Received: 21 October 2020 | Accepted: 1 March 2021

DOI: 10.1111/1365-2745.13646

RESEARCH ARTICLE

Functional Traits 2.0: The Power of Metabolomics for Plant Ecology

Journal of Ecology | BRITISH ECOLOGICAL SOCIETY

The role of plant secondary metabolites in shaping regional and local plant community assembly

María-José Endara[1,2] | Abrianna J. Soule[3] | Dale L. Forrister[3] | Kyle G. Dexter[4] | R. Toby Pennington[5,6] | James A. Nicholls[7] | Oriane Loiseau[4] | Thomas A. Kursar[3] | Phyllis D. Coley[3]

[1]Grupo de Investigación en Biodiversidad, Medio Ambiente y Salud-BIOMAS- Universidad de las Américas, Quito, Ecuador; [2]Centro de Investigación de la Biodiversidad y Cambio Climático (BioCamb) e Ingeniería en Biodiversidad y Recursos Genéticos, Facultad de Ciencias de Medio Ambiente, Universidad Tecnológica Indoamérica, Quito, Ecuador; [3]Department of Biology, University of Utah, Salt Lake City, UT, USA; [4]School of Geosciences, University of Edinburgh, Edinburgh, UK; [5]Department of Geography, University of Exeter, Exeter, UK; [6]Royal Botanic Garden Edinburgh, Edinburgh, UK and [7]The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australian National Insect Collection (ANIC), Black Mountain, ACT, Australia

Correspondence
María-José Endara
Email: majo.endara@utah.edu

Funding information
National Science Foundation USA, Grant/Award Number: DEB-0640630 and DEB-1135733

Handling Editor: Roberto Salguero-Gómez

Abstract

1. The outstanding diversity of Amazonian forests is predicted to be the result of several processes. While tree lineages have dispersed repeatedly across the Amazon, interactions between plants and insects may be the principal mechanism structuring the communities at local scales.
2. Using metabolomic and phylogenetic approaches, we investigated the patterns of historical assembly of plant communities across the Amazon based on the Neotropical genus of trees *Inga* (Leguminosae) at four, widely separated sites.
3. Our results show a low degree of phylogenetic structure and a mixing of chemotypes across the whole Amazon basin, suggesting that although biogeography may play a role, the metacommunity for any local community in the Amazon is the entire basin. Yet, local communities are assembled by ecological processes, with the suite of *Inga* at a given site more divergent in chemical defences than expected by chance
4. *Synthesis*. To our knowledge, this is the first study to present metabolomic data for nearly 100 species in a diverse Neotropical plant clade across the whole Amazonia. Our results demonstrate a role for plant–herbivore interactions in shaping the clade's community assembly at a local scale, and suggest that the high alpha diversity in Amazonian tree communities must be due in part to the interactions of diverse tree lineages with their natural enemies providing a high number of niche dimensions.

KEYWORDS

Amazon, chemical defences, community assembly, *Inga*, local scale, metabolomics, regional scale, tropical rainforests

1 | INTRODUCTION

Amazonian forests are considered one of world's richest plant assemblages, with an estimated 16,000 species of trees for the whole region (ter Steege et al., 2020), and more than 650 woody species in a single hectare (Valencia et al., 2004). At a regional scale, recent studies have highlighted the role of dispersal across the Amazon in assembling tree communities (Dexter et al., 2017; Fine et al., 2014). At a local scale, there is still much debate regarding the ecological and evolutionary mechanisms that determine the co-occurrence of large numbers of species at a site, many of which are congeners. Some studies argue that niche differentiation may arise through competition for resources or adaptation to abiotic niches (Chesson, 2000; Kraft, Adler, et al., 2015; Kraft, Godoy, et al., 2015), while others claim that biotic factors such as natural enemy damage may facilitate coexistence (Coley & Kursar, 2014). The central premise of the latter is that the myriad of defences against herbivores may generate key additional niche axes that allow coexistence of a greater diversity of species (Levi et al., 2019).

The idea that the interactions between plants and their insect herbivores may contribute to the assembly of communities has received considerable recent attention. Specifically, this theory suggests that specialist pests may play a main role in maintaining the high local diversity of rainforests by preventing most plant species from becoming abundant (Comita et al., 2014; Connell, 1971; Janzen, 1970). Species do not share herbivores with their nearby neighbours if they have divergent defences (Becerra, 2007; Endara et al., 2017), which gives a species the advantage of reduced damage or 'enemy release' (Yguel et al., 2011). This in turn may promote the coexistence of species that are defensively divergent, increasing local plant species diversity (Becerra, 2007; Coley & Kursar, 2014; Fine et al., 2013; Forrister et al., 2019; Janzen, 1970; Salazar et al., 2016a, 2016b; Sedio & Ostling, 2013). Kursar et al. (2009) reported that co-occurring species of *Inga* in the Peruvian Amazon were more closely related yet differed more in their defences than expected by chance. Studies with other genera in the tropics reveal the same patterns (e.g. *Bursera*, *Ficus*, *Piper*, *Protium*, *Psychotria*; Becerra, 1997; Becerra et al., 2009; Coley & Kursar, 2014; Kursar et al., 2009; Salazar et al., 2018; Sedio, 2013; Wills et al., 2016). Because plants have many types of defences that can evolve independently from one another (Endara et al., 2017), defensive traits may provide a large number of niche dimensions among which a very large number of co-occurring species might sort in ecological time. Thus, plant–herbivore interactions may be key to understanding the high local diversity in tropical forest communities.

Relevant progress towards understanding the local and regional processes that underlie the assembly of communities has been made in recent years, though largely focused on the evolutionary attributes of species (phylogenetic history). These studies are based on the premise that historical species interactions and environmental conditions of communities are reflected in phylogenies, and that phylogeny is a good proxy for functional trait data that are difficult to obtain (Mace et al., 2003), especially at the large scale that is necessary for such studies. Yet, if phylogeny is only a proxy for species traits, and some traits may show low or no phylogenetic signal, an ideal approach would be to directly compare the explanatory power of traits and phylogeny (Pearse et al., 2014). Recent advances in analytical techniques have greatly enhanced the potential of researchers to characterize trait diversity at unprecedented scales. One such exciting new development is in the area of metabolomics. Specifically, mass spectrometry-based metabolomics is a powerful tool to characterize the chemical composition of complex biological samples containing tens to hundreds of individual compounds at the community or macroevolutionary scale (Sedio et al., 2017). In particular, tandem mass spectrometry (MS/MS) facilitates the structural comparison of unknown compounds and their comparison to global databases of known chemical structures (Treutler et al., 2016; Wang et al., 2016).

Here, we use metabolomic and phylogenetic approaches to investigate the patterns of assembly of plant communities across the whole Amazon basin. We focus our study on the speciose (>300 species), ecologically important and abundant Neotropical genus of trees, *Inga* (Leguminosae). Our previous studies with *Inga* show that defences diverge rapidly and that divergent defences may contribute to coexistence in neighbourhoods (~metres; Kursar et al., 2009). In this study, we examine community assembly at the regional (the Amazon basin) and local scale (within a site, ~100 ha) and build on previous work by incorporating a larger number of *Inga* species (37 in Kursar et al., 2009 versus. 91 in this study) collected over their entire geographic range, as well as a more resolved phylogeny and a more comprehensive chemistry dataset. Taken together, we aim to provide a more robust test of the ideas proposed by Kursar et al. (2009) and to extend the spatial scale from metres to kilometres.

At four widely separated sites, we characterize the chemical composition of 91 species, which represents roughly one third of known *Inga* species. We follow an untargeted approach that lets us obtain the entire chemical profile of a species rather than quantifying a subset of metabolites. In doing so, we can determine how many compounds are produced by each species and how many compounds are shared between them.

A critical component of our analyses is to determine the chemical similarity between all pairwise combinations of *Inga* species. However, this presents an apple/orange comparison challenge as few compounds are shared between species. We therefore have developed methods to account for the fact that two species may have different compounds that are structurally similar (Coley et al., 2019; Endara et al., 2018; Forrister et al., 2019). We join other ecological researchers pioneering metrics to classify chemical structure based on MS/MS spectra in order to quantify differences between species (Sedio et al., 2018). Our untargeted methods provide data on hundreds of compounds per species, and we can generate a matrix of MS/MS-based structural similarity between every pair of compounds (Wang et al., 2016), which can allow for a calculation of chemical similarity even when no compounds are shared between

a pair of species. This in turn allows us to better quantify both the chemical similarity among plant populations and to understand how plant–herbivore interactions may play a role in the assembly of plant communities. Specifically, we expect a lack of phylogenetic and chemical structure in the assembly of *Inga* communities at a regional scale, suggesting that the metacommunity for any regional community in the Amazon is the entire Amazon basin (Dexter et al., 2017). In contrast, the observation that the suite of *Inga* at a given local site is more overdispersed with respect to defences would suggest that local communities are assembled by ecological processes.

2 | MATERIALS AND METHODS

2.1 | Sampling

We sampled 91 *Inga* species across the Amazonia between July 2010 and September 2014. Sampling was focused at four sites (~100 ha each) that include a wide range of soils along with a large fraction of *Inga* diversity throughout the Amazon (Figure 1). At each site, we sampled all the known *Inga* species: Nouragues, French Guiana, 4°N 53°W, with 46 species; Tiputini in the Yasuní National Park, Ecuador, 0°N, 75°W, 41 species; Los Amigos in Madre de Dios, Peru, 13°S, 70°W, 39 species; and Km 41 near Manaus, Brazil, 2°S, 60°W, 29 species. The four sites are lowland moist forests with no pronounced dry season. For simplicity in the text, each site will be referred by the country only.

At each site, sampling was performed over 6 months and at the same time of the year. We focused on expanding leaves of 0.5–4 m tall understorey saplings, a key stage in the life cycle of a tree (Green et al., 2014). More than 40 km of trails were walked regularly to search for plants, and collections are widely separated. We focused our study on the chemical defences of young leaves because during this ephemeral stage they receive more than 75% of the herbivore damage accrued during the lifetime of a leaf (Brenes-Arguedas et al., 2008; Coley & Aide, 1991; Kursar & Coley, 2003), and the chemistry of expanding leaves has been shown to be very important for shaping associations between plants and their insect herbivores (Endara et al., 2017, 2018).

FIGURE 1 Map of study sites at (1) Nouragues, French Guiana, (2) Tiputini, Ecuador, (3) Los Amigos, Peru and (4) Manaus, Brazil. Size is weighted by the sample size (*n*) of *Inga* trees present at each site

2.2 | Phylogenetic reconstruction of *Inga*

A phylogenetic tree for 165 *Inga* accessions, including all the taxa sampled at each site, was reconstructed using a newly generated targeted enrichment (HybSeq) dataset of 810 genes. These 810 loci include those presented in Nicholls et al. (2015), supplemented with a subset of the loci from Koenen et al. (2020). DNA library preparation, sequencing and the informatics leading to final sequence alignments follow protocols in Nicholls et al. (2015). We used IQtree 2 (Minh et al., 2020) to infer a phylogenetic tree from the complete dataset of 810 genes. We performed a partitioned analysis (Chernomor et al., 2016) after inferring the best-partition scheme for the 810 genes and the best substitution model for each partition using ModelFinder module implemented in IQtree 2 (Kalyaanamoorthy et al., 2017). The resulting phylogenetic tree was subsequently time-calibrated using penalized likelihood implemented in the program treePL (Smith & O'Meara, 2012). We used cross-validation to estimate the best value of the smoothing parameter. We implemented a secondary calibration point on the crown age of *Inga* with a minimum age of 6 Ma and a maximum age of 10 Ma following previous estimates (Pennington et al., 2006; Richardson, 2001). Finally, the complete phylogeny was pruned to include only the 91 species for which chemistry data were available.

2.3 | Characterization of leaf defensive chemistry

2.3.1 | Secondary metabolites

For leaf defence analyses, expanding leaves were dried on silica gel at ambient temperature immediately after collection in the field, and then stored at −20°C. Samples consisted of whole leaves with little or no damage in order to control for potential defence induction, although induction is rare in tropical trees like *Inga* (Bixenmann et al., 2016). The defence metabolome for each species was determined using untargeted metabolomics methods. Defensive compounds were extracted from dried leaf samples in the Coley/Kursar lab at the University of Utah using a solution of (60:40, v/v) ammonium acetate buffered water, pH 4.8:acetonitrile, resulting in 2 ml of retained supernatant from 100mg (±2.5 mg) of sample for chromatographic analysis (Wiggins et al., 2016). Small molecules (50–2,000 Da) of intermediate polarity were analysed using ultraperformance liquid chromatography (Waters Acquity I-Class, 2.1 × 150 mm BEH C18 and 2.1 × 100 mm BEH Amide columns) and mass spectrometry (Waters Xevo G2 QToF) (UPLC-MS) in

negative ionization mode. Additionally, MS/MS spectra were acquired for each species by running DDA (Data Dependent Acquisition Mode), whereby MS/MS data were collected for all metabolites that were ionized above a set threshold (Total ion current/TIC of 5,000).

2.3.2 | L-Tyrosine

Some *Inga* species invest in the overexpression of the essential amino acid L-tyrosine as an effective chemical defence (Coley et al., 2019). Tyrosine is insoluble in our extraction buffer, so a different protocol was used to determine the percentage of leaf dry weight. Following Lokvam et al. (2006), extractable nitrogenous metabolites were extracted from a 5 mg subsample of each leaf using 1 ml of aqueous acetic acid (pH 3) for 1 hr at 85°C. Fifteen microlitres of the supernatant was injected on a 4.6 × 250 mm amino-propyl HPLC column (Microsorb 5u, Varian). Metabolites were chromatographed using a linear gradient (17%–23%) of aqueous acetic acid (pH 3.0) in acetonitrile over 25 min. The mass of solutes in each injection was measured using an evaporative light scattering detector (SEDERE S.A., Alfortville, France). Tyrosine concentrations were determined by reference to a 4-point standard curve (0.2–3.0 mg tyrosine/ml, $r^2 = 0.99$) prepared from pure tyrosine.

2.4 | Data analysis

We employed a compound-based molecular networking approach, where we first group related features into compounds and then we generate (a) a species-by-compound abundance matrix and (b) a compound-by-compound MS/MS cosine similarity matrix. We combine these data into a pairwise species similarity matrix, which accounts for both shared compounds between species and the MS/MS structural similarity of unshared compounds, following a similar approach to the one developed by Sedio et al. (2017). All scripts from this study are deposited in a github repository (Forrister & Soule, 2020; https://gitlab.chpc.utah.edu/01327245/evolution_of_inga_chemistry).

2.4.1 | Creation of species-by-compound matrix

Raw UPLC-MS data files were converted to mzXML format using the 'RAW2MZML' package in Python (Schmitt, 2016). Converted files were processed by species within each site (accession) and for MS level 1 peak detection using the XCMS package in R (Smith et al., 2006), which combined chromatographic features into features based on the mass/charge (m/z) ratio and retention time (RT) of individual ions. We then grouped features into putative compounds using CAMERA (Kuhl et al., 2012), which groups features that co-elute and have correlated abundance traces between scans, identifying likely adducts and related features within compounds. Finally, we removed from the analysis known contaminants and surfactants, as well as features with an abundance less than three times greater than the abundance of that feature in a blank (pure organic solvent).

After initial peak detection, features were aligned across accessions based on kernel density clustering of m/z and RT, and putative compounds grouped based on the cosine similarity of aligned feature abundance, resulting in a list of unique compounds across all samples. Here, abundance is considered the intensity or total ion current (TIC) for each feature. Each sample was then re-examined for all compounds to avoid data skewing during peak detection by accession. Finally, in an effort to remove temporal variance in UPLC-MS performance, compound abundance was normalized by the average abundance of a standard retention time index run the same day. This produced a data frame containing the normalized abundance of each compound within each sample, which was converted to a wide format to create a sample-by-compound matrix where the normalized abundance of each compound was assigned to a unique row (sample) and column (compound). In order to create a species-level comparison of compound abundance, all replicates (minimum of five) per accession were combined into a single species-level chemical profile by averaging the abundance of each compound across all replicates for a given species.

It is important to note that while we consider our method of grouping features into putative compounds to be fairly conservative, there remains the possibility of over- or under-splitting features into distinct compounds, with the former being more common. To address this issue in our method, the incorporation of MS/MS structural similarity (see Section 2.4.2) of distinct compounds allows the overall chemical similarity of samples (see Section 2.4.4) containing pseudo-replicated compounds to remain mathematically the same.

2.4.2 | Creation of compound-by-compound matrix

MS compounds (grouped chromatographic features) were matched to their associated MS/MS spectra based on the mz/RT of the parent ion isolated by DDA. A consensus MS/MS spectrum for each compound was generated by averaging all scans matched to that compound. A single MS/MS spectrum for each compound was then submitted to the Global Natural Products Social Molecular Networking in.mgf format (GNPS; https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp; Wang et al., 2016) for molecular networking. In R, the resulting network was used to create a pairwise compound-by-compound similarity matrix based on the similarity of their MS/MS fragmentation spectra. Here, the shortest through-network path between each compound pair was calculated, and a similarity score was assigned using the cosine scores along that path:

$$\text{Similarity}_{A,B} = \left(\sum_{1=}^{n} \frac{1}{i} \right)^{-1}, \quad (1)$$

where n is the number of edges separating compound A and compound B, and i is the cosine score of the current edge. The score ranges from 0 (completely dissimilar) to 1 (identical).

2.4.3 | Compound annotation

Our analysis yielded 6,217 compounds from 91 *Inga* species and one species in its sister genus, *Zygia mediana* (156 accessions including the same species from different sites). In order to annotate compounds, we performed MS/MS spectral matching to all publicly available datasets in GNPS as well as in silico fragmentation of the Universal Natural Products Database (Allard et al., 2016; Gu et al., 2013) and our own in-house database built from compounds found in *Inga* (Lokvam & Kursar, 2005). We further enumerated the library using in silico combinatorial chemistry to generate ~75,000 plausible structures using the 'scaffold' and 'building block' structures within the CLEVER application (Song et al., 2009). These enumeration structures were chosen based on the patterns of biosynthesis that we have observed in *Inga*. All compounds in this in silico database were uploaded to GNPS as a spectral library after performing in silico fragmentation using CFM-ID to predict MS/MS spectra (Allen et al., 2014). We also used Network Annotation Propagation (da Silva et al., 2018) to further annotate unknown compounds. Library hits and in silico prediction suggest that these compounds consist primarily of phenylpropanoids, flavonoid monomers, flavan-3-ol polymers and saponin glycosides, which are all classes known for their defensive function. These results confirm previous work done classifying *Inga* chemistry (Kursar et al., 2009).

2.4.4 | Chemical similarity between *Inga* species

Following Endara et al. (2018) with some modifications, we estimated chemical similarity between species using the species-by-compound and compound-by-compound matrices. After creating these matrices, compounds were grouped into saponins and phenolics based on m/z, RT and residual mass defect (RMD), and the species-by-compound matrix was separated based on this grouping. Abundances in each matrix were then normalized such that total abundance of all compounds in any given species was equal to 1.0.

Pairwise similarity for each species pair was calculated by quantifying the degree to which two species contain compounds that overlap in the molecular network. This includes the degree to which two species invest in the same compounds (species-by-compound), and the structural similarity of compounds that are not shared between the two species (compound-by-compound). These parameters are calculated as follows:

$$\begin{aligned}\text{Total similarity} = {} & \text{TIC overlap in shared compounds} \\ & + \text{similarity of unshared compounds}\end{aligned}$$

To calculate the TIC (abundance) overlap in shared compounds, the minimum TIC values for all compounds that are shared between the two samples are summed. The similarity of unshared compounds is calculated in a similar manner, by pairing the most similar compounds, taking the minimum TIC value for those two compounds, and multiplying by the through-network similarity score. For shared compounds, through-network similarity becomes mathematically obsolete as similarity for the same compound is always equal to 1. Thus, the overall similarity score results as a sum of the investment (TIC) in the same or structurally similar defences between two samples.

The pairwise similarity calculation for each species pair was repeated separately for phenolics and for saponins, resulting in a separate pairwise similarity matrix for each compound class. The similarity matrices from each compound class were combined with tyrosine data to produce an overall chemical similarity score for each sample pair according to the dry weight investment in each of the three compound classes. For further details, please review our gitlab repository (Forrister & Soule, 2020).

2.4.5 | Leaf defensive chemistry and phylogenetic signal

Phylogenetic signal was estimated for the principal coordinates of the chemical similarity matrix using Blomberg's *K* (Blomberg et al., 2003). *K* is close to 0 for traits lacking phylogenetic signal, but higher than 1 when close relatives are more similar than expected under the Brownian motion model of character evolution. We used the function *phylosignal* in the R package *picante* v.1.8.2 (Kembel et al., 2020).

2.4.6 | Analysis of community assembly

We analysed the assembly of *Inga* communities both at the local scale and at the level of the Amazon basin (regional scale, including the whole Amazon basin). Using incidence data (presence/absence), through a Bayesian approach with GLMM in the R package *MCMCglmm* v.2.29 (Hadfield, 2019), we determined patterns of the phylogenetic/chemical structure across all the assemblages simultaneously. We partitioned variance in the *Inga* species-by-site matrix into the effects of phylogenetic relatedness (termed phylogenetic effect) and chemical similarity between *Inga* species (a chemical effect). The magnitude of the effect of each term is determined by the magnitude of the variance associated with it. The phylogenetic effect determines the contribution of the main effect of the *Inga* phylogeny to the covariance and captures the variation in the *Inga* co-occurrence data explained by pairwise phylogenetic distances between *Inga* species. The chemical effect is the contribution of the main effect of *Inga* defensive chemistry to the covariance and captures the variation in the *Inga* co-occurrence data explained by the chemical similarity between *Inga* species. Thus, if the structuring of the communities is due to phylogenetic sorting, then the phylogenetic effect would show the greatest variance in the model. In contrast, if the assembly of *Inga* is mainly due to the occurrence of species with different chemistry, then the chemical effect would contribute the greatest to the variance in the model. Because the *Inga* occurrence data are collected from several sites across the

Amazon basin, rather than consolidate the data across sites, we analyse the site-specific incidence matrices as the geographic region information effect. In the model, this effect is termed Geographical region (see Table S1).

Phylogeny and chemistry were incorporated into the model as variance–covariance matrices of relatedness and similarity, respectively, in the random effect structure of the GLMM. Region effects were also fitted as random in the model. We compared models that included between-site effects (analyses at the level of the whole Amazon basin, as a random factor) with models that ignored between-site effects to assess patterns within sites (hence, analyses at small spatial scales). For the analyses, parameter-expanded priors were used for all variance components. The chain was run for 500,000 iterations with a burn-in of 50,000 and a thinning interval of 450. Because the response variable was incidence data, a Bernoulli error distribution was applied.

We also used classic dispersion metrics to determine whether a local *Inga* assemblage is a phylogenetically biased subset of the species that could coexist in that assemblage (Pearse et al., 2014). We estimated whether the mean pairwise distance (MPD, mean of the phylogenetic distance between all the members in a community) and the mean nearest taxon index (MNTD, mean of the phylogenetic distance between a species and its closest relative or neighbour in the community), were under- or overdispersed compared to the null expectation derived from a random assembly of same-size assemblages from the regional pool (Webb et al., 2002). To assess uncertainty, we repeated this process 9,999 times using the functions *ses.mpd* and *ses.mntd*, respectively, in the R package *picante* v.1.8.2 (Kembel et al., 2020).

Within-site chemical dissimilarity was estimated following Vleminckx et al. (2018). Observed dissimilarities between *Inga* species at each site were compared to the null expectation of a lack of divergence or convergence for trait expression. For this, the species-by-compound matrix (see above under the Chemical similarity between *Inga* species section) was randomized by reshuffling the compounds and species equiprobably, preserving differences in the abundance and presence/absence of compounds among species (Gotelli, 2000). Departure from the null expectation was estimated as the mean of the difference between the observed and expected dissimilarities between species at each site. This procedure was repeated 1,000 times. A *p*-value was obtained as the proportion of mean values above (overdispersion) or below (underdispersion) zero.

3 | RESULTS

3.1 | Leaf defensive chemistry in *Inga* shows low phylogenetic signal

We sampled young leaves from a minimum of five individual plants per species per site. A compound accumulation curve shows that five plants capture on average ~75% of the compounds encountered if more individuals are sampled (see Figure S1).

We determined chemical similarity between *Inga* species based on the similarity of chemical structure and relative abundance of compounds. In general, closely related species of *Inga* in the Amazonia tend to have different chemical defences. Principal coordinates of the chemical similarity matrix show low phylogenetic signal (PCO1 $K = 0.57$, $p = 0.001$; PCO2 $K = 0.28$, $p = 0.06$), with estimates of K that are substantially lower than the expected value of 1 under Brownian motion evolution.

3.2 | Low geographic signal of phylogeny and chemistry at regional scales

Because phylogeny is a poor predictor for chemistry in *Inga*, it was possible for us to separate the effect of chemistry and phylogeny in the analyses. Thus, we investigated the relative role of phylogeny and chemical defences against herbivores in the assembly of *Inga* communities at different scales. Our community structure models at the regional and local scales incorporating phylogenetic and chemical effects showed a differential role for both terms. At large spatial scales (models with between-site information), the phylogenetic effect was larger than the chemical effect, with 12% of the variation in the incidence of *Inga* species across the Amazon region attributed solely to phylogeny, versus 6% attributed to chemistry (Table S1, Figure 2). In fact, there is little regional selectivity based on chemistry, with all sites showing strong overlap in chemical space (Figure 3). Geographic information showed a large effect in the model (Table S1).

3.3 | Chemistry is more important than phylogeny at structuring local communities

To determine if chemistry or phylogeny influenced the assembly of species co-occurring at a single site, we fitted community-level

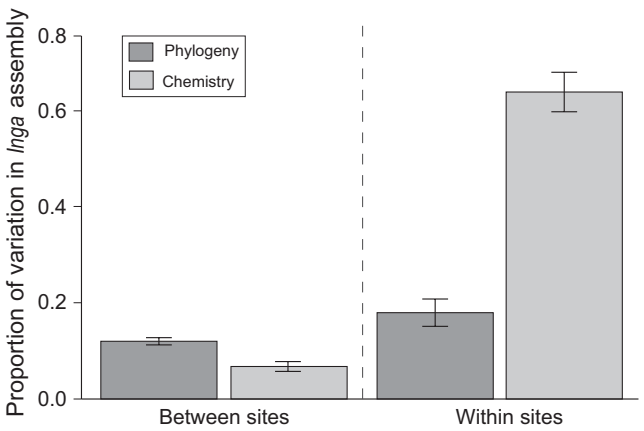

FIGURE 2 Proportion of variation in *Inga* community assembly explained by phylogeny and chemistry at the regional (between sites) and local (within sites) scales. Bars represent the mean ± the standard error

structure models at small spatial scales (without between-site information). There was some phylogenetic sorting, but the chemical effect contributed the greatest variation, with more than 60% of the *Inga* occurrence data explained by chemistry (Table S1, Figure 2). Thus, at small spatial scales, coexistence of *Inga* species is mainly due to the occurrence of species with dissimilar chemical defences.

We further evaluated phylogenetic structure within a community by estimating dispersion metrics and compared the observed values with a null expectation generated by randomly assembling same-size assemblages from the regional pool. None of the four Amazonian communities showed phylogenetic structure (Table 1, Figure 4).

In contrast, trait dispersion analyses showed significant chemical overdispersion for *Inga* communities in the Amazonia. When similarity in all chemical classes was considered, the chemical distance among all the *Inga* species within Peru, French Guiana and Ecuador is significantly larger than the null expectation (Table 2, Figure 5). This effect was maintained for phenolics and for saponins (except for Peru and French Guiana, Table 2). Brazil showed significant chemical overdispersion only for saponins (Table 2, Figure 5).

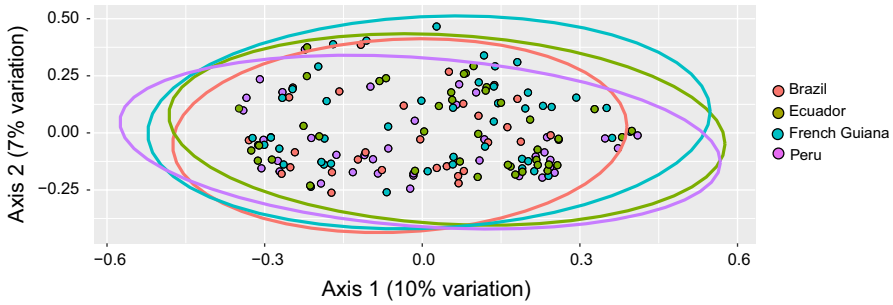

FIGURE 3 Principal coordinate analysis (PCoA) of chemical distance between species accessions (estimated as 1 minus our chemical similarity score). Accessions are coloured by site, and ellipsoids for each site represent the 95% confidence interval around their mean position in chemical space

Site	*N*	Metric	Observed	Observed Z	*p*-value
Peru	41	MPD	0.0334	−0.1683	0.471
		MNTD	0.0152	0.5350	0.710
French Guiana	43	MPD	0.0347	0.5464	0.713
		MNTD	0.0161	1.5045	0.929
Ecuador	41	MPD	0.0348	0.6032	0.735
		MNTD	0.0163	1.3836	0.913
Brazil	29	MPD	0.0336	−0.0603	0.535
		MNTD	0.0163	0.0686	0.534

TABLE 1 Results for the phylogenetic community structure analyses for each *Inga* community sampled. *N* = number of taxa in each community. MPD = mean pairwise distance, MNTD = mean nearest taxon distance. Observed Z = standardized effect size of mean pairwise distance versus null model

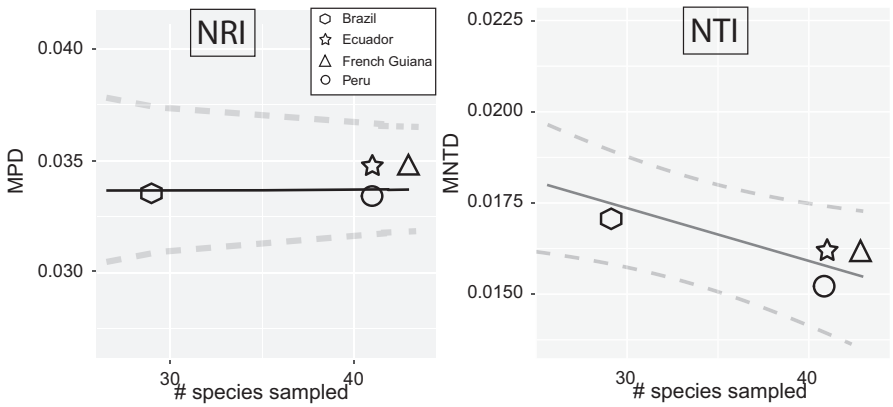

FIGURE 4 Relationship between the number of *Inga* species sampled and the mean pairwise phylogenetic distance (MPD, left), and the mean nearest taxon distance (MNTD, right) in the Amazon. Solid line represents the null expectation for MPD and MNTD respectively. Dotted line represents the 95% confidence interval of the null expectation. Sites are differentiated by shape, and significance is indicated by shading (no fill = non-significant, grey = $p < 0.10$)

TABLE 2 Within-site chemical dissimilarity analyses separated by compound class. Significant values are bolded. O–E indicates the average difference in observed chemical similarity values compared to a randomized null model

Site	Compound Class	O–E	95% CI (lower)	95% CI (upper)	*p*-value
Peru	All Chemistry	**0.061**	**0.05300958**	**0.070**	**2.20E−16**
	Phenolics	**0.063**	**0.05517029**	**0.072**	**2.20E−16**
	Saponins	**−0.008**	**−0.0180129614**	**0.000**	**0.05**
Brazil	All Chemistry	0.003	−0.009744713	0.016	0.6088
	Phenolics	**−0.39**	**−0.3999554**	**−0.380**	**2.20E−16**
	Saponins	**0.04**	**0.03395981**	**0.055**	**2.10E−15**
French Guiana	All Chemistry	**0.021**	**0.01112708**	**0.031**	**4.29E−05**
	Phenolics	**0.150**	**0.1409909**	**0.160**	**2.20E−16**
	Saponins	**−0.119**	**−0.1288449**	**−0.110**	**2.20E−16**
Ecuador	All Chemistry	**0.131**	**0.1218379**	**0.140**	**2.20E−16**
	Phenolics	**0.248**	**0.2385003**	**0.258**	**2.20E−16**
	Saponins	**0.070**	**0.05877019**	**0.083**	**2.20E−16**

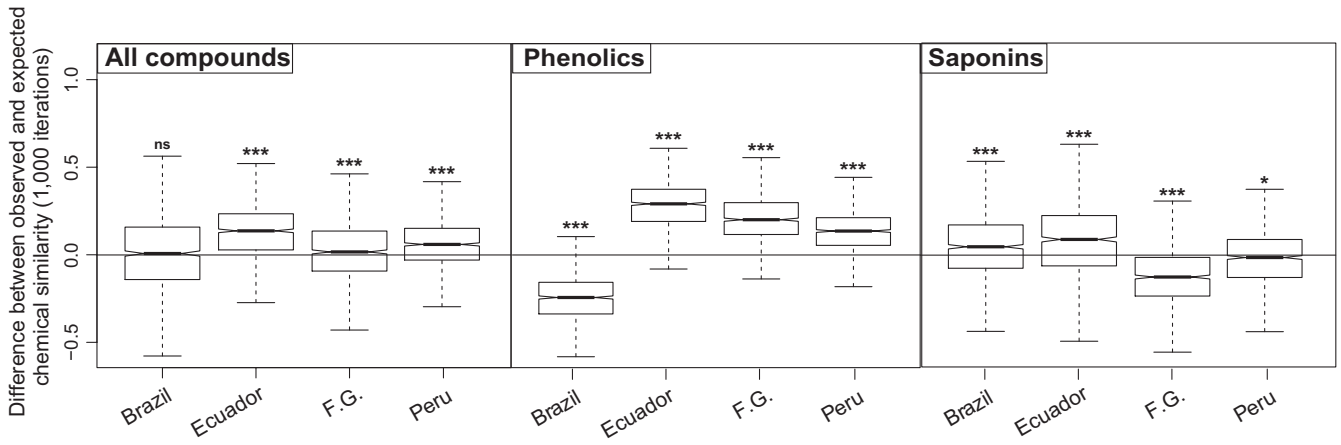

FIGURE 5 Within-site chemical similarity analysis. Boxplots represent 1,000 bootstrap iterations of the difference between observed (real data) and expected (null model) chemical similarity values at each site, separated by compound class. Significance is indicated by asterisks (ns = non-significant; $^{*}p < 0.05$; $^{***}p < 0.001$). All country names excepting French Guiana (F. G.) are spelled out

4 | DISCUSSION

We have argued that at a regional level, there is essentially no limitation in the dispersal of species across the Amazon such that the metacommunity for any regional community is the entire Amazon basin (Dexter et al., 2017). In contrast, interactions between plants and insects could be a principal mechanism structuring community assembly at a local scale (Coley & Kursar, 2014; Kursar et al., 2009). Results from our analyses are consistent with these hypotheses. At a large scale, we found a lack of chemical structure in the assembly of *Inga* communities, with low, but significant geographic filtering based on ancestry. In contrast, at each of four widely separated sites in the Amazon, co-occurring species of *Inga* are more different in defence chemistry than expected by chance, implying that species with similar defensive traits are less likely to coexist in the same community. Thus, herbivores may have a key role in niche differentiation of their host plants promoting local diversity.

4.1 | Low geographic signal for phylogeny and chemistry at regional scales

Consistent with the hypothesis that regional tree communities in the Amazon are influenced by historical processes of widespread dispersal (Dexter et al., 2017), we found a low signal for phylogeny and almost no signal for chemistry in the assembly of *Inga* communities across the Amazon when between-region information was included (Table S1, Figure 2). Nevertheless, geographic region had a large effect in the model, implying that biogeography might play a role in *Inga* community assembly at regional level (Table S1). Thus, although *Inga* lineages have dispersed repeatedly across the Amazon (Dexter et al., 2017), the detected signal of regional phylogenetic structure together with the geographic region term effect imply that closely related species might be co-occurring within some regions, and that there are some differences in the lineage composition between regions. These differences could be mediated by environmental filtering at regional scale, such as the gradient in soils observed across the

Amazon basin (Tuomisto et al., 2019). For chemistry, the extremely low signal in the assembly of *Inga* communities at the regional level (Figure 2) suggests that local assemblages are drawn from a metacommunity representing the full chemical space exhibited by the genus (Figure 3).

4.2 | Chemistry not phylogeny structures local communities

In contrast to regional patterns, analyses of community structure at a local scale showed that chemistry better explained variation in the incidence of *Inga* at a single site than plant relatedness (Figure 2). Thus, defensive chemistry plays a key role in determining which plant species can coexist in each community at small spatial scales. Analyses with phylogenetic dispersion metrics and within-site functional similarity agreed with this hypothesis. Although our community composition models suggest a degree of phylogenetic sorting in species composition (Table S1), dispersion-trait analyses for the four *Inga* communities sampled showed no significant phylogenetic clustering (Table 1, Figure 4). Meanwhile, the species of *Inga* that are co-occurring in Peru, French Guiana, Ecuador and Brazil are more different in their defensive chemistry than expected by chance (Table 2, Figure 5). Except for Brazil, this effect was more pronounced for phenolics than for saponins (Table 2, Figure 5). Phenolics are the most structurally diverse and common compound class for the genus *Inga* (D. Forrister, unpublished results), which is the most divergent among close relatives (Endara et al., 2015). This suggests that phenolics might be under stronger selective pressure to diverge among co-occurring species than other defence classes or that phenol biosynthesis is more easily modified. Given that for *Inga*, each defence class varies independently of the others (Endara et al., 2017), defensive chemistry may represent many axes of trait divergence.

Interactions of plant species with their enemies are likely the mechanism responsible for the co-occurrence of species with divergent chemotypes. Specialist herbivores might be foraging on species with similar defensive chemotypes. Within a site, this would allow defensively distinct species to coexist and increase local plant diversity (Sedio & Ostling, 2013). In contrast, species with similar defences may share herbivores and suffer greater attack, making it more difficult for them to colonize or to coexist in the same community. Thus, herbivores might be regulating the structure of communities through negative-density dependence interactions at scales ranging from metres to kilometres (Becerra, 2007; Forrister et al., 2019; Strauss & Lau, 2008), linking local systems to regional processes (Ricklefs, 2007).

An essential component of this proposition is that plant defences influence host choice. Previously, we found that at a given site, lepidopteran herbivores preferentially feed on subsets of *Inga* species with similar defensive profiles and that different families of herbivores chose hosts based on different defensive traits (Endara et al., 2017). In addition, we have shown that high chemical similarity and shared herbivore communities are associated with a decrease in survival and growth for neighbouring plants at the 5- to 10-m scale (Forrister et al., 2019). In this study, we provide evidence that the antagonistic interactions with enemies are playing out across the entire community, not just spatially proximal neighbours. Thus, the composition of plant species within a community appears to respond to the entire community of herbivores that could potentially attack them.

Because phylogeny is a synthetic measure for phylogenetically conserved traits, the low phylogenetic structure in *Inga* at four widely separated communities suggests that other mechanisms than herbivore pressure might not be contributing as much to their assembly. For example, phylogenetically conserved traits associated with resource use, pollination and dispersal are quite similar across *Inga* species (Endara et al., 2015; Kursar et al., 2009; Pennington et al., 1997). Thus, it is hard to see how they would provide sufficient niche differentiation to explain the coexistence of so many species. Alternatively, if we consider the almost infinite number of possible defence profiles, there could be an enormous number of niches with respect to herbivores (Coley & Kursar, 2014; Levi et al., 2019; Singer & Stireman, 2005). For *Inga*, anti-herbivore defences fall into at least six different independent axes of defence variation (Endara et al., 2017). It clearly provides a multidimensional, if not hyperdimensional niche space for coexistence (Hutchinson, 1957).

Are there parallels in other tropical regions? Several studies have shown that neighbours growing within metres of each other differ in defences, including the genera *Eugenia*, *Ocotea* and *Psychotria* in Panama (Sedio et al., 2017), *Bursera* in Mexico (Becerra, 2007), *Piper* in Costa Rica (Salazar et al., 2016a, 2016b) and *Protium* in Peru (Vleminckx et al., 2018). Here we extend this concept and show that these patterns of defence divergence hold true across a much larger community of plants, not just immediate neighbours. It is quite striking that these patterns are consistent even when we included in our analyses the *Inga* community in Panama, a site with a different biogeographic history that is isolated from the Amazonian study sites (data not shown). Similarly, community structure and trait dispersion analyses showed significant overdispersion of defensive chemistry at the local scale (Figures S2 and S3). Thus, the similarity of secondary metabolite profiles among species may play a large role in shaping community assembly beyond the tropical forest in Amazonia.

5 | CONCLUSIONS

A number of recent, independent studies suggest that herbivore pressure contributes to the high local plant diversity, or coexistence, that is typical of plant communities in tropical rainforests (Becerra, 2007; Forrister et al., 2019; Kursar et al., 2009; Salazar et al., 2016a, 2016b; Sedio et al., 2017; Vleminckx et al., 2018). Our phylogenetic and metabolomic approach provides evidence for the key role that natural enemies play in the assembly of these local communities. Although *Inga* species have dispersed freely across the Amazon, with some recent regional in-situ speciation events, what

seems to determine which species are allowed to coexist within a single community are natural enemies.

Our results expand the spatial scale over which negative-density dependence mechanisms mediate community assembly and bring into play processes related to ecological interactions between populations at larger spatial scales. The fact that coexistence of closely related species is allowed by divergence in defensive traits on scales ranging from metres to kilometres brings the time-scale of species sorting and species diversification close to each other (Ricklefs, 2007). This leads us to hypothesize that herbivore pressure might be one of the drivers of species diversification. Thus, divergent selection by herbivores could potentially be one of the main factors behind both the maintenance and the origin of diversity in tropical forests.

ACKNOWLEDGEMENTS

The authors thank the Ministry of Environment and Water of Ecuador and the Ministry of Agriculture of Peru for granting the research and exportation permits. Valuable field assistance was provided by Wilder Hidalgo, Zachary Benavidez, Allison Thompson, Yamara Serrano and Mayra Ninazunta. This work was supported by grants from the National Science Foundation (DEB-0640630 and DIMENSIONS of Biodiversity DEB-1135733) and Nouragues Travel Grants Program, CNRS, France, to T.A.K. and P.D.C., and the Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación del Ecuador (SENESCYT) to M.-J.E.

AUTHORS' CONTRIBUTIONS

M.-J.E., D.L.F., T.A.K. and P.D.C. designed and conducted the research; M.-J.E. designed and performed the data analysis; D.L.F. and A.J.S. contributed to the metabolomic analysis; J.A.N., R.T.P., K.G.D. and O.L. contributed the phylogeny of *Inga*; M.-J.E., D.L.F., A.J.S. and P.D.C. wrote the manuscript, with input from K.G.D., O.L. and R.T.P.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/1365-2745.13646.

DATA AVAILABILITY STATEMENT

Chemical data and scripts to estimate chemical similarity are deposited in a gitlab repository (Forrister & Soule, 2020; https://gitlab.chpc.utah.edu/01327245/evolution_of_inga_chemistry).

ORCID

María-José Endara https://orcid.org/0000-0002-8805-1456
Abrianna J. Soule https://orcid.org/0000-0001-9164-7317
Dale L. Forrister https://orcid.org/0000-0001-8170-7187
Kyle G. Dexter https://orcid.org/0000-0001-9232-5221
R. Toby Pennington https://orcid.org/0000-0002-8196-288X
James A. Nicholls https://orcid.org/0000-0002-9325-563X
Oriane Loiseau https://orcid.org/0000-0002-9852-857X
Phyllis D. Coley https://orcid.org/0000-0001-7718-597X

REFERENCES

Allard, P.-M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., Pham, V. C., Roussi, F., Litaudon, M., & Wolfender, J.-L. (2016). Integration of molecular networking and *in* silico MS/MS fragmentation for natural products dereplication. *Analytical Chemistry*, *88*(6), 3317–3323. https://doi.org/10.1021/acs.analchem.5b04804

Allen, F., Pon, A., Wilson, M., Greiner, R., & Wishart, D. (2014). CFM-ID: A web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Research*, *42*(W1), W94–W99. https://doi.org/10.1093/nar/gku436

Becerra, J. X. (1997). Insects on plants: Macroevolutionary chemical trends in host use. *Science*, *276*(5310), 253–256. https://doi.org/10.1126/science.276.5310.253

Becerra, J. X. (2007). The impact of herbivore–plant coevolution on plant community structure. *Proceedings of the National Academy of Sciences of the United States of. America*, *104*(18), 7483–7488. https://doi.org/10.1073/pnas.0608253104

Becerra, J. X., Noge, K., & Venable, D. L. (2009). Macroevolutionary chemical escalation in an ancient plant-herbivore arms race. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(43), 18062–18066. https://doi.org/10.1073/pnas.0904456106

Bixenmann, R. J., Coley, P. D., Weinhold, A., & Kursar, T. A. (2016). High herbivore pressure favors constitutive over induced defense. *Ecology and Evolution*, *6*(17), 6037–6049. https://doi.org/10.1002/ece3.2208

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, *57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Brenes-Arguedas, T., Rios, M., Rivas-Torres, G., Blundo, C., Coley, P. D., & Kursar, T. A. (2008). The effect of soil on the growth performance of tropical species with contrasting distributions. *Oikos*, *117*, 1453–1460. https://doi.org/10.1111/j.2008.0030-1299.16903.x

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Chesson, P. (2000). General Theory of Competitive Coexistence in spatially-varying environments. *Theoretical Population Biology*, *58*(3), 211–237. https://doi.org/10.1006/tpbi.2000.1486

Coley, P. D., & Aide, T. M. (1991). Comparison of herbivory and plant defenses in temperate and tropical broad-leaved forests. In P. W. Price, T. M. Lewinsohn, G. W. Fernandes, & W. W. Benson (Eds.), *Plant-animal interactions: Evolutionary ecology in tropical and temperate regions* (pp. 25–49). Wiley & Sons.

Coley, P. D., Endara, M., Ghabash, G., Kidner, C. A., Nicholls, J. A., Pennington, R. T., Mills, A. G., Soule, A. J., Lemes, M. R., Stone, G. N., & Kursar, T. A. (2019). Macroevolutionary patterns in overexpression of tyrosine: An anti-herbivore defence in a speciose tropical tree genus, *Inga* (Fabaceae). *Journal of Ecology*, *107*(4), 1620–1632. https://doi.org/10.1111/1365-2745.13208

Coley, P. D., & Kursar, T. A. (2014). On tropical forests and their pests. *Science*, *343*(6166), 35–36. https://doi.org/10.1126/science.1248110

Comita, L. S., Queenborough, S. A., Murphy, S. J., Eck, J. L., Xu, K., Krishnadas, M., Beckman, N., & Zhu, Y. (2014). Testing predictions of the Janzen-Connell hypothesis: A meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *Journal of Ecology*, *102*(4), 845–856. https://doi.org/10.1111/1365-2745.12232

Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rainforest trees. In P. J. Den Boer, & G. R. Gradwell (Eds.), *Dynamics of populations* (pp. 298–312). Centre for Agricultural Publishing and Documentations.

da Silva, R. R., Wang, M., Nothias, L.-F., van der Hooft, J. J. J., Caraballo-Rodríguez, A. M., Fox, E., Balunas, M. J., Klassen, J. L., Lopes, N. P., & Dorrestein, P. C. (2018). Propagating

annotations of molecular networks using *in silico* fragmentation. *PLOS Computational Biology*, *14*(4), e1006089. https://doi.org/10.1371/journal.pcbi.1006089

Dexter, K. G., Lavin, M., Torke, B. M., Twyford, A. D., Kursar, T. A., Coley, P. D., Drake, C., Hollands, R., & Pennington, R. T. (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*(10), 2645–2650. https://doi.org/10.1073/pnas.1613655114

Endara, M.-J., Coley, P. D., Ghabash, G., Nicholls, J. A., Dexter, K. G., Donoso, D. A., Stone, G. N., Pennington, R. T., & Kursar, T. A. (2017). Coevolutionary arms race versus host defense chase in a tropical herbivore–plant system. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(36), E7499–E7505. https://doi.org/10.1073/pnas.1707727114

Endara, M.-J., Nicholls, J. A., Coley, P. D., Forrister, D. L., Younkin, G. C., Dexter, K. G., Kidner, C. A., Pennington, R. T., Stone, G. N., & Kursar, T. A. (2018). Tracking of host defenses and phylogeny during the radiation of neotropical *Inga*-feeding sawflies (Hymenoptera; Argidae). *Frontiers in Plant Science*, *9*, 1237. https://doi.org/10.3389/fpls.2018.01237

Endara, M.-J., Weinhold, A., Cox, J. E., Wiggins, N. L., Coley, P. D., & Kursar, T. A. (2015). Divergent evolution in antiherbivore defences within species complexes at a single Amazonian site. *Journal of Ecology*, *103*(5), 1107–1118. https://doi.org/10.1111/1365-2745.12431

Fine, P. V. A., Metz, M. R., Lokvam, J., Mesones, I., Zuñiga, J. M. A., Lamarre, G. P. A., Pilco, M. V., & Baraloto, C. (2013). Insect herbivores, chemical innovation, and the evolution of habitat specialization in Amazonian trees. *Ecology*, *94*(8), 1764–1775. https://doi.org/10.1890/12-1920.1

Fine, P. V. A., Zapata, F., & Daly, D. C. (2014). Investigating processes of neotropical rainforest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae): Evolution and historical biogeography of Protieae. *Evolution*, *68*(7), 1988–2004. https://doi.org/10.1111/evo.12414

Forrister, D. L., Endara, M.-J., Younkin, G. C., Coley, P. D., & Kursar, T. A. (2019). Herbivores as drivers of negative density dependence in tropical forest saplings. *Science*, *363*(6432), 1213–1216. https://doi.org/10.1126/science.aau9460

Forrister, D. L., & Soule, A. J. (2020). *Evolution of Inga Chemistry* [Gitlab Repository]. Retrieved from https://gitlab.chpc.utah.edu/01327245/evolution_of_inga_chemistry

Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, *81*(9), 2606–2621. https://doi.org/10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2

Green, P. T., Harms, K. E., & Connell, J. H. (2014). Nonrandom, diversifying processes are disproportionately strong in the smallest size classes of a tropical forest. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(52), 18649–18654. https://doi.org/10.1073/pnas.1321892112

Gu, J., Gui, Y., Chen, L., Yuan, G., Lu, H.-Z., & Xu, X. (2013). Use of natural products as chemical library for drug discovery and network pharmacology. *PLoS One*, *8*(4), e62839. https://doi.org/10.1371/journal.pone.0062839

Hadfield, J. (2019). *R Package MCMCglmm: MCMC Generalised Linear Mixed Models* (2.29) [R].

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Janzen, D. H. (1970). Herbivores and the number of trees in tropical forests. *The American Naturalist*, *104*(940), 501–529.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kembel, S. W., Ackerly, D. D., Blomberg, S. P., Cornwell, W. K., Cowan, P. D., Helmus, M. R., Morlon, H., & Webb, C. O. (2020). *picante: Integrating phylogenies and ecology* (1.8.1) [R]. Retrieved from https://cran.r-project.org/web/packages/picante/picante.pdf

Koenen, E. J. M., Kidner, C., Souza, É. R., Simon, M. F., Iganci, J. R., Nicholls, J. A., Brown, G. K., Queiroz, L. P., Luckow, M., Lewis, G. P., Pennington, R. T., & Hughes, C. E. (2020). Hybrid capture of 964 nuclear genes resolves evolutionary relationships in the mimosoid legumes and reveals the polytomous origins of a large pantropical radiation. *American Journal of Botany*, *107*(12), 1710–1735. https://doi.org/10.1002/ajb2.1568

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, *29*(5), 592–599. https://doi.org/10.1111/1365-2435.12345

Kraft, N. J. B., Godoy, O., & Levine, J. M. (2015). Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(3), 797–802. https://doi.org/10.1073/pnas.1413650112

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., & Neumann, S. (2012). CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Analytical Chemistry*, *84*(1), 283–289. https://doi.org/10.1021/ac202450g

Kursar, T. A., & Coley, P. D. (2003). Convergence in defense syndromes of young leaves in tropical rainforests. *Biochemical Systematics and Ecology*, *31*(8), 929–949. https://doi.org/10.1016/S0305-1978(03)00087-5

Kursar, T. A., Dexter, K. G., Lokvam, J., Pennington, R. T., Richardson, J. E., Weber, M. G., Murakami, E. T., Drake, C., McGregor, R., & Coley, P. D. (2009). The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(43), 18073–18078. https://doi.org/10.1073/pnas.0904786106

Levi, T., Barfield, M., Barrantes, S., Sullivan, C., Holt, R. D., & Terborgh, J. (2019). Tropical forests can maintain hyperdiversity because of enemies. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(2), 581–586. https://doi.org/10.1073/pnas.1813211116

Lokvam, J., Brenes-Arguedas, T., Lee, J. S., Coley, P. D., & Kursar, T. A. (2006). Allelochemic function for a primary metabolite: The case of L-tyrosine hyper-production in *Inga umbellifera* (Fabaceae). *American Journal of Botany*, *93*(8), 1109–1115. https://doi.org/10.3732/ajb.93.8.1109

Lokvam, J., & Kursar, T. A. (2005). Divergence in structure and activity of phenolic defenses in young leaves of two co-occurring Inga species. *Journal of Chemical Ecology*, *31*(11), 2563–2580. https://doi.org/10.1007/s10886-005-7614-x

Mace, E. S., Buhariwalla, K. K., Buhariwalla, H. K., & Crouch, J. H. (2003). A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Molecular Biology Reporter*, *21*(4), 459–460. https://doi.org/10.1007/BF02772596

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Nicholls, J. A., Pennington, R. T., Koenen, E. J. M., Hughes, C. E., Hearn, J., Bunnefeld, L., Dexter, K. G., Stone, G. N., & Kidner, C. A. (2015). Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2015.00710

Pearse, W. D., Purvis, A., Cavender-Bares, J., & Helmus, M. R. (2014). Metrics and models of community phylogenetics. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 451–464). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43550-2_19

Pennington, R. T., Richardson, J. E., & Lavin, M. (2006). Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist*, *172*(4), 605–616. https://doi.org/10.1111/j.1469-8137.2006.01902.x

Pennington, T. D., Revelo, N., Linklater, R., & Wise, R. (1997). *El género Inga en el Ecuador: Morfología, distribución y usos*. Royal Botanic Gardens.

Richardson, J. E. (2001). Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science*, *293*(5538), 2242–2245. https://doi.org/10.1126/science.1061421

Ricklefs, R. E. (2007). History and diversity: Explorations at the intersection of ecology and evolution. *The American Naturalist*, *170*(S2), S56–S70. https://doi.org/10.1086/519402

Salazar, D., Jaramillo, A., & Marquis, R. J. (2016a). The impact of plant chemical diversity on plant–herbivore interactions at the community level. *Oecologia*, *181*(4), 1199–1208. https://doi.org/10.1007/s00442-016-3629-y

Salazar, D., Jaramillo, M. A., & Marquis, R. J. (2016b). Chemical similarity and local community assembly in the species rich tropical genus *Piper*. *Ecology*, *97*(11), 3176–3183. https://doi.org/10.1002/ecy.1536

Salazar, D., Lokvam, J., Mesones, I., Vásquez Pilco, M., Ayarza Zuñiga, J. M., de Valpine, P., & Fine, P. V. A. (2018). Origin and maintenance of chemical diversity in a species-rich tropical tree lineage. *Nature Ecology & Evolution*, *2*(6), 983–990. https://doi.org/10.1038/s41559-018-0552-0

Schmitt, U. (2016). *Raw2mzML* (0.0.3) [Python]. Retrieved from https://pypi.org/project/raw2mzml/#description

Sedio, B. E. (2013). *Trait evolution and species coexistence in the hyperdiverse tropical forest genus* Psychotria [PhD]. University of Michigan.

Sedio, B. E., Boya P., C. A., & Rojas Echeverri, J. C. (2018). A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. *Applications in Plant Sciences*, *6*(3), e1033. https://doi.org/10.1002/aps3.1033

Sedio, B. E., & Ostling, A. M. (2013). How specialised must natural enemies be to facilitate coexistence among plants? *Ecology Letters*, *16*(8), 995–1003. https://doi.org/10.1111/ele.12130

Sedio, B. E., Rojas Echeverri, J. C., Boya P., C. A., & Wright, S. J. (2017). Sources of variation in foliar secondary chemistry in a tropical forest tree community. *Ecology*, *98*(3), 616–623. https://doi.org/10.1002/ecy.1689

Singer, M. S., & Stireman, J. O. (2005). The tri-trophic niche concept and adaptive radiation of phytophagous insects. *Ecology Letters*, *8*(12), 1247–1255. https://doi.org/10.1111/j.1461-0248.2005.00835.x

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry*, *78*(3), 779–787. https://doi.org/10.1021/ac051437y

Smith, S. A., & O'Meara, B. C. (2012). treePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics*, *28*(20), 2689–2690. https://doi.org/10.1093/bioinformatics/bts492

Song, C. M., Bernardo, P. H., Chai, C. L. L., & Tong, J. C. (2009). CLEVER: Pipeline for designing *in silico* chemical libraries. *Journal of Molecular Graphics and Modelling*, *27*(5), 578–583. https://doi.org/10.1016/j.jmgm.2008.09.009

Strauss, S. Y., Lau, J. A., Schoener, T. W., & Tiffin, P. (2008). Evolution in ecological field experiments: Implications for effect size. *Ecology Letters*, *11*(3), 199–207. https://doi.org/10.1111/j.1461-0248.2007.01128.x

ter Steege, H., Prado, P. I., Lima, R. A. F. D., Pos, E., de Souza Coelho, L., de Andrade Lima Filho, D., Salomão, R. P., Amaral, I. L., de Almeida Matos, F. D., Castilho, C. V., Phillips, O. L., Guevara, J. E., de Jesus Veiga Carim, M., Cárdenas López, D., Magnusson, W. E., Wittmann, F., Martins, M. P., Sabatier, D., Irume, M. V., … Pickavance, G. (2020). Biased-corrected richness estimates for the Amazonian tree flora. *Scientific Reports*, *10*(1), 10130. https://doi.org/10.1038/s41598-020-66686-3

Treutler, H., Tsugawa, H., Porzel, A., Gorzolka, K., Tissier, A., Neumann, S., & Balcke, G. U. (2016). Discovering regulated metabolite families in untargeted metabolomics studies. *Analytical Chemistry*, *88*(16), 8082–8090. https://doi.org/10.1021/acs.analchem.6b01569

Tuomisto, H., Van doninck, J., Ruokolainen, K., Moulatlet, G. M., Figueiredo, F. O. G., Sirén, A., Cárdenas, G., Lehtonen, S., & Zuquim, G. (2019). Discovering floristic and geoecological gradients across Amazonia. *Journal of Biogeography*, *46*(8), 1734–1748. https://doi.org/10.1111/jbi.13627

Valencia, R., Foster, R. B., Villa, G., Condit, R., Svenning, J.-C., Hernandez, C., Romoleroux, K., Losos, E., Magard, E., & Balslev, H. (2004). Tree species distributions and local habitat variation in the Amazon: Large forest plot in eastern Ecuador. *Journal of Ecology*, *92*(2), 214–229. https://doi.org/10.1111/j.0022-0477.2004.00876.x

Vleminckx, J., Salazar, D., Fortunel, C., Mesones, I., Dávila, N., Lokvam, J., Beckley, K., Baraloto, C., & Fine, P. V. A. (2018). Divergent secondary metabolites and habitat filtering both contribute to tree species coexistence in the Peruvian Amazon. *Frontiers in Plant Science*, *9*, 836. https://doi.org/10.3389/fpls.2018.00836

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., … Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837. https://doi.org/10.1038/nbt.3597

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*(1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Wiggins, N. L., Forrister, D. L., Endara, M.-J., Coley, P. D., & Kursar, T. A. (2016). Quantitative and qualitative shifts in defensive metabolites define chemical defense investment during leaf development in *Inga*, a genus of tropical trees. *Ecology and Evolution*, *6*(2), 478–492. https://doi.org/10.1002/ece3.1896

Wills, C., Harms, K. E., Wiegand, T., Punchi-Manage, R., Gilbert, G. S., Erickson, D., Kress, W. J., Hubbell, S. P., Gunatilleke, C. V. S., & Gunatilleke, I. A. U. N. (2016). Persistence of neighborhood demographic influences over long phylogenetic distances may help drive post-speciation adaptation in tropical forests. *PLoS One*, *11*(6), e0156913. https://doi.org/10.1371/journal.pone.0156913

Yguel, B., Bailey, R., Tosh, N. D., Vialatte, A., Vasseur, C., Vitrac, X., Jean, F., & Prinzing, A. (2011). Phytophagy on phylogenetically isolated trees: Why hosts should escape their relatives. *Ecology Letters*, *14*(11), 1117–1124. https://doi.org/10.1111/j.1461-0248.2011.01680.x

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Endara M-J, Soule AJ, Forrister DL, et al. The role of plant secondary metabolites in shaping regional and local plant community assembly. *J Ecol*. 2022;110:34–45. https://doi.org/10.1111/1365-2745.13646