same information for each habitat. Numbers in parentheses correspond to the standard deviation of the mean.

Species diversity was characterized in each study sub-region using species richness, as well as the effective number of species expected from a random sample of two individuals, to weight for species abundance (Dauby and Hardy [2011]; Table 1).

environmental (R Development Core Team 2020). The larger number of soil variables (nine) compared to the unique climate variable (DSI) was taken into account by analyzing the relative effect of each variable (see the Data analysis section).

Environmental data

Soil conditions were characterized in each plot using nine physicochemical properties: texture (percentages of sand, silt, and clay), bioavailable cations content (Ca, Mg, and K), available phosphorus content (AP), organic matter (OM), and carbon (OC) contents, total N content (TN), and C:N ratio. Variables were measured from bulked soil cores collected at 0–15-cm depth within each plot. Cores were mixed into a 500-g sample that was dried to constant mass (at 25°C) and sieved (2-mm mesh). Samples were shipped to the University of California, Davis DANR laboratory for physical and chemical analyses (see Baraloto et al. 2011 for full details).

We calculated environmental data including a dry season index (DSI), which was calculated for each plot, as the sum (over 12 months) of the ratios between the mean monthly temperature and the mean monthly rainfall. This provided an estimate of the potential hydric stress accumulated during the dry seasons. Rainfall and temperature data were extracted from worldclim data¹⁵ via the raster package (Hijmans 2018) in R statistical

Data analysis

Taxonomic and phylogenetic heterogeneity.—To investigate Question 1 (*How different are eastern and western Amazonian forests; and to what extent do taxonomic and phylogenetic turnover differ across Amazonia?*), a non-metric multidimensional scaling analysis (NMDS) was performed on taxonomic dissimilarity among plots calculated at the species, genus, family, and order levels. For each level, taxonomic dissimilarity was calculated using Euclidean distance on Hellinger-transformed abundances (Legendre and Gallagher 2001). The latter transformation prevented the floristic composition variation from being overly weighted by the most abundant species. We also considered taxonomic identities only, by calculating the Jaccard dissimilarity on the plot × presence-absence of taxa. As a complementary approach, we computed phylogenetic dissimilarity using metrics that provided relatively more weight to either recent diversification events (the unique fraction of branch length *UniFrac* [Lozupone and Knight 2005]) or to relatively deeper evolutionary differences among taxa (B_{st} and Π_{st} statistics [Hardy and Senterre 2007]). Finally, we used the geographical node divergence (GND) score to

¹⁵<http://www.worldclim.com/>

evaluate the degree of geographical differentiation of (1) each node separating taxa from the family level to (2) the most basal branch separation (Borregaard et al. 2014).

Environmental heterogeneity.—The environmental heterogeneity among plots was decomposed using a principal component analysis, after normalizing (using a Box-Cox transformation) then standardizing (z -score transformation) all environmental variables. We also used a two-way ANOVA to test (1) the mean difference of each environmental variable across habitats and regions, and (2) the interaction between regions and habitats.

Relative influence of habitat vs. spatial distance effects on taxonomic and phylogenetic composition.—To investigate Question 2 (*What is the relative strength of environmental factors [climate, soil and flooding] and spatial distance [both regional and local] in explaining floristic and phylogenetic differences among tree communities?*), we used multiple regression on distance matrices (MRM [Lichstein 2007]). MRM is an adaptation of multivariate regression to distance data, in which dissimilarity matrices are vectorized for each predictor and the response variable. Four different MRM models were performed, each one integrating different combinations of predictors (habitat, environmental, and/or spatial dissimilarity). For each model, we quantified the adjusted coefficient of determination (R^2) corresponding to the effect of all predictors combined, as well as the relative effects of each predictor using partial coefficients of determination (R^2) averaged over orderings using the Lindeman, Merenda and Gold (LMG) statistic of Lindeman et al. (1980). More specifically, the LMG value of a predictor is calculated as the increment in the coefficient of determination (R^2) obtained when adding this predictor in the model, averaged over all possible sequences of ordering among the predictors.

In the first MRM model (Model 1), we calculated the relative effects of habitat, region, and subregion difference on taxonomic and phylogenetic dissimilarity. Differences in habitat, region, and subregion within region were characterized using dummy variable columns containing 0 or 1 (= same or different habitats, regions or subregions). In the second MRM (Model 2), we quantified the relative effects of each habitat comparison (TF-SF, WS-TF, and WS-SF) to test further whether compositional turnover was stronger or weaker between flooded vs. nonflooded habitats compared to white-sand vs. non-white-sand habitats. Model 3 was similar to Model 1, but instead of analyzing habitat difference, we calculated the relative effects of quantitative environmental dissimilarities, corresponding to climatic dissimilarity (in altitude and DSI) and soil dissimilarity (in the nine physico-chemical variables described in Environmental Data) separately. A fourth model (Model 4) tested the relative effects of each environmental variable.

To take spatial autocorrelation into account, the adjusted R^2 values of the four MRM models were tested by comparing their observed values with 4,999 null values obtained after performing Moran spectral randomizations (Wagner and Dray 2015) of each composition data matrix. The MSR is a spatially constrained permutation procedure allowing an explicit consideration of the multiscale spatial autocorrelation structures in any quantitative variable. The MSR uses information on the spatial connectivity among sampling points obtained when selecting Moran's eigenvector maps (MEMs [Dray et al. 2006]), which are commonly used to model multiscale spatial structures in ecological data. Connections among plots were defined using a Gabriel's graph, which has been shown to be appropriate even in the case of a highly irregular sampling design (Bauman et al. 2018a). The selection of the MEMs was optimized following a forward selection procedure (Blanchet et al. 2008) that provides correct Type I error rates for the selection of spatial eigenvectors (Bauman et al. 2018b). In the MSR, the connectivity information among plots, obtained during the selection of MEMs, is then used in a constrained randomization algorithm to reproduce variables that accurately mimic the observed spatial structures of the randomized variable(s) (Wagner and Dray 2015).

Habitat differentiation and proportion of specialist species in each habitat.—Several analyses were performed to address our third question (*How strongly are species associated with different habitats [white-sand, terra firme and seasonally flooded forests]? Is habitat specialization more frequent in less seasonal forests with a steeper gradient of soil resource availability [Peru versus French Guiana]?*). These analyses were performed for each region separately, because of the low percentage of shared species between FG and Peru (8.8%). Only the species represented by at least four individuals were analyzed (621 and 566 species in FG and Peru, respectively), a threshold representing the minimum abundance to detect a significant habitat indicator value (*indval*) (Dufrene and Legendre 1997).

First, overall species differentiation in habitat preferences was tested in each region using a procedure derived from the D statistic described in Vleminckx et al. (2015), calculated as follows:

$$D = \sum_i \sum_h \left[(O_{ih} - E'_{ih})^2 / E'_{ih} \right],$$

where E'_{ih} is the expected relative abundance of species i in habitat h , in the absence of niche differentiation among species. E'_{ih} is calculated as

$$E'_{ih} = (O_{.h} O_{i.}) / O_{..},$$

where $O_{.h}$, $O_{i.}$ and $O_{..}$ correspond, respectively, to the total number of individuals (all species) sampled in habitat h , the abundance of species i (across all habitats), and

the total number of individuals in the data set. The observed D value was compared with 4,999 values obtained after randomizing the overall community structure using the MSR method described in the previous section to account for spatial autocorrelation in species distributions.

Species were considered *specialists* if displaying a significant indval score for one of the three habitats (white-sand, terra firme, or seasonally flooded forests). Indval scores for each species were compared with 4,999 score values obtained with the MSR method.

We further examined whether specialization for one of the three habitats occurred in certain clades (families and higher levels) more often than expected by chance, following the approach of Fine and Kembel (2011). For each clade (from the family level to the most basal lineages) and each habitat, we compared the observed number of descendent species (not weighting for their abundances) with the number of species obtained after randomly shifting the tips of the phylogeny (null model with 4,999 randomizations). We considered that clades with observed values for one habitat type higher than at least 95% of null values had more descendants than expected by chance in this habitat.

All statistical analyses were performed in R statistical environment version 3.5.1 (R Development Core Team 2020), using packages referenced in the provided R code. Species abundance, environmental data sets and the R code are provided in Baraloto et al. (2021).

RESULTS

Floristic diversity and dominance across subregions and habitats

We inventoried 26,386 trees, of which we identified 25,397 (96%) to the genus level. We assigned them to 2,745 named species or putative novel species with distinct standardized names, from 473 genera and 90 families (Baraloto et al. 2021).

Floristic diversity was lower overall in FG plots than in Peruvian plots: the average number of species at the plot level reached 81.2 (\pm standard error = 3.5) and 101.6 (\pm 5.0), and the average expected number of species based on random samplings of two individuals (to give more weight to the most abundant species) reached 49.5 (\pm 5.2) and 61.4 (\pm 6.8) in FG and Peru, respectively (Table 1). Terra firme hosted the highest species diversity, followed by seasonally flooded and white-sand, and within each habitat, species diversity was also lower in FG than in Peru (Table 1).

Only 2 of the 10 most abundant species across both study regions occurred in both study regions (Appendix S3: Table S1). *Eschweilera coriacea* (Lecythidaceae) dominated both TF and SF forests in Peru, and SF forests in French Guiana (Appendix S3: Table S2). *Chrysophyllum sanguinolentum* (Sapotaceae) was among the most abundant taxa across both regions

(Appendix S3: Table S1) but occurred in moderate densities in Peru and in low densities in FG (Appendix S3: Table S2). The other dominant species were generally dominant in a single habitat in one of the two regions. For example, *Euterpe oleracea* (Arecaceae; 2.83% of all stems) was dominant in SF habitats in FG, *Pachira brevipes* (Malvaceae; 2.79%) was dominant in WS forests in Peru, and *Clusia fockeana* (Clusiaceae; 1.50%) was dominant in WS forests in FG. Moreover, very few species dominant in each habitat actually occurred in both regions, and none, other than *Eschweilera coriacea*, was dominant in both regions (Table S2 in Appendix S4). For example, *Guaerea pubescens* (Meliaceae) was dominant in French Guiana TF forests (Table S2 in Appendix S4), was also found occasionally in other habitats in FG, but was relatively rare in Peru (only 12 of 203 individuals observed). Similarly, *Anaxagorea dolichocarpa* (Annonaceae) was dominant in French Guianan SF forests (Appendix S3: Table S2), also found occasionally in other habitats in FG, but was very rare in Peru (only 6 of 182 individuals observed).

Floristic and phylogenetic heterogeneity

Overall, forests in Peru and FG had very little overlap in species composition. Only 8.8% of all observed species ($n = 231$) were present in both regions (Appendix S4). This percentage reached 52.9%, 80.0%, and 92.9% for genera, families, and orders, respectively. This strong difference in species composition between Peru and FG was observed along the first axis of the NMDS analysis, with a marked contrast between WS and the two other habitats (TF and SF) observed along Axis 2 (Fig. 2a, b). Compositional differences among plots decreased when analyzing higher taxonomic levels, although differences remained clearly distinguishable between regions and across habitats at the genus level (Fig. 2c, d), and to a lesser extent when using family abundance data (Fig. 2e, f). Only the WS plots from FG and Peru remained well separated on all NMDS plans (Fig. 2a–h). The relatively high abundance of *Euterpe oleracea* (Arecaceae), a species not present in Peru and representing 4.95% of all stems in FG, likely explained that SF plots from this region were set distinctly apart from the other plots on the NMDS plan when using abundance instead of presence–absence data for species (Fig. 2a, b). The NMDS obtained using phylogenetic composition calculated with the Unifrac index, which gives more weight to phylogenetic distance among recent nodes, produced a similar ordination to the one obtained with genus composition data (Appendix S5). When using B_{st} and Π_{st} values, which give more weight to phylogenetic distance among older nodes, ordination patterns rather resembled those obtained with family presence–absence or order composition data (Appendix S5). Phylogenetic differences across regions were further emphasized by relatively high geographic node divergence (GND) values (>0.6)

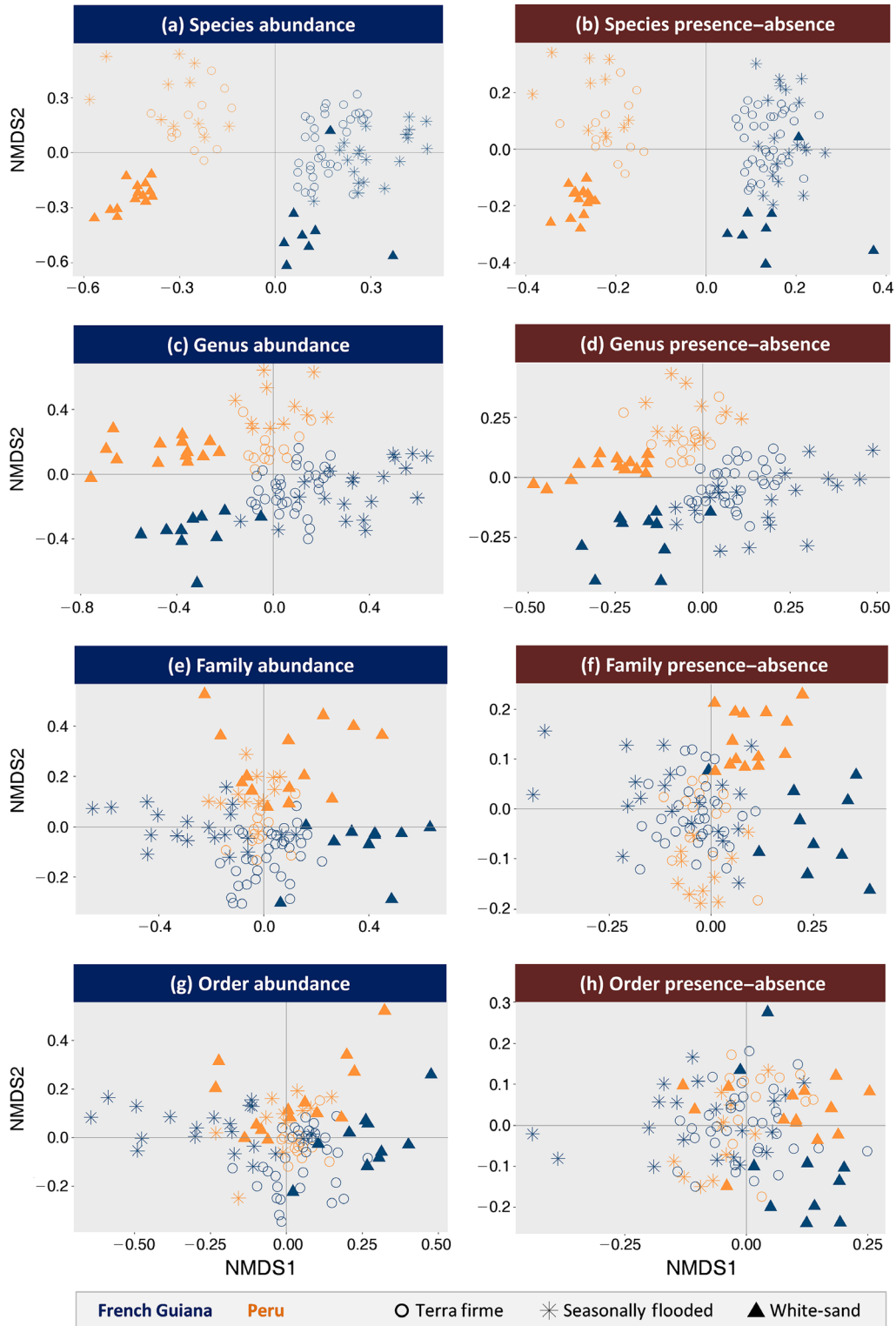


FIG. 2. Projection of plot scores on Axes 1–2 of a nonmetric multidimensional scaling (NMDS) analysis performed on the abundance (left) and presence–absence (right) of different taxonomic levels (species, genus, family, and order). The projection of plots emphasizes both the two study regions using colours and the different habitats using symbols (seasonally flooded = SF, terra firme = TF, white-sand = WS). A complementary analysis using phylogenetic distance is provided in Appendix S5.

observed in both basal lineages and near the tips of the phylogeny (Appendix S6).

Environmental heterogeneity

Differences in soil texture and nutrient availability were associated with substantial shifts in species composition. The first axis of the PCA (accounting for 46.3% of the overall environmental variation) was the only axis that explained more variation than expected under a broken stick model. This axis separated plots with relatively high soil nutrient contents and clayey texture from plots located on relatively more nutrient-depleted and sandy soils (Fig. 3a, c). The variables most significantly associated with this axis (r -Pearson between variables and PCA plot scores ≥ 0.7 ; $P \leq 0.05$; t -test of Pearson's product moment correlation) were TN, Mg, OC, and percent clay (see Baraloto et al. 2021 for details). The second axis (14.7%) was mostly associated with orthogonal variation of percent sand and P content ($r = 0.56$ and 0.64 , respectively; $P \leq 0.001$). The third axis (10.4%) was not clearly associated to any particular variable, whereas the fourth axis (9.2%) represented a gradient of DSI ($r = -0.78$; $P \leq 0.001$; Appendix S8 in Baraloto et al., 2021).

Environmental conditions did not show any clear difference between regions in the PCA (Fig. 3b), which was not surprising, given that the experimental design aimed to capture the breadth of habitat and soil conditions found in each region. WS habitats were markedly distinguishable from the two other habitats along the first PC axis (Fig. 3c), and TF and SF exhibited similar soil fertility. The distinct conditions of WS plots were further illustrated by the significant differences for each soil variable between WS and the two other habitats (Table 2). Differences among habitats were highly significant for all soil variables (Table 2), especially when comparing WS with TF or SF.

Environmental vs. spatial effects on species and phylogenetic turnover

Habitat, region, and subregion differences together explained 46.2 and 51.5% of the variation in species abundance and presence-absence dissimilarity ($P \leq 0.001$; Model 1 in Table 3), respectively, with regional difference representing the strongest effect among predictors (LMG = 0.249 and 0.319), followed by habitat difference (0.144 and 0.109), and subregion difference within region (0.069 and 0.088). Conversely, the composition dissimilarity of higher-ranked taxa (\geq genus) was better explained by habitat than by spatial differences (using both abundance and presence-absence data), although all three effects progressively decreased from the genus to the order level. The latter observation was consistent when comparing Model 1 built using response variables quantifying phylogenetic dissimilarity while weighting recent (UniFrac index) vs. deeper evolutionary

differences among taxa (B_{st} and Π_{st} statistics [Baraloto et al. 2021]).

Examining the relative effect of each pair of different habitats (Model 2) showed that the habitat effect on species abundance dissimilarity was stronger when comparing WS with another habitat (LMG = 0.090 and 0.069 for the WS-TF and the WS-SF pairs, respectively) than when comparing TF and SF (LMG = 0.025; Table 3). A consistent trend was observed when using presence-absence data. Within each region, comparing two plots from a same habitat (TF-TF, SF-SF, or WS-WS) generally led to lower species turnover than when comparing different habitats (see Baraloto et al. [2021]). At the same time, intra- vs. interhabitat differences in composition dissimilarity were almost completely masked by regional effects when comparing plots from Peru and FG.

In Model 3, we found that, in addition to the region and subregion difference effect for species dissimilarity (LMG = 0.204 and 0.059, respectively), the continuous measure of soil dissimilarity represented a relatively substantial driver of composition turnover from the species up to the family level (LMG > 0.10) and relatively less at the order level (Table 3). Subregion differences within region and climate dissimilarity had generally weaker relative effects. The environmental variables that best explained composition dissimilarity (LMG > 0.045) were Ca and C:N ratio (Model 4; Table 3), with DSI also showing relatively high effect (LMG > 0.035).

In agreement with the NMDS (Fig. 3) and the regression results (Table 3), phylogenetic dissimilarity was generally more weakly associated with region, subregion, habitat and environmental (soil, climate) dissimilarity across plots than taxonomic dissimilarity, with phylogenetic composition distance among plots showing intermediate levels when using the UniFrac index (Baraloto et al. 2021).

Habitat specialization across subregions and regions

A significant differentiation in habitat preferences was found at the whole community level, both in Peru and FG ($P \leq 0.001$; D tests of habitat differentiation; not shown in a table). Habitat specialist species were represented by at least four individuals, which was the minimum number of individuals with a significant indval ($P \leq 0.05$). The percentage of specialists among all species represented by at least four individuals (minimum number of individuals for which a significant signal of specialization could be detected) within SF, TF, and WS habitats reached, respectively, 7.76, 10.58, and 13.76% in Peru, and 3.40, 11.72, and 11.34% in FG (Table 4). The five specialists most strongly associated with each habitat and each region are presented in Table 4. In Peru, the overall five highest indval values were obtained for WS specialists, while in FG they were obtained for two SF and three WS specialists. All indval values for each species represented by at least four individuals are available

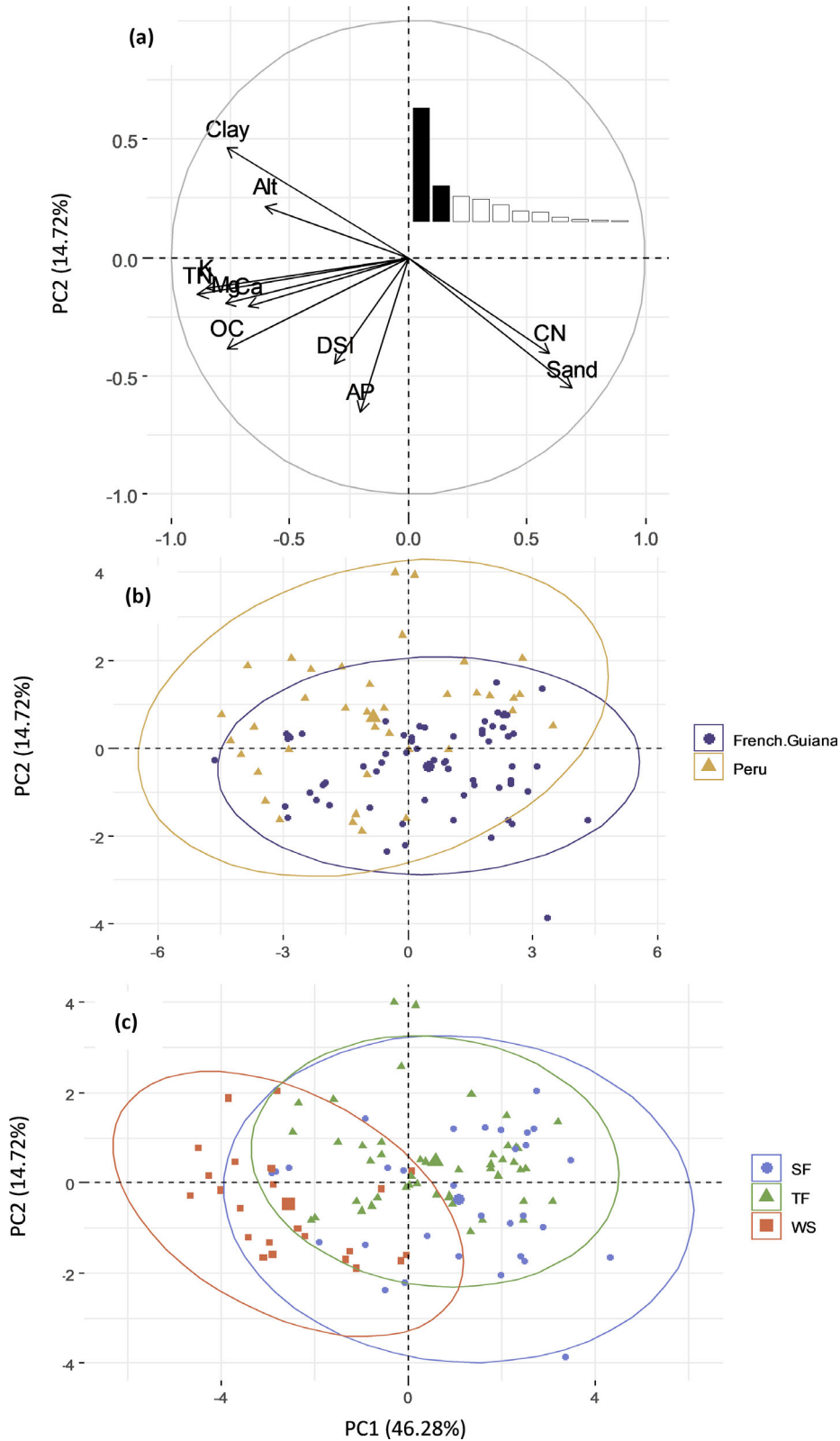


FIG. 3. Projection of (a) variables and (b), (c) plot scores on Axes 1–2 of a principal component analysis performed on the environmental data. The projection of plots emphasizes either (b) the two study regions (Peru vs. French Guiana) or (c) the different habitats characterized on the field (seasonally flooded = SF, terra firme = TF, white-sand = WS).

TABLE 2. Mean (\pm standard deviation) of each environmental variable in each habitat and each study region, and *P* values of two-way ANOVA tests to test the effect of region, habitat (H-ab), and the interaction of region and habitat (R:H).

Habitat	French Guiana			Peru			ANOVA			Tukey HSD		
	SF	TF	WS	SF	TF	WS	Region	Hab	R:H	TF-SF	WS-SF	WS-TF
Altitude (m)	165.4 (104.67)	218.94 (137.55)	100.3 (92.62)	117.45 (20.81)	128.62 (12.93)	137.86 (16.4)	0.863	0.028	0.013	0.247	0.590	0.033
Dry season ($^{\circ}$ C/mm)	1.88 (0.43)	1.92 (0.43)	1.75 (0.33)	1.54 (0.13)	1.51 (0.1)	1.56 (0.16)	0.001	0.842	0.485	0.981	0.929	0.838
Clay (%)	20.6 (17.72)	47.32 (22.59)	4.21 (5.04)	45.45 (19.44)	27.77 (12.44)	8.14 (13.17)	0.089	0.001	0.001	0.005	0.001	0.001
Sand (%)	64.53 (24.89)	41.84 (22.66)	92.27 (6.86)	20.91 (12.96)	46.23 (24.84)	64 (22.19)	0.013	0.001	0.001	0.528	0.001	0.001
OC (%)	4.73 (4.48)	3.22 (1.52)	2.24 (2.97)	2.05 (0.83)	1.46 (0.55)	1.48 (1.04)	0.001	0.001	0.171	0.513	0.001	0.004
TN (%)	2.29 (3.25)	1.65 (1.35)	0.17 (0.22)	0.23 (0.09)	0.14 (0.05)	0.07 (0.06)	0.001	0.001	0.142	0.999	0.001	0.001
C:N	15 (4.04)	12.84 (1.46)	18.16 (7.1)	9.34 (1.78)	10.46 (1.68)	20.42 (3.59)	0.257	0.001	0.001	0.770	0.001	0.001
Avail. P (mg/kg)	10.78 (15.51)	5.07 (5.81)	3.67 (4.11)	7.4 (5.75)	2.25 (1.54)	8.4 (8.01)	0.501	0.003	0.002	0.002	0.388	0.189
Ca (mg/kg)	1.26 (2.37)	0.53 (1.02)	0.16 (0.17)	5.96 (7.23)	0.84 (2.58)	0.04 (0.03)	0.001	0.001	0.024	0.004	0.001	0.003
Mg (mg/kg)	0.81 (0.92)	0.32 (0.31)	0.22 (0.24)	1.35 (1.45)	0.31 (0.6)	0.05 (0.04)	0.001	0.001	0.025	0.001	0.001	0.022
K (mg/kg)	0.13 (0.11)	0.1 (0.04)	0.05 (0.05)	0.18 (0.07)	0.06 (0.02)	0.06 (0.03)	0.312	0.001	0.007	0.002	0.001	0.001

Notes: The last three columns represent the *P*-values from Tukey honestly significant difference tests comparing each pair of habitats (TF = terra firme, SF = seasonally flooded and WS = white-sand).

in Appendix S4. The latter values are also represented in a radial phylogenetic tree, in which we also show lineages having more descendant species in one of the three studied habitats than expected by chance (Fig. 4). We detected that Malpighiales, Fabales, and Monocots had more descendant species than expected by chance in SF habitats, whereas the Magnoliales, Ericales, Oxalidales, Rosales, and Sapindales were overrepresented in TF, and the Campanulidae, Lamiidae, and especially the Gentianales, contained significantly more species in WS.

DISCUSSION

Diversity and dominance—our results compared with previous studies

Several studies have compared plant diversity across different regions of the Amazon basin and evaluated the relative importance of geographic distance and environmental variables in driving compositional differences (Tuomisto et al. 2003, ter Steege et al. 2006, 2013, Stropp et al. 2009). Our effort expands upon their work in three important ways. First, our plot size and diameter cutoff (2.5 cm) for trees were more inclusive than that of previous large-scale Amazonian studies, almost all of which have used a minimum size threshold of 10-cm diameter at breast height (dbh), surveyed over 1 ha. We sampled smaller trees using modified Gentry plots characterized by a herringbone arrangement of 50-m transects over a 2-ha area (Baraloto et al. 2013). Compared to other 1-ha samples, our plots included a wider range of size classes, and thus include small understory trees (and some shrubs) and juveniles of large trees in addition to the largest emergent canopy trees. In addition, the fact that our transects are spread out over a larger area means that the most common tree species (which often can have clumped distributions [Condit et al. 2000]) are less likely to show strong dominance patterns and are more likely to capture rare species than in analyses from 1-ha plot networks (see Draper et al. 2021).

A second contribution of our study is that we sampled terra firme, white-sand, and seasonally flooded forests within subregions in each region, so that we could optimally disentangle geographic distance from environmental effects. Other large plot networks that have investigated floristic turnover in Amazonia generally have located plots along rivers or have sampled a large forest area with many plots evenly distributed in latitude or longitude (Tuomisto et al. 1995, 2003, Ruokolainen et al. 1997, Pitman et al. 2001, 2008, ter Steege et al. 2013, Stropp et al. 2009). In such designs, because habitats are often spatially autocorrelated (i.e., two nearby plots are likely to occupy more similar habitats than expected by chance), it may be difficult to discern the extent to which changes in species composition are driven by environmental factors versus dispersal limitation. Within each region, we sampled several subregions, and in (almost) each subregion we sampled terra firme,

TABLE 3. R^2 values quantifying the effects of different assemblages of predictors (Models 1–4) on taxonomic dissimilarity among plots, and Lindeman, Merenda and Gold values corresponding to the decomposition of the model R^2 and quantifying the relative effect of each predictor.

	SD1	GD1	FD1	OD1	SD2	GD2	FD2	OD2
Model 1	0.462***	0.355***	0.203***	0.113***	0.515***	0.309***	0.122***	0.024*
Habitat	0.144	0.228	0.156	0.088	0.109	0.139	0.101	0.019
Region	0.249	0.098	0.036	0.016	0.319	0.132	0.016	0.003
Subregion	0.069	0.029	0.011	0.009	0.088	0.037	0.005	0.002
Model 2	0.184***	0.29***	0.228***	0.107***	0.154***	0.188***	0.187***	0.039**
TF-SF	0.025	0.038	0.015	0.028	0.014	0.021	0.007	0.001
WS-SF	0.069	0.146	0.121	0.063	0.064	0.101	0.101	0.028
WS-TF	0.09	0.106	0.092	0.015	0.076	0.065	0.078	0.01
Model 3	0.422***	0.324***	0.203***	0.116***	0.511***	0.308***	0.168***	0.059*
Soil	0.129	0.207	0.147	0.076	0.134	0.151	0.137	0.04
Climate	0.031	0.023	0.029	0.031	0.039	0.026	0.02	0.017
Region	0.204	0.07	0.021	0.007	0.262	0.1	0.008	0.001
Subregion	0.059	0.023	0.006	0.003	0.075	0.031	0.003	0.001
Model 4	0.26***	0.321***	0.233***	0.138***	0.301***	0.271***	0.201***	0.077*
Altitude	0.008	0.01	0.006	0.009	0.01	0.006	0.002	0.003
Dry season Index	0.035	0.015	0.021	0.016	0.048	0.029	0.019	0.015
Clay	0.022	0.032	0.027	0.014	0.018	0.016	0.02	0.002
Sand	0.015	0.015	0.006	0.003	0.016	0.01	0.009	0.004
OC	0.008	0.018	0.018	0.008	0.01	0.018	0.021	0.006
TN	0.025	0.021	0.026	0.026	0.03	0.021	0.018	0.003
C:N ratio	0.045	0.093	0.058	0.013	0.048	0.056	0.065	0.017
P	0.014	0.019	0.012	0.019	0.016	0.019	0.013	0.017
Ca	0.05	0.06	0.036	0.021	0.063	0.063	0.011	0.003
Mg	0.031	0.024	0.01	0.005	0.034	0.023	0.005	0.001
K	0.007	0.014	0.013	0.004	0.008	0.009	0.019	0.006

Notes: SD/GD/FD/OD = Species/genus/family/order composition dissimilarity (1 = Hellinger-transformed abundance; 2 = presence-absence data). Asterisks indicate the P values of the MSR test of the adjusted R^2 of each model: *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$.

white-sand forest, and seasonally flooded plots that were relatively evenly spaced from one another. Thus, plots of a given habitat were rarely closer to each other in space than plots of different habitats.

Third, our attention to taxonomic precision is critical to emphasize. Previous efforts by networks including the Amazon Tree Diversity Network (ter Steege et al. 2013) and RAINFOR (Malhi et al. 2002) and taxonomic specialists such as Cardoso et al. (2017) have compiled species lists from plot inventories into a single database and cleaned the taxonomic names by checking published sources and eliminated synonyms and errors. However, in none of those efforts have botanists actually reviewed the voucher specimens of all of the plots at the same time and in the same place to determine if, for example, individuals identified as *Protium opacum* in one plot matched those identified as *P. opacum* from every other plot. Exercises to check the accuracy of the taxonomy of species from three common and diverse genera in subsamples from across a large existing plot network (RAINFOR) uncovered error rates of 20–45% (Baker et al. 2017). In our plots, each species (or putative novel species) was collected at least once per plot, and all voucher specimens were

evaluated multiple times by several botanists in the home herbaria (AMAZ in Iquitos and CAY in French Guiana). Importantly, six botanists were involved in both Peru and French Guiana to collect plants and identify species. In addition, four of these six botanists with field and herbarium knowledge of both floras visited Missouri Botanical Garden (MO) with all specimens from both countries and worked for weeks comparing these vouchers with existing collections from across the Amazon. The end result of this careful work is a data set that gives a more accurate estimate of plot-level and region-wide diversity than any previous published work of Amazonian tree plots.

Our vouchered database of species and morphospecies (Baraloto et al. 2021 and accompanying digital voucher library) represents an extremely valuable resource. First, our collections have been compared to reference collections at MO, the best Neotropical plant collection on the planet, so we have high confidence in our species determinations. Second, when we did not find a species in MO to match our voucher, we believe that this represents strong evidence for a putative novel species, which systematists can pursue in future research to describe new tree species. The fact that our collections come from

TABLE 4. Species displaying the five highest Indicator values (*Indval*) within each habitat (columns 4–6) in each region.

Species	Family	Ab	SF	TF	WS	<i>P</i>
Peru						
<i>Brosimum lactescens</i>	Moraceae	52	0.629	0.037	0.016	0.001
<i>Campsiandra angustifolia</i>	Fabaceae	20	0.609	0.003	0	0.001
<i>Tapura acreana</i>	Dichapetalaceae	19	0.549	0.021	0	0.001
<i>Euterpe precatorea</i>	Arecaceae	18	0.520	0.004	0	0.001
<i>Garcinia madruno</i>	Clusiaceae	16	0.518	0	0.004	0.001
<i>Rinorea racemosa</i>	Violaceae	43	0	0.615	0	0.001
<i>Tetrastylidium peruvianum</i>	Olacaceae	12	0	0.615	0	0.001
<i>Nealchornea yapurensis</i>	Euphorbiaceae	31	0.060	0.539	0	0.002
<i>Leonia cymosa</i>	Violaceae	26	0	0.538	0	0.001
<i>Ophiocaryon heterophyllum</i>	Sabiaceae	20	0	0.529	0.034	0.001
<i>Matayba inelegans</i>	Sapindaceae	58	0	0.001	0.981	0.001
<i>Chrysophyllum sanguinolentum</i>	Sapotaceae	136	0	0.008	0.877	0.001
<i>Roucheria punctata</i>	Linaceae	26	0	0	0.857	0.001
<i>Macrobium microcalyx</i>	Fabaceae	128	0	0.003	0.843	0.001
<i>Macrobium bifolium</i>	Fabaceae	90	0.001	0.004	0.825	0.001
% specialists			7.76	10.58	13.76	
French Guiana						
<i>Euterpe oleracea</i>	Arecaceae	664	0.898	0	0	0.001
<i>Pterocarpus officinalis</i>	Fabaceae	122	0.796	0	0	0.001
<i>Symphonia globulifera</i>	Clusiaceae	37	0.541	0	0	0.001
<i>Anaxagorea dolichocarpa</i>	Annonaceae	176	0.504	0.047	0	0.003
<i>Inga cylindrica</i>	Fabaceae	58	0.463	0.011	0	0.002
<i>Cordia nodosa</i>	Boraginaceae	42	0.008	0.515	0	0.001
<i>Unonopsis rufescens</i>	Annonaceae	61	0.005	0.474	0.020	0.005
<i>Siparuna decipiens</i>	Siparunaceae	51	0.024	0.433	0.006	0.006
<i>Geissospermum argenteum</i>	Apocynaceae	34	0	0.412	0	0.002
<i>Pouteria gongrijpii</i>	Sapotaceae	32	0	0.412	0	0.004
<i>Pachira flaviflora</i>	Malvaceae	209	0.004	0	0.781	0.001
<i>Clusia fockeana</i>	Clusiaceae	518	0.005	0	0.772	0.001
<i>Calycolpus revolutus</i>	Myrtaceae	54	0.002	0.002	0.653	0.001
<i>Cassipourea guianensis</i>	Rhizophoraceae	80	0.013	0.003	0.642	0.001
<i>Manilkara bidentata</i>	Sapotaceae	74	0	0.004	0.580	0.001
% specialists			3.40	3.72	11.3	

Notes: Ab: total abundance of the species in the region. *P*: *P* value for the test of *Indval* differences among habitats. For each region, the last line shows the percentage of specialist species among all species represented by at least four individuals in Peru ($n = 566$) and French Guiana ($n = 621$).

a georeferenced plot that can easily be revisited is also a benefit (Baker et al. 2017). Finally, many of the putative novel species in our data set belong to species complexes, some of which have been agglomerated in regional checklists of species, such as in ter Steege et al. (2020). For example, *Protium heptaphyllum*, the 12th most abundant species in ter Steege et al. (2013), is now known to represent at least seven different taxonomic entities (Damasco et al. 2019, 2021).

How different is the diversity and composition of eastern and western Amazonian tree communities?

Our attention to taxonomic detail means we have the first reliable estimate of the overlap in species composition across Amazonia. The 8.8% overlap we report between Peru and French Guiana therefore is not substantially inflated by the conflation of species complexes

into single species. The most common tree species in our plots in both Peru and French Guiana do sometimes match the list of Amazonian hyperdominants of ter Steege et al. (2013). In Peru, 7 of our top 15 species are ATDN hyperdominants, and in French Guiana, 4 of our top 15 species are ATDN hyperdominants. Nevertheless, some species in our most-common list are white-sand or seasonally flooded specialists, which were perhaps underrepresented in the ATDN (and overrepresented in our data because of our sampling design). Some species are also likely to have been misidentified in networks such as ATDN. For example, some of our common species are listed as extremely rare in the ATDN. Notably, *Eschweilera coriacea*, the only common species found in both countries in our data set, is now known to include many species and the complex itself does not represent a monophyletic group (O. Vargas, *personal communication*).

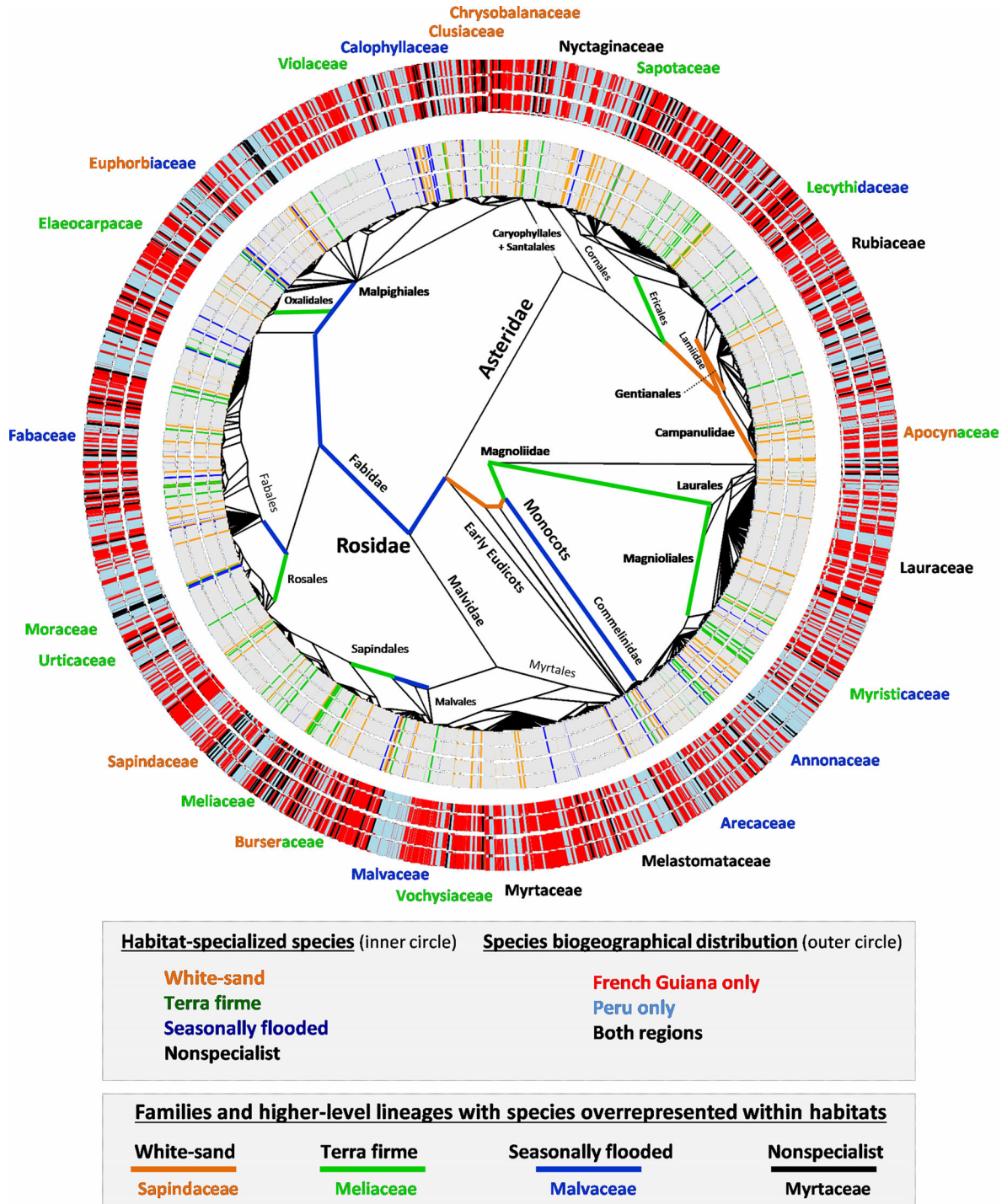


FIG. 4. Radial phylogenetic tree representing all species inventoried in French Guiana and Peru, with colors indicating whether they are significantly indicative of one of the three studied habitats (inner ring) and whether they are represented in Peru, French Guiana, or both regions (outer circle; see legend in the upper rectangle). Colored branches and colored family names indicate, respectively, lineages (>families) and families having more descendant species than expected by chance in one (or two) of the three studied habitats, as indicated by the legend in the lower rectangle. Only the position of families represented by at least 200 individuals is shown for clarity.

Interestingly, only 1 of our top 15 species from Peru was shared with the Pitman et al. (2008) lists of oligarchs from Yasuní National Park in Ecuador and Manu National Park in Peru. Yasuní and Manu lie on fertile clay terra firme forests, but our sampling of these habitats in Peru was rather limited. Our plots included a mosaic of terra firme soils, corresponding to older geologic sediments for the most part (Pitman et al. 2008). Similarly, Pitman et al. (2008) found a strong discontinuity of tree species composition along the Napo River between Ecuador and Peru, and they reported that plots in the east did not contain the Yasuní/Manu oligarchs. Our study thus represents an important complement to existing work and paves the way for the integration of phylogeographic analyses of widespread species to resolve species complexes (Prata et al. 2018).

To what extent are there similar versus contrasting patterns of taxonomic and phylogenetic turnover across Amazonia?

The phylogenetic beta-diversity analyses towards the tips showed significant, nonrandom patterns of tree composition with respect to both region and habitat (Fig. 4). In other words, both countries harbor many habitat specialists only found in that region. The fact that we found nonrandom phylogenetic signal means that different genera and/or families dominate the three habitats in each region. Terborgh and Andresen (1998) reported that seasonally flooded forests were dominated by Chrysobalanaceae and Lecythidaceae in eastern Amazonia, but Moraceae and Myristicaceae in western Amazonia. We report a similar signal, with regional floras being distinct from each other and habitats within regions being composed of assemblages from each regional flora. As detailed in Fig. 4, we also found that different lineages were more represented than expected in each habitat, with Arecaceae, Annonaceae, Fabaceae, and Malvaceae diversifying particularly in SF habitats; Meliaceae, Moraceae, and Sapotaceae overrepresented in TF; and Chrysobalanaceae, Clusiaceae, and Sapindaceae in WS forests.

Yet, the variance in phylogenetic signal across the regions suggests that deeper-time events influenced dispersal and speciation for some tree lineages. One series of events that influenced Peru differently than French Guiana was the Andean uplift, the concomitant erosion of volcanic sediments, and the formation of large lakes and/or marine incursions (Hoornt et al. 2010). These events likely caused the formation of large areas of fertile soils and massive changes in the area and location of seasonally flooded habitat. This, in turn, would have promoted speciation in the late Miocene in western Amazonia (Fine et al. 2014), while habitat heterogeneity in eastern Amazonia probably remained more stable over this time frame. In contrast to strong taxonomic turnover and tip-weighted phylogenetic beta diversity (Fig. 2a–d), we found no clear basal-weighted

phylogenetic structure across regions and habitats (Fig. 2e–f). This result is consistent with previous broader analyses of the Neotropical flora (Guevara et al. 2016, Dexter et al. 2017).

These phylogenetic beta-diversity differences became blurred when taking species abundance into account, maybe because of the presence of widespread and locally abundant genera that comprise numerous closely related species (Guevara et al. 2016). Examples of these patterns can be found within the genera *Protium* (Burseraceae) and *Pachira* (Malvaceae). In both genera, relatively few species are shared across regions, although both genera contain widespread generalists (e.g., *Protium opacum*, *Pachira insignis*). Nevertheless, both lineages are characterized by specialists, not only among habitats within regions (Fine and Baraloto 2016), but also specialists of the same habitat 2,000 km apart (e.g., white-sand specialists *Pachira brevipes* in Peru and *P. flaviflora* in French Guiana). The existence of specialists of contrasting habitats in the same genera in both regions further suggests recent habitat-mediated speciation in each region is a common phenomenon. Moreover, the absence of strong habitat difference and soil dissimilarity on phylogenetic turnover among plots may indicate that habitat specialization occurred relatively recently and repeatedly within clades.

What is the relative strength of environmental factors and spatial distance in explaining floristic and phylogenetic differences among tree communities?

We found a very strong effect of habitat in driving tree species composition (Fig. 2). It is worth noting that our study design replicated the three habitats within subregions for each country. Thus, the strong separation of WS from SF and TF is likely due to environmental differences rather than geographic distance or dispersal limitation. Similarly, environmental distance induced stronger species turnover than dispersal limitation in ferns and Melastomataceae shrubs (Ruokolainen et al. 1997, Tuomisto et al. 2003). Analyzing 55 Amazonian RAINFOR plots, Honorio Colorado et al. (2009) concluded that at very large spatial scales (encompassing several regions), spatial distance was the main determinant of species turnover, followed by habitat and climate differences within regions. We found that spatial distance between subregions was also a significant factor, yet clearly much less important than environmental variables within regions. Our design allows us to report this with great confidence. Nevertheless, the relatively small number of plots we assessed within regional blocks does not allow us to disentangle whether dispersal limitation or sampling error within such a diverse tree flora explains such low turnover between plots of a similar habitat type.

The spatial scale encompassed by our study also allowed us to detect species composition variation driven by spatially structured climatic variables (DSI and

altitude). The seasonality effect suggests that, independently from spatial processes, hydraulic stress represents a major factor influencing species performance, corroborating other recent analyses (Comita and Engelbrecht 2014, Fortunel et al. 2014, Dexter et al. 2017) and underlining the potential impacts of a changing climate on community composition (Esquivel-Muelbert et al. 2018).

How strongly are species associated with different habitats? Is habitat specialization more frequent in less seasonal forests with a steeper gradient of soil resource availability (Peru vs. French Guiana)?

Within regions, species composition was substantially more dissimilar between different habitats than within similar habitats, with WS differing from TF and SF (Fig. 2; see Baraloto et al. [2021]). Moreover, within each region, white-sand forests hosted a higher proportion of specialists compared to the two other habitats. These results further confirm that white-sand soils (with extremely poor nutrient content and low water retention capacity) are particularly strong filters favoring plant species investing in resource conservation and protection against enemies (Fortunel et al. 2014, Fine and Baraloto 2016). In addition to their strong contribution to Amazonian biodiversity (Frasier 2008), white-sand endemics also represent a potential major species pool that may display greater resilience in response to more intense drought events predicted in eastern Amazonia (Fortunel et al. 2014, Guevara et al. 2016, Esquivel-Muelbert et al. 2018).

The higher similarity in the floristic composition between TF and SF could be explained by relatively rare flooding events and the substantial overlap of environmental conditions across these two habitats, allowing many TF species to establish on soils that are not constantly water saturated, and conversely enabling many swamp species to be competitive on soils with stronger drainage. Species distributed on TF and SF soils may also exhibit greater plasticity than WS species, which allocate costly adaptation mechanisms to survive on these relatively extreme habitats (Fortunel et al. 2020).

Consistent with overall patterns of diversity and composition, and in accordance with our predictions, the proportion of habitat specialists was higher in all three habitats in Peru than in French Guiana (Table 4). In particular, Peruvian forests comprised nearly twice the proportion of terra firme specialists as French Guianan forests (10.3% vs. 5.8%). White-sand specialists were more frequent than for other habitats in both regions, with relatively consistent proportions in Peru (13.8%) and French Guiana (10.6%; Table 4). Why would habitat specialization be more common in Peru? One reason may be that biotic interactions can interact with environmental gradients to promote habitat specialization in trees (Fine et al. 2004). The more aseasonal climate and more fertile soils in Peru are linked to more productive

forests with higher disturbance rates (with faster generation times) (Baker et al. 2014). These highly productive western Amazonian forests harbor greater insect populations with less seasonal dormancy, which should cause more intense and more consistent herbivore pressure in Peru compared to French Guiana. Moreover, with greater climatic stability over the past several million years in the west vs. the east, we would expect a higher potential for habitat specialization by trees in western Amazonia than in the east (Fine 2015).

CONCLUSIONS

Our unique and comprehensive experimental design and analyses represent an advance in the understanding of diversity, dominance, and composition in lowland forest tree communities of Amazonia. We refine previously reported patterns of family and genus dominance, and we report an overlap of only 8.8% of species composition on opposite ends of the Basin. Together, these results emphasize the importance of recent processes in the diversification of tree lineages across these communities, and the role of environmental filtering in generating and maintaining the diversity and composition of these forests.

ACKNOWLEDGMENTS

Research was supported by collaborative National Science Foundation (NSF) grant DEB-0743103/0743800 to CB and PVAF; by ANR Blanc NEBEDIV (Projet ANR-13-BSV7-009) and an INRAE Package grant to CB; and by NSF DEB 1254214 to PVAF. This work also benefited from an “Investissement d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA ANR-10-LABX-25-01). The authors thank the Réserve Naturelle de la Trinité¹⁶ and the Parc Amazonien de Guyane¹⁷ for authorizations and access to the inventory sites and additional funding. Part of the material was collected during the “Our Planet Reviewed” Guyane-2015 expedition in the Mitaraka range, in the core area of the French Guiana Amazonian Park, organized by the MNHN and Pro-Natura international. The expedition was funded by the European Regional Development Fund (ERDF), the Conseil régional de Guyane, the Conseil général de Guyane, the Direction de l’Environnement, de l’Aménagement et du Logement and by the Ministère de l’Éducation nationale, de l’Enseignement supérieur et de la Recherche. We thank the many colleagues who participated in field work in French Guiana, especially Seth Kauppinen, Alec Baxt, Benjamin Leudet, Benoit Burban, and Greg Lamarre; and in Peru, especially Julio Sanchez, Julio Grandez Ríos, Fabio and Milka Casado, and Tracy Misiewicz. Authors’ contributions: CB and PVAF were co-principal investigators for the project. JV led the data analysis. PVAF, JV, and CB wrote the manuscript. The collection and identification of tree species was coordinated by CB and PVAF, with help of JE, PP, ND, MR, EV, IM, JG, CF, EA, TP, AD, J-YG, OV, FD, and JV. All authors contributed to the preparation of the final manuscript. CB and JV contributed equally to this work.

¹⁶ <http://www.reserve-trinite.fr/>

¹⁷ <http://www.parc-amazonien-guyane.fr>

LITERATURE CITED

- Antonelli, A., J. A. Nylander, C. Persson, and I. Sanmartín. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* 106:9749–9754.
- Baker, T. R., et al. 2014. Fast demographic traits promote high diversification rates of Amazonian trees. *Ecology Letters* 17:527–536.
- Baker, T. R., et al. 2017. Maximising synergy among tropical plant systematists, ecologists, and evolutionary biologists. *Trends in Ecology & Evolution*, 32:258–267.
- Baldeck, C. A., et al. 2013. Soil resources and topography shape local tree community structure in tropical forests. *Proceedings of the Royal Society B* 280. <https://doi.org/10.1098/rspb.2012.2532>
- Baraloto, C., et al. 2012. Using functional traits and phylogenetic trees to examine the assembly of tropical tree communities. *Journal of Ecology* 100:690–701.
- Baraloto, C., et al. 2021. Data from: Biogeographic history and habitat specialisation shape floristic and phylogenetic composition across Amazonian forests. Dryad, data set. <https://doi.org/10.5061/dryad.9s4mw6mg0>
- Baraloto, C., et al. 2011. Disentangling stand and environmental correlates of aboveground biomass in Amazonian forests. *Global Change Biology* 17:2677–2688.
- Baraloto, C., Q. Molto, S. Rabaud, B. Hérault, R. Valencia, L. Blanc, P. V. A. Fine, and J. Thompson. 2013. Rapid simultaneous estimation of aboveground biomass and tree diversity across Neotropical forests: a comparison of field inventory methods. *Biotropica* 45:288–298.
- Bauman, D., T. Drouet, M. J. Fortin, and S. Dray. 2018a. Optimizing the choice of a spatial weighting matrix in eigenvector-based methods. *Ecology* 99:2159–2166.
- Bauman, D., T. Drouet, S. Dray, and J. Vleminckx. 2018b. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. *Ecography* 41:1–12.
- Blanchet, F. G., P. Legendre, and D. Borcard. 2008. Forward selection of explanatory variables. *Ecology* 89:2623–2632.
- Borregaard, M. K., C. Rahbek, J. FjeldsAa, J. L. Parra, R. J. Whittaker, and C. H. Graham. 2014. Node-based analysis of species distributions. *Methods in Ecology and Evolution* 5 (11):1225–1235.
- Cardoso, D., et al. 2017. Amazon plant diversity revealed by a taxonomically verified species list. *Proceedings of the National Academy of Sciences of the United States of America* 114:10695–10700.
- Comita, L. S., and B. M. J. Engelbrecht. 2014. Drought as a driver of tropical tree species regeneration dynamics and distribution patterns. Pages 261–308 in D. A. Coomes, W. D. Simonson, and D. F. Burslem, editors. *Forests and global change*. Cambridge University Press, Cambridge, UK.
- Condit, R., et al. 2000. Spatial patterns in the distribution of tropical tree species. *Science* 288:1414–1418.
- Condit, R., B. M. J. Engelbrecht, D. Pino, R. Perez, and B. L. Turner. 2013. Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *Proceedings of the National Academy of Sciences of the United States of America* 110:5064–5068.
- Damasco, G., C. Baraloto, A. Vicentini, D. C. Daly, B. G. Baldwin, and P. V. Fine. 2021. Revisiting the hyperdominance of Neotropical tree species under a taxonomic, functional and evolutionary perspective. *Scientific Reports* 11(1):1–11.
- Damasco, G., D. C. Daly, A. Vicentini, and P. V. A. Fine. 2019. Reestablishment of *Protium cordatum* (Burseraceae) based on integrative taxonomy. *Taxon* 68:34–46.
- Dauby, G., and O. J. Hardy. 2011. Sampled-based estimation of diversity sensu stricto by transforming Hurlbert diversities into effective number of species. *Ecography* 34:001–012.
- Dexter, K. G., M. Lavin, B. M. Torke, A. D. Twyford, T. A. Kursar, P. D. Coley, C. Drake, R. Hollands, and R. T. Pennington. 2017. Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of the National Academy of Sciences of the United States of America* 114:2645–2650.
- Dick, C. W., K. Abdul-Salim, and E. Bermingham. 2003. Molecular systematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. *American Naturalist* 162:691–703.
- Draper, F. C., et al. 2021. Amazon tree dominance across forest strata. *Nature Ecology & Evolution* 5(6):757–767.
- Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* 196:483–493.
- Dufrène, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* 67:345–366.
- Esquivel-Muelbert, A., et al. 2018. Compositional response of Amazon forests to climate change. *Global Change Biology* 25:39–56.
- Fine, P. V. A. 2015. Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Reviews of Ecology, Evolution and Systematics* 46:369–392.
- Fine, P. V. A., and C. Baraloto. 2016. Habitat endemism in white-sand forests: insights into the mechanisms of lineage diversification and community assembly of the Neotropical flora. *Biotropica* 48:24–33.
- Fine, P. V. A., and S. W. Kembel. 2011. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* 34:553–656.
- Fine, P. V., F. Zapata, and D. C. Daly. 2014. Investigating processes of neotropical rain forest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae). *Evolution* 68(7):1988–2004.
- Fortunel, C., C. E. T. Paine, P. V. A. Fine, N. J. B. Kraft, and C. Baraloto. 2014. Environmental factors predict community functional composition in Amazonian forests. *Journal of Ecology* 102:145–155.
- Fortunel, C., C. E. T. Paine, P. V. A. Fine, I. Mesones, J.-Y. Goret, B. Burban, J. Cazal, and C. Baraloto. 2016. There's no place like home: seedling mortality contributes to the habitat specialisation of tree species across Amazonia. *Ecology Letters* 19:1256–1266.
- Fortunel, C., C. Stahl, P. Heuret, E. Nicolini, and C. Baraloto. 2020. Disentangling the effects of environment and ontogeny on tree functional dimensions for congeneric species in tropical forests. *New Phytologist* 226:385–395.
- Frasier, C. 2008. Amazonian lowland, white sand areas as ancestral regions for South American biodiversity: biogeographic and phylogenetic patterns in Potalia (Angiospermae: Gentianaceae). *Organisms Diversity & Evolution* 8:44–57.
- Gentry, A. H. 1981. Distributional patterns and an additional species of the *Passiflora vitifolia* complex: Amazonian species diversity due to edaphically differentiated communities. *Plant Systematics and Evolution* 137:95–105.
- Gentry, A. H. 1986. Endemism in tropical versus temperate plant communities. Pages 153–181 in M. E. Soule, editor. *Conservation biology: the science of scarcity and diversity*. Sinauer, Sunderland, Massachusetts, USA.

- Gerhold, P., M. B. Carlucci, Ş. Procheş, and A. Prinzing. 2018. The deep past controls the phylogenetic structure of present, local communities. *Annual Review of Ecology, Evolution, and Systematics* 49:477–497.
- Gourlet-Fleury, S., J. M. Guehl, and O. Laroussinie, editors. 2004. Ecology and management of a neotropical rainforest: lessons drawn from Paracou, a long-term experimental research site in French Guiana. Page 326.
- Graham, C. H., and P. V. A. Fine. 2008. Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, 11:1265–1277.
- Guevara, J. E., et al. 2016. Low phylogenetic beta diversity and geographic neo-endemism in Amazonian white-sand forests. *Biotropica* 48:34–46.
- Hardy, O. J., and B. Senterre. 2007. Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology* 95:493–506.
- Hijmans, R. J. 2018. raster: Geographic data analysis and modeling. R package version 2.8-4. <https://cran.r-project.org/web/packages/raster/index.html>
- Honorio Coronado, E. N., et al. 2009. Multi-scale comparisons of tree composition in Amazonian terra firme forests. *Biogeosciences* 6:2719–2731.
- Hoorn, C., et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927–931.
- Hopkins, M. J. 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. *Journal of Biogeography* 34:1400–1411.
- Hubbell, S. P. 2001. A unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.
- Kraft, N. J. B., et al. 2011. Disentangling the drivers of beta-diversity along latitudinal and elevational gradients. *Science* 333:1755–1758.
- Legendre, P., and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129:271–280.
- Lichstein, J. W. 2007. Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology* 188:117–131.
- Lindeman, R. H., P. F. Merenda, and R. Z. Gold. 1980. Introduction to bivariate and multivariate analysis. Scott, Foresman, Glenview, Illinois, USA.
- Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environment Microbiology* 71:8228–8235.
- Malhi, Y., et al. 2002. An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). *Journal of Vegetation Science* 13:439–450.
- Mc Fadden, I. R., B. Sandel, C. Tsirogiannis, N. Morueta-Holme, J.-C. Svenning, B. J. Enquist, and N. J. B. Kraft. 2019. Temperature shapes opposing latitudinal gradients of plant taxonomic and phylogenetic β diversity. *Ecology Letters* 22:1126–1135.
- Nicholls, J. A., R. T. Pennington, E. J. M. Koenen, C. E. Hughes, J. Hearn, L. Bunnefeld, K. G. Dexter, G. N. Stone, and C. A. Kidner. 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). *Frontiers in Plant Science* 6:710.
- Phillips, O. L., R. Vásquez Martínez, P. Núñez Vargas, A. Lorenzo Monteagudo, M.-E. Chuspe Zans, W. Galiano Sánchez, A. Peña Cruz, M. Timaná, M. Yli-Halla, and S. Rose. 2003. Efficient plot-based floristic assessment of tropical forests. *Journal of Tropical Ecology* 19:629–645.
- Pitman, N. C., J. W. Terborgh, M. R. Silman, V. P. Núñez, D. A. Neill, C. E. Cerón, W. A. Palacios, and M. Aulesia. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* 82(8):2101–2117.
- Pitman, N. C., et al. 2008. Tree community change across 700 km of lowland Amazonian forest from the Andean foothills to Brazil. *Biotropica* 40(5):525–535.
- Prata, E. M., C. Sass, D. P. Rodrigues, F. M. Domingos, C. D. Specht, G. Damasco, C. Rivas, P. V. A. Fine, and A. Vicentini. 2018. Towards integrative taxonomy in Neotropical botany: disentangling the *Pagamea guianensis* species complex (Rubiaceae). *Botanical Journal of the Linnean Society* 188:213–231.
- R Development Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org
- Richardson, J. E., R. T. Pennington, T. D. Pennington, and P. M. Hollingsworth. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293:2242–2245.
- Ruokolainen, K., A. Linna, and H. Tuomisto. 1997. Use of Melastomataceae and pteridophytes for revealing phytogeographical patterns in Amazonian rain forests. *Journal of Tropical Ecology* 13:243–256.
- Stropp, J., H. Ter Steege, Y. Malhi, ATDN, and RAINFOR. 2009. Disentangling regional and local tree diversity in the Amazon. *Ecography* 32:46–54.
- Terborgh, J., and E. Andresen. 1998. The composition of Amazonian forests: patterns at local and regional scales. *Journal of Tropical Ecology* 14(5):645–664.
- ter Steege, H., et al. 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* 443:444–447.
- ter Steege, H., et al. 2013. Hyperdominance in the Amazonian tree flora. *Science* 342:1243092.
- ter Steege, H., et al. 2020. Biased-corrected richness estimates for the Amazonian tree flora. *Scientific Reports* 10:1–13.
- Tuomisto, H., K. Ruokolainen, R. Kalliola, A. Linna, W. Danjoy, and Z. Rodriguez. 1995. Dissecting Amazonian biodiversity. *Science* 269:63–66.
- Tuomisto, H., K. Ruokolainen, and M. Yli-Halla. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. *Science* 299:241–244.
- van Breugel, M., D. Craven, H. Ran Lai, M. Baillon, B. L. Turner, J. S. Hall. 2019. Soil nutrients and dispersal limitation shape compositional variation in secondary tropical forests across multiple scales. *Journal of Ecology* 107:566–581.
- Vargas, O. M., and C. W. Dick. 2020. Diversification history of Neotropical Lecythidaceae, an ecologically dominant tree family of Amazon rain forest. Pages 791–809 in *Neotropical diversification: patterns and processes*. Springer, Cham, Switzerland.
- Vlemminckx, J., T. Drouet, C. Amani, J. Lisingo, J. Lejoly, and O. J. Hardy. 2015. Impact of fine-scale edaphic heterogeneity on tree species assembly in a central African rainforest. *Journal of Vegetation Science* 26:134–144.
- Wagner, H. H., and S. Dray. 2015. Generating spatially constrained null models for irregularly spaced data using Moran spectral randomization methods. *Methods in Ecology & Evolution* 6:1169–1178.
- Wittmann, F., W. J. Junk, and M. T. F. Piedade. 2004. The varzea forests in Amazonia: flooding and the highly dynamic geomorphology interact with natural forest succession. *Forest Ecology and Management* 196:199–212.

Wittmann, F., J. Schöngart, and W. J. Junk. 2010. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplain forests. *In* W. Junk, M. Piedade, F. Witt-

mann, J. Schöngart, and P. Parolin, editors. Amazonian floodplain forests. Ecological studies (analysis and synthesis). Volume 210. Springer, Dordrecht, The Netherlands.

SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecm.1473/full>

OPEN RESEARCH

Data (Baraloto et al. 2021) are available in the Dryad digital repository: <https://doi.org/10.5061/dryad.9s4mw6mg0>