Received: 19 May 2023 | Accepted: 21 November 2023

DOI: 10.1111/1365-2745.14255

Journal of Ecology
BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Functional genomics and co-occurrence in a diverse tropical tree genus: The roles of drought- and defence-related genes

Jie Yang[1,2] | Huan Fan[3] | Yunyun He[1,4] | Gang Wang[1] | Min Cao[1] | Nathan G. Swenson[5]

[1]CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, China

[2]National Forest Ecosystem Research Station at Xishuangbanna, Mengla, China

[3]Centre for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, China

[4]University of Chinese Academy Sciences, Beijing, China

[5]Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA

Correspondence
Nathan G. Swenson
Email: nswenson@nd.edu

Funding information
Chinese Academy of Sciences Youth Innovation Promotion Association, Grant/Award Number: Y202080; Distinguished Youth Scholar of Yunnan, Grant/Award Number: 202001AV070016; West Light Foundation of the Chinese Academy of Sciences; Ten Thousand Talent Plans for Young Top-Notch Talents of Yunnan Province, Grant/Award Number: YNWR-QNBJ-2018-309; National Natural Science Foundation of China, Grant/Award Number: 31870410; 14th Five-Year Plan of the Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Grant/Award Number: E3ZKFF1K; International Postdoctoral Exchange Fellowship Program and Yunnan Postdoc Directed Training Program; NSF US-China Dimensions of Biodiversity, Grant/Award Number: DEB 2124466 and DEB 32061123003

Handling Editor: Peter Vesk

Abstract

1. Tropical tree communities are among the most diverse in the world. A small number of genera often disproportionately contribute to this diversity. How so many species from a single genus can co-occur represents a major outstanding question in biology. Niche differences are likely to play a major role in promoting congeneric diversity, but the mechanisms of interest are often not well-characterized by the set of functional traits generally measured by ecologists.
2. To address this knowledge gap, we used a functional genomic approach to investigate the mechanisms of co-occurrence in the hyper-diverse genus *Ficus*. Our study focused on over 800 genes related to drought and defence, providing detailed information on how these genes may contribute to the diversity of *Ficus* species.
3. We find widespread and consistent evidence of the importance of defence gene dissimilarity in co-occurring species, providing genetic support for what would be expected under the Janzen-Connell mechanism. We also find that drought-related gene sequence similarity is related to *Ficus* co-occurrence, indicating that similar responses to drought promote co-occurrence.
4. *Synthesis*. We provide the first detailed functional genomic evidence of how drought- and defence-related genes simultaneously contribute to the local co-occurrence in a hyper-diverse genus. Our results demonstrate the potential of community transcriptomics to identify the drivers of species co-occurrence in hyper-diverse tropical tree genera.

1 | INTRODUCTION

Tropical tree communities are among the richest assemblages on the planet (Newman, 2002; Zuidema et al., 2013). This diversity can be partly described by two key components. First, there are large numbers of rare species that disproportionally contribute to this richness. This has caused many to investigate topics such as rare species advantages (Bachelot & Kobe, 2013), the spatial dispersion of individuals (Pacifici et al., 2016) and plant-pollinator interactions (Aizen, 2021). A second less-well-understood phenomenon is that there are a few, but very diverse, genera that make a disproportionate contribution to local-scale tropical tree richness (Gentry, 1982, 1989). These genera often contribute tens of species to local tree richness in the most diverse forests world-wide (Cottee-Jones et al., 2016; Harrison, 2005). For example, the most diverse forests in the Neotropics can contain up to tens of co-occurring species from a genus such as *Inga*, *Psychotria*, *Piper* and *Miconia* (Kursar et al., 2009; Messeder et al., 2021; Richards et al., 2015; Sedio et al., 2012). Similarly, the most diverse Paleotropical forests contain mega-genera such as *Shorea*, *Diospyros*, *Cola* and *Syzygium* (Ashton, 1969; Duangjai et al., 2009; LaFrankie et al., 2006; Nair, 2017). These genera challenge the view of community assembly that evolutionary niche conservatism and niche overlap prevent coexistence. Determining what promotes and maintains the local diversity of these handful of genera or species swarms (Gentry, 1989), therefore, is critical for understanding the promotion and maintenance of tropical tree richness, in general.

The local-scale richness of large genera can be attributed to differential responses to abiotic and biotic conditions, which likely acting in combination. Examples of the abiotic environment sorting diverse genera locally can be seen in the local-scale distributions of *Psychotria* species along soil water gradients in Panama (Sedio et al., 2012) and the sorting of species along steep soil nutrient gradients in Malaysia (Russo et al., 2013). Examples of the biotic environment (e.g., natural enemies) determining the co-occurrence of species can be found in genus *Inga* (Coley & Barone, 1996; Coley & Kursar, 2014) and *Piper* (Dyer & Palmer, 2004) by examining their spatial distributions and phytochemical similarity. Interestingly, these abiotic and biotic drivers may interact to influence the spatial distributions of species within genera. Specifically, it has been hypothesized that the strength of pest or pathogen-driven conspecific negative density dependence is higher in wetter forests (Fricke et al., 2014; Forrister et al., 2019; but see Gripenberg et al., 2014), and it has been shown in *Inga* that traits related to abiotic and biotic interactions are non-randomly distributed in local assemblages such that the abiotic environment constrains and the biotic environment expands trait space (Coley et al., 2018; Endara et al., 2017).

Trait data have been used extensively for over a decade in tropical tree community ecology to elucidate the abiotic and biotic drivers of species distributions and co-occurrence (e.g., Swenson & Enquist, 2009; Worthy & Swenson, 2019). A limitation of this work has been the measurement of a handful of traits that, while important, do not always convey information about the processes of interest (Yang et al., 2018). Among the processes that are not well-characterized by the traits typically measured are the roles of water limitation and natural enemies on species co-occurrence. Measurements of plant-water relations and plant defence are possible, but traditional approaches can be challenging to broadly employ. This had led to important innovations that permit measures of traits like leaf osmotic potentials (Bartlett et al., 2012) or the metabolomic profiles of species on a community scale (Sedio et al., 2012; Wang et al., 2022, 2023). An alternative, but under-utilized approach, has been functional genomics or transcriptomics (Swenson et al., 2017, 2023; Swenson & Jones, 2017). Transcriptomics is the study of the genes being expressed in a sampled tissue. Thus, one may study the number of transcripts per gene in a sample (i.e., gene expression) as well as the sequence similarity of the expressed genes. For example, research has shown that the study of leaf gene expression in response to drought is a superior predictor of tree distributions than commonly measured functional traits (Swenson et al., 2017) and leaf defence gene sequence similarity in neighbourhoods is a strong predictor of individual tree survival (Zambrano et al., 2017). These two transcriptomic approaches have different strengths. Gene expression studies allow for the quantification of the dynamic functional response of plants to environmental stresses. This is particularly advantageous because many of the traits commonly measured by community ecologists are static and do not adequately characterize these dynamic responses that are likely more important for tree performance and co-occurrence. Specifically, Swenson et al. (2017) showed that in drier environments species with similar gene expression responses to drought co-occur, but in wetter environments dissimilar gene expression profiles co-occur. However, one downside of such studies is that they require far greater experimental control and much higher RNA sequencing costs due to the number of individuals that must be sequenced. In contrast, measures of functional gene sequence similarity, like those in Zambrano et al. (2017), are of interest because they are more logistically feasible and cost-effective. Specifically, since such studies only consider sequence similarity, and not expression levels, one individual per species can feasibly be sampled and less control over the timing of sampling is needed. Moreover, there is evidence that sequence dissimilarity strongly correlates with gene expression dissimilarity (Yang & Wang, 2013) and this approach can be easily integrated into existing analytical tools used in community phylogenetics (Swenson, 2019). A final strength of each of these approaches is that, while gene annotation databases and methods are still improving, gene annotations for non-model trees using model (e.g., *Arabidopsis*) databases are surprisingly robust (Swenson et al., 2017). This allows for studies targeting specific groups of genes of interest such as genes related to drought or defence.

Here, we focus on functional gene sequence similarity and not gene expression levels. Analysing the dissimilarity of drought- or defence-related genes in locally co-occurring species can reveal co-occurrence patterns among species in response to drought or defence. Gene expression similarity or dissimilarity may reflect how species respond to similar or different environmental conditions, respectively. For instance, gene sequence similarity may indicate that related species have comparable responses to the prevailing environment, while dissimilar gene sequences may suggest different response strategies, such as tolerators and avoiders of drought or herbivores. Therefore, the similarity of functional gene sequences in locally co-occurring species may provide insights into the drivers of species co-occurrence. This would be particularly interesting in the context of hyper-diverse genera in tropical forests.

In this work, we present an initial transcriptomic study of co-occurring tropical tree species from a hyper-diverse genus. We focus on species from the tropical tree genus *Ficus* and employ community transcriptomic approaches to investigate the roles of drought- and defence-related genes in determining the spatial distributions and co-occurrence of conspecifics on a local scale. Specifically, we sequenced and assembled the leaf transcriptomes of 13 locally co-occurring *Ficus* in a tropical rainforest in southwest China and used these data to quantify whether genetic distance between species in genes related to drought and defence were linked to patterns of co-occurrence within and between habitats varying in their soil water content. We asked a series of four main questions in this study. First, do *Ficus* species partition soil moisture habitats? Second, does the degree of sequence similarity in genes related to drought and defence response predict the degree of plot-wide pairwise species co-occurrence? Third, are defence-related genes more dissimilar than expected in locally co-occurring species as would be expected if shared enemies limited the potential for co-occurrence? Lastly, are there opposing patterns in drought and defence-related genes across habitats as would be expected if the relative importance of water limitation and natural enemies shifted along a soil water gradient?

2 | MATERIALS AND METHODS

2.1 | Study site

This study was conducted at the 20-ha tropical forest dynamics plot (FDP) in Xishuangbanna, Yunnan Province, China (21°37′08″ N, 101°35′07″ E). This plot is dominated by *Parashorea chinensis* (Dipterocarpaceae) and is characterized as a seasonal tropical rainforest influenced by the Indian summer monsoon. The mean annual temperature is 21.8°C. The mean annual precipitation is 1493 mm, and 84% of it occurs during wet months from May to October (Cao et al., 2006). The elevation ranges from 708 to 869 m a.s.l. in the plot. Three perennial streams traverse the plot resulting in strong soil–water habitat heterogeneity (Lan et al., 2009). All trees ≥1 cm dbh in the plot have been mapped, tagged and measured. The data used in this study come from the second census of the plot that was carried out in 2012. Fieldwork at this site was permitted by the Chinese Academy of Sciences.

2.2 | Study species

We focus on species from the genus *Ficus* (Moraceae) found in the plot as a case study for the following reasons: (1) *Ficus* is one of the largest angiosperm genera with a pantropical distribution containing more than 800 species (Harrison, 2005); (2) *Ficus* is the most abundant genus with respect to their importance value among genera in the plot, including 3712 individuals and 6.4% of the total basal area (Asefa et al., 2021); (3) *Ficus* species in the plot are distributed across elevations (i.e., from the valley to the ridge habitats), which coincides with variation in soil moisture (Cao et al., 2008); and (4) While *Ficus* are famous for their specialized pollinator interactions, which no doubt have spurred net diversification in the lineage, most species of *Ficus* also have extraordinary defence chemistry in latex, such as tannin, alkaloid content, that likely plays a key role in their distributions with respect to herbivores and other natural enemies (Zhao et al., 2021). In addition, divergence time estimates and biogeographic distributions suggest that *Ficus* originated in either Eurasia or the Neotropics roughly 63.20 (58.70–71.15) Mya (Wang et al., 2021), followed by a rapid (~1 million years) divergence between the Neotropical section Pharmacosycea and all other figs (Figure S1; Wang et al., 2021). Thus, the genus has undergone a spectacular radiation that has resulted in high levels of regional and local diversity that demand further investigation. While more than 13 species of *Ficus* occur in the forest plot, we focused only on the 13 species with an abundance of at least 30 individuals (Table 1) to assure we had enough power for spatial analyses. All of these *Ficus* species in our study are native to southwestern China, and previous work has suggested that drought stress likely plays an important role in maintaining *Ficus* species at our study site (Lasky et al., 2014).

2.3 | Habitat categorization

To test whether species were associated with different soil moisture habitats, we divided the plot into three habitats according to soil moisture content. Specifically, soil moisture was measured using a soil conductivity probe meter (Theta probe MPM-160B, ICT International Proprietary Limited, Armidale, Australia) both at the end of rainy season and dry season at three points near the centre of each 20 m × 20 m quadrat in the forest plot. During each measurement, we removed the litter and humus layer on the ground, then inserted the soil moisture probe 5 cm below the ground and recorded the reading (Song et al., 2017). The three soil moisture readings in a quadrat were then averaged. The 500 quadrats were assigned to three habitat types: high soil moisture (Habitat A), medium soil moisture (Habitat B) and low soil moisture (Habitat C) using hierarchical clustering by the *hclust()* function in R. The distribution of focal *Ficus* species in three soil moisture habitats is displayed in Figure 1 and Table S1.

TABLE 1 Thirteen *Ficus* species in this study, their abundances and the results of their habitat association analyses across the three habitat types in the plot.

Species	Code	Abundance	Habitat A	Habitat B	Habitat C
Ficus auriculata	FICUAU	105	-	-	N
Ficus beipeiensis	FICUBEI	214	-	-	N
Ficus esquiroliana	FICUES	79	-	-	-
Ficus fistulosa	FICUFI	917	P	N	N
Ficus fulva	FICUFU	103	-	-	-
Ficus glaberrima	FICUGL	79	-	-	-
Ficus hirta	FICUHIR	30	-	N	-
Ficus hispida	FICUHI	501	-	-	N
Ficus langkokensis	FICULA	1412	-	-	-
Ficus oligodon	FICUOL	30	-	-	N
Ficus semicordata	FICUSE	39	P	N	N
Ficus subincisa	FICUSU	31	-	-	-
Ficus variegata	FICUVA	172	-	N	-

Note: 'P' represents a positive association, 'N' represents a negative association, and '-' represents no corresponding association. Habitat A: highest soil moisture, Habitat B: medium soil moisture and Habitat C: lowest soil moisture.

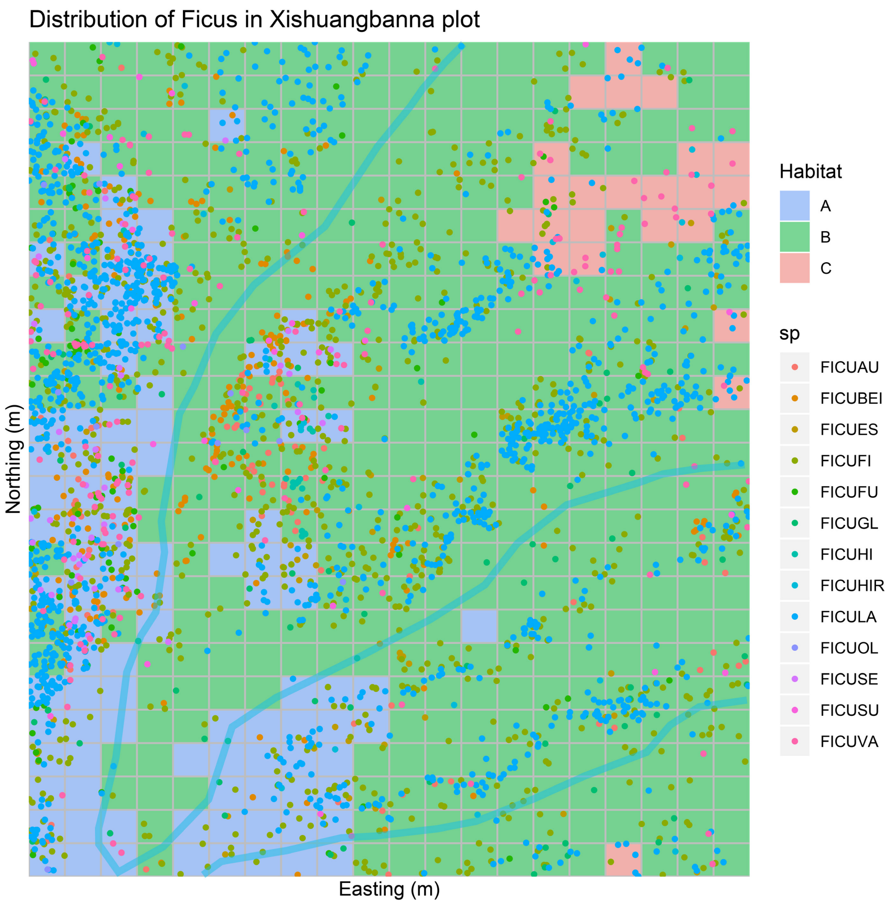

FIGURE 1 Distribution of the focal 13 *Ficus* species in Xishuangbanna 20-ha forest dynamics plot. The species codes in the legend correspond to the Latin binomials provided in Table 1. Habitat A is the wettest habitat followed by Habitat B and then Habitat C.

2.4 | Quantifying species–habitat associations

To quantify the habitat associations of tree species, we used a torus-translation test, which calculates correlations between soil habitats and spatial distributions of individuals in a species while controlling for spatial autocorrelation in both the plant and soil data (Harms et al., 2001). Specifically, we rotated (i.e., translated) the real map into four directions (up, down, left and right) in 20-meter increments, 20 times in the horizontal direction and 25 times in the vertical direction. Three maps were generated by 180° rotation, mirroring and 180° rotating mirroring. For each of these increments, the relative density of trees in a species and each habitat was recorded. This

Journal of Ecology BRITISH ECOLOGICAL SOCIETY

was done 1999 times to generate a null distribution for each habitat. The relative density of individuals from the species in each habitat was then compared with the null distributions for test of significance with an alpha of 0.05. Thus, if the true relative density of the species in the focal habitat was greater than 97.5% of the expected relative density (i.e., the 0.975 quantile in the null distribution), it was determined that the species and habitat are positively correlated. If the density was less than 2.5% of the expected relative density (i.e., the 0.025 quantile in the null distribution), a negative correlation between the species and habitat was determined. When the true relative density was between 2.5% and 97.5%, there we determined the species and habitat had no association.

2.5 | Transcriptome sequencing and community functional phylogenomics

In this work, we utilized a community transcriptomics workflow to assess the functional similarity of co-occurring *Ficus* species (Figure S2). We will discuss the workflow in the following, but the end products from the workflow that were used for analyses are homologous gene trees. The branch lengths on these gene trees are used to calculate a distance matrix that represents the genetic similarity of species for a given gene and given the substitution model fit to the sequence data. Therefore, they are approximations of the similarity of species for a gene that is linked to an annotation of interest.

We performed functional phylogenomic inference using the techniques and tools that recently published in Yang et al. (2015), and then estimated the homologous functional gene trees of *Ficus* species in the community, in particular, the distance separating species in a gene tree represents the degree of their functional gene similarity. This workflow begins with sampling, sequencing and assembling the transcriptome of each species.

To generate transcriptome assemblies for each of our 13 study species, we sampled fully expanded and undamaged leaf material in the field during the morning on a single day in the dry season from a single individual per species and immediately flash-froze the tissue in liquid nitrogen. We note that this study focused on genetic similarity and not expression. However, we temporally constrained sampling to the morning of a single day to reduce any variation in the presence or absence of a gene being expressed. While some genes are expressed less than others, our approach focuses on sequence similarity and not expression levels. Thus, subtle or major changes in the environment that cause changes in the expression level, but not presence, of a gene were not consequential to our study design. Finally, functional phylogenomic studies frequently rely on a single individual per species, as they are not focused on expression levels per se (e.g., Han et al., 2017; Yang et al., 2015; Zambrano et al., 2017). There are clear downsides of this approach that should be considered in future studies. First, there may be genes expressed at other times and in other tissues or individuals that may be critically important.

RNA was extracted from the tissue using a Qiagen RNAeasy Plant Mini Kit (Qiagen, Valencia, California, USA). The quality and quantity of RNA was measured using a NanoDrop 2000 spectrophotometer (NanoDrop Products, Wilmington, Delaware, USA) and an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Poly-A tail enrichment library preparation and sequencing were done at the Beijing Genomics Institute using the Illumina HiSeq 2000 platform that generates 150-bp paired-end reads. Reads were assembled de novo into transcriptomes using Trinity v2.6.6 (Grabherr et al., 2011). The quality of assemblies was checked using N50 values and, more importantly, BUSCO scores (Waterhouse et al., 2018) using a standard set of genes known from the Eudicots (Table S2). We also quantified the compactness of the transcriptome assemblies using the EX90N50 metric employed in Trinity, which also indicated a good-quality assembly.

The assembled transcriptomes were then used to conduct functional phylogenomic inference following the pipeline developed by Yang et al. (2015) and used in previous work investigating the functional genomics of tree communities (Zambrano et al., 2017). In brief, we used TransDecoder v5.0.2 to translate the assembled contigs and cd-hit v4.6.8 to infer putative peptides. Next, we performed an all-by-all *blastn* of the cds of the transcriptome assemblies using a cut-off *E*-value of 10^{-5} with max_target_seqs set 1000 and a hit fraction of 0.4. Next, contigs were clustered based upon their blastn results using MCL (van Dongen, 2000). This produced putative homologous gene clusters that could contain zero to many sequences per species. The sequences within clusters were then aligned using MAFFT with trimmed alignments produced by phyutility v2.2.6 (Smith & Dunn, 2008). Gene trees were inferred using these alignments using a GTRCAT substitution model implemented in RAxML (Stamatakis, 2006). Thus, the resulting gene trees with branch lengths consider transitions, transversions and synonymous versus nonsynonymous substitutions. This makes distance matrices calculated from these branch lengths an indicator of functional similarity. The tips of gene trees were removed if the tip length was less than 0.3 or more than five times the length of its sister branch. Gene tree tips from the same exact taxa were masked in the tree, keeping the tip with the most un-ambiguous and well-aligned characters in the trimmed alignment. Lastly, we realigned the sequences that resulted from this process using MAFFT and inferred the final gene trees using RAxML using the same parameters used in the first inferences. We retained all homologue gene trees that contained at least five tips. The homologue gene trees were annotated using blastp against the Swiss-Prot database in 2018, and the corresponding Gene Ontology (GO) terms for each UniProt identifier were looked up accordingly. We retained all gene trees that were annotated with GO terms related to drought (i.e., response to water deprivation [GO:0009414]; response to osmotic stress [GO:0006970]) or defence (i.e., defence response [GO:0006952]; defence response to bacterium [GO:1900425]; negative regulation of defence response to oomycetes [GO:1902289]; and regulation of defence response to fungus [GO:1900150]). We stress that gene ontologies are hypotheses and do change as databases improve and some caution is urged. However, they do provide the only feasible and reasonable approach for a comparative study such as the present one attempting to focus on specific functions. We also note that ontologies can be challenging to lump or split into discrete ecological categories of interest. For example, individual GOs that we have lumped into the 'drought' category could be investigated individually or one could expand their definition of what is related to drought. Furthermore, genes may be linked to stress responses that encompass responses to both drought and defence. Lastly, many sequences have no match to annotation databases and are not studied here, but they may still be related to drought and defence and not known. In sum, the use of GOs and annotations will remain a challenge for fields such as transcriptomics and metabolomics and the caveats listed above should be considered when interpreting any future analyses.

The analyses that follow focused on the genetic distance between species in drought and defence genes. The genetic distance between two species was the average similarity between species calculated using the gene tree phylogenetic distance matrices. These same methods have been used previously in community transcriptomic studies of functional genetic distance (Zambrano et al., 2017). We also note that, while gene expression varies based upon ecological conditions and studies of differential expression utilize multiple individuals per species and condition, functional phylogenomic investigations such as that used here typically utilize a single sample per species (Swenson et al., 2017; Waterhouse et al., 2018) as they are not focused on differential gene expression.

2.6 | Transcriptomic similarity and co-occurrence

To investigate whether the degree of sequence similarity in drought and defence genes was related to the degree of pairwise species co-occurrence in the plot, we conducted a statistical approach that involved correlating the genetic distance between two species with their co-occurrence at the 20 m × 20 m scale in the forest dynamics plot. Specifically, we first quantified the observed co-occurrence of species using Schoener's index of co-occurrence (Schoener, 1970), a measure of co-occurrence that takes into account both the presence or absence of species in each plot and their proportional abundances. We calculated observed co-occurrence values at the 20 m × 20 m scale, then compared them to a null distribution containing 999 values. To generate the null distribution, we randomized the community data matrix using the independent swap algorithm (Swenson, 2014). Next, we then calculated a standardized effect size (SES) by quantifying the difference between the observed values and the mean of the null distribution and dividing this value by the standard deviation of the null distribution (Swenson, 2014). Positive SES values indicated more than expected co-occurrence between species and negative values indicated less than expected co-occurrence between species.

Because we focused on drought- and defence-related genes it may be unclear whether the results are exceptional. That is, are these genes more closely tied to co-occurrence or habitats than other functional genes? Thus, in order to generate a baseline expectation, we randomly selected 1000 homologous genes that were not linked to defence or drought and re-ran all of our analyses. Specifically, we used these gene trees to calculate the SES value for each 20 m × 20 m subplot and then compared these observed SES values to a null distribution to test for significant deviations from random co-occurrence patterns. These SES values were used in all of the following analyses along with the SES values from the drought and defence gene analyses.

To test whether the SES co-occurrence values could be predicted using drought or defence genes, we used Mantel tests (Mantel, 1967) to estimate the correlation between two matrices, the genetic distance matrices and the species co-occurrence distance matrix. The genetic distance matrices were calculated based on the difference in genetic distance between species in a gene tree. The species co-occurrence distance matrix was the species-by-species SES co-occurrence matrix calculated in the previous step. Given the large number of statistical tests performed increases the false discover rate, we adjusted the *p*-values from the Mantel tests were adjusted using a Holm correction (Holm, 1979) to control for a false discovery rate of 5%.

2.7 | Transcriptomic dispersion in local assemblages along a soil moisture gradient

The second set of analyses quantified the dispersion of species in 20 m × 20 m quadrats using genetic distance data. Specifically, the distance matrix for each gene tree was used to quantify the dispersion of *Ficus* species in a quadrat. This was done using a mean pairwise distance and a mean nearest neighbour distance (Swenson, 2014). These observed values were then compared with a null distribution to calculate SES values using the same SES calculation used in the co-occurrence analyses. Here, for consistency with the phylogenetic community ecology literature, we use the terms Net Relatedness Index (NRI) and Nearest Taxon Index (NTI) for the SES of the mean pairwise distance and mean nearest neighbour distance, respectively (Swenson, 2014). This resulted in one SES value per quadrat for each drought or defence gene (Table S3). Significant positive deviations indicated that a gene is overdispersed (i.e., more dissimilar than expected), while significant negative deviations indicate that a gene is clustered (i.e., more similar than expected). This was tested using a non-parametric Wilcoxon test as the distributions were not normal. As with the Mantel tests, the *p*-values from the Wilcoxon tests were adjusted with a Holm correction (Holm, 1979) to account for a 5% expected false discovery rate.

Next, for each gene, we were interested in whether the SES values changed along the soil moisture gradient. We used a linear model to examine the association between the SES value and soil moisture gradient. We also conducted a similar analysis using the SES values for seven commonly measured functional traits—leaf area, specific leaf area, leaf thickness, leaf toughness, leaf carbon content, leaf nitrogen content and leaf potassium content and soil moisture gradient that we had collected in a prior study (Asefa et al., 2021). This was done to determine whether genetic information was more closely linked to the moisture gradient than commonly measured functional traits. All analyses were implemented in R.

3 | RESULTS

3.1 | Ficus species co-occurrence and habitat preferences

We first tested for the existence of habitat preferences in the 13 *Ficus* species in our study. The results showed that there were several significant positive and negative associations with soil moisture habitats in the forest plot (Table 1). Specifically, for the 13 *Ficus* species, eight species (61.5%) were negatively associated with at least one soil habitat type. Six of these species were negatively associated with Habitat C (i.e., the driest habitat) and four of the species were negatively associated with Habitat B (i.e., the mid-level soil moisture). Conversely, only two species *F. fistulosa* and *F. semicordata* had significant positive associations with a habitat which has the highest soil moisture, and those species were negatively associated with the two other habitats (Habitats B and C). These results demonstrate non-random habitat preferences of *Ficus* species in the plot largely related to avoidance of the drier and higher regions of the plot.

3.2 | Transcriptomic similarity and species co-occurrence—Plot scale

The transcriptomic analysis generated 240 drought- and 570 defence-related homologous gene trees. We then quantified the observed co-occurrence of species using SES values generated from the observed and null distribution of Schoener's co-occurrence index values. Overall, after correcting for multiple tests, we found nine drought-related genes (Table 2) and 18 defence-related genes (Table 3) that were correlated with pairwise species co-occurrence patterns. Five of the drought-related genes were negatively correlated with the SES Schoener's index values indicating that species that tend to co-occur have lower distances between them in these genes. Conversely, four drought-related genes had a positive correlation indicating that species divergent in these genes tended to co-occur. In other words, we found similarity and dissimilarity in drought-related genes in co-occurring *Ficus* species. The defence-related genes had a more consistent pattern. Specifically, 16 of the 18 genes had a positive correlation with the SES Schoener's index (Table 3). This indicated that co-occurring species tended to be more divergent in these genes. The other two genes indicated that sequence similarity in these genes was associated with co-occurrence. We also considered whether any of the 1000 randomly selected genes were related to Schoener's index values. After correcting for multiple tests, we found that none of these genes was related to the index.

3.3 | Transcriptomic similarity and species co-occurrence—Local scale

The prior analyses investigated species co-occurrence at the whole forest plot scale. We also investigated local co-occurrence on the scale of 20 m × 20 m quadrats using genetic distance matrices and the NRI and NTI metrics. Overall, the distribution of NRI and NTI values in local assemblages deviated from zero in the majority of drought- and defence-related genes (Figure 2; Table 4). Specifically, ~30% of the defence-related genes and ~50% of the drought-related genes tended to show clustering locally (i.e., co-occurring species with smaller than expected genetic distances separating them in those genes). Conversely, ~50% of the defence-related genes and ~30% of the drought-related genes were overdispersed locally (i.e., co-occurring species with larger than expected genetic distances separating them in those genes). Thus, only a small proportion (~20%) of the drought- and defence-related genes studied had no association with species co-occurrence. Interestingly, ~30% of the random, but functional (i.e., expressed), genes were locally clustered, but ~20% of these genes were overdispersed. Thus, the clustering of drought-related genes is on par with that of other functional genes, but the degree

TABLE 2 Drought-related genes where sequence similarity was significantly associated with (i.e., $p < 0.05$ after a Holm correction) pairwise species co-occurrence via Mantel tests.

Gene	Mantel *r*	Protein name	Organism
PUB23_ARATH	0.03605431	E3 ubiquitin-protein ligase PUB23	*Arabidopsis thaliana (Mouse-ear cress)*
LSU2_ARATH	−0.6699423	Protein RESPONSE TO LOW SULFUR 2	*Arabidopsis thaliana (Mouse-ear cress)*
PMI1_ARATH	0.26530418	Protein PLASTID MOVEMENT IMPAIRED 1	*Arabidopsis thaliana (Mouse-ear cress)*
CIP1_ARATH	−0.0084521	COP1-interactive protein 1	*Arabidopsis thaliana (Mouse-ear cress)*
STP13_ARATH	0.1701683	Sugar transport protein 13	*Arabidopsis thaliana (Mouse-ear cress)*
SUD1_ARATH	−0.0615809	Probable E3 ubiquitin ligase SUD1	*Arabidopsis thaliana (Mouse-ear cress)*
ZAT10_ARATH	0.0269102	Zinc finger protein ZAT10	*Arabidopsis thaliana (Mouse-ear cress)*
HA22C_ARATH	−0.2672513	HVA22-like protein c	*Arabidopsis thaliana (Mouse-ear cress)*

Note: Positive Mantel *r* values indicate dissimilar species tend to co-occur and negative Mantel *r* values indicate similar species tend to co-occur. The organism name corresponds to the organism associated with the protein annotation in UniProt.

TABLE 3 Defence-related genes where sequence similarity was significantly associated with (i.e., $p < 0.05$ after a Holm correction) pairwise species co-occurrence via Mantel tests.

Gene	Mantel r	Protein name	Organism
TLP_ACTDE	0.13674796	Thaumatin-like protein	*Actinidia deliciosa (Kiwi)*
C93A1_SOYBN	0.783756	3,9-dihydroxypterocarpan 6A-monooxy	*Glycine max (Soybean) (Glycine hispida)*
RGA4_SOLBU	0.7557265	Putative disease resistance protein RGA4	*Solanum bulbocastanum (Wild potato)*
LSU2_ARATH	0.6699423	Protein RESPONSE TO LOW SULFUR 2	*Arabidopsis thaliana (Mouse-ear cress)*
LRK82_ARATH	0.6235554	L-type lectin-domain containing receptor kinase VIII.2	*Arabidopsis thaliana (Mouse-ear cress)*
DRL21_ARATH	0.381585	Putative disease resistance protein At3g14460	*Arabidopsis thaliana (Mouse-ear cress)*
CRSP_ARATH	0.2793561	CO(2)-response secreted protease	*Arabidopsis thaliana (Mouse-ear cress)*
R13L4_ARATH	0.2744614	Disease resistance RPP13-like protein 4	*Arabidopsis thaliana (Mouse-ear cress)*
IRE1B_ARATH	0.232015	Serine/threonine-protein kinase/endoribonuclease IRE1b	*Arabidopsis thaliana (Mouse-ear cress)*
FAS1_ARATH	0.1108338	Chromatin assembly factor 1 subunit FAS1	*Arabidopsis thaliana (Mouse-ear cress)*
SGT1B_ARATH	0.0989354	Protein SGT1 homologue B	*Arabidopsis thaliana (Mouse-ear cress)*
IOS1_ARATH	0.0949381	LRR receptor-like serine/threonine-protein kinase IOS1	*Arabidopsis thaliana (Mouse-ear cress)*
R13L2_ARATH	0.0684092	Putative disease resistance RPP13-like protein 2	*Arabidopsis thaliana (Mouse-ear cress)*
9DC3_SOLPI	0.039446	Receptor-like protein 9 DC3	*Solanum pimpinellifolium (Currant tomato) (Lycopersicon pimpinellifolium)*
PUB23_ARATH	−0.03605431	E3 ubiquitin-protein ligase PUB23	*Arabidopsis thaliana (Mouse-ear cress)*
SNC1_ARATH	−0.14675819	Protein SUPPRESSOR OF npr1-1, CONSTITUTIVE 1	*Arabidopsis thaliana (Mouse-ear cress)*

Note: Positive Mantel *r* values indicate dissimilar species tend to co-occur and negative Mantel *r* values indicate similar species tend to co-occur. The organism name corresponds to the organism associated with the protein annotation in UniProt.

of overdispersion in defence and drought genes far exceeds that found in other functional genes (Figure 2; Table 4).

Next, we investigated the correlation between the distribution of NRI and NTI values (i.e., is a given habitat more likely to have a clustered versus overdispersed pattern in these genes) and soil moisture. The results from these analyses found that the NRI and NTI metrics exhibited significant differences across the soil moisture gradient in the forest plot (Figure 3). We found that the NRI values of both drought and defence genes were significantly associated with soil moisture, while the NTI value did not exhibit a significant relationship. Moreover, we observed that the SES values of drought and defence genes increased with increasing soil moisture content, causing a shift from a clustered pattern to an overdispersed pattern. We also found that the SES values from genetic data were more strongly correlated with the moisture gradient than the SES. values for the functional traits (Table S4). Finally, we found that the dispersion patterns in the 1000 random genes were not related to soil moisture using either the NRI or NTI metrics and, like functional traits, there was no correlation between the moisture gradient and the SES values (Figure 3).

4 | DISCUSSION

This study investigated 13 species of *Ficus* that occur within a 20-ha forest dynamics plot using spatially explicit soil data and species-level transcriptomic data. We first asked whether pairs of species co-occur non-randomly within this forest plot. There were a handful of species pairs that non-randomly co-occur in the forest plot (Table S2). These co-occurrences contained either *F. auriculata*, with *F. beipeiensis*, *F. hirta* and *F. semicordata*, or *F. hirta*, with *F. auriculata* and *F. fulva*. The co-occurrence of these species may be explained by niche differences and/or reduced demographic performance differences in a preferred habitat. Conversely, *F. langkokensis* co-occurred less than expected with nine of the other 13 *Ficus* species in the forest. The other two species, *F. fistulosa* and *F. fulva*, were randomly associated with *F. langkokensis*. This may indicate that *F. langkokensis* dominates a given habitat via large demographic performance differences.

To investigate the drivers of species distributions and co-occurrence levels further, we conducted correlative analyses of soil habitat preferences using a torus-translation null modelling approach that preserved spatial autocorrelation in habitats and individuals (Harms et al., 2001). We found that six of the 13 species were negatively associated with the driest and most elevated habitat (i.e., Habitat C, Table 1). This is, perhaps, not surprising as *Ficus* species are known to prefer wetter habitats and the habitat gradient in the forest dynamics plot is steep (Lasky et al., 2014). A total of three of the 13 species were negatively associated with Habitat B, which has a moderate soil moisture, and two of these three species were positively associated with Habitat A, which is the wettest habitat in the forest (Table 1).

From these results, we can draw a few conclusions. First, *Ficus* in this forest generally do not recruit successfully into the driest habitat. Given the dispersal syndrome of *Ficus* and the distances involved (Nason et al., 1998), we can be confident that the observed

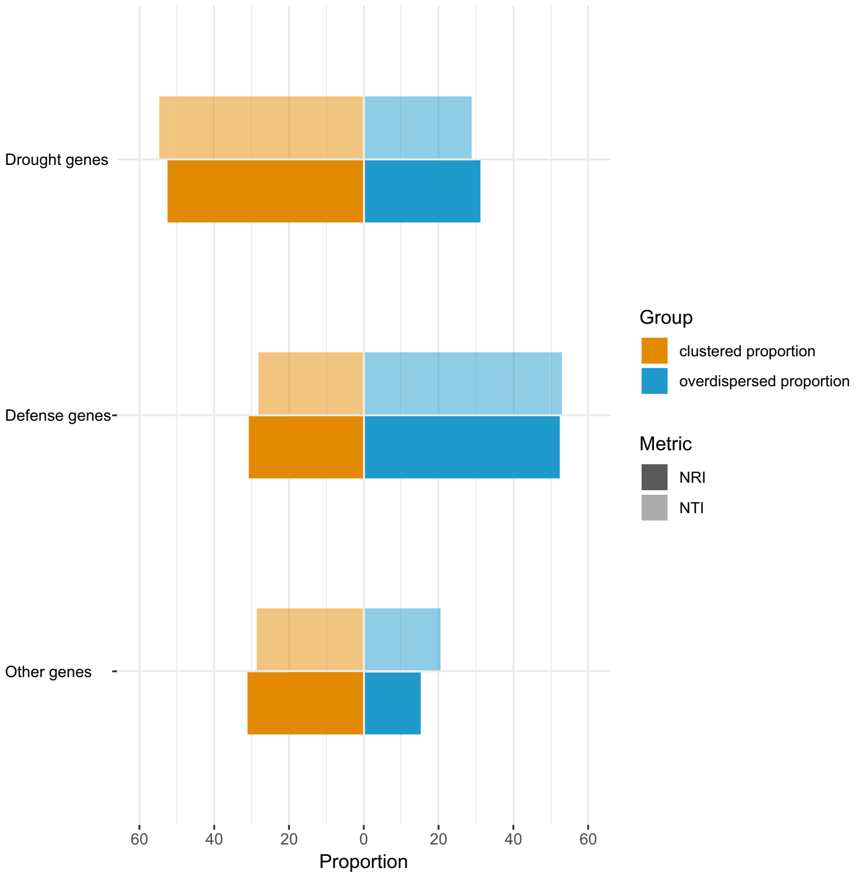

FIGURE 2 Proportion of genes that had clustered or overdispersed results in 20 m × 20 m quadrats in the forest dynamics plot. Clustered results indicate that the species co-occurring in quadrats were more similar in their gene sequence, for a given gene, than expected by chance. Overdispersed results indicated co-occurring species were more dissimilarity in their gene sequence, for a given gene, than expected by chance. The top bars show the genes related to drought. The middle bards show the genes related to defence and the bottom bars depict the results from the random sample of functional genes from the transcriptome. We used two metrics: the net relatedness index (NRI) and the Nearest Taxon Index (NTI), which are darker and lighter shaded bars, respectively.

TABLE 4 Percentage of the drought-related, defence-related genes and randomly selected, but functional, genes that were on average more similar than expected (i.e., clustered) or dissimilar (i.e., overdispersed) than expected in 20 m × 20 m quadrats.

Index	Clustered (%)	Overdispersed (%)	Function
NRI	53	31	Drought
NTI	55	29	Drought
NRI	31	52	Defence
NTI	28	53	Defence
NRI	31	15	Randomly
NTI	29	21	Randomly

Note: Two standardized effect size metrics were used: the Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI).

negative associations with the driest habitat are not a matter of dispersal limitation and are best explained by a failure to establish in an unpreferred habitat. Second, very few species have non-random associations with the wettest and moderately wet habitats. Thus, there is clear evidence of spatial structure in the *Ficus* populations in this forest (Figure 1) and some species pairs non-randomly cluster or appear to partition habitats between one another. However, outside of avoiding the driest habitat, superior demographic performance of a species in the wetter habitats cannot explain their co-occurrence per se and niche differences may be of greater importance. A final possibility is that the broad-scale habitat categories utilized in this study do not capture important fine-scale information regarding soil moisture and other important nutrients.

4.1 | Transcriptomic dissimilarity and species co-occurrence

Although it is common to detect phenotypic divergence between closely related species occupying tightly intertwined habitats (Cannon & Petit, 2020), cases of genetic divergence over local environmental gradients are rarer (Brousseau et al., 2021; Tysklind et al., 2020). Here, we present the first transcriptomic investigation that quantifies the relative importance of functional gene differences for species co-occurrence in tropical trees. Our first question using these data was whether species co-occurrence, represented as SES of Schoener's co-occurrence index, was correlated with the sequence dissimilarity of individual drought or defence-related genes. We identified nine drought-related genes that were correlated with species co-occurrence (Table 2). Four of these genes indicated that sequence dissimilarity was associated with co-occurrence while five of these genes indicated that sequence similarity was associated with species co-occurrence. Thus, the levels of co-occurrence of species pairs were related to multiple genes associated with drought, but in opposing directions. This is interesting as it parallels some results from functional trait-based studies of tropical forest community structure that have uncovered convergence and divergence

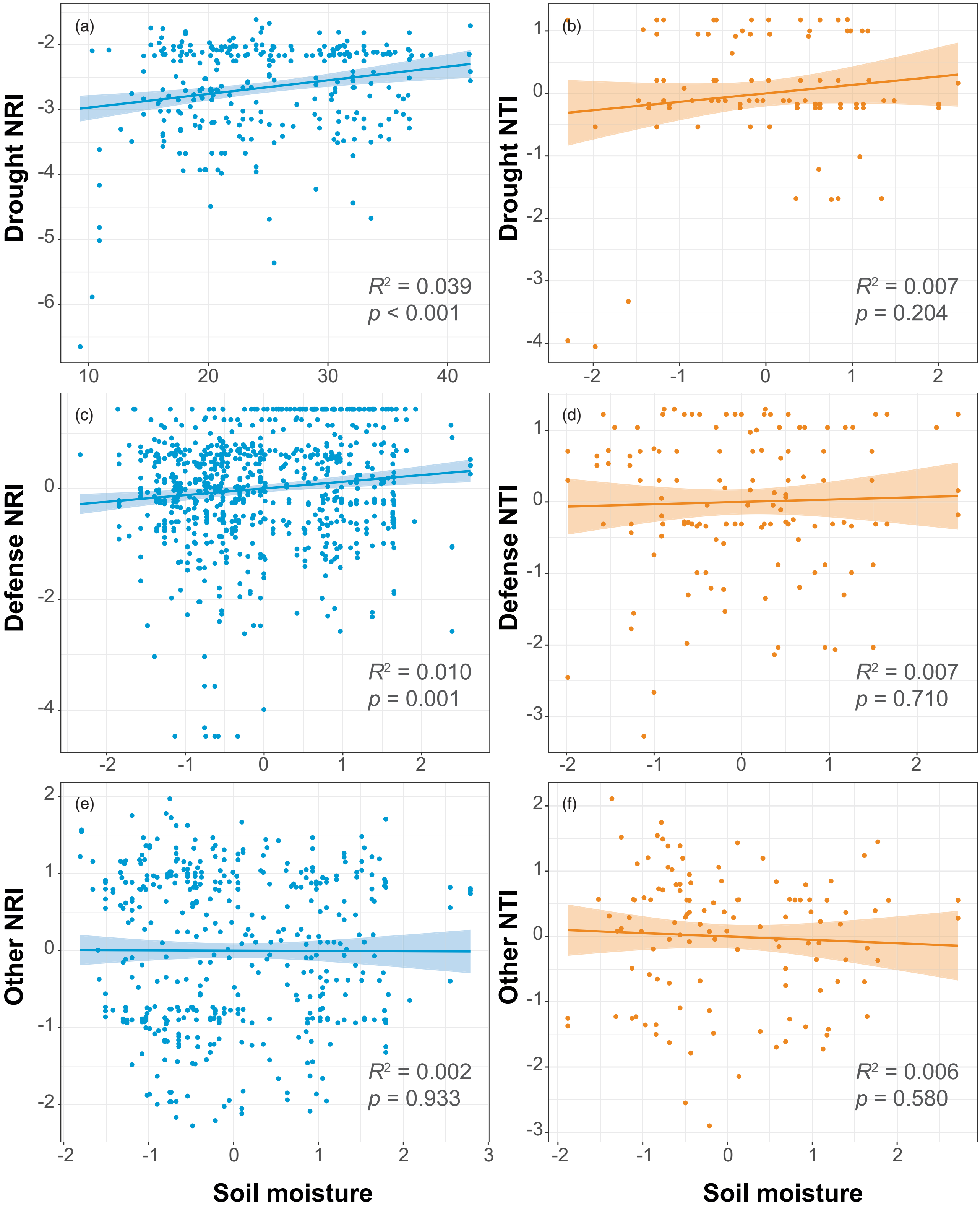

FIGURE 3 Regression analyses depicting the relationship between metrics of dispersion (*y*-axes) and the soil moisture gradient (*z*-scaled *x*-axis) with each point representing a 20 m × 20 m quadrat. Two metrics of dispersion were used: the Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI), which are the left and right columns, respectively. Negative NRI or NTI values indicate co-occurring species in the quadrat are more similar than expected in their gene sequence (i.e., they are clustered) and positive values indicate they are more dissimilar than expected (i.e., they are overdispersed). (a, b) The results for genes related to drought. (c, d) The results for the genes related to defence. (e, f) The results for the genes were randomly sampled from the transcriptome.

in the traits of co-occurring species (Swenson et al., 2007; Umaña et al., 2016; Yang et al., 2014). However, in this instance it is a similarity and dissimilarity in the genes associated with a particular function, which, probably underpins adaption and divergence of congeneric species, has not been documented in tropical trees previously.

Like the drought gene analyses, we uncovered multiple defence genes that correlated with species co-occurrence patterns. Specifically, we identified 18 defence genes of importance. Interestingly, 16 of the 18 indicated that sequence dissimilarity was associated with co-occurrence and 2 of the 18 indicated that sequence similarity was associated with co-occurrence (Table 2). Thus, while there is some evidence of defence gene similarity being important, most of the non-random correlations pointed to the importance of defence gene dissimilarity. Finally, we were interested in whether the co-occurrence results from the drought and defence genes were unusual compared with other functional (i.e., expressed) genes. When considering the correlations between the Schoener index values with the 1000 randomly selected genes, we found no significant correlations after correcting for multiple tests. Thus, groups of drought and defence-related genes tend to have an unusual proportion of genes related to *Ficus* co-occurrence in this forest.

Measures of co-occurrence can, at times, provide only a coarse overview of the spatial structure of communities. Thus, we also investigated the role of drought and defence gene similarity in this forest by quantifying the sequence similarity of species co-occurring in each of the 20 m × 20 m quadrats in the forest using each gene. This was done using a mean pairwise distance and a mean nearest neighbour metric standardized using null models to produce an NRI and NTI for each gene. The results were striking. Specifically, >50% of the drought-related genes were non-randomly clustered in 20 m × 20 m quadrats. This indicates species that were similar in these genes tended to non-randomly co-occur on this scale, which may be indicative of the importance of performance differences driving co-occurrence (i.e., superior performance of similar species excluding species with inferior performance via dissimilarity in their genes). We do note, however, ~30% of the drought-related genes were locally overdispersed. This indicates that for some genes, dissimilarity (i.e., niche differences) promotes co-occurrence. The NRI and NTI results for defence-related genes were also striking, but in the opposite direction. Specifically, ~50% of these genes were locally overdispersed and ~30% were clustered. Thus, dissimilarity (i.e., niche differences) with respect to defence appears to play a larger role in promoting co-occurrence while similarity in many fewer genes appears to be associated with co-occurrence.

The levels of clustering in the defence genes were similar to the proportion of randomly selected, but functional, genes that were significantly clustered in quadrats. This indicates that roughly 30% of all functional genes are clustered locally and that drought-related genes are not unusual in this regard. It is, perhaps, not surprising that species with similar gene sequences locally co-occur on these spatial scales. Functional trait clustering is typical on these spatial scales as patterns are likely related to broader gradients in the abiotic environment and less so to biotic interactions on fine scales (e.g., Swenson & Enquist, 2009). The relatively few randomly sampled genes that are locally overdispersed indicate that drought and defence-related genes are usually overdispersed, particularly at the spatial scale of 20 m × 20 m. This indicates that these functions are likely critical for determining the local-scale co-occurrence of species from this hyper-diverse genus.

From these results across spatial scales, we can draw a series of conclusions. First, the local-scale analyses provided more detailed information regarding how drought and defence genes impact *Ficus* co-occurrence on a local scale in this forest. Second, we find evidence for the importance of both genetic similarity and dissimilarity impacting co-occurrence. Interestingly, for the majority of the genes analysed, similarity in drought-related genes and dissimilarity in defence-related genes are found in local assemblages. The dissimilarity of defence genes is likely best understood, as noted above, in the context of a Janzen-Connell where shared pests and pathogens likely prevent species co-occurrence while driving species with dissimilar pest and pathogens and, therefore, dissimilar defence strategies to co-occur (Connell, 1971; Janzen, 1970). The drought gene similarity results are intuitive in that we may expect sorting along an abiotic gradient to select for similar species locally co-occurring and excluding dissimilar species that do not perform as well on that portion of the gradient. The drought gene dissimilarity results may appear counter-intuitive as drought is frequently considered a filter leading to a single type of functional strategy. However, the functional solutions to drought likely may not always be distilled into a single axis. On the most basic level, this diversity can be seen in plant assemblages ranging from drought-avoiding and tolerant plants in deserts to seasonal tropical forests, and on a molecular level the diversity of drought responses is grander (Ratzmann et al., 2019). Furthermore, the only other functional genomic investigation of a tree community that considers drought (Swenson et al., 2017) demonstrated that species with dissimilarity gene expression patterns in response to experimental drought tended to naturally co-occur in wet habitats in a temperate forest while similar expression patterns were found in drier habitats.

4.2 | Soil moisture and transcriptomic dissimilarity

The final part of this study was designed to ask whether functional genetic dissimilarity varied with respect to soil moisture. Previous research in a temperate forest has indicated that drought-related traits or genes may be overdispersed in wetter habitats and clustered in drier habitats (Swenson et al., 2017). In addition, there is conflicting evidence of the strength of conspecific negative density dependence along precipitation or soil moisture gradients with some inferring stronger Janzen-Connell effects in wetter habitats and others inferring stronger effects in drier habitats (Gilbert et al., 2001; Lin et al., 2012; Song et al., 2020, 2021). At the community level, habitat adaptation may take place as seedling mortality is high during the dry season in this plot (Lin et al., 2012; Song et al., 2020).

The results from this study show that overall assemblage genetic dissimilarity in both drought and defence-related genes increases with soil moisture (Figure 3). That is, the average pairwise genetic dissimilarity between species increases in wetter sites. However, when looking at nearest genetic neighbour distances, there was no relationship with soil moisture (Figure 3). The increase in defence gene dissimilarity in wetter locations aligns with previous evidence from this forest showing that conspecific negative density dependence is strongest during the wetter portion of the year (Song et al., 2020). That is natural enemies (e.g., pathogenic fungi and herbivores) may exert more pressure in wetter places and times and we might, therefore, expect more dissimilarity in co-occurring species. Conversely, we may expect more similar species in drier locations. The increase in dissimilarity in drought-related genes in wetter locations is similar to previous work by Swenson et al. (2017) using gene expression data in a temperate forest. Specifically, they found that species with dissimilar gene expression responses to drought are found in wetter locations and species with similar gene expression responses are found in drier locations. It is likely that abiotic filtering and/or performance differences are driving an increase in interspecific similarity in drier locations. Finally, as with the Schoener's index analyses, the dispersion patterns found in the 1000 randomly selected, but functional, genes were not correlated with the soil moisture gradient. This is important as we have no functional reason to expect a random sampling of genes to be correlated with this abiotic gradient, but there is a set of prior mechanistic expectations linking drought- and defence-related genes to soil moisture.

Next, we explored whether the pairwise functional gene dispersion patterns were more strongly linked to the soil moisture gradient than pairwise functional trait dispersion patterns. The results show that the soil moisture gradient was not related to dispersion levels in any of the traits measured, but it was related to both drought and defence-related genes (Table S4). This mirrors a similar result in Swenson et al. (2017) showing spatial distributions of species along soil moisture gradients are better predicted by functional genomic data as compared to functional trait data. This evidence does not necessarily indicate that the functional traits measured are uninformative. Rather, it indicates that commonly measured traits do not capture all relevant functional information and the addition of functional genomic information likely will improve our understanding of the drivers of species distributions and co-occurrence.

4.3 | Caveats

Here, we have employed a community transcriptomics (Swenson & Jones, 2017) approach for studying the co-occurrence and distributions of species from a hyper-diverse genus. We encourage future research that builds off this novel integration of transcriptomic information into community ecology, but we believe it is appropriate to highlight some pros and cons of the approaches taken. Perhaps one of the biggest challenges with omic approaches in ecology is our ability to robustly annotate the genes or metabolites under study. We rely on growing, but still sparse, annotation databases that are heavily reliant upon a few model systems. Thus, the annotations used in this, or other studies are hypotheses of function that may be incorrect and, in the future, should be validated experimentally. This will be challenging generally and particularly challenging for quickly evolving groups of defence genes that are responding to a potentially extensive group of pests and pathogens that would be challenging to exhaustively assess in an experimental context. A second major challenge is that, here, we have considered only genes found across all species. There are many genes that will be unique to species and there are duplication events and resulting paralogues. These types of data do not fit nicely into the common analytical pipelines used in phylogenetic or trait analyses of communities. These genes hold the potential to be very important for the interactions of interest, but they may often go unanalysed as will the 1000's of genes that have no annotations. Lastly, while the costs of sequencing are constantly falling and are now to the point where community transcriptomic investigations are possible, comprehensive transcriptome atlases for multiple species that are inclusive of multiple tissues and environmental contexts are still beyond the reach of an ecological budget for a non-model system. Thus, there will invariably be important rare transcripts that go unsequenced, tissue-specific transcripts that go unnoticed and large dynamic changes in expression through space and time that cannot be appreciated. These missing pieces of information are no doubt important and future studies are needed to identify the magnitude of this variation such that we can understand where robust inferences can or cannot be made.

5 | SUMMARY

Hyper-diverse genera disproportionately contribute to the spectacular species richness in tropical tree communities. The co-occurrence of large numbers of congeners presents a conundrum for community ecology that may be best understood in the light of functional biology. In particular, species differences in how they respond to water availability and the degree to which they share natural enemies are leading hypotheses to explain species co-occurrence. Here, we utilized detailed data on the functional genetic similarity of species to assess whether co-occurring congeners are dissimilar in drought and defence-related genes. We find substantial evidence of gene sequence dissimilarity between co-occurring *Ficus* species, particularly with respect to defence genes. Furthermore, the degree of genetic dispersion is linked to an underlying soil moisture gradient thereby indicating that functional genetic differentiation is related to congeneric co-occurrence patterns and that the way in which genetic differentiation impacts

Journal of Ecology BRITISH ECOLOGICAL SOCIETY

co-occurrence varies with the underlying abiotic template. Thus, the co-occurrence of a large number of *Ficus* species in this forest can be linked to functional genetic differentiation in drought and defence, which is not easily captured by commonly measured functional traits.

AUTHOR CONTRIBUTIONS

Nathan Swenson and Jie Yang designed the study. Huan Fan, Jie Yang, Yunyun He and Nathan G. Swenson performed the analysis. Jie Yang, Min Cao and Yunyun He collected and measured the data. Jie Yang and Nathan G. Swenson wrote the manuscript. All authors provided comments.

ACKNOWLEDGEMENTS

This research was supported by the NSFC China-US Dimensions of Biodiversity grant (DEB 32061123003), the Chinese Academy of Sciences Youth Innovation Promotion Association (Y202080), the Distinguished Youth Scholar of Yunnan (202001AV070016), the West Light Foundation of the Chinese Academy of Sciences and the Ten Thousand Talent Plans for Young Top-Notch Talents of Yunnan Province (YNWR-QNBJ-2018-309), the National Natural Science Foundation of China (31870410) and the 14th Five-Year Plan of the Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences (E3ZKFF1K). HF was supported by the International Postdoctoral Exchange Fellowship Program and Yunnan Postdoc Directed Training Program. NGS was funded by an NSF US-China Dimensions of Biodiversity grant (DEB 2124466). We are grateful for support from Xishuangbanna Station for Tropical Rain Forest Ecosystem Studies.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/1365-2745.14255.

DATA AVAILABILITY STATEMENT

Transcriptome sequencing data are available in NCBI under BioProject ID: PRJNA1045068.

The remaining data are available via Dryad: https://doi.org/10.5061/dryad.kprr4xhbf (Swenson et al., 2024).

ORCID

Gang Wang https://orcid.org/0000-0003-1834-9561
Min Cao https://orcid.org/0000-0002-4497-5841
Nathan G. Swenson https://orcid.org/0000-0003-3819-9767

REFERENCES

Aizen, M. A. (2021). Pollination advantage of rare plants unveiled. *Nature*, *597*, 638–639.

Asefa, M., Song, X., Cao, M., Lasky, J. R., & Yang, J. (2021). Temporal trait plasticity predicts the growth of tropical trees. *Journal of Vegetation Science*, *32*, e13056.

Ashton, P. S. (1969). Speciation among tropical forest trees: Some deduction in the light of recent evidence. *Biological Journal of the Linnean Society*, *1*, 155–196.

Bachelot, B., & Kobe, R. K. (2013). Rare species advantage? Richness of damage types due to natural enemies increases with species abundance in a wet tropical forest. *Journal of Ecology*, *101*, 846–856.

Bartlett, M. K., Scoffoni, C., Ardy, R., Zhang, Y., Sun, S., Cao, K., & Sack, L. (2012). Rapid determination of comparative drought tolerance traits: Using an osmometer to predict turgor loss point. *Methods in Ecology and Evolution*, *3*, 880–888.

Brousseau, L., Fine, P. V. A., Dreyer, E., Vendramin, G. G., & Scotti, I. (2021). Genomic and phenotypic divergence unveil microgeographic adaptation in the Amazonian hyperdominant tree *Eperua falcata* Aubl. (Fabaceae). *Molecular Ecology*, *30*, 1136–1154.

Cannon, C. H., & Petit, R. J. (2020). The oak syngameon: More than the sum of its parts. *New Phytologist*, *226*(4), 978–983.

Cao, M., Zhu, H., & Wang, H. (2008). *Xishuangbanna tropical seasonal rainforest dynamics plot: Tree distribution maps, diameter tables and species documentation*. Yunnan Science and Technology Press.

Cao, M., Zou, X. M., Warren, M., & Zhu, H. (2006). Tropical forests of Xishuangbanna, China. *Biotropica*, *38*, 306–309.

Coley, P. D., & Barone, J. A. (1996). Herbivory and plant defenses in tropical forests. *Annual Review of Ecology and Systematics*, *27*, 305–335.

Coley, P. D., Endara, M. J., & Kursar, T. A. (2018). Consequences of interspecific variation in defenses and herbivore host choice for the ecology and evolution of Inga, a speciose rainforest tree. *Oecologia*, *187*, 361–376.

Coley, P. D., & Kursar, T. A. (2014). On tropical forests and their pests. *Science*, *343*, 35–36.

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 population* (pp. 298–312). Centre for Agricultural Publishing and Documentation.

Cottee-Jones, H. E. W., Bajpai, O., Chaudhary, L. B., & Whittaker, R. J. (2016). The importance of *Ficus* (Moraceae) trees for tropical forest restoration. *Biotropica*, *48*, 413–419.

Duangjai, S., Samuel, R., Munzinger, J., Forest, F., Wallnoefer, B., Barfuss, M. H. J., Fischer, G., & Chase, M. W. (2009). A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. *Molecular Phylogenetics and Evolution*, *52*, 602–620.

Dyer, L. A., & Palmer, A. D. (2004). *Piper: A model genus for studies in phytochemistry, ecology, and evolution*. Springer.

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*, E7499–E7505.

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*, 1213–1216.

Fricke, E. C., Tewksbury, J. J., & Rogers, H. S. (2014). Multiple natural enemies cause distance-dependent mortality at the seed-to-seedling transition. *Ecology Letters*, *17*, 593–598.

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

Gentry, A. H. (1989). Tropical forests: Botanical dynamics, speciation, and diversity. In L. B. Holm-Nielsen, I. C. Nielsen, & H. Balslev (Eds.), *Speciation in tropical forests* (1st ed., pp. 113–134). Academic Press.
Gilbert, G., Harms, K., Hamill, D., & Hubbell, S. P. (2001). Effects of seedling size, El Niño drought, seedling density, and distance to nearest conspecific adult on 6-year survival of Ocotea whitei seedlings in Panamá. *Oecologia*, *127*, 509–516.
Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652.
Gripenberg, S., Bagchi, R., Gallery, R. E., Freckleton, R. P., Narayan, L., & Lewis, O. T. (2014). Testing for enemy-mediated density-dependence in the mortality of seedlings: Field experiments with five Neotropical tree species. *Oikos*, *123*, 185–193.
Han, B., Umaña, M. N., Mi, X., Liu, X., Chen, L., Wang, Y., Liang, Y., Wei, W., & Ma, K. (2017). The role of transcriptomes linked with responses to light environment on seedling mortality in a subtropical forest, China. *Journal of Ecology*, *105*, 592–601.
Harms, K. E., Condit, R., Hubbell, S. P., & Foster, R. B. (2001). Habitat associations of trees and shrubs in a 50-ha Neotropical forest plot. *Journal of Ecology*, *89*, 947–959.
Harrison, R. D. (2005). Figs and the diversity of tropical rainforests. *Bioscience*, *55*, 1053–1064.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.
Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, *104*, 501–528.
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*, 18073–18078.
LaFrankie, J. V., Ashton, P. S., Chuyong, G. B., Co, L., Condit, R., Davies, S. J., Foster, R., Hubbell, S. P., Kenfack, D., Lagunzad, D., Losos, E. C., Nor, N. S. M., Tan, S., Thomas, D. W., Valencia, R., & Villa, G. (2006). Contrasting structure and composition of the understory in species-rich tropical rain forests. *Ecology*, *87*, 2298–2305.
Lan, G., Zhu, H., Cao, M., Hu, Y., Wang, H., Deng, X., Zhou, S., Cui, J., Huang, J., He, Y., Liu, L., Xu, H., & Song, J. (2009). Spatial dispersion patterns of trees in a tropical rainforest in Xishuangbanna, Southwest China. *Ecological Research*, *24*, 1117–1124.
Lasky, J. R., Yang, J., Zhang, G., Cao, M., Tang, Y., & Keitt, T. H. (2014). The role of functional traits and individual variation in the co-occurrence of Ficus species. *Ecology*, *95*, 978–990.
Lin, L. X., Comita, L. S., Zheng, Z., & Cao, M. (2012). Seasonal differentiation in density-dependent seedling survival in a tropical rain forest. *Journal of Ecology*, *100*, 905–914.
Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.
Messeder, J. V. S., Silveira, F. A. O., Cornelissen, T. G., Fuzessy, L. F., & Guerra, T. J. (2021). Frugivory and seed dispersal in a hyperdiverse plant clade and its role as a keystone resource for the Neotropical fauna. *Annals of Botany*, *127*, 577–595.
Nair, K. N. (2017). *The genus Syzygium: Syzygium cumini and other underutilized species* (1st ed.). CRC Press.
Nason, J. D., Herre, E. A., & Hamrick, J. L. (1998). The breeding structure of a tropical keystone plant resource. *Nature*, *391*, 685–687.
Newman, A. (2002). *Tropical rainforest: Our most valuable and endangered habitat with a blueprint for its survival into the third millennium* (2nd ed.). Checkmark Books.
Pacifici, K., Reich, B. J., Dorazio, R. M., & Conroy, M. J. (2016). Occupancy estimation for rare species using a spatially-adaptive sampling design. *Methods in Ecology and Evolution*, *7*, 285–293.
Ratzmann, G., Meinzer, F. C., & Tietjen, B. (2019). Iso/Anisohydry: Still a useful concept. *Trends in Plant Science*, *24*, 191–194.
Richards, L. A., Dyer, L. A., Forister, M. L., Smilanich, A. M., Dodson, C. D., Leonard, M. D., & Jeffrey, C. S. (2015). Phytochemical diversity drives plant-insect community diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 10973–10978.
Russo, S. E., Kochsiek, A., Olney, J., Thompson, L., Miller, A. E., & Tan, S. (2013). Nitrogen uptake strategies of edaphically specialized Bornean tree species. *Plant Ecology*, *214*, 1405–1416.
Schoener, T. W. (1970). Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, *51*, 408–418.
Sedio, B. E., Joseph Wright, S., & Dick, C. W. (2012). Trait evolution and the coexistence of a species swarm in the tropical forest understorey. *Journal of Ecology*, *100*, 1183–1193.
Smith, S. A., & Dunn, C. W. (2008). Phyutility: A phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*, *24*, 715–716.
Song, X., Hogan, J. A., Brown, C., Cao, M., & Yang, J. (2017). Snow damage to the canopy facilitates alien weed invasion in a subtropical montane primary forest in southwestern China. *Forest Ecology and Management*, *391*, 275–281.
Song, X., Lim, J. Y., Yang, J., & Luskin, M. S. (2021). When do Janzen-Connell effects matter? A phylogenetic meta-analysis of conspecific negative distance and density dependence experiments. *Ecology Letters*, *24*, 608–620.
Song, X., Zhang, W., Johnson, D. J., Yang, J., Asefa, M., Deng, X., Yang, X., & Cao, M. (2020). Conspecific negative density dependence in rainy season enhanced seedling diversity across habitats in a tropical forest. *Oecologia*, *193*, 949–957.
Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*, 2688–2690.
Swenson, N. G. (2014). *Null models. Functional and phylogenetic ecology in R. Use R!* Springer.
Swenson, N. G. (2019). *Phylogenetic ecology: A history, critique, and remodeling*. University of Chicago Press.
Swenson, N. G., & Enquist, B. J. (2009). Opposing assembly mechanisms in a Neotropical dry forest: Implications for phylogenetic and functional community ecology. *Ecology*, *90*, 2161–2170.
Swenson, N. G., Enquist, B. J., Thompson, J., & Zimmerman, J. K. (2007). The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. *Ecology*, *88*, 1770–1780.
Swenson, N. G., Iida, Y., Howe, R., Wolf, A., Umana, M. N., Petprakob, K., Turner, B. L., & Ma, K. (2017). Tree co-occurrence and transcriptomic response to drought. *Nature Communications*, *8*(1), 1996.
Swenson, N. G., & Jones, F. A. (2017). Community transcriptomics, genomics and the problem of species co-occurrence. *Journal of Ecology*, *105*, 563–568.
Swenson, N., Yang, J., Fan, H., He, Y., Wang, G., & Cao, M. (2024). Functional genomics and co-occurrence in a diverse tropical tree genus: The roles of drought and defense related genes [Dataset]. *Dryad*. https://doi.org/10.5061/dryad.kprr4xhb
Swenson, N. G., Zambrano, J., Howe, R., & Wolf, A. (2023). Biogeographic context is related to local scale tree demography, co-occurrence and functional differentiation. *Ecology Letters*, *26*, 1212–1222.
Tysklind, N., Etienne, M. P., Scotti-Saintagne, C., Tinaut, A., Casalis, M., Troispoux, V., Cazal, S. O., Brousseau, L., Ferry, B., & Scotti, I. (2020). Microgeographic local adaptation and ecotype distributions: The role of selective processes on early life-history traits in sympatric, ecologically divergent symphonia populations. *Ecology and Evolution*, *10*, 10735–10753.
Umaña, M. N., Forero-Montana, J., Muscarella, R., Nytch, C. J., Thompson, J., Uriarte, M., Zimmerman, J., & Swenson, N. G.

(2016). Interspecific functional convergence and divergence and intraspecific negative density dependence underlie the seed-to-seedling transition in tropical trees. *American Naturalist*, *187*, 99–109.

van Dongen, S. (2000). *Graph clustering by flow simulation* [PhD dissertation]. University of Utrecht.

Wang, G., Zhang, X., Herre, E. A., McKey, D., Machado, C. A., Yu, W. B., Cannon, C. H., Arnold, M. L., Pereira, R. A. S., Ming, R., Liu, Y. F., Wang, Y., Ma, D., & Chen, J. (2021). Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism. *Nature Communications*, *12*, 718.

Wang, X., He, Y., Sedio, B. E., Jin, L., Ge, X., Glomglieng, S., Cat, M., Yang, J., Swenson, N. G., & Yang, J. (2023). Phytochemical diversity impacts herbivory in a tropical rainforest tree community. *Ecology Letters*, *26*, 1898–1910.

Wang, X., Sun, S., Sedio, B. E., Glomglieng, S., Cao, M., Cao, K. F., Yang, J. H., Zhang, J. L., & Yang, J. (2022). Niche differentiation along multiple functional-trait dimensions contributes to high local diversity of Euphorbiaceae in a tropical tree assemblage. *Journal of Ecology*, *110*, 2731–2744.

Waterhouse, R. M., Seppey, M., Simao, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, *35*, 543–548.

Worthy, S. J., & Swenson, N. G. (2019). Functional perspectives on tropical tree demography and forest dynamics. *Ecological Processes*, *8*, 1–11.

Yang, J., Cao, M., & Swenson, N. G. (2018). Why functional traits do not predict tree demographic rates. *Trends in Ecology & Evolution*, *33*, 326–336.

Yang, J., Zhang, G., Ci, X., Swenson, N. G., Cao, M., Sha, L., Li, J., Baskin, C. C., Slik, J. W. F., & Lin, L. (2014). Functional and phylogenetic assembly in a Chinese tropical tree community across size classes, spatial scales and habitats. *Functional Ecology*, *28*, 520–529.

Yang, R., & Wang, X. (2013). Organ evolution in angiosperms driven by correlated divergences of gene sequences and expression patterns. *Plant Cell*, *25*, 71–82.

Yang, Y., Moore, M. J., Brockington, S. F., Soltis, D. E., Wong, G. K.-S., Carpenter, E. J., Zhang, Y., Chen, L., Yan, Z., Xie, Y., Sage, R. F., Covshoff, S., Hibberd, J. M., Nelson, M. N., & Smith, S. A. (2015). Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution*, *32*, 2001–2014.

Zambrano, J., Iida, Y., Howe, R., Lin, L., Umana, M. N., Wolf, A., Worthy, S. J., & Swenson, N. G. (2017). Neighbourhood defence gene similarity effects on tree performance: A community transcriptomic approach. *Journal of Ecology*, *105*, 616–626.

Zhao, J., Segar, S. T., McKey, D., & Chen, J. (2021). Macroevolution of defense syndromes in *Ficus* (Moraceae). *Ecological Monographs*, *91*, e01428.

Zuidema, P. A., Baker, P. J., Groenendijk, P., Schippers, P., van der Sleen, P., Vlam, M., & Sterck, F. (2013). Tropical forests and global change: Filling knowledge gaps. *Trends in Plant Science*, *18*, 413–419.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Maximum likelihood phylogenetic tree constructed from genome-wide SNPs.

Figure S2. Community functional phylogenomic workflow used in the study.

Table S1. Thirteen *Ficus* species in this study and their abundances in each habitat types in the plot.

Table S2. Transcriptome assembly and annotation quality statistics for the 13 *Ficus* species used in this study.

Table S3. The standardized effect size of Schoener's index for the 13 species of *Ficus* in our study using a 20 m × 20 m spatial scale.

Table S4. Relationship between community dispersion and soil water content in the forest dynamics plot.

Table S5. Mean and standard error of soil volumetric water content (cm^3/cm^3) in three habitats in forest dynamic plot.

How to cite this article: Yang, J., Fan, H., He, Y., Wang, G., Cao, M., & Swenson, N. G. (2024). Functional genomics and co-occurrence in a diverse tropical tree genus: The roles of drought- and defence-related genes. *Journal of Ecology*, *112*, 575–589. https://doi.org/10.1111/1365-2745.14255