



# Leaf Endophytes Relationship with Host Metabolome Expression in Tropical Gymnosperms

Adriel M. Sierra<sup>1,2</sup> · Omayra Meléndez<sup>3,5</sup> · Rita Bethancourt<sup>3</sup> · Ariadna Bethancourt<sup>3</sup> · Lilisbeth Rodríguez-Castro<sup>4,5</sup> · Christian A. López<sup>5,6</sup> · Brian E. Sedio<sup>5,6</sup> · Kristin Saltonstall<sup>5</sup> · Juan Carlos Villarreal A.<sup>1,2,5,7</sup>

Received: 4 December 2023 / Revised: 7 May 2024 / Accepted: 17 May 2024  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

## Abstract

Plant–microbe interactions play a pivotal role in shaping host fitness, especially concerning chemical defense mechanisms. In cycads, establishing direct correlations between specific endophytic microbes and the synthesis of highly toxic defensive phytochemicals has been challenging. Our research delves into the intricate relationship between plant–microbe associations and the variation of secondary metabolite production in two closely related *Zamia* species that grow in distinct habitats; terrestrial and epiphytic. Employing an integrated approach, we combined microbial metabarcoding, which characterizes the leaf endophytic bacterial and fungal communities, with untargeted metabolomics to test if the relative abundances of specific microbial taxa in these two *Zamia* species were associated with different metabolome profiles. The two species studied shared approximately 90% of the metabolites spanning diverse biosynthetic pathways: alkaloids, amino acids, carbohydrates, fatty acids, polyketides, shikimates, phenylpropanoids, and terpenoids. Co-occurrence networks revealed positive associations among metabolites from different pathways, underscoring the complexity of their interactions. Our integrated analysis demonstrated to some degree that the intraspecific variation in metabolome profiles of the two host species was associated with the abundance of bacterial orders Acidobacteriales and Frankiales, as well as the fungal endophytes belonging to the orders Chaetothyriales, Glomerellales, Heliotiales, Hypocreales, and Sordariales. We further associate individual metabolic similarity with four specific fungal endophyte members of the core microbiota, but no specific bacterial taxa associations were identified. This study represents a pioneering investigation to characterize leaf endophytes and their association with metabolomes in tropical gymnosperms, laying the groundwork for deeper inquiries into this complex domain.

**Keywords** Cycads · Defensive traits · Phyllosphere · Biotic interactions · Fungal endophytes · Zamiaceae · Secondary metabolites

## Introduction

Plants have evolved a myriad of adaptive traits to cope with biotic stressors (Kursar & Coley 2003). Through evolutionary spans, species have engaged in interactions exhibiting reciprocal evolutionary responses to adapt to each other's changes—a phenomenon known as coevolution (Ehrlich & Raven 1964; Wink 2003). Among these adaptive mechanisms, the production of specialized small organic molecules (secondary metabolites) stands out as a crucial defense mechanism against natural enemies such as herbivores and pathogens (Sedio 2017; Sedio et al. 2021; Walker et al. 2022). The synthesis of secondary metabolites can be modulated by

multiple biotic interactions, including those with herbivores and inhabiting endophytic microbes, leading to broad intraspecific chemical variation (Salazar et al. 2018; Endara et al. 2017; Fine et al. 2023; Christian et al. 2020).

Mutualistic endophytic microbes, comprising bacteria and fungi, play a pivotal role in inducing plant defensive secondary metabolites such as alkaloids, phenolics, and terpenoids (U'Ren et al. 2019; Walker et al. 2022). For instance, experiments manipulating the plant microbial community have elucidated the patterns observed in plants' chemical defense conferring tolerances to abiotic and biotic stressors (Arnold et al. 2003; Estrada et al. 2013; Mejía et al. 2014; Yamaji et al. 2016; Christian et al. 2020). Advances in high-throughput genomic sequencing and untargeted metabolomic techniques enable the identification

Extended author information available on the last page of the article

of thousands of microbial taxa and metabolites within plant tissues respectively, facilitating an integrative study of the mechanisms behind this mutualistic relationship (Sedio 2017; Christian et al. 2020).

The plant phyllosphere, the encompassing microenvironment of the outer surface (epiphyte) and internal tissues of leaves (endophytes) harbors a diverse microbiome comprising all living domains, including archaea, bacteria, fungi, and viruses (Rosado et al. 2018). Endophytes contribute significantly to various plant functions, such as nitrogen fixation, methanol emission, and biosynthesis of phytohormones (Vorholt 2012; Moyes et al. 2016; Dorokhov et al. 2018). Indeed, the leaf microbiome modulates many plants' functional traits and primary metabolic processes, synergistically promoting host fitness (Lambais et al. 2017; Rosado et al. 2018; Christian et al. 2019; Lajoie et al. 2020). Empirical evidence confirms that plant defensive secondary metabolites are modulated by plant–microbe reciprocal interactions (Christian et al. 2020). In this interaction, secondary metabolites produced by the plant influence the assembly of endophytic microbial species in the leaf tissue (Arnold et al. 2003; Christian et al. 2020), leading to a certain host-distinct associated microbiome. This endophytic microbiota then induces and generates secondary metabolites that influence the viability of both the host and its microbiome (Christian et al. 2020). The diverse functions of the phyllosphere microbiota (Lambais et al. 2017; Lajoie et al. 2020), coupled with the spatial delimitation of the leaf provide an ideal system to explore how plant–microbe interactions induces host defensive metabolites (Estrada et al. 2013; Christian et al. 2015). Despite advances in methodologies toward understanding plant–microbe complexity, only 2% of plant species have been screened for endophytes (Khare et al. 2018; Laforest-Lapointe & Whitaker 2019), and metabolome work has been done in diverse genera of angiosperms and conifers, but not in tropical cycads (Salazar et al. 2018; Sedio et al. 2021).

Cycads are long-lived gymnosperms, that engage in highly obligate plant–insect interactions with phylogenetically unrelated specialized herbivores (i.e., Coleoptera and Lepidoptera) (Salzman et al. 2018; Robbins et al. 2021; Sierra-Botero et al. 2023). Despite possessing neurotoxic and carcinogenic anti-herbivore toxins like methylazoxymethanol acetate (MAM),  $\beta$ -methylamino-L-alanine (BMAA), and azoxyglycosides, cycad leaves may suffer substantial damage from herbivores (Prado et al. 2014; Sierra-Botero et al. 2023). These defensive compounds are the main conserved chemical defensive traits in all cycad genera (Schneider et al. 2002; Whitaker & Salzman 2020). Moreover, the genera *Cycas* and *Zamia* possess diverse bioflavonoids exhibiting cytotoxic, antioxidant, and antipathogenic properties against microbes and protozoans (El-Seadawy et al. 2023a, El-Seadawy 2023b).

Research on the gut microbiome of cycad insect herbivores has shed light on their abilities to tolerate the plant-toxic phytochemicals (Salzman et al. 2018; Robbins et al. 2021; Gutierrez-García et al. 2023).

Cycads have evolved specialized symbiotic associations with microbes, making them essential to understanding the origin and maintenance of plant–microbe interactions and their role in shaping the secondary chemistry across the evolution of land plants (Zheng & Gong 2019; Whitaker & Salzman 2020). Notably, nitrogen-fixing cyanobacteria are the dominant functional group within the coralloid roots of all cycads (Gehring et al. 2010; Yamada et al. 2012; Gutiérrez-García et al. 2018; Zheng et al. 2018; Suárez-Moo et al. 2019; Bell-Doyon et al. 2020). Yet to our knowledge, no studies have addressed the leaf microbiota of cycads, thereby limiting our understanding of the interactions between endophytes and their host plants. Therefore, further investigations of the role of the host microbial endophytes would help to understand the significance of multitrophic interactions in the evolution of cycad defensive metabolomic traits (Whitaker & Salzman 2020).

Here, we investigated the potential role of host-microbe co-interaction on the foliar defensive metabolites of cycads. To study plant–microbe interactions, we selected two phylogenetically close *Zamia* species, with similar leaf anatomical traits (Calonje et al. 2019; Glos et al. 2022). This approach allows us to account for foliar metabolome variation associated with specific differences in host-microbial endophyte in plant species with similar evolutionary history and leaf traits. Moreover, the two focal species differed in their habitat preferences, *Zamia nana* A. Lindstr., Calonje, D.W. Stev. & A.S. Taylor being a terrestrial plant, and *Zamia pseudoparasitica* J. Yates an epiphyte, offering the opportunity to explore differences in leaf endophyte and metabolome in contrasting habitats. Here we ask whether plant chemical variations in cycads are related to host plant species or predominantly driven by microbial endophytic communities (bacteria and fungi). Using a combined approach involving metabarcoding and untargeted metabolomics we aim to answer the question: Does the metabolomic expression of host leaf chemistry, associated with defense, correlate with the microbial endophyte community composition? Given the specialized plant–herbivore association in cycads, we predict that mutualistic endophytes enhance the production of host secondary metabolites, thereby accounting for intraspecific variation attributable to differences in microbial communities. We expect that this variation is more pronounced in the epiphytic species because individuals will be prone to vertical microclimate variation. Consequently, we propose that the composition of the endophytic community significantly influences the diversity and abundance of secondary metabolites within the genus *Zamia*.

## Methods

### Study System

We focus on two species of *Zamia* (Zamiaceae) from Panama. Individual plants of the Neotropical genus *Zamia* are characterized by their long-lived palm-like compound leaves. *Zamia nana* and *Zamia pseudoparasitica* are two endemic species restricted to the premontane to montane forest on the Caribbean slope of western Panama (50–1000 m.a.s.l), under similar environmental conditions with daily temperatures of 18–29 °C and annual rainfall of 2500–4000 mm, with average monthly rain < 60 to 353 mm (Bell-Doyon & Villarreal 2020). *Zamia nana* is a terrestrial species that occurs in a single native population in Cerro Gaital (Coclé) (Lindström et al. 2013), nearly ~40 km apart from the closest population of *Z. pseudoparasitica*. *Zamia pseudoparasitica* is the only known epiphytic gymnosperm in the world and it grows attached to the trunks and nestled in lower forks of large canopy trees at heights of 7 to 20 m (Bell-Doyon & Villarreal 2020). Both species are nested within the isthmian clade that recently diverged ~3 Ma years ago (Calonje et al. 2019). A total of 73 individuals ( $n=30$  of *Z. nana* and  $n=43$  of *Z. pseudoparasitica*) were selected for this study. We collected one leaflet per plant and were stored at 4 °C in the field for less than three hours and then taken to the laboratory for processing. Macroepiphytes from leaflet samples were removed by scratching the surface with a sterile sponge. Then leaflet rectangular punches of 10 × 5 mm were surface sterilized using a 70% ethanol wash for 2 min, 1% Sodium hypochlorite for 3 min, and sterile distilled water for 1 min. Two leaflet punches from individual plants were separated and stored at -80 °C for untargeted metabolomics, and DNA extraction for bacterial and fungal metabarcoding analyses (Supplementary Table 1). The leaflet tissue was flash frozen in liquid nitrogen and pulverized using a tissue homogenizer with metallic beads.

### Untargeted Metabolomics

Metabolites were extracted from 10 mg pulverized leaflet tissue from nine individuals of *Z. nana* and 42 individuals of *Z. pseudoparasitica* (Supplementary Table 1), using 1.8 ml (90:10 v/v) methanol:water pH 5 overnight at 4 °C and shaking at 300 rpm, then centrifuged at 14000 rpm for 10 min (Sedio et al. 2018). The supernatant was transferred and filtered for LC–MS analysis (Sedio et al. 2021). Extracts were separated using a Vanquish Horizon Duo ultra-high performance liquid chromatography (UHPLC)

system (Thermo Fisher, Waltham, MA, United States) with an Accucore C18 column with 150 mm length, 2.1 mm internal diameter, and 2.6 μm particle size. UHPLC buffer A (0.1% v/v formic acid in water) and buffer B (0.1% v/v formic acid in methanol) were employed in a solvent gradient from 5 to 100% buffer B over 18 min. The UHPLC–MS parameters were optimized to detect and fragment metabolites representing a wide range in polarity and mass (Sedio et al. 2018). Separation of metabolites by UHPLC was followed by heated electrospray ionization (HESI) in positive mode using full scan MS1 and data-dependent acquisition of MS2 (dd-MS2) on a Q Exactive hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher). The structural similarity of each unique metabolite was calculated based on the mass-to-charge ratio ( $m/z$ ) of the fragments for every pair of compounds found in the *Zamia* species. Consensus spectra in individual chemical extracts were inferred using The Global Natural Products (GNPS) Molecular Networking tool by clustering the MS/MS spectra into respective unique structures (Wang et al. 2016). The GNPS method compares the structural similarity of every molecule pair using the cosine of the angle between vectors defined by the  $m/z$  values of their constituent fragments with a cosine threshold of  $\cos = 0.7$  (Wang et al. 2016). We calculated metabolite concentration based on peak areas/mg, the number of metabolites per individual, and chemical variation among the samples. Chromatograms were aligned using MZmine2 (Pluskal et al. 2010), and then we inferred chemical molecular formulas using Sirius (Dührkop et al. 2019) and predicted their structures with CSI:fingerID (Dührkop et al. 2015).

To compare intra- and interspecific host chemical variation between the two species, first, we tested for differences in the number of compounds observed using richness metrics (Chao, Shannon, and Simpson indexes). Secondly, we used Qemistree (Tripathi et al. 2021) to build a hierarchical dendrogram that reflects the metabolite structural similarity between the compounds observed. The phylogenetic diversity Faith index was calculated using the *picante* package (Kembel et al. 2010) used to represent structural diversity among the metabolites. Statistical differences for the four indexes were addressed using the Permutation Welch Two Sample *t*-test, a reliable test between group means with unequal population variances and sample sizes. Additionally, we searched for differentially expressed metabolites between the two host species using the DESeq2 package (Love et al. 2014) and used the student's *t*-test to evaluate significance when  $P < 0.05$  and fold change  $> 10$ .

We described metabolite co-occurrence using the software *Cytoscape* and *CoNet* co-occurrence network analyses (Shannon et al. 2003; Faust & Raes 2016). Significant interactions (threshold  $P < 0.05$ ) were estimated using Pearson

and Spearman correlations, a mutual information similarity, a Bray–Curtis dissimilarity, and the Kullback–Liebler dissimilarity index. A co-occurrence network was generated based on 100 iterations with renormalization and null distribution computing. Estimated metrics were compared to a null distribution by bootstrapping the metrics under 1000 iterations to assess statistical significance. Correlation metrics were combined with the Brown method, and Benjamin Hochberg multiple-test corrections were applied to retain edges with  $q$ -values  $< 0.05$ .

To compare structural metabolic composition between host species we calculated the chemical structural-compositional similarity (CSCS) metric, which accounts for the structural similarity of compounds between pairs of individuals or species (Sedio et al. 2017). This metric weight the similarity of pairwise comparison of the metabolite composition between the two species based on the product of their proportional ion intensities in each species. The CSCS metric was calculated for all individual pairs of the two species. We also calculated the Bray–Curtis dissimilarity metric to infer specific variations in shared metabolic compounds between species. Using the two metrics we tested if chemical composition significantly varies between species using a one-way analysis of variance (ANOVA) with species as the explanatory variable using the *car* package (Fox & Weisberg 2011). Chemical variation was visualized using non-metric multidimensional scaling (NMDS) with functions from the *MASS* package (Venables & Ripley 2002). Permutation tests for homogeneity of multivariate dispersion (PERMDISP2) were used to test for chemical intraspecific variation with the *vegan* package (Oksanen et al. 2016). Significance values were obtained by permuting the model residuals of a null model analysis of variance (ANOVA) 1000 times.

### DNA Extraction and Amplicon Sequencing Library

Genomic DNA was extracted from ~30 mg of leaflet tissue using a cetyl trimethylammonium bromide (CTAB) extraction protocol (Doyle & Doyle 1987) and DNA concentrations were evaluated using a Nanodrop 1000 spectrophotometer (ThermoFisher, Waltham, USA).

A two-step dual-indexed PCR approach specifically designed for Illumina instruments was carried out to amplify two molecular markers. In the first stage, bacterial and fungal leaf endophytes were amplified using primers that included the Illumina sequencing primer on the 5' end. The 16S ribosomal rRNA gene V3-V4 region was targeted to amplify bacteria using the primer set V3-V4\_F\_341F/V3-V4\_R\_805R (Klindworth et al. 2013). The fungal microbiota was characterized with the internal transcribed spacer (ITS1) region using the primer set ITS1F/ITS2R (White et al. 1990). A second PCR reaction was performed to add dual-indexes and Illumina flow cell adaptors to all

amplicons. PCR products from the previous steps were purified with magnetic microbeads and successive washes in 80% ethanol on a magnetic rack. The complete and detailed PCR protocol followed is provided in the Supplementary methods. Final products were quantified spectrophotometrically using a Spark 10 M plate reader (Tecan US Inc., Raleigh, USA), and pooled in equimolar concentration ratio, for sequencing on the Illumina Miseq (Illumina Inc., San Diego, USA). Sequencing was performed at the platform of genomic analysis of the Institute of Integrative Biology and Systems, Université Laval, Québec, Canada (16S; 2 × 300 bp run), and the Naos Molecular Lab (ITS1; 2 × 250 bp run) at the Smithsonian Tropical Research Institute, Panama.

### Amplicon Sequence Assembly and Filtering

Sequence adapters and primers were removed for the 16S and ITS1 markers using the software *cutadapt* (Martin 2011). With the package *DADA2* (Callahan et al. 2016; Callahan et al. 2017) raw reads were filtered based on the quality profile (Phred quality score) and trimmed as follows: forward reads at 280 bp and reverse reads at 210 bp for 16S, and forward and reverse reads at 220 bp for ITS1. Expected errors were set for 16S (maxEE = 4), while ITS1 at (maxEE = 2). For both markers default settings follow for ambiguous nucleotide (maxN = 0) and probability of erroneous assignment (truncQ = 2). Sequences were dereplicated and merged by aligning forward and reverse reads with an overlap of a minimum of 20 base pairs. Amplicon Sequence Variants (ASVs) were inferred for each contig, and chimera sequences were removed (Callahan 2016). Taxonomy was assigned to bacterial ASVs with the bacterial SILVA reference database version 132 (Pruesse et al. 2007), and fungal ASVs against the UNITE community database of fungi (Nilsson et al. 2018) with the *DADA2* taxonomic implementation of the naive Bayesian RDP classifier. ASV sequences of 16S were aligned using the *DECIPHER* package with default parameters (Wright 2016). A maximum likelihood phylogenetic tree was built under the GTR + GAMMA + I DNA evolution model, with a neighbor-joining tree with the *phangorn* package (Schliep 2011). Due to the difficulties of aligning ITS1 amplicon sequences from different fungal genera, phylogenetic-based methods were not employed for the fungal endophytes.

All ASV sequences from the host plant, those identified as chloroplast, mitochondria, eukaryotes, or any ASVs that remained unidentified (i.e., “NA”) at the kingdom level were excluded from the dataset. The bacterial (Bdellovibrionota, Elusimicrobiota, Gemmatimonadota, and NB1-j), and fungal (Basidiobolomycota, Chytridiomycota, Glomeromycota, Kickxellomycota, and Rozellomycota) phyla, each of which included fewer than 10 ASVs, in the dataset were excluded. Laboratory external contaminants sequenced in

PCR negative controls were identified in our study samples by performing ASVs frequency-based and prevalence-based methods with the *decontam* package (Davis et al. 2018). We excluded any sample that after filtering remained with low sequencing depth presenting less than 50 reads per sample (Supplementary Table 1). The final dataset resulted in 1129 ASVs from 32 samples for the bacterial marker 16S (*Z. nana* 594 ASVs and *Z. pseudoparasitica* 588 ASVs) and 7945 fungal ASVs from 64 samples for the fungal marker ITS1 (*Z. nana* 3111 ASVs and *Z. pseudoparasitica* 5342 ASVs). These final datasets were used to characterize the alpha and beta diversity in the endophyte communities.

### Leaf Endophyte Community Diversity

For alpha and beta-diversity analysis, we normalized our data by calculating the relative abundance of each taxon based on the sample sequence coverage (McMurdie & Holmes 2014; Weiss et al. 2017). Alpha diversity was assessed using Hill numbers with three richness indices that accounted for differences in the weight of common taxa: (1) the Chao index, (2) the Shannon exponential, which weights the ASVs by their frequency, and 3) the Inverse Simpson's diversity index which considers the presence and relative abundances of ASVs present in a sample. The Pielou index was used as another measure of species' evenness between plant hosts. The phylogenetic diversity with the Faith index was estimated for the bacterial communities, using the *picante* package (Kembel et al. 2010). The assumptions of normality and homogeneity of variance were tested with the Shapiro–Wilk and Levene's tests, respectively, for all alpha diversity measurements. As all alpha diversity metrics were not normally distributed according to Shapiro–Wilk tests ( $P < 0.001$ ) and Levene's test was not significant for the five measures of alpha diversity ( $P > 0.05$ ), indicating that there were equal variances in the diversity metrics between groups (host species), significant differences between host species were assessed using the Kruskal–Wallis nonparametric test, and *post-hoc* pairwise comparisons were performed with the exact sum of Wilcoxon ranks, and applying a  $< 5\%$  Bonferroni correction.

Principal Coordinate Analysis (PCoA), with two non-phylogenetic metrics and three phylogenetic distance metrics, was used to visualize bacterial community variation within and between *Zamia* species. The multiple distance metrics allow us to explore differences in endophytic microbial composition between *Zamia* species by accounting for the presence of ASVs, their relative abundances, and the different phylogenetic relationships of taxa in the dataset. The two non-phylogenetic metrics allow us to examine how communities differ according to the presence or absence of ASVs (Jaccard), and according to the relative abundance of species present in the community (Bray–Curtis). The

phylogenetic metrics of Unifrac were calculated using the package GUniFrac (Chen et al. 2012). The Unifrac metrics consider the phylogenetic relatedness of ASV lineages present in the community (Unweighted Unifrac), accounting for the relative abundances of the ASV lineages present (Weighted Unifrac) or controlling the weight of abundant lineages of ASV with the parameter  $\alpha = 0.5$  (Generalized Unifrac). PCoA for the fungal endophytic community was restricted to non-phylogenetic metrics. Permutation tests for homogeneity of multivariate dispersion (PERMDISP2) were used to test the hypothesis that the *Zamia* endophytic communities are more variable between individuals of each species than across species with the *vegan* package (Oksanen et al. 2016). The  $P$  values were obtained by permuting the model residuals of a null model analysis of variance (ANOVA) 1000 times with the *car* package (Fox & Weisberg 2011). Significant compositional differences in the microbial community between *Zamia* species were tested by analyzing sums of squares from the distance matrix using the *vegan* package (Oksanen et al. 2016). The analysis of permutational multivariate of variance (PERMANOVA) was computed with 1000 permutations and  $P$  values were calculated according to  $< 5\%$  Bonferroni correction.

Core microbiota was identified using the threshold-based frequency method (Salonen et al. 2012) in the *microbiome* package (Lahti & Shetty 2017). We used variable minimal threshold percentages of taxon prevalence from 50%, 75%, and 90% to identify the core members of the bacterial and fungal community. For each member identified as a member of the core microbiota, we estimated the relative abundance by sample.

### Correlation of Metabolome and Leaf Endophytic Microbiota

We integrated the metabolomic and metabarcoding data to investigate the relationship of host secondary metabolites with the microbial endophyte community. Of the 32 samples in the bacterial dataset, 19 had the corresponding metabolomic data, while from the 64 samples in the fungal dataset, 44 had metabolomic data. Seventeen samples had all three of the datasets included in this study.

To investigate whether similarity in host endophyte communities will correspond to equally similar metabolite composition, we correlated the ordination Axis1 of the chemical ordination with the Axis1 of bacteria ( $n = 19$ ) and fungal ( $n = 44$ ) ordinations. We fitted multiple linear regression using Axis1 of chemical NMDS ordinations (CSCS and Bray–Curtis metrics) with Axis1 of the PCoA ordination with the Bray–Curtis for bacteria and fungi, and Weighted Unifrac metrics for bacteria. A significant relationship was considered when  $P < 0.05$ . Since we aimed to test for intraspecific chemical variations related to the presence and

abundance of specific microbial endophytic taxa, our statistical correlation model did not use host species identity as an explanatory variable.

Using the ordinations generated with CSCS and Bray–Curtis, we regressed plot-level measures of individual host plant bacterial and fungal diversity, as well as the abundance of predominant taxa orders and members of the core microbiota. Microbiota variables were fit onto the metabolome ordinations using the *vegan* package (Oksanen et al. 2016). Relationships were considered moderate  $P < 0.1$  to strongly  $P < 0.001$  significant. Since the integration of the datasets resulted in some samples with foliar metabolome comprising only one data for the two associated microbiota (bacteria and fungi), we explored the correlative association of the dataset between metabolome-bacteria ( $n = 19$ ), metabolome-fungi ( $n = 44$ ) and metabolome-bacteria-fungi ( $n = 17$ ). All analyses were done using the software R v. 03.1 + 446 (R Core Team 2023) with the functions in the packages cited in the methods.

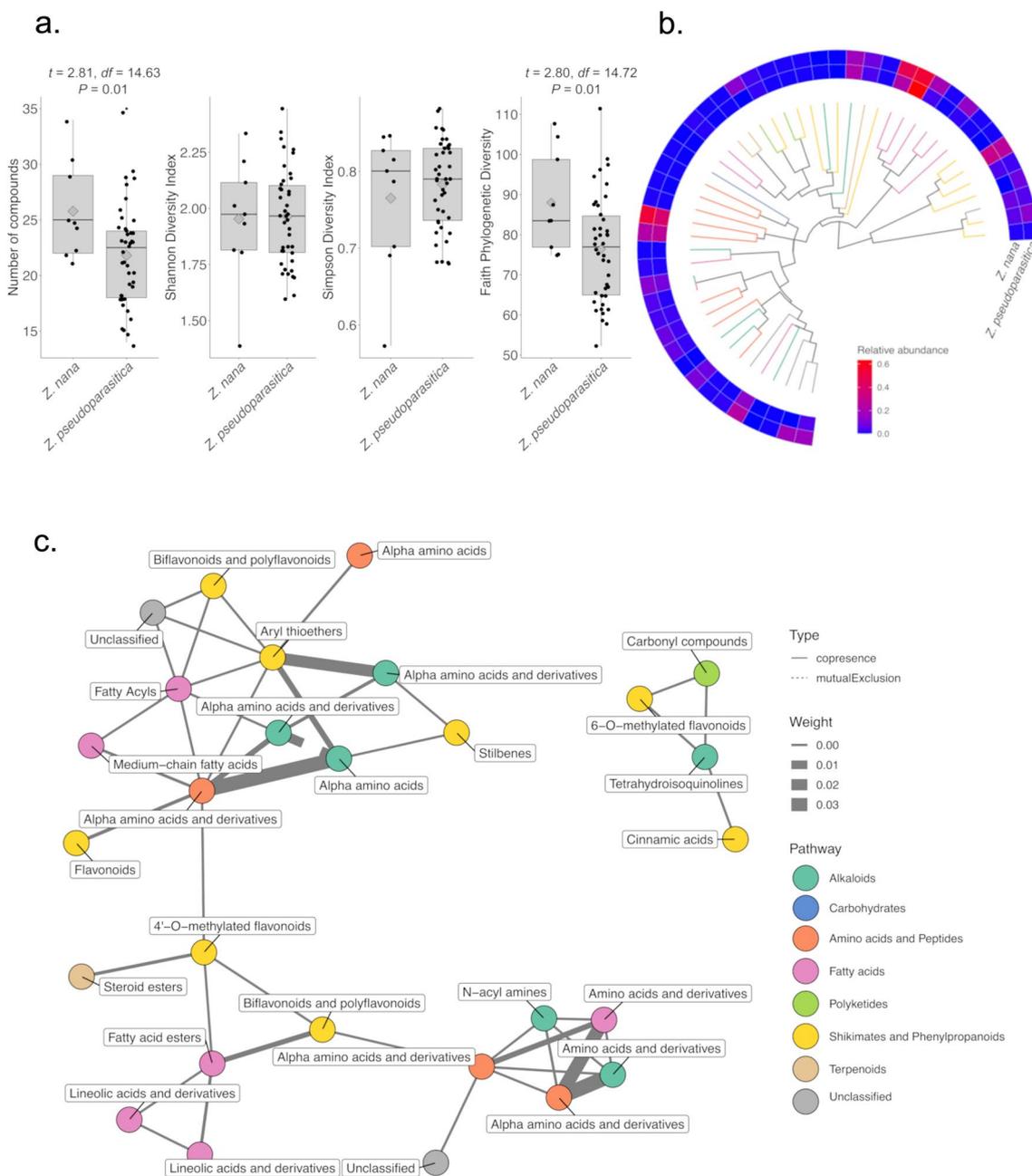
## Results

Forty-nine metabolites, ranging from 137.06 to 595.47 Daltons (Da), were found in 51 individuals of the two *Zamia* species. We observed five unique metabolites in *Z. pseudoparasitica*, that were not detected in *Z. nana*. The molecular structure of 44 and the chemotaxonomic classification of 25 of these compounds were predicted, that based on the NPClassifier corresponded to the following biosynthetic pathways: alkaloids (7), amino acids and peptides (9), carbohydrates (1), fatty acids (11), polyketides (2), shikimates and phenylpropanoids (11), and terpenoids (3). These classified compounds correspond to the following classes: carboxylic acids and derivatives (16), fatty acyls (8), and flavonoids (7), among others (Supplementary Fig. 1; Supplementary Table 2).

The mean number of chemical compounds observed was significantly higher in *Z. nana* ( $\bar{X} = 26 \pm 1.3$ ) compared to *Z. pseudoparasitica* ( $\bar{X} = 21.8 \pm 0.7$ ) (Fig. 1a). However, the Shannon entropy (*Z. nana*:  $\bar{X} = 1.92 \pm 0.08$ ; *Z. pseudoparasitica*:  $\bar{X} = 1.96 \pm 0.03$ ), and Simpson evenness (*Z. nana*:  $\bar{X} = 0.75 \pm 0.03$ ; *Z. pseudoparasitica*:  $\bar{X} = 0.78 \pm 0.01$ ) diversity indices showed no differences in chemical diversity (Shannon:  $t = -0.38$ ,  $df = 11.81$ ;  $P > 0.05$ ; Simpson:  $t = -0.09$ ,  $df = 10.64$ ;  $P > 0.05$ ). In contrast, the phylogenetic diversity index, representing structural diversity, was significantly higher in *Z. nana* ( $\bar{X} = 88.50 \pm 3.8$ ) than in *Z. pseudoparasitica* ( $\bar{X} = 76.39 \pm 2.0$ ). The shared metabolites between the two species were observed in similar relative abundance (Fig. 1b). The most abundant compounds were phenyl sulfoxides, aryl-phenylketones, and tetrahydroisoquinolines in

both species and 1,2-diacylglycerols and dipeptides were more abundant in *Z. pseudoparasitica*. The co-occurrence network showed three main significant interactions between the 45 nodes (of the 49 metabolites) joined by 28 edges (Supplementary Table 3). Forty-four of the nodes showed a co-occurrence interaction (solid edges in Fig. 1c), and there was only one mutually exclusive interaction (dashed edge in Fig. 1c). The network showed that the more abundant tetrahydroisoquinolines, interact with two flavonoids and a polyketide compound (Fig. 1c). Alpha amino acids with the alkaloid pathway, strongly interact with four flavonoids (shikimates and phenylpropanoids) and fatty acids compounds; while two other alkaloids (N-acyl amines, and amino acids and derivatives), were positively associated with two amino acids and peptides, a fatty acid, and an unclassified compound. One group of nodes in the network of fatty acids (esters, linoleic acids, and derivatives) had a strong positive interaction with flavonoid compounds (Fig. 1c). Five compounds were observed to be upregulated in *Z. pseudoparasitica* in contrast to *Z. nana* (Fold change = 30,  $P < 0.001$ ) (Supplementary Fig. 2)). The five upregulated compounds correspond to the shikimate and phenylpropanoids, fatty acid, and amino acid biosynthetic pathways. Based on the most specific classification from ClassyFire, these metabolites correspond to Benzodioxoles, Biflavonoids and polyflavonoids, 1,2-diacylglycerols, N-acyl-alpha amino acids and Glutamic acid and derivatives.

We then examined leaf microbial endophytes and compared them to the host chemical similarity. A detailed description of the alpha and beta-diversity patterns found in the endophyte microbial community is presented as supplementary results. We found that host species generally showed similar variation in bacterial alpha-diversity (Supplementary Fig. 4a; Supplementary Table 6) and no bacterial ASVs were prevalent in all individuals to be considered a member of a core microbiota either within or across the two *Zamia* species. Similar bacterial species richness (alpha-diversity) and overlap of bacterial species composition (beta-diversity) were observed between the two species, with *Z. nana* showing higher intraspecific variation in their endophyte community, based on the Jaccard and Bray–Curtis metrics (Supplementary Fig. 5; Supplementary Tables 7, 8). The fungal endophytic community showed significant differences in both alpha- and beta-diversity between the two species, with fungal endophytes being more diverse in *Z. pseudoparasitica* despite the average read count per sample being half that of *Z. nana* (Supplementary Figs. 4b, 6; Supplementary Tables 7, 8). We identified 34 fungal ASVs as members of the core microbiota with prevalence in at least 50%–75% of the individuals sampled within each *Zamia* species. These fungal endophytes were classified into eleven orders of Ascomycota and three of Basidiomycota (Table 1).



**Fig. 1** Comparative leaf metabolomics of *Zamia* species. **a.** Individual chemical diversity: mean number of compounds observed, Shannon diversity index, evenness Simpson diversity index, and chemical structural variation based on phylogenetic diversity (Faith Index). Statistically significant differences are given with their respective *P* values. **b.** Hierarchical dendrogram of 49 *Zamia* foliar metabolites classified by their biochemical pathway and branches coded by color (NPC classification). Qemistree illustrates the structural similarity of unique metabolites in a phylogenetic tree, a hierarchical dendrogram

in which structurally related metabolites form sister pairs and clades. A heatmap corresponding to terminal branches illustrates the relative abundances of each leaf metabolite for each species. **c.** Co-occurrence chemical network of significant co-occurrences (solid line) or mutually exclusive (dashed line) of metabolites observed in *Zamia*. Branch color in the phylogenetic tree and node color on the network graph indicate their biochemical pathway. Node labels indicate the core biosynthetic pathway from which each metabolite is derived, based on the NPCClassifier chemical classification scheme

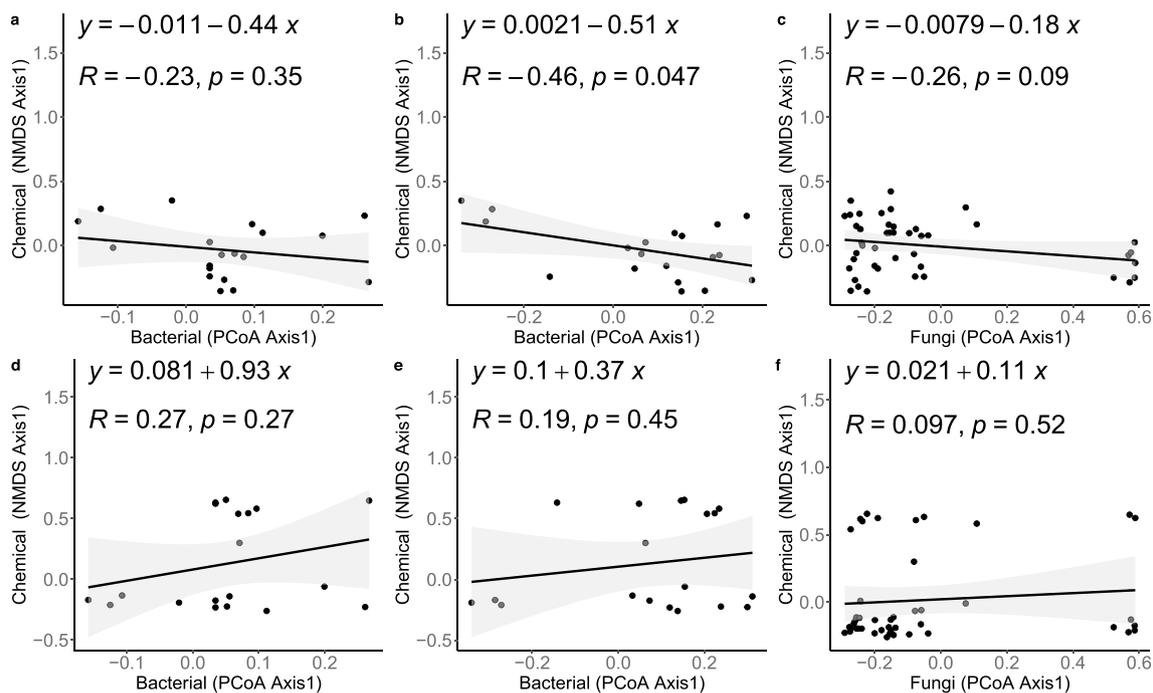
A single fungal taxon, AVS20, of the genus *Colletotrichum* (Ascomycota), was included in the core microbiota of both species in 50% of the individuals.

We investigated the relationships between host chemical variation and microbial endophyte community variation for bacteria ( $n = 19$ ) and fungi ( $n = 44$ ). The linear regression

**Table 1** Core fungal endophytes of *Z. nana* and *Z. pseudoparasitica*, as identified using a threshold-based frequency method. The complete taxonomic classification for each ASV and the minimal threshold

detection point is given. NA indicates that the ASV was not classified at a given taxonomic level

| <i>Zamia nana</i>             |         |               |                 |                   |                       |                         |                            |           |
|-------------------------------|---------|---------------|-----------------|-------------------|-----------------------|-------------------------|----------------------------|-----------|
| ASV                           | Kingdom | Phylum        | Class           | Order             | Family                | Genus                   | Species                    | Threshold |
| ASV1                          | Fungi   | Ascomycota    | Dothideomycetes | Capnodiales       | Cladosporiaceae       | <i>Cladosporium</i>     | <i>delicatulum</i>         | 75%       |
| ASV18                         | Fungi   | Ascomycota    | Sordariomycetes | Glomerellales     | Glomerellaceae        | <i>Colletotrichum</i>   | <i>gigasporum</i>          | 50%       |
| ASV20                         | Fungi   | Ascomycota    | Sordariomycetes | Glomerellales     | Glomerellaceae        | <i>Colletotrichum</i>   | <i>annellatum</i>          | 50%       |
| ASV22                         | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Saccharomycetales     | <i>Candida</i>          | <i>pseudojiufen-gensis</i> | 50%       |
| ASV29                         | Fungi   | Ascomycota    | Dothideomycetes | Pleosporales      | NA                    | NA                      | NA                         | 50%       |
| ASV59                         | Fungi   | Basidiomycota | Agaricomycetes  | Polyporales       | Meruliaceae           | <i>Scopuloides</i>      | NA                         | 50%       |
| ASV62                         | Fungi   | Ascomycota    | Sordariomycetes | Hypocreales       | NA                    | NA                      | NA                         | 50%       |
| ASV68                         | Fungi   | Ascomycota    | NA              | NA                | NA                    | NA                      | NA                         | 50%       |
| ASV73                         | Fungi   | Ascomycota    | Dothideomycetes | Capnodiales       | Cladosporiaceae       | <i>Cladosporium</i>     | NA                         | 50%       |
| ASV112                        | Fungi   | Ascomycota    | NA              | NA                | NA                    | NA                      | NA                         | 50%       |
| ASV129                        | Fungi   | Basidiomycota | Agaricomycetes  | Auriculariales    | Exidiaceae            | <i>Heterochaete</i>     | <i>shearii</i>             | 50%       |
| ASV130                        | Fungi   | Ascomycota    | Dothideomycetes | Pleosporales      | NA                    | NA                      | NA                         | 50%       |
| ASV145                        | Fungi   | Ascomycota    | Dothideomycetes | Capnodiales       | Cladosporiaceae       | <i>Cladosporium</i>     | <i>delicatulum</i>         | 50%       |
| ASV272                        | Fungi   | Ascomycota    | Sordariomycetes | Hypocreales       | Nectriaceae           | <i>Nectria</i>          | <i>bactridioides</i>       | 50%       |
| <i>Zamia pseudoparasitica</i> |         |               |                 |                   |                       |                         |                            |           |
| ASV                           | Kingdom | Phylum        | Class           | Order             | Family                | Genus                   | Species                    | Threshold |
| ASV2                          | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Debaryomycetales      | <i>Debaryomyces</i>     | <i>hansenii</i>            | 75%       |
| ASV4                          | Fungi   | Ascomycota    | Sordariomycetes | Xylariales        | Sporocadaceae         | <i>Pestalotiopsis</i>   | <i>coffea-arabicae</i>     | 75%       |
| ASV5                          | Fungi   | Ascomycota    | Sordariomycetes | Glomerellales     | Glomerellaceae        | <i>Colletotrichum</i>   | <i>cymbidiicola</i>        | 50%       |
| ASV20                         | Fungi   | Ascomycota    | Sordariomycetes | Glomerellales     | Glomerellaceae        | <i>Colletotrichum</i>   | <i>annellatum</i>          | 50%       |
| ASV21                         | Fungi   | Basidiomycota | Tremellomycetes | Tremellales       | Rhynchogastremataceae | <i>Papiliotrema</i>     | <i>flavescens</i>          | 75%       |
| ASV40                         | Fungi   | Ascomycota    | Sordariomycetes | Xylariales        | NA                    | NA                      | NA                         | 50%       |
| ASV48                         | Fungi   | Ascomycota    | Eurotiomycetes  | Chaetothyriales   | NA                    | NA                      | NA                         | 50%       |
| ASV49                         | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Debaryomycetales      | <i>Debaryomyces</i>     | <i>nepalensis</i>          | 75%       |
| ASV50                         | Fungi   | Ascomycota    | Dothideomycetes | Pleosporales      | Cucurbitariaceae      | <i>Pyrenochaetopsis</i> | <i>leptospora</i>          | 75%       |
| ASV54                         | Fungi   | Ascomycota    | Leotiomycetes   | Helotiales        | NA                    | NA                      | NA                         | 50%       |
| ASV60                         | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Saccharomycetales     | <i>Candida</i>          | <i>haemulonis</i>          | 75%       |
| ASV90                         | Fungi   | Ascomycota    | Sordariomycetes | Sordariales       | NA                    | NA                      | NA                         | 50%       |
| ASV100                        | Fungi   | Ascomycota    | Sordariomycetes | NA                | NA                    | NA                      | NA                         | 60%       |
| ASV104                        | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Metschnikowiaceae     | <i>Kodamaea</i>         | <i>ohmeri</i>              | 50%       |
| ASV105                        | Fungi   | Ascomycota    | Sordariomycetes | Hypocreales       | Nectriaceae           | <i>Neonectria</i>       | NA                         | 70%       |
| ASV113                        | Fungi   | Ascomycota    | Dothideomycetes | Botryosphaeriales | Botryosphaeriaceae    | <i>Lasiodiplodia</i>    | <i>brasiliensis</i>        | 50%       |
| ASV137                        | Fungi   | Basidiomycota | Tremellomycetes | NA                | NA                    | NA                      | NA                         | 50%       |
| ASV152                        | Fungi   | Ascomycota    | Sordariomycetes | Hypocreales       | Nectriaceae           | <i>Ilyonectria</i>      | <i>destructans</i>         | 50%       |
| ASV311                        | Fungi   | Ascomycota    | Sordariomycetes | Hypocreales       | Stachybotryaceae      | <i>Myxospora</i>        | <i>aprootii</i>            | 50%       |
| ASV495                        | Fungi   | Ascomycota    | NA              | NA                | NA                    | NA                      | NA                         | 50%       |
| ASV1034                       | Fungi   | Ascomycota    | Saccharomycetes | Saccharomycetales | Saccharomycetales     | <i>Candida</i>          | <i>spencermartinsiae</i>   | 50%       |



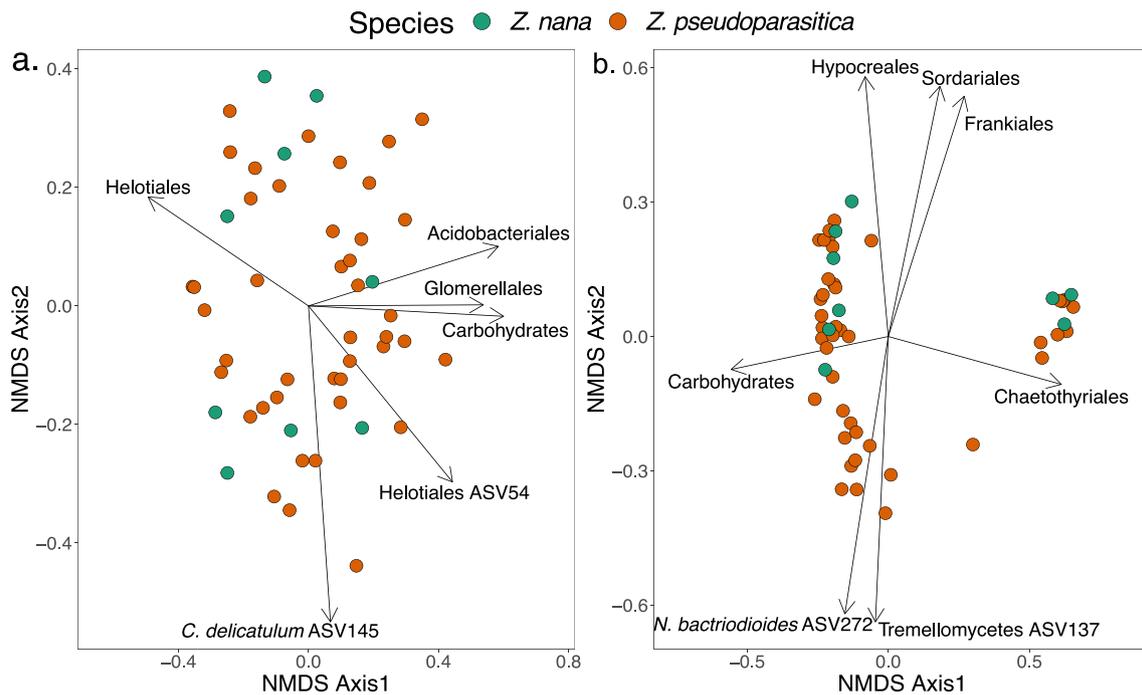
**Fig. 2** Linear correlation between chemical variation summarized in the NMDS Axis1 from the CSCS chemical similarity metrics with endophytic microbial community variation. based on the PCoA Axis1 for (a) bacterial Bray–Curtis dissimilarity index, (b) bacterial Weighted Unifrac and (c) fungal Bray–Curtis dissimilarity index respectively. The second row shows chemical similarity based on the

Bray–Curtis dissimilarity relationship with the previous endophytic microbial community variation metrics (d) bacteria using Bray–Curtis, (e) bacteria using Weighted Unifrac, and (f) fungal Bray–Curtis. Linear regression equation,  $R^2$ , and  $P$  values are given for each regression

showed a relationship between the chemical variation along the NMDS Axis1 with the endophyte bacterial and fungal community variation along the PCoA Axis1 (Fig. 2). Between 19 to 46% of *Zamia* leaf chemical variation was explained by bacterial community variation, however, it was only significant when NMDS Axis1 from Bray–Curtis (chemical) was regressed against the PCoA Axis1 from Weighted Unifrac metric ( $P = 0.05$ ). On the other hand, chemical variation was negatively related when the NMDS Axis1 from CSCS chemical similarity was regressed against the PCoA Axis1 of the fungal community (Bray–Curtis) explaining 26% of the variation ( $P = 0.10$ ). When chemical variation was summarized with the NMDS Axis1 using Bray–Curtis metrics, it was not related to the fungal endophyte community along the PCoA Axis1 (Bray–Curtis).

Host species chemical similarity was summarized in the non-metrical distance space ordination (Fig. 3) represented by the 51 individuals. Chemical composition variation indicated high metabolomic similarity between *Zamia* species (Supplementary Table 4. PERMANOVA: CSCS  $F = 0.90$ ,  $P = 0.70$ ; Bray–Curtis  $F = 1.13$ ,  $P = 0.31$ ). Ordination showed that within-plant intraspecific variation of secondary metabolites composition diverged along both NMDS axes, with two diverging clusters along Axis2 (Fig. 3a) and Axis1 (Fig. 3b). These two clusters using the

CSCS metrics, which considered the structural similarity of metabolites, were significantly related to differences in the abundance of one bacterial order (Acidobacteriales) and two fungal orders (Helotiales, and Glomerellales). Specific taxa of the core endophytic fungal community were correlated to chemical variation between the two clusters along the NMDS Axis2 (Fig. 3a). The taxa classified as Helotiales ASV54 and *Cladosporium delicatulum* ASV145 were correlated along the negative coordinates of Axis2 of the ordination. This analysis with a subset dataset with samples that have the complete integrated metabolomic and microbiota ( $n = 17$ ), showed the two clusters and significant correlations for the order Acidobacteriales and *C. delicatulum* ASV145 (Supplementary Fig. 7a). When the Bray–Curtis metric was used, the two clusters were significantly related to the fungal orders Sordariales and Hypocreales, and the bacterial order Frankiales along the positive coordinates of NMDS Axis2 (Fig. 3b), while along the NMDS Axis1 the samples with two upregulated metabolites of the superclass amino acid and peptide, were enriched with ASVs of the fungal order Chaetothyriales. Two members of the fungal core microbiota (*Nectria bactriodioides* ASV272 and ASV137 in the class Tremellomycetes) were correlated to chemically similar individuals of *Z. pseudoparasitica* along the negative coordinates of the NMDS Axis2. Some



**Fig. 3** Leaf metabolome non-metric multidimensional scaling (NMDS) from pairwise CSCS (**a**) and Bray-Curtis (**b**) metrics showing chemical similarity of 51 individuals of two *Zamia* species (green points: *Zamia nana*; and orange points: *Zamia pseudoparasitica*). Line segments correspond to plot-level regression of significant rela-

tionships ( $P < 0.1$ ) chemical similarity to chemical pathway abundances, leaf endophyte microbial order abundances, and fungal core microbiota identified based on their prevalence in 50% of samples (Table 1)

samples in this cluster show upregulation of one metabolite with the fatty acid pathway of the superclass glycerolipids. This pattern was also observed when a subset of 17 samples with the complete integrated dataset was analyzed, yet only the fungal orders Sordariales and Chaetothyriales were significant correlated, and the core taxa *N. bactriodioides* ASV272 and Tremellomycetes ASV137 remained significant (Supplementary Fig. 7b).

## Discussion

The coevolution of cycads and associated microbes has yielded a highly specialized biotic association within their plant tissues (Suárez-Moo et al. 2019; Bell-Doyon et al. 2020; Zheng & Gong 2019). Yet, connecting specific plant-microbe interactions to the highly toxic and diverse defensive metabolites observed in cycads has been significantly challenging due to the difficulties of integrating different data types (Cox et al. 2003; Marler et al. 2010; Mantas et al. 2022). We relate intraspecific variation of host leaf secondary chemistry to differences in the composition of the endophytic community. The individual intraspecific metabolomic similarity in the two *Zamia* species was associated with an increased abundance of bacterial and fungal

orders, while four ASVs core members of the fungal microbiota were enriched in individuals with similar metabolomic composition.

Untargeted metabolomics revealed that both species of *Zamia* shared metabolites in similar abundance and biosynthetic pathways: alkaloids, amino acids and peptides, carbohydrates, fatty acids, polyketides, shikimates, and phenylpropanoids, and terpenoids. Co-occurrence networks of metabolites in the two species suggested a complex association, leading to the co-expression among and between metabolic pathways. This highlights the complexity of inducible defensive secondary metabolites, where different biosynthetic pathways are codependent. Therefore, plant defensive metabolites could result from potential synergistic interactions among different chemical pathways (Piasecka et al. 2015). Notably, the seven flavonoids in *Zamia* leaves might act as potential precursors or signaling molecules for other defensive metabolites within the alkaloids and amino acids pathways. However, the regulation of secondary metabolites by either the host genome, the microbiome, or the interactions of biomolecules remains largely underexplored (Piasecka et al. 2015).

Both *Zamia* species shared 90% of the detected metabolites which also had similar relative abundances, suggesting no significant differences between their chemical profiles as

observed with the ordinations and corresponding statistical test. For example, the most frequent nitrogenous metabolite in both species corresponds to an amino acid, indicating similar metabolite repertoires for defending against obligate herbivores (Schneider et al. 2002; Whitaker & Salzman 2020; Sierra-Botero et al. 2023). In addition, *Zamia* species seem to produce similar defensive phytochemistry, with comparable metabolites in the flavonoid class across several species (El-Seadawy et al. 2023b). Despite this similarity between species, ordinations with both the CSCS similarity index and Bray–Curtis dissimilarity index, describing the chemical variation in a two-dimensional space, revealed wide intraspecific variation in both species with no sign of host-specific differentiation. However, as our study only sampled a single site for each species, these results should be taken with caution as individuals in other populations of these *Zamia* species could show different patterns in their chemical profile variation across their distribution ranges as observed in other plant groups (Sedio et al. 2021; Christian et al. 2020).

Our hypothesis posits that endophytic microbes enhance the production of host secondary metabolites, accounting for intraspecific variation due to differences in microbial communities. In *Z. pseudoparasitica*, we identified five upregulated metabolites associated with samples enriched with specific prevalent fungal endophytes (core fungal microbiota). This suggests that specific fungal endophyte members can modulate host secondary metabolites, influencing the plant's chemical intraspecific variation (Arnold et al. 2003; Christian et al. 2020), rather than only being host modulated. The two components of the endophytic microbial community indicate a stronger host affinity with the fungal endophytes than with bacteria. Host specificity of fungal endophytes could be the result of differential host chemistry affecting fungal colonization of the endosphere, or the metabolite production may be driven by the endophytic community after colonization (Christian et al. 2020). Host endophyte affinity could also be driven by the host leaf anatomical resistance traits such as the leaf cell walls, flavonoids, anthocyanins, and terpenoids (González-Teuber et al. 2019). However, the species studied here do not differ in leaf traits such as the presence of mesophyll fibers and sclereids (Glos et al. 2022). Fungal endophyte host specificity could also stem from their contrasting habitats (terrestrial vs epiphytic), exposing the plants to distinct fungal propagules coming from the soil or the air (Arnold et al. 2003; Christian et al. 2017; Zheng & Gong 2019). For example, Christian et al. (2017) observed that plants experimentally placed in high sites within the forest canopy showed a lower number of fungal isolates and a higher variation in endophytic composition when compared to plants placed in the understory. Donald et al. (2020) suggested that endophyte assembly in leaves in the canopy was driven by neutral processes, where

the host identity but also dispersal limitation, lead to dissimilar fungal composition within distant branches.

We correlated both bacterial and fungal community variation with the host chemical variation along the ordination Axis1 (Fig. 2). The endophytic bacteria community variation (based on the Weighted Unifrac) was significantly correlated ( $P=0.05$ ) with foliar chemical variation (based on the CSCS metrics). Regression analysis of bacterial orders' relative abundance with intraspecific chemicals indicated that ASVs in the orders Acidobacteriales and Frankiales explained to some degree the variation of host secondary metabolites. While previous studies have suggested that cyanobacteria enhance BMAA production in cycads (Cox et al. 2003) via an unknown biosynthetic pathway (Mantas et al. 2022), our analysis did not directly link the diverse community of heterocyst-forming cyanobacteria (85 ASVs of the genus *Nostoc*) to the observed intraspecific chemical differences and upregulated metabolites. In contrast, the overall endophytic fungal community variation did not correlate with the foliar chemical variation based on the CSCS and Bray–Curtis metrics. However, fungal endophyte orders (Chaetothiales, Glomerellales, Heliotiales, Hypocreales, and Sordariales) explained the intraspecific variation of host secondary metabolites in the two clusters assessed with the CSCS and Bray–Curtis metrics. Additionally, four specific fungal endophyte members of the core microbiota were correlated with a cluster of individuals of *Z. pseudoparasitica* differing in their chemical profile from the rest of the individuals, based on the Bray–Curtis metric. Moreover, a recent genomic study has uncovered many secondary metabolite gene clusters in saprophytic and non-pathogenic fungal endophytes of Xylariaceae (Sordariomycetes), including phenylpropanoids (Franco et al. 2022). Phenylpropanoids act as inducible physical and chemical barriers against pathogen infection and as a signaling molecule involved in the plant immune system (Dixon et al. 2002). Despite the order Xylariales is abundant and diverse in *Zamia*'s endophytic mycobiota (903 ASVs), correlations with metabolite variation were not significant.

## Conclusion

Our study marks the first attempt to describe cycad leaf endophytes and their relationship with secondary metabolites from untargeted metabolome, providing new insights for our understanding of the coevolution of cycad leaf toxicity and specific leaf herbivores with host microbiota (Schneider et al. 2002; Gutierrez-García et al. 2023). Our data suggest that leaf fungal endophytes may play an important role inducing host defensive traits,

such as plant secondary metabolites (Estrada et al. 2013; Walker et al 2022; Christian et al. 2020) as we observed that individuals with similar leaf chemical profiles were enriched by certain fungal taxa prevalent in most of the individual studied. Future identification of a core fungal microbiota across the cycad species phylogeny, and its relation to metabolite expression on below- and above-ground plant organs is a promising research direction in this group of gymnosperms (Pecundo et al. 2021). Moreover, control experiments with synthetic microbial communities will aid to determine direct relationship of mutualistic microbes inducing host foliar metabolites (Emmenegger et al. 2023).

## Benefit-Sharing Statement

The present work results from an international scientific partnership that was developed with scientists from the countries providing data and genetic samples, included as co-authors. The work was done in partnership between research institutions and results will be shared with the provider communities and the broader scientific community.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10886-024-01511-z>.

**Acknowledgements** This project was funded by SENACYT contract No. 12-2018-4-FID16-237 awarded to KS and JCVA, a Packard Foundation Fellowship (2016-65130), a grant from the National Science Foundation (NSF BCS 1514174), and the Cycad Society Grant awarded to AMS. AMS acknowledges financial support from doctoral scholarships of the NSERC CREATE program in Biodiversity, Ecosystem Services and Sustainability (BESS) and the Canada Research Chair (950-232698), Canada. JCVA acknowledges support from the Canada Research Chair (950-232698), the CRNSG-RGPIN 05967-2016, and the Canadian Foundation for Innovation (projects 36781, 39135), and KS and LR-C acknowledges support by SENACYT-UP scholarship for the Master's Program in Environmental Microbiology. BES and CAL are supported by NSF DEB Award 2240430 and a Corvea Agrisciences grant to STRI. The authors thank Jorge Mendieta, Maycol Madrid, Marta Vargas, and Armando A. Durant-Archibold, for their assistance during this work.

**Author Contribution** JCVA, BES, and KS conceived and designed the research. JCVA, OM, RB, AB, and LR-C identified the host species and collected leaf samples. JCVA, KS, OM, LR-C, and AMS prepared the specimens, carried out the DNA extraction, and prepared the amplicon sequencing library. AMS processed microbial metabarcoding raw data. BES and CAL prepared the specimens and carried out extractions and chemical analysis. BES processed metabolomic raw data. AMS compiled and analyzed the complete dataset with the assistance from CAL, BES, KS, and JCVA. AMS prepared all figures. AMS and JCVA lead the writing of the manuscript with the assistance of all co-authors.

**Funding** This project was funded by SENACYT contract No. 12-2018-4-FID16-237, a Packard Foundation Fellowship (2016-65130), a grant from the National Science Foundation (NSF BCS 1514174) and the Canada Research Chair (950-232698).

**Data Accessibility** Raw sequence data were deposited in the NCBI Sequence Read Archive (SRA) with their respective accession numbers under the BioProject: PRJNA1061598.

The process data with its corresponding metadata and the scripts used for the analyses that support the findings of this study are openly available in: [https://github.com/adrielsierra/Zamia\\_Phyllosphere](https://github.com/adrielsierra/Zamia_Phyllosphere).

## Declarations

**Competing Interests** The authors declare no competing interests.

## References

- Arnold AE, Mejía LC, Kyllo D, Rojas EI, Maynard Z, Robbins N, Herre EA (2003) Fungal endophytes limit pathogen damage in a tropical tree. *Proceedings of the National Academy of Sciences, USA* 100:15649–15654. <https://doi.org/10.1073/pnas.253348310>
- Bell-Doyon P, Villarreal AJC (2020) New Notes on the Ecology of the Epiphytic gymnosperm and Panamanian endemic *Zamia pseudo-parasitica*. *Neotropical Naturalist* 2:1–7
- Bell-Doyon P, Laroche J, Saltonstall K, Villarreal Aguilar JC (2020) Specialized bacteriome uncovered in the coralloid roots of the epiphytic gymnosperm, *Zamia pseudo-parasitica*. *Environmental DNA* 2:418–428. <https://doi.org/10.1002/edn3.66>
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13(7):581–583. <https://doi.org/10.1038/nmeth.3869>
- Callahan B, McMurdie P, Holmes S (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 11:2639–2643. <https://doi.org/10.1038/ismej.2017.119>
- Calonje M, Meerow AW, Griffith MP, Salas-Leiva D, Vovides AP, Coiro M, Francisco-Ortega J (2019) A time-calibrated species tree phylogeny of the New World cycad genus *Zamia* L. (Zamiaceae, Cycadales). *Int J Plant Sci* 180:286–314. <https://doi.org/10.1086/702642>
- Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, Collman RG, Bushman FD, Li H (2012) Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* 28(16):2106–2113. <https://doi.org/10.1093/bioinformatics/bts342>
- Christian N, Herre EA, Mejía LC (1858) Clay K (2017) Exposure to the leaf litter microbiome of healthy adults protects seedlings from pathogen damage. *Proc R Soc B* 284:20170641. <https://doi.org/10.1098/rspb.2017.0641>
- Christian N, Whitaker BK, Clay K (2015) Microbiomes: unifying animal and plant systems through the lens of community ecology theory. *Front Microbiol* 6:1–15. <https://doi.org/10.3389/fmicb.2015.00869>
- Christian N, Herre EA, Clay K (2019) Foliar endophytic fungi alter patterns of nitrogen uptake and distribution in *Theobroma cacao*. *New Phytol* 222:1573–1583. <https://doi.org/10.1111/nph.15693>
- Christian N, Sedio BE, Florez-Buitrago X, Ramírez-Camejo LA, Rojas EI, Mejía LC, Palmedo S, Rose A, Schroeder JW, Herre EA (2020) Host affinity of endophytic fungi and the potential for reciprocal interactions involving host secondary chemistry. *Am J Bot* 107:219–228. <https://doi.org/10.1002/ajb2.1436>
- Cox AP, Anne Banack S, Murch SJ (2003) Biomagnification of cyanobacterial neurotoxins and neurodegenerative disease among the Chamorro people of Guam. *Proceedings of the National Academy of Sciences of the United States of America*. <https://www.pnas.org/cgi/doi/https://doi.org/10.1073/pnas.2235808100>

- Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ (2018) Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6(1):226. <https://doi.org/10.1186/s40168-018-0605-2>
- Dixon RA, Achnine L, Kota P, Liu CJ, Reddy MS, Wang L (2002) The phenylpropanoid pathway and plant defense—a genomics perspective. *Mol Plant Pathol* 3(5):371–390. <https://doi.org/10.1046/j.1364-3703.2002.00131.x>
- Donald J, Roy M, Suescun J, Iribar A, Manzi S, Péllissier L, Gaucher P, Chave J (2020) A test of community assembly rules using foliar endophytes from a tropical forest canopy. *J Ecol* 108:1605–1616. <https://doi.org/10.1111/1365-2745.13344>
- Dorokhov YL, Sheshukova EV, Komarova TV (2018) Methanol in Plant Life. *Front Plant Sci* 9:1623. <https://doi.org/10.3389/fpls.2018.01623>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11–15
- Dührkop K, Shen H, Meusel M, Rousu J, Böcker S (2015) Searching molecular structure databases with tandem mass spectra using CSI: FingerID. *Proc Natl Acad Sci* 112:12580–12585. <https://doi.org/10.1073/pnas.1509788112>
- Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, Dorrestein PC, Rousu J, Böcker S (2019) SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods* 16:299–302. <https://doi.org/10.1038/s41592-019-0344-8>
- Ehrlich PR, Raven PH (1964) Butterflies and plants: A study in plant coevolution. *Evolution* 18:586–608. <https://doi.org/10.2307/2406212>
- El-Seadawy H, Abo El-Seoud K, El-Aasr M, Ragab A (2023b) Phytochemical profile, ethnobotanical and biological impacts of various *Zamia* species: A mini-review. *Journal of Advanced Medical and Pharmaceutical Research* 4(2):63–73
- El-Seadawy H, Abo El-Seoud K, El-Aasr M, Ragab A (2023a) Cycadaceae: An important source for biflavonoids and various pharmacological effects of different *Cycas* species. *Journal of Advanced Medical and Pharmaceutical Research* 0(0): 0–0. <https://doi.org/10.21608/jampr.2023.202016.1051>
- Emmenegger B, Massoni J, Pestalozzi CM, Bortfeld-Miller M, Maier BA, Vorholt JA (2023) Identifying microbiota community patterns important for plant protection using synthetic communities and machine learning. *Nat Commun* 14:7983. <https://doi.org/10.1038/s41467-023-43793-z>
- Endara M-J, Coley PD, Ghabash G, Nicholls JA, Dexter KG, Donoso DA et al (2017) Coevolutionary arms race versus host defense chase in a tropical herbivore–plant system. *Proceedings of the National Academy of Sciences, USA* 114:E7499–E7505. <https://doi.org/10.1073/pnas.1707727114>
- Estrada C, Weislo WT, Van Bael SA (2013) Symbiotic fungi alter plant chemistry that discourages leaf-cutting ants. *New Phytol* 198:241–251. <https://doi.org/10.1111/nph.12140>
- Faust K, Raes J (2016) CoNet app: inference of biological association networks using Cytoscape. *F1000Research*, 5: 1519. <https://doi.org/10.12688/f1000research.9050.1>
- Fine PVA, Baccaro FB, Lokvam J, Mesones I, Vásquez Pilco M, Ayarza Zuñiga JM, Merkel E, Sanches M, Nogueira CA, Salazar D (2023) A test of the Geographic Mosaic Theory of Coevolution: investigating widespread species of Amazonian *Protium* (Burseraceae) trees, their chemical defenses, and their associated herbivore faunas. *Front Ecol Evol* 11:1180274. <https://doi.org/10.3389/fevo.2023.1180274>
- Fox J, Weisberg S (2011) *An R Companion to Applied Regression*. Sage Publishing, Thousand Oaks
- Franco MEE, Wisecaver JH, Arnold AE, Ju Y-M, Slot JC, Ahrendt S, Ahrendt S, Moore LP, Eastman KE, Scott K, Konkel Z, Mondo SJ, Kuo A, Hayes RD, Haridas S, Andreopoulos B, Riley R, LaButti K, Pangilinan J, Lipzen A, Amirebrahimi M, Yan J, Adam C, Keymanesh K, Ng V, Louie K, Northen T, Drula E, Henrissat B, Hsieh H-M, Youens-Clark K, Lutzoni F, Miadlikowska J, Eastwood DC, Hamelin RC, Grigoriev IV, U'Ren JM (2022) Ecological generalism drives hyperdiversity of secondary metabolite gene clusters in xylarialean endophytes. *New Phytol* 233:1317–1330. <https://doi.org/10.1111/nph.17873>
- Gehring MM, Pengelly JL, Cuddy WS, Fieker C, Forster PI, Neilan BA (2010) Host selection of symbiotic cyanobacteria in 31 species of the Australian cycad genus: *Macrozamia* (Zamiaceae). *Mol Plant Microbe Interact* 23:811–822. <https://doi.org/10.1094/MPMI-23-6-0811>
- Glos RAE, Salzman S, Calonje M, Vovides AP, Coiro M, Gandolfo MA, Specht CD (2022) Leaflet Anatomical Diversity in *Zamia* (Cycadales: Zamiaceae) Shows Little Correlation with Phylogeny and Climate. *Bot Rev* 88(4):437–452. <https://doi.org/10.1007/s12229-021-09272-0>
- González-Teuber M, Vilo C, Guevara-Araya MJ, Salgado-Luarte C, Gianoli E (2019) Leaf resistance traits influence endophytic fungi colonization and community composition in a South American temperate rainforest. *J Ecol* 108:1019–1029. <https://doi.org/10.1111/1365-2745.13314>
- Gutiérrez-García K, Bustos-Díaz ED, Corona-Gómez JA, Ramos-Aboites HE, Sélem-Mojica N, Cruz-Morales P, Pérez-Farrera MA, Barona-Gómez F, Cibrián-Jaramillo A (2018) Cycad coraloid roots contain bacterial communities including cyanobacteria and *Caulobacter* spp. that encode niche-specific biosynthetic gene clusters. *Genome Biol Evol* 11:319–334. <https://doi.org/10.1093/gbe/evy266>
- Gutiérrez-García K, Whitaker MRL, Bustos-Díaz ED, Salzman S, Ramos-Aboites HE, Reitz ZL, Pierce NE, Cibrián-Jaramillo A, Barona-Gómez F (2023) Gut microbiomes of cycad-feeding insects tolerant to  $\beta$ -methylamino-L-alanine (BMAA) are rich in siderophore biosynthesis. *ISME Communication* 3:122. <https://doi.org/10.1038/s43705-023-00323-8>
- Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>
- Khare E, Mishra J, Arora NK (2018) Multifaceted interactions between endophytes and plant: Developments and Prospects. *Frontiers in Microbiology*, 9. <https://doi.org/10.3389/fmicb.2018.02732>
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41(1). <https://doi.org/10.1093/nar/gks808>
- Kursar TA, Coley PD (2003) Convergence in defense syndromes of young leaves in tropical rainforests. *Biochem Syst Ecol* 21:929–949. [https://doi.org/10.1016/S0305-1978\(03\)00087-5](https://doi.org/10.1016/S0305-1978(03)00087-5)
- Laforest-Lapointe I, Whitaker BK (2019) Decrypting the phyllosphere microbiota: progress and challenges. *Am J Bot* 106:171–173. <https://doi.org/10.1002/ajb2.1229>
- Lahti L, Shetty S (2017) *microbiome* R package. *Bioconductor*. <https://doi.org/10.18129/B9.bioc.microbiome>
- Lajoie G, Maglione R, Kembel SW (2020) Adaptive matching between phyllosphere bacteria and their tree hosts in a neotropical forest. *Microbiome* 8:70. <https://doi.org/10.1186/s40168-020-00844-7>
- Lindström AJ, Calonje M, Stevenson D, Husby C, Taylor A (2013) Clarification of *Zamia acuminata* and a new *Zamia* species from Coclé Province, Panama. *Phytotaxa* 98:27–42. <https://doi.org/10.11646/phytotaxa.98.2.1>
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>

- Lambais MR, Barrera SE, Santos EC, Crowley DE, Jumpponen A (2017) Phyllosphere metaproteomes of trees from the Brazilian Atlantic Forest show high levels of functional redundancy. *Microb Ecol* 73:123–134. <https://doi.org/10.1007/s00248-016-0878-6>
- Mantas MJQ, Nunn PB, Codd GA, Barker D (2022) Genomic insights into the biosynthesis and physiology of the cyanobacterial neurotoxin 3-N-methyl-2,3-diaminopropanoic acid (BMAA). *Phytochemistry* 200:113198. <https://doi.org/10.1016/j.phytochem.2022.113198>
- Marler TE, Snyder LR, Shaw CA (2010) *Cycas micronesica* (Cycadales) plants devoid of endophytic cyanobacteria increase in  $\beta$ -methylamino-L-alanine. *Toxicol* 56:563–568. <https://doi.org/10.1016/j.toxicol.2010.05.015>
- Martin M (2011) *Cutadapt* removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17: 10–12. <https://doi.org/10.14806/ej.17.1.200>
- McMurdie PJ, Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 10:e1003531. <https://doi.org/10.1371/journal.pcbi.1003531>
- Mejía LC, Herre EA, Sparks JP, Winter K, García MN, Van Bael SA, Stitt J, Shi Z, Zhang Y, Guiltinan MJ, Maximova SN (2014) Pervasive effects of an endophytic fungus on host genetic and phenotypic expression in a tropical tree. *Front Microbiol* 5:1–15. <https://doi.org/10.3389/fmicb.2014.00479>
- Moyes AB, Kueppers LM, Pett-Ridge J, Carper DL, Vandehey N, O'Neil J, Frank AC (2016) Evidence for foliar endophytic nitrogen fixation in a widely distributed subalpine conifer. *New Phytol* 210:657–668. <https://doi.org/10.1111/nph.13850>
- Nilsson RH, Larsson K-H, Taylor AFS, Bengtsson-Palme J, Jeppesen TS, Schigel D, Kennedy P, Picard K, Glöckner FO, Tedersoo L, Saar I, Kõljalg U, Abarenkov K (2018) The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gky1022>
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn F, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner E (2016) *vegan*: Community Ecology Package. R package version 2.4–4. <https://CRAN.R-project.org/package=vegan>
- Pecundo MH, Dela Cruz TEE, Chen T, Notarte KI, Ren H, Li N (2021) Diversity, Phylogeny and Antagonistic Activity of Fungal Endophytes Associated with Endemic Species of *Cycas* (Cycadales) in China. *J Fungi (basel)* 7(7):572. <https://doi.org/10.3390/jof7070572>
- Piasecka A, Jedrzejczak-Rey N, Bednarek P (2015) Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytol* 206:948–964. <https://doi.org/10.1111/nph.13325>
- Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395. <https://doi.org/10.1186/1471-2105-11-395>
- Prado A, Sierra A, Windsor D, Bede JC (2014) Leaf traits and herbivory levels in a tropical gymnosperm, *Zamia stevensonii* (Zamiaceae). *Am J Bot* 101:437–447. <https://doi.org/10.3732/ajb.1300337>
- Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35(21):7188–7196. <https://doi.org/10.1093/nar/gkm864>
- R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Robbins RK, Cong Q, Zhang J, Shen J, Quer Riera J, Murray D, Busby RC, Faynel C, Hallwachs W, Janzen DH, Grishin NV (2021) A switch to feeding on cycads generates parallel accelerated evolution of toxin tolerance in two clades of *Eumaeus* caterpillars (Lepidoptera: Lycaenidae). *Proceedings of the National Academy of Sciences of the United States of America* 118(7). <https://doi.org/10.1073/pnas.2018965118>
- Rosado BHP, Almeida LC, Alves LF, Lambais MR, Oliveira RS (2018) The importance of phyllosphere on plant functional ecology: a phyllo trait manifesto. *New Phytol* 219:1145–1149. <https://doi.org/10.1111/nph.15235>
- Salazar D, Lokvam J, Mesones I, Pilco MV, Zuñiga JMA, de Valpine P, Fine PVA (2018) Origin and maintenance of chemical diversity in a species-rich tropical tree lineage. *Nature Ecology and Evolution* 2(6):983–990. <https://doi.org/10.1038/s41559-018-0552-0>
- Salonen A, Salojärvi J, Lahti L, de Vos WM (2012) The adult intestinal core microbiota is determined by analysis depth and health status. *Clin Microbiol Infect* 18:16–20. <https://doi.org/10.1111/j.1469-0691.2012.03855.x>
- Salzman S, Whitaker M, Pierce NE (2018) Cycad-feeding insects share a core gut microbiome. *Biol J Lin Soc* 123(4):728–738. <https://doi.org/10.1093/biolinnean/bly017>
- Schliep KP (2011) *phangorn*: phylogenetic analysis in R. *Bioinformatics* 27(4):592–593. <https://doi.org/10.1093/bioinformatics/btq706>
- Schneider D, Wink M, Sporer F (2002) Cycads: their evolution, toxins, herbivores and insect pollinators. *Naturwissenschaften* 89:281–294. <https://doi.org/10.1007/s00114-002-0330-2>
- Sedio BE (2017) Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification. *New Phytol* 214:952–958. <https://doi.org/10.1111/nph.14438>
- Sedio BE, Boya CA, Rojas Echeverri JC (2018) A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. *Applications in Plant Sciences* 6(12):e1033. <https://doi.org/10.1002/aps3.1033>
- Sedio BE, Spasojevic MJ, Myers JA, Wright SJ, Person MD, Chandrasekaran H, Dwenger JH, Prechi ML, López CA, Allen DN, Anderson-Teixeira KJ, Baltzer JL, Bourg NA, Castillo BT, Day NJ, Dewald-Wang E, Dick CW, James TY, Kueneman JG, LaManna J, Lutz JA, McGregor IR, McMahon SM, Parker GG, Parker JD, Vandermeer JH (2021) Chemical Similarity of Co-occurring Trees Decreases with Precipitation and Temperature in North American Forests. *Frontiers in Ecology and Evolution*, 9. <https://doi.org/10.3389/fevo.2021.679638>
- Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Sierra-Botero L, Calonje M, Robbins RK, Rosser N, Pierce NE, López-Gallego C, Valencia-Montoya WA (2023) Cycad phylogeny predicts host plant use of *Eumaeus* butterflies. *Ecol Evol* 13:e09978. <https://doi.org/10.1002/ece3.9978>
- Suárez-Moo PI, Vovides AP, Griffith MP, Barona-Gómez F, Cibrián-Jaramillo A (2019) Unlocking a high bacterial diversity in the coralloid root microbiome from the cycad genus *Dioon*. *PLoS ONE* 14(3):e0211271. <https://doi.org/10.1371/journal.pone.0211271>
- Tripathi A, Vázquez-Baeza Y, Gauglitz JM, Wang M, Dührkop K, Nothias-Espósito M (2021) Chemically informed analyses of metabolomics mass spectrometry data with Qemistree. *Nat Chem Biol* 17:146–151. <https://doi.org/10.1038/s41589-020-00677-3>
- U'Ren JM, Lutzoni F, Miadlikowska J, Zimmerman NB, Carbone I, May G, Arnold AE (2019) Host availability drives distributions of fungal endophytes in the imperilled boreal realm. *Nature Ecology & Evolution* 3:1430–1437. <https://doi.org/10.1038/s41559-019-0975-2>
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, Fourth edition. Springer, New York. ISBN 0–387–95457–0, <https://www.stats.ox.ac.uk/pub/MASS4/>.

- Vorholt J (2012) Microbial life in the phyllosphere. *Scientific Reviews* 10:828–840. <https://doi.org/10.1038/nrmicro2910>
- Walker TW, Alexander JM, Allard P-M, Baines O, Baldy V, Bardgett RD, Coley PD, David B, Defosses E, Endara M-J, Ernst M, Fernandez C, Forrister D, Gargallo-Garriga A, Jassey VEJ, Marr S, Neumann S, Pellissier L, Salguero-Gómez R (2022) Functional Traits 2.0: The power of the metabolome for ecology. *J Ecol* 110:4–20. <https://doi.org/10.1111/1365-2745.13826>
- Wang MX, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y et al (2016) Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat Biotechnol* 34:828–837. <https://doi.org/10.1038/nbt.3597>
- Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5(1):27. <https://doi.org/10.1186/s40168-017-0237-y>
- Whitaker MRL, Salzman S (2020) Ecology and evolution of cycad-feeding Lepidoptera. *Ecol Lett* 23(12):1862–1877. <https://doi.org/10.1111/ele.13581>
- White TJ, Bruns TD, Lee SB, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. *PCR Protocols* 315–322. <https://doi.org/10.1016/B978-0-12-372180-8.50042-1>
- Wink M (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* 64(1):3–19. [https://doi.org/10.1016/S0031-9422\(03\)00300-5](https://doi.org/10.1016/S0031-9422(03)00300-5)
- Wright ES (2016) Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *The R Journal* 8:352–359. <https://doi.org/10.32614/RJ-2016-025>
- Yamada S, Ohkubo S, Miyashita H, Setoguchi H (2012) Genetic diversity of symbiotic cyanobacteria in *Cycas revoluta* (Cycadaceae). *FEMS Microbiol Ecol* 81:696–706. <https://doi.org/10.1111/j.1574-6941.2012.01403.x>
- Yamaji K, Watanabe Y, Masuya H, Shigeto A, Yui H, Haruma T (2016) Root fungal endophytes enhance heavy-metal stress tolerance of *Clethra barbinervis* growing naturally at mining sites via growth enhancement, promotion of nutrient uptake and decrease of heavy-metal concentration. *PLoS ONE* 11:e0169089. <https://doi.org/10.1371/journal.pone.0169089>
- Zheng Y, Gong X (2019) Niche differentiation rather than biogeography shapes the diversity and composition of microbiome of *Cycas panzhihuaensis*. *Microbiome* 7:152. <https://doi.org/10.1186/s40168-019-0770-y>
- Zheng Y, Chiang TY, Huang CL, Gong X (2018) Highly diverse endophytes in roots of *Cycas bifida* (Cycadaceae), an ancient but endangered gymnosperm. *J Microbiol* 56:337–345. <https://doi.org/10.1007/s12275-018-7438-3>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Adriel M. Sierra<sup>1,2</sup> · Omayra Meléndez<sup>3,5</sup>  · Rita Bethancourt<sup>3</sup> · Ariadna Bethancourt<sup>3</sup> · Lilisbeth Rodríguez-Castro<sup>4,5</sup>  · Christian A. López<sup>5,6</sup>  · Brian E. Sedio<sup>5,6</sup>  · Kristin Saltonstall<sup>5</sup>  · Juan Carlos Villarreal A.<sup>1,2,5,7</sup> 

✉ Adriel M. Sierra  
adriel-michel.sierra-pinilla.1@ulaval.com

✉ Juan Carlos Villarreal A.  
jcvil9@ulaval.ca

Omayra Meléndez  
omayraimeemp@gmail.com

Rita Bethancourt  
bethancourtrita61@gmail.com

Ariadna Bethancourt  
ariadna.bethancourt@up.ac.pa

Lilisbeth Rodríguez-Castro  
lili\_0990@outlook.es

Christian A. López  
christianlopezs@utexas.edu

Brian E. Sedio  
sedio@utexas.edu

Kristin Saltonstall  
SaltonstallK@si.edu

<sup>1</sup> Département de Biologie, Université Laval, Québec, (QC) G1V 0A6, Canada

<sup>2</sup> Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, (QC) G1V 0A6, Canada

<sup>3</sup> Departamento de Microbiología y Parasitología, Facultad de Ciencias Naturales, Exactas y Tecnología, Universidad de Panamá, Panamá, Panamá

<sup>4</sup> Departamento de Microbiología, Facultad de Ciencias Naturales, Exactas y Tecnología, Universidad de Panamá, Panamá, Panamá

<sup>5</sup> Smithsonian Tropical Research Institute, Ancón, Panamá

<sup>6</sup> Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA

<sup>7</sup> Canada Research Chair in Genomics of Tropical Symbioses, Department of Biology, Université Laval, Québec G1V 0A6, Canada