Spatial patterns of fungal endophytes in a subtropical montane rainforest of northern Taiwan

Daniel Thomasa, Roo Vandegrifta, B. A. Roya, Huei-Mei Hsiehb,Yu-Ming Jub

a Institute of Ecology and Evolution. 272 Onyx Bridge, 5289 University of Oregon, Eugene, OR 97403-5289

b Institute of Plant and Microbial Biology, Academia Sinica. 128 Academia Road, Section 2, Nankang, Taipei 11529, Taiwan

# Abstract

Fungal endophytes of plants are ubiquitous and important to host plant health. Wood-inhabiting and foliar endophyte communities from multiple tree hosts were sampled at multiple spatial scales across the Fushan forest dynamics plot in northern Taiwan, using culture-free, community DNA amplicon sequencing methods. Fungal endophyte communities were distinct between leaves and wood, but the mycobiomes were highly variable across and within tree species. Despite this, host tree species was an important predictor of mycobiome community-composition. Within a single common tree species, “core” mycobiomes were characterized using co-occurrence analysis. The spatial coocurrence patterns of these few species of fungal endophytes appear to explain the strong host effect. For wood endophytes, a consistent core mycobiome coexisted with the host across the extent of the study. For leaf endophytes, the core fungi resembled a more dynamic, “gradient” model of the core microbiome, changing across the topography and distance of the study.

# Keywords:

Fungal Endophytes, Microbial ecology, ITS, Plant-fungal interactions, Fushan Subtropical Forest Dynamics Plot, Mycobiome, Core microbiome, PCNM/dbMEM spatial analysis

# Introduction

Microbial community assembly and geographic patterns in microbes remain poorly understood, despite nearly a century of discussion (Baas-Becking, 1934; De Wit and Bouvier, 2006; Green and Bohannan, 2006; Martiny et al., 2006; Peay et al., 2010; Hanson et al., 2012; Nemergut et al., 2013;). Rich microbial communities appear to be associated with all multicellular organisms (Hoffman and Arnold, 2010; Rosenberg et al., 2010). Host-associated microbes present additional complexity in modeling microbial community assembly, and raise questions concerning fidelity of host-microbe interactions. Study of microbial communities takes on a new urgency in the discussion of plant microbiomes and plant health in a changing planet (Woodward et al., 2012).

The potential importance of microbes in adding ecological functions to their hosts (Rodriguez et al., 2009; Johnson and Versalovic, 2012; Woodward et al., 2012) has led some to suggest that multicellular organisms may host core microbiomes (Hamady and Knight, 2009; Shade and Handelsman, 2012; Vandenkoornhuyse et al., 2015), which are subsets of important and consistent microbial partners. Initial explorations of plant core microbiomes have been highly controlled (Lundberg et al., 2012; Edwards et al., 2015). Studies of plant-associated microbiomes in natural settings have rarely been framed in terms of core microbiomes (Kim et al., 2012; Zimmerman and Vitousek, 2012; Bodenhausen et al., 2013; Higgins et al., 2014; Kembel and Mueller, 2014). This is not a coincidence: outside of experimental settings, the prospect of detecting a cadre of microorganisms absolutely loyal to their host in the face of a complex and dynamic natural environment is daunting. This definition of the core microbiome, known as either a “substantial” or “minimal” core(Hamady and Knight, 2009) may be useful when carefully applied to long-studied symbioses such as ruminant gut communities (Liggenstoffer et al., 2010) or mycorrhizal relationships (Malloch et al., 1980; Van Der Heijden and Horton, 2009). This definition may not always serve for describing the other numerous and labyrinthine microbe-host interactions that occur between hosts and microbes. However, other definitions of core microbiomes exist that may be more useful for ecologically modeling microbiomes (Hamady and Knight, 2009).

Fungal endophytes, or fungi that live internally in plant tissues causing incurring disease symptoms (Wilson, 1995), are an important component of the plant microbiome. They are widespread and important to plant health (Arnold et al., 2003; Mejía et al., 2008, 2007; Rodriguez et al., 2009; Porras-Alfaro and Bayman, 2011). The endophytic compartment in which they reside is a distinct ecological space, in the sense that very different communities of microbes are observed outside vs. inside plant tissues (Santamaría and Bayman, 2005; Lundberg et al., 2012; Bodenhausen et al., 2013), at least partly due to host-microbe preferences (Schulz et al., 1999; Oldroyd, 2013; Venkateshwaran et al., 2013). Plant organs host distinct communities of endophytes (Bodenhausen et al., 2013; Peršoh, 2013; Edwards et al., 2015; Tateno et al., 2015). Endophyte communities are also influenced by environmental conditions (Carroll and Carroll, 1978; Arnold and Herre, 2003; Zimmerman and Vitousek, 2012), despite presumed buffering from environmental stresses by host tissues. Fungal communities are also subject to spatial processes such as dispersal limitation (Peay et al., 2010; Higgins et al., 2014) at multiple scales (Mummey and Rillig, 2008; Norros et al., 2012; Tedersoo et al., 2014). Fungal endophytes, therefore, make ideal systems for studying the interplay of host-microbe interactions, environmental influences, and spatial patterning of both host and microbes in natural settings.

It must be acknowledged that plant hosts exert strong influence on community membership of their endophyte community. However, we hypothesized that even the most faithful fungal associates will uncouple from their hosts with changing environmental conditions and dispersal constraints. We predicted, on the scale of the present study, that plant mycobiomes resemble “gradient” core microbiomes (Hamady and Knight, 2009). Under this model, microbiomes can totally change across a landscape, with host-interactions mitigating, but ultimately not preventing, environmentally- and spatially-driven changes in the microbiome. In other words, we hypothesized that a persistent “core” of microbes shared among all members of a plant species does not truly exist in nature, on any meaningful scale. To test this, we compared community composition between wood and leaf fungal endophytes in multiple species of plant host across the landscape of the Fushan Forest Dynamics plot. We examined patterns in the total detected endophyte community of several plant hosts. Delving further within these data, we focused on a single tree host and the spatial patterns of its most strongly associated endophytic fungi.

# Materials and methods

Background/Site:

Sampling occurred in summer of 2013 at Fushan forest, in Northeastern Taiwan (24º 45' 40" N, 121º 33' 28" E), which hosts a 25-ha Smithsonian-associated Forest Dynamics Plot (FDP) (Losos and Leigh, 2004; Su et al., 2007). Fushan is a humid subtropical old-growth montane site that receives 4.27 m of rain each year. Most of this precipitation falls during rainy, cool winters, though a significant fraction of this rain is due to typhoons, the main agent of disturbance in this system, during warm summer months. The flora is diverse, characterized by many evergreen broadleaf tree species and a diverse understory of lianas, ferns, tree ferns, and other herbs, graminoids, and shrubs. Vegetative communities can be broadly categorized into four community types described by dominant tree species combinations (Fig. 1). Topography is highly variable, with a maximum elevation of 733 m above sea level at an approximately central hilltop within the FDP, and a minimum of 600 m, though the present study sampled areas only as low as 650 m. The complex topography of Fushan has been summarized by classification of each 20 m x 20 m quadrant of the FDP into one of seven habitat types, based on aspect, slope, convexity, and elevation (Figs.1 and 2), which are found to influence vegetative communities (Su et al., 2010).

## Field methods

Fushan FDP was divided into 9 sub-plots, and subplots were sampled using a nested logarithmic scheme intended to detect dispersal limitation and community turnover (Rodrigues et al., 2013) (Fig. 2). Sampling of each set of nested points was undertaken in random order. Once sampling of a single set of nested squares had begun, all points within that set of nested points were sampled prior to beginning another. Only six out of nine sets of nested squares were sampled, due to time constraints. For each sampling point, we located the tree with the largest DBH with canopy above the point and collected the three lowest “healthy” appearing leaves that were safely reachable. Leaves and accompanying woody stems were obtained using a 3 m collapsible pole pruner. Identification of host tree was determined survey data from ongoing ecological research at Fushan FDP (Su et al., 2007). All plant material was carried to a nearby field station and stored at 4°C for no longer than 5 days before processing.

## Lab methods

Leaf Endophyte Metabarcode library

Samples of leaves were processed to allow for DNA extraction and next-generation sequencing of the ITS region of fungal endophytes. We did all leaf DNA extractions in the lab at Academia Sinica in Taipei, Taiwan. First, the surfaces of fresh leaves were washed gently with tap water to reduce epiphytes. Then, one square centimeter leaf segments were cut from each of the three leaves collected per sampling plot and surface-sterilized by immersion in 70% ethanol for 30 sec, full-strength bleach (5% sodium hypochlorite) for 1 min, an additional 30 sec in ethanol, then rinsed thoroughly in sterile deionized water. Leaf tissues were disrupted via bead beating using three 5 mm stainless steel beads for an 80 s agitation cycle at 3450 oscillations/minute. DNA was extracted from homogenized leaf tissues using a Qiagen DNeasy 96 Plant Kit following the manufacturer's instructions. Extracted DNA was shipped overnight on wet ice to our lab at the University of Oregon, where a metabarcode sequencing library of the fungal internal transcribed spacer (ITS) region of the rRNA gene was prepared. Library preparation followed Meadow et al. (2013)(2013) , with slight modifications. Briefly, the ITS region was amplified using a modified fungal specific ITS1F/ITS2 primer set adapted from Mueller et al. (2014) (5"-CTTGGTCATTTAGAGGAAGTAA-3" / 5"-GCTGCGTTCTTCATCGATGC-3") (Gardes and Bruns, 1993) through a two-step custom Illumina preparation protocol. We used a split-barcode system, with unique combinations of six base pair barcodes appended to both the forward and reverse primers; this allowed for fewer total primers to be synthesized, while maintaining a large number of unique possible combinations (Gloor et al., 2010). Primer secondary structures were validated using PrimerProspector (Walters et al., 2011). The first PCR step used forward and reverse primers that contained barcodes and partial Illumina adapters; the second PCR step appended the rest of the Illumina adapters, and barcodes were combined into unique 12 base-pair sequences in silico using paired-end reads. All first-step PCRs were amplified in triplicate, and then pooled before second-step PCR. First-step PCR (25 μL total reaction volume) was performed using 2.5 μL 10X high fidelity PCR buffer (Thermo Fisher Scientific), 0.125 μL dNTPs (10 mM, Sigma-Aldrich), 1.25 μL MgCl2 (50 mM, Thermo Fisher Scientific), 0.25 μL Platinumª Taq high fidelity polymerase (Thermo Fisher Scientific), 14.875 μL certified nucleic-acid free water, 0.5 μL forward primer, 0.5 μL reverse primer, and 5 μL template DNA using the following conditions: initial denaturation for 2 min at 98 °C; 20 cycles of 30 s at 98 °C, 30 s at 60 °C, and 45 s at 72 °C; and 72 °C for 5 min for final extension. The products of first-step PCR triplicates were pooled and cleaned with DNA Clean & Concentrator (Zymo Research, Irvine, CA) following the manufacturer's instructions; 10 μL of 3M NaOAc (pH 5.2) was added to decrease the pH of the pooled reactions and facilitate efficient binding to the spin column, and all samples were eluted using 10 μL of the provided elution buffer. Second-step PCR reactions used a single primer pair to add the remaining Illumina adaptor sequence to the ends of the concentrated amplicons from the first-step PCR. Second-step PCR (25 μL total reaction volume) included the same reagents as above, and used 5 μL of the pooled and concentrated first-step PCR products as template; the conditions were as follows: 2 min denaturation at 98 °C; 14 cycles of 30 s at 98 °C, 30 s at 58°C, and 45 s at 72 °C; and 3 min at 72 °C for final extension. Equal volumes of each sample were then pooled, and the library was size-selected by gel electrophoresis: the wide gel bands centered at ~275bp (175-400bp were removed, to account for the variation present at the ITS1 locus across the kingdom Fungi) were extracted and concentrated using the ZR-96 Zymoclean Gel DNA Recovery Kit (ZYMO Research, Irvine, CA), following manufacturer's instructions. DNA concentration was quantified using a Qubit Fluoromoeter (Invitrogen, NY). Samples were sent to the IBEST Genomics Resources Core at the University of Idaho (Moscow, ID; http://www.ibest.uidaho.edu/), and sequenced on the Illumina MiSeq platform as paired-end reads after qPCR validation with Illumina-specific primers.

Wood-inhabiting Endophyte Metabarcode library

Wood was debarked and phloem and sapwood was collected using tools that were ethanol- and flame-sterilized between cuts (Gazis and Chaverri, 2010). Approximately 0.5 grams of wood tissue was disrupted via bead beating using three 5 mm stainless steel beads for 3x30 second agitation cycles (3450 oscillations/minute), followed by an additional 30s cycle with two additional 3 mm stainless steel beads. DNA was extracted from homogenized leaf tissues using a Qiagen DNeasy 96 Plant Kit following the manufacturer’s instructions.

Samples were tested for presence of endophytic fungi using a preliminary PCR amplification and gel visualization of full ITS region with fungal specific primers (Gardes and Bruns, 1993). 91 samples that amplified successfully and 3 controls were then re-amplified in triplicate PCRs using ITS1F forward and ITS2 reverse primers, covering the ITS1 region (Blaalid et al., 2013), with illumina adapter sequences and dual-indexed barcodes appended (Integrated DNA Technologies, Coralville, IA), as described above. Samples were identified using 94 unique combinations of twelve forward and eight reverse 8 bp barcodes (full primer sequences are available in the Supplemental Materials). PCR protocols: Initial denature of 94 ºC for 5 min, followed by 30 amplifications cycles of 94 ºC for 30 s, 55 degrees C for 1 min, 72 ºC for 30 sec each, and a final elongation of 72 ºC for 7 min. Triplicate PCRs were done in 20 µL volumes. Triplicate PCRs were done in three 20 µL volumes using the following PCR recipe: foward and reverse primers, 0.6 µL each (10 µM), additional MgCl2 (25 nM) 0.8 µL, template DNA 2.5 µL, water 5.5 µL, and 10 µL 2X PCR Super Master Mix, which contains Taq polymerase, dNTPs and MgCl2 (Biotool©, now Bimake©, Houston, TX). Triplicate PCR products were combined and cleaned with MagBind© Rxn PurePlus (OMEGA bio-tek©, Norcross, GA) beads, in equal volumes to the PCR product. Preparation of PCR plates were undertaken in a Purifier Logic+ Class II biological safety cabinet (Labconco©, Kansas City, MO).

Illumina© MiSeq library preparation, after cleaning, was done using the services of the Genomics and Cell Characterization Core Facility of the Institute of Molecular Biology of the University of Oregon (Eugene, OR). Samples were normalized and pooled, along with samples from another study for a shared Illumina run. The amount of DNA being pulled from each sample was 10.45 ng (maximum allowed by the lowest concentration sample), with 258 x 10.45 ng = 2696.1 ng total, in a final volume of 384.47 µL = 7.013 ng/µL final pool concentration. Size selection was done using a Blue Pippen system with a 1.5% agarose cassette (Sage Science, Inc., Beverly, MA) to exclude DNA fragments with less than 250 bp lengths. Average ITS1 fragment length was 343 bp. Fragments larger than expected ITS1 lengths were removed bioinformatically after sequencing. Final DNA concentration within 250-1200 bp range was 5.213 nM, eluted in approximately 30 μl.

Illumina MiSeq platform sequencing of wood endophyte ITS library occurred at the Center for Genome Research and Biocomputing at Oregon State University (Corvallis, OR) using a 600 cycle (2x300 bp) v3 MiSeq reagent kit and including a 10% PhiX spike-in. Quantification of the shared library using qPCR was also done at the Center for Genome Research and Biocomputing facility. Reads from the shared run totaled to approximately 23 x 106 sequences, of which approximately 5.5 x 106 were from the present study.

Mock community construction

In addition to ecological samples, a pure-water negative control and two positive controls (in the form of “mock communities”, as suggested by Nguyen 2015) were included with the wood fungal endophyte library. To construct the positive controls, purified genomic DNA from 23 species of fungi from three phyla (19 Ascomycota, 3 Basidiomycota, and 1 Mucoromycota) were quantified using a NanoDrop 1000 UV-Vis Spectrophotometer (Thermo Scientific, NanoDrop products¸ Wilmington, DE) and diluted to a mean concentration of 9.44 ng/µl (SD = 2.35), then combined into a single sample for inclusion in the multiplexed wood fungal endophyte library. An ITS-region-only positive control was also generated using these same 23 species of fungi, using ITS1F and ITS4 primers (Gardes 1993) to amplify the full ITS region of each fungal species. PCR reagents were, per 20 µL rxn: 0.8 µL MgCl2, 0.6 µL each of forward and reverse primers, 4.0 µL H2O, 4.0 µL template DNA, and 10 µL 2x PCR Super Master Mix (Bimake, Houston, TX). PCR protocols were as follows: 5 min denaturation at 95 °C; 34 cycles of 60 s at 95 °C, 60 s at 55°C, and 60 s at 72 °C; and 10 min at 72 °C for final extension. PCR products were purified with Zymo© Clean and Concentrator column kits (Zymo Research Corp., Irvine CA). Full ITS PCR product from each fungal species was then diluted to a mean concentration of 24.30 ng/µL (SD=1.74) and combined to provide a second, ITS-region-only positive control. Full ITS region PCR product from each member of the mock community were sequenced using Sanger sequencing at Functional Biosciences, Inc (Madison, Wisconsin) on ABI 3730xl instruments using Big Dye V3.1 (ThermoFisher Scientific, Waltham, MA), to provide sequence information for UNITE database taxonomy assignments and to provide reference sequences for downstream recovery of these fungal sequences when examining positive controls (see below). All mock communities were prepared in a physically separate location from PCR preps of ecological samples to avoid cross-contamination. Taxonomic identities of positive control members are shown in Table 5.1.

## Bioinformatics

Full scripts of bioinformatic pipelines with extensive annotation are available in supplementary information, and as a jupyter notebook at: <https://nbviewer.jupyter.org/github/danchurch/taiwan_combined_biom/blob/master/combo_biome-revived.ipynb> . G,eneral bioinformatics protocols followed the USEARCH/UPARSE pipeline version 8.1 (Edgar, 2013) wherever possible. Libraries of leaf and wood fungal endophyte DNA were prepared separately, to maximize comparability, the reads from both libraries were combined as early as possible in the bioinformatics pipeline, following merging of paired ends. Reads were trimmed according to remove regions of low quality base calls. Wood forward and reverse reads were trimmed to 255 bp and 210 bp lengths, respectively. Leaf reads were trimmed to 170 and 263 bp lengths. OTUs were generated using a 3% dissimilarity cutoff. Variance stabilization, the process of maximizing comparability among samples and studies while considering differences in sampling depth, was done using using the DESeq2 and phyloseq packages in R (Love et al., 2014; McMurdie and Holmes, 2013), using leaf/wood as the design variable. Positive controls were used to calibrate OTU similarity radius and minimum cutoffs, which were subtracted from all observations to reduce error from index-misassignment and artificial splitting of OTUs. Large differences in abundances remained among positive control OTUs even after variance stabilization, so all statistical analyses were conducted with incidence (presence/absence)-transformed community matrices.

Initial taxonomic assignments were assigned to all possible reads in the USEARCH 8.1 UTAX algorithm against the UNITE fungal ITS database (Edgar, 2013; Kõljalg et al., 2013). This is a high-throughput method, used here to create an aggregate, “big picture” of the endophyte communities, summarized in Figure 3. Because of the low confidence of many of the high-throughput identifications, classes containing less than 1% of total OTUs were not included in this figure, but are available for inspection in scripts. Manual curation of taxonomic calls was then required for members of the core microbiomes. For higher confidence taxonomic calls, members of the wood and leaf core microbiome were assigned taxonomy using the BLAST (Altschul et al., 1990), against the UNITE database. See scripts (<https://nbviewer.jupyter.org/github/danchurch/taiwan_combined_stats/blob/master/CSrev/CSrev.ipynb#leafManTax> ) for details of this process .

## Statistical methods

### Overview

Ecological patterns of the entire fungal community of leaves and wood of all hosts were examined first. Analyses were then focused on patterns in the mycobiome of the most commonly-sampled host tree, *Helicia* *formosana*. Finally, host-fungus co-occurrence patterns were used to define a core mycobiome that was also examined for ecological patterns. Statistical analysis was conducted in R Statistical Software, version 3.3.1 (R Core Team, 2017), with the vegan (Oksanen et al., 2017), phyloseq (McMurdie and Holmes, 2013), cooccur (Griffith et al., 2016), igraph (Csardi and Nepusz, 2006) and ecodist (Goslee and Urban, 2007) packages. The statistical pipeline is available as a jupyter notebook, viewable at:

<https://nbviewer.jupyter.org/github/danchurch/taiwan_combined_stats/blob/master/CSrev/CSrev.ipynb>

All read abundances in fungal community matrices were transformed to presence/absence (see bioinformatics, above), so all downstream community analyses were based on this incidence data. Tests and visualizations that required the use of dissimilarity included PERMANOVA and Mantel tests, NMS and Bray-Curtis map visualizations. In all cases endophyte community comparisons were conducted using Bray-Curtis distances (Bray and Curtis, 1957; McCune et al., 2002).

### Mycobiome of all hosts

Dissimilarity of leaf and wood endophyte communities were, as a first step, modeled and visualized using non- parametric multivariate analysis of variance (NPMANOVA or PERMANOVA) (Anderson, 2017, 2001), and non-metric multidimensional scaling (NMS). Following initial comparisons of community composition, all analyses of wood and leaf endophytes were conducted separately, in parallel. Host and environmental variables of vegetative community and topography (Fig. 1) as predictors of endophyte communities were modeled individually using PERMANOVA. However, PERMANOVA tests as employed here are not sensitive to the complexities of spatial patterns on the variable landscape of Fushan FDP. For this reason spatially-explicit tests for autocorrelation and correlations with endophyte community composition and environmental data were used.

Spatial trends in endophyte communities were first explored using multivariate Mantel tests (Mantel, 1967; Legendre and Fortin, 1989) of community dissimilarity matrices against physical distance matrices, and visualized with Mantel multivariate correlograms. For greater resolution of spatial trends, distance-based Moran’s eigenvector maps analysis, also known as Principal Components of Neighbor Matrices (PCNM) analysis, was conducted on our sampling scheme. Following the general statistical pipeline recommended by Legendre et al. (Borcard et al., 2011; Legendre and Legendre, 2012), endophyte community matrices were Hellinger-transformed (Legendre and Gallagher, 2001), and “regressed” using Redundancy analysis (RDA) (Legendre and Gallagher, 2001; Buttigieg and Ramette, 2014) against all eigenvectors (“PCNM vectors”) resulting from dbMEM analysis. Stepwise model selection was then used to filter the ecologically informative eigenvectors (Oksanen et al., 2017).

Eigenanalysis of sampling schemes such as dbMEM/PCNM used here give a portrait of what patterns can be confidently tested by a particular sampling scheme. They are useful in that they can approximate classical linear spatial patterns, but also model other, more complex spatial patterns. These eigenvectors represent a range of spatial patterns that are periodic functions of wavelengths of varying size and direction. When combined with constrained direct gradient analyses such as Canonical Correspondence Analysis (CCA) or Redundancy analysis (RDA), eigenanalyses of sampling schemes become a very sensitive tool for detecting important spatial patterns in biological communities.

As a first filter, only PCNM vectors that are in some way correlated with changes in the community matrix are considered ecologically meaningful and are retained. Following this, if a correlation is found with these PCNM vectors with either between differences in fungal community or with patterns in environmental data, we can infer that these environmental or community differences behave on the landscape somewhat like the PCNM vectors. This is especially true if a large amount of variance in the community matrix or environmental data is explained by a PCNM (i.e. R2 is high). The wavelengths and shapes of PCNM vectors are often visualized using “bubble” plots, 2-D scatter plot with proportionally sized and colored symbols (Borcard et al., 2011). Ecological patterns of interest detected in spatial analysis were further visualized by mapping Bray-Curtis distance of all wood or leaf samples from a single point of interest on maps of the reserve and in NMS ordinations.

### Mycobiome of a single host, *Helicia* *formosana*

To exclude variation in fungal communities resulting from differing host tree species, above analyses were repeated for the fungal endophytes of a single host tree, *Helicia* *formosana* This was the host tree for which the most samples (leaves, n=31; wood n=22) were available. Bray-Curtis dissimilarity values resulting from comparison were then plotted onto a map of Fushan FDP.

### Core fungi of *Helicia* *formosana*

To test for the presence of a core mycobiome, co-occurrence analysis was conducted on the all-host, all- endophyte species-using a pairwise, probabilistic model (Veech, 2013; Griffith et al., 2016). Core mycobiomes of hosts were defined as the subset of fungi that showed strong co-occurrence associations with a host. Strong associations were defined as those with probabilities under null models of random association corrected to a Benjamini-Hochberg false discovery rate (FDR) of 0.05 or less. Fungal OTUs found to be strongly cooccurring with *H. formosana* were used to populate a species composition matrix of just these core species as columns, with rows of just sites where *H. formosana* was sampled. Patterns of this subset of core fungi were then visualized by first calculating Bray-Curtis dissimilarity distance of each sample (row) of this subsetted “core matrix” from an idealized core mycobiome row that contained all members of the core fungi. These values were then mapped on the Fushan FDP plot.

# Results

## Mycobiome of all hosts

After variance-stabilization, in the the leaf endophyte library 1302 OTUs were detected, and 2025 OTUs in the the wood library. Both leaf and wood samples were dominated by Ascomycota (Fig. 3), but a larger proportion of remaining reads in wood OTUs matched to Basidiomycota (10% of leaf OTUs vs. 17% of wood OTUs) than in in leaves. This larger percentage of OTUs identified to Basidiomycota was due mostly to a larger percentage of OTUs identified to Tremellomycetes (<1% of leaf OTUs vs. 5.6% of wood OTUs). Within Ascomycota, both leaf and wood samples contained high percentages of OTUs identified to Sordariomycetes (24% of leaf OTUs vs. 16% of wood OTUs), Dothideomycetes (20% of leaf OTUs vs. 27% of wood OTUs), and Eurotiomycetes (7% of leaf OTUs vs. 14% of wood OTUs). Significant observations of Lecanoromycete fungi also occurred, especially in leaves (%18.5 of leaf OTUs vs. %4 of wood OTUs). At the class level 16% of leaf OTUs were unidentified, compared to 10% of wood.

Host species was the strongest single predictor of similarity within leaf endophyte communities (PERMANOVA, F(33, 89) = 1.95, p < 0.01, R2 =0.42, permutations = 10000). Wood endophyte communities were also most strongly predicted by host (PERMANOVA, F(29,61) = 1.49, p < 0.01, R2 = 0.42, permutations = 10000), see NMS visualizations (Fig. 4). Surrounding vegetative community (Fig. 1a) was a weak predictor of similarity in both leaf (PERMANOVA, F(3, 119) = 2.04, p < 0.01, R2 = .05, permutations = 10000) and wood endophyte community (PERMANOVA, F(3,87) = 1.71, p < 0.01, R2 = 0.055, permutations = 10000). Micro-topographic conditions (Fig. 1b) were also weak predictors of similarity in both leaf (PERMANOVA, F(6, 116) = 1.23, p < 0.01, R2 = 0.06, permutations = 10000) and wood endophyte community (PERMANOVA, F(6,84) = 1.36, p < 0.01, R2 = 0.09, permutations = 10000). PERMANOVA tests cannot easily account for complex spatial patterns, some environmental correlations were uncovered in more sensitive spatial analyses summarized below.

Wood endophyte community displayed a weak pattern of community-turnover/distance-decay over the entire study area (global Mantel's r = 0.10, p < 0.01) (Supp Fig. 1). Leaf communities displayed no global distance decay relationship (global Mantel's r = -0.01, p = 0.57), but displayed local negative autocorrelation in comparisons of samples approximately 200 m apart (Mantel correlogram, local Mantel's r = -0.10, p < 0.05). In both wood and leaf endophyte communities of all hosts, Mantel’s r approached zero in both leaf and wood samples at comparisons around 150 m apart, indicating that positive autocorrelation was undetected beyond this.

For leaves, our sampling scheme yielded 7 biologically significant PCNM vectors, explaining a combined total of 7.1% of endophyte community variation (Redundancy analysis, constrained inertia = 0.06, Unconstrained inertia = 0.82, F(7,117) = 1.25, P < 0.01) (Supp. Fig. 2). Of these, the smallest scale vectors were uncorrelated with environmental data, and probably indicative of endogenous autocorrelation (Borcard et al., 2011), with wavelengths up to 50 m. Two mid-range (300m) vectors represented north-south and east-west surface trends, unexplained by environmental data, each predicting approximately 1 % of fungal species variation. Two mid-range leaf PCNM vectors of interest correlated with environmental variables, most strongly with the presence of the steep habitat zone present mostly on the central hill of the plot (Linear model/multiple regression, adj-R2=0.23, F(9,113)=4.95, p < 0.01). The largest scale (500 m) PCNM vector ran in a NE-SW direction, resulting in a partial contrast between the southwest valley and the rest of the plot. It predicted 1.2% of leaf endophyte community variation, and was most strongly explained by the presence of the two vegetation communities dominated by the tree *Helicia formosana* (Linear model/multiple regression, adj-R2=0.33, F(9,113)=7.68, p < 0.01).

For wood, our sampling scheme also yielded 5 biologically significant PCNM vectors, explaining 7.5% of variation (Redundancy analysis, constrained inertia = 0.07, Unconstrained inertia = 0.80, F(5,85) = 1.38, P < 0.01). Of these, the smallest scale vectors were probably indicative of endogenous autocorrelation, at a wavelength of 70 m or less. The three other PCNM vectors are large- to mid-range, with lengths from 300m to 500m. They correlated with environmental variables, and generally point to a contrast between upland and lowland habitat (Supp. Fig. 3). For more detailed results and interpretation of all-host PCNM results, see scripts (supplementary text, also viewable online as a Jupyter Notebook at: <https://nbviewer.jupyter.org/github/danchurch/taiwan_combined_stats/blob/master/CSrev/CSrev.ipynb>)

## Mycobiome of a single host, *Helicia* *formosana*

When the community of endophytes was constrained to one host tree species, environmental variables were not found to directly explain any variance in endophyte community structure (PERMANOVA tests, permutations=10000. Leaf community ~ topography: F(4,26) = 0.93, R2 = 0.12, p = 0.66. Leaf community ~ vegetative community: F(3,27) = 1.07, R2 = 0.11, p = 0.27. Wood community ~ topography: F(4,17) = 1.06, R2 = 0.20, p = 0.20. Wood community ~ Vegetative community: F(3,18) = 1.08, R2 = 0.15, p = 0.19.).

Leaf endophyte community yielded three ecologically significant PCNMs explaining 13% of community variation (RDA, leaves: constrained inertia = 0.011, Unconstrained inertia = 0.67, F(2,19) = 1.32, P < 0.01). Wood endophyte community yielded two ecologically significant PCNMs explaining 12% of community variation (RDA, leaves: constrained inertia = 0.10, Unconstrained inertia = 0.70, F(1,29) = 1.78, P < 0.01). As with the non-spatial PERMANOVA model of *Helicia* endophytes, these PCNMs also did not correlate with any available environmental data, in either wood or leaves. However, both wood and leaf communities showed ecologically meaningful PCNM vectors that center on the southwestern valley of the FDP as a place of difference in the *Helicia* mycobiome (Fig. 5). This pattern of dissimilarity was particularly pronounced in the *Helicia* leaf endophyte community when visualized with Bray-Curtis comparisons, (Fig. 6).

## Core fungi of *Helicia* *formosana*

Out of 1302 possible fungal OTUs observed in leaves of all hosts, 426 OTUs were found in *Helicia* leaves*.* Of these 12 showed strong patterns of co-occurrence with *Helicia* *formosana.* Out of 2025 possible fungal OTUs observed in woody tissue of all hosts, 731 OTUs were found in *Helicia* wood (Supp. Fig. 4)*.* Of these 7 OTUs showed strong patterns of co-occurrence with *Helicia* *formosana* (Supp. Fig. 4)*.* These fungi were considered members of the *H. formosana* core mycobiome (Table 1).Visual inspection of patterns of dissimilarity show that only leaves within the southern valley of the plot contained relatively high proportions of core fungi (Fig. 7). Wood samples retained most of their core fungi consistently throughout the plot (Fig. 7).

## Summary comparison

The above analysis compared patterns of community dissimilarity at several levels (Table 2). Wood and leaf endophytes of all host-trees showed a similar pattern, very high levels of dissimilarity among all samples (all-host leaf endophyte mean BC=0.90, sd=0.09; wood endophyte mean BC=0.87, sd=0.07). *Helicia* *formosana* samples show a lower average level of dissimilarity (leaf mean BC=0.80, sd =0.11; wood mean BC = 0.80, sd=0.06). This variation can then be partitioned into two groups: (1) non-core fungi, which showed a high average level of dissimilarity (leaf mean BC=0.89, sd =0.08 ; wood mean BC = 0.83, sd=0.06), and (2) core fungi, which showed a lower mean BC (leaf mean BC=0.38, sd =0.17; wood mean BC = 0.39, sd=0.17).

# Discussion

Contrary to our original predictions, we found evidence for a consistent core of fungi in the wood of *Helicia formosana*. Closer to our predictions, we found in leaves a subset of fungi that cooccurred with *H. formosana* in just one area of the plot (Fig. 7). In leaves, these core fungi were most consistently present in the southern valley, and were often completely missing in other areas of the study. In wood, they were more "loyal", and coexisted more reliably with *H*. *formosana* throughout the plot. Applying terminology proposed by Hamady and Knight (2009), core woody endophytes here may be best described by the “minimal” core model: they were few in number among a large and highly variable microbiome, but were consistently present throughout the study. Leaves might be considered to have lacked a core mycobiome, or in the terminology of Hamady and Knight (2009), their core mycobiomes resembled “gradient” or “subpopulation” cores. These terms refer to microbes strongly associated with a host, but whose presence is highly conditional on environment and spatial scale.

When all host trees were compared, the average dissimilarity between any two trees was extremely high, (Fig. 8, Table 2). Samples were slightly more similar on average when constrained to a single host, for wood and leaves, seemingly a result of the strong effects of host (Figs. 4 and 8). The taxonomic core of fungi behaved differently from the fungal microbiome at large. Removing these fungi from consideration brought the mycobiome of their host, *H. formosana*, nearly back to the high background levels of dissimilarity among samples of the entire study, indicating that these may be the species through which host effects are manifested (Fig. 8).

To fungal symbionts, wood represents a clearly distinct set of ecological challenges and rewards. The absence of a consistent core mycobiome may perhaps be due to the more dynamic environment of leaves. Leaves are flushed mostly sterile (Arnold and Herre, 2003), and are shed within 1 to several years, in contrast with the longer lifespan of woody tissues. Access to internal tissues of leaves is relatively more abundant - high concentrations of stoma exist in leaves vs. lenticels, wounds or other openings in bark (Melotto et al., 2008). Bark as a protective tissue is much thicker and persistent than the cuticle of leaves. Woody tissues as a whole also contain relatively lower simple-sugar reward than leaves, high moisture levels inhibit cellulose decomposition activity (Chapela and Boddy, 1988), and often lignified structures are often present. Despite the challenges, extensive endophyte colonization and early decomposition has been detected (Boddy and Rayner, 1983; Oses et al., 2008), and significant diversity of aggressive decomposer species from Xylariaceae (Whalley, 1996) and white-rot clade basidiomycetes have been consistently detected as endophytes in wood (Martin et al., 2015). It has been suggested that mycelial networks of wood endophytes may converge from multiple distal origins, in roots, wounds, or branch tips (Boddy, 1994). Many of the fungi described as endophytes from woody tissue are therefore thought to be “patient”, latent saprotrophs (Boddy, 1994; Oses et al., 2008; Parfitt et al., 2010), that utilize the endophyte life stage to gain priority in decomposition. Some of these latent saprotrophs appear somewhat specialized in their substrates, if only by presence of macroscopic symptoms (Parfitt et al., 2010). All of this paints a portrait of the woody tissues of trees as an system that accrues its microbial partners more selectively and perhaps more slowly.

Most of the candidate core fungal species of the tree *Helicia formosana* were unidentified (Table 1), often even to the phylum level. This makes ecological interpretation of the core fungi we observed difficult or impossible. It also speaks to an urgent and daunting challenge facing mycologists today: the need for more high-quality accessions of lesser known fungi in fungal barcode databases and herbaria, to keep pace with the ongoing technical revolution of high-throughput sequencer technologies.

Within the candidates for core fungi of *H. formosana* that were indentified, three members of the genus *Phyllosticta* were present as leaf endophytes. This prevalence of *Phyllosticta* is not surprising, as the genus is commonly observed as leaf-associated endophytes, pathogens and saprotrophs (Promputtha et al., 2007). One of these *Phyllosticta* species was *P. capitalensis*, a well-known endophyte of tropical woody plants (Baayen et al., 2002; Okane et al., 2003, 2001), that may have been transported throughout the world via nursery trade (George Carroll, pers. Com.). Here we defined the fungi of a core mycobiome as those most reliably cooccurring with a particular plant host *within the sampled area.* In the case of *P. capitalensis*, this resulted in the inclusion of a cosmopolitan fungal species which has been observed in multiple hosts, but that selectively inhabited *H. formosana* within the study area. Whether this is appropriate is a topic for discussion, but does highlight the need for a mature definition of a taxonomic core microbiome beyond a purely bioinformatic one.

The presence of a core taxonomic group of microbes in a host might be considered a kind of stabilization or structuring of a portion of a host’s microbiome, possibly as a result of interactions among hosts and select microbes. When defining core microbiomes as we have here, it may be important to consider the different organs of hosts as very different refugia for microbes: here the woody tissues appeared to host a more consistent assemblage of core fungi. Similarly, the leaves of *Helicia* *formosana* trees in the more sheltered southwestern valley held more a consistent microbial core than those in more exposed areas of the plot. We are limited here in our ability to examine the importance of neutral spatial processes, given our coarse environmental data. However, these patterns suggest that even strong biological interactions between microbe and host can be disrupted or prevented. This disruption could stem from neutral processes such as obstacles to dispersal from topography, or environmental changes such as fierce summer storm events resetting community assembly in leaves. Regardless, it may be that for a consistent taxonomic core to develop in a plant microbiome, either local habitat or more persistent host tissue may need to provide some measure of stability from change. High rates of dispersal and disturbance can disrupt the tendency to local structure in communities and gene pools (Wright, 1940; Cadotte, 2006; Vellend, 2010). A parallel logic may apply for taxonomic microbiomes of large hosts.

# Acknowledgements

This project was jointly funded by the USA National Science Foundation and the National Science Council in Taiwan, under the 2013 East Asia and Pacific Summer Institutes program (EAPSI program 12-498). Additional funds for laboratory work were contributed by the Cascade Mycological Society, and the Mycological Society of America. We give our immense thanks to Dr. Sheng-Hsin Su and colleagues in the 2013 Fushan tree census crew, for supporting our field work. The Fushan Forest Dynamics Plot is a collaborative project of the Taiwan Forestry Research Institute, Taiwan Forestry Bureau, and National Taiwan University, and was funded by the Council of Agriculture and National Science Council in Taiwan.

# Works cited Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson, M.J., 2017. Permutational multivariate analysis of variance (PERMANOVA), in: Wiley StatsRef: Statistics Reference Online. American Cancer Society, pp. 1–15. https://doi.org/10.1002/9781118445112.stat07841

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Arnold, A.E., Herre, E.A., 2003. Canopy cover and leaf age affect colonization by tropical fungal endophytes: Ecological pattern and process in *Theobroma cacao* (Malvaceae). Mycologia 95, 388–398. https://doi.org/10.1080/15572536.2004.11833083

Arnold, A.E., Mejía, L.C., Kyllo, D., Rojas, E.I., Maynard, Z., Robbins, N., Herre, E.A., 2003. Fungal endophytes limit pathogen damage in a tropical tree. PNAS 100, 15649–15654. https://doi.org/10.1073/pnas.2533483100

Baas-Becking, L.G.M., 1934. Geobiologie; of inleiding tot de milieukunde. WP Van Stockum & Zoon NV.

Baayen, R.P., Bonants, P.J.M., Verkley, G., Carroll, G.C., van der Aa, H.A., de Weerdt, M., van Brouwershaven, I.R., Schutte, G.C., Maccheroni, W., de Blanco, C.G., Azevedo, J.L., 2002. Nonpathogenic isolates of the citrus black spot fungus, *Guignardia citricarpa*, identified as a cosmopolitan endophyte of woody plants, *G. mangiferae* (*Phyllosticta capitalensis*). Phytopathology 92, 464–477. https://doi.org/10.1094/PHYTO.2002.92.5.464

Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M., Kauserud, H., 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. Mol Ecol Resour 13, 218–224. <https://doi.org/10.1111/1755-0998.12065>

Boddy, L., 1994. Latent Decay Fungi: The Hidden Foe? Arboricultural Journal 18, 113–135. https://doi.org/10.1080/03071375.1994.9747007

Boddy, L., Rayner, A.D.M., 1983. Ecological roles of basidiomycetes forming decay communities in attached oak branches. New Phytologist 93, 77–88.

Bodenhausen, N., Horton, M.W., Bergelson, J., 2013. Bacterial Communities Associated with the Leaves and the Roots of Arabidopsis thaliana. PLOS ONE 8, e56329. https://doi.org/10.1371/journal.pone.0056329

Borcard, D., Gillet, F., Legendre, P., 2011. Spatial analysis of ecological data, in: Numerical Ecology with R. Springer, pp. 227–292.

Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological monographs 27, 325–349.

Buttigieg, P.L., Ramette, A., 2014. A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. FEMS microbiology ecology 90, 543–550.

Cadotte, M.W., 2006. Dispersal and Species Diversity: A Meta‐Analysis. The American Naturalist 167, 913–924. https://doi.org/10.1086/504850

Carroll, G., 1988. Fungal endophytes in stems and leaves: from latent pathogen to mutualistic symbiont. Ecology 69, 2–9.

Carroll, G.C., Carroll, F.E., 1978. Studies on the incidence of coniferous needle endophytes in the Pacific Northwest. Canadian Journal of Botany 56, 3034–3043.

Chapela, I.H., Boddy, L., 1988. Fungal colonization of attached beech branches. New Phytologist 110, 47–57. https://doi.org/10.1111/j.1469-8137.1988.tb00236.x

Crous, P.W., 1998. Mycosphaerella spp. and their anamorphs associated with leaf spot diseases of Eucalyptus. American Phytopathological Society (APS Press).

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695, 1–9.

De Wit, R., Bouvier, T., 2006. ‘Everything is everywhere, but, the environment selects’; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8, 755–758. https://doi.org/10.1111/j.1462-2920.2006.01017.x

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10, 996–998. https://doi.org/10.1038/nmeth.2604

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., Sundaresan, V., 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. PNAS 112, E911–E920. https://doi.org/10.1073/pnas.1414592112

Fang, F.C., Casadevall, A., 2011. Reductionistic and Holistic Science. Infect Immun 79, 1401–1404. https://doi.org/10.1128/IAI.01343-10

Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. Molecular Ecology 2, 113–118. https://doi.org/10.1111/j.1365-294X.1993.tb00005.x

Gazis, R., Chaverri, P., 2010. Diversity of fungal endophytes in leaves and stems of wild rubber trees (Hevea brasiliensis) in Peru. Fungal Ecology 3, 240–254. https://doi.org/10.1016/j.funeco.2009.12.001

Glienke, C., Pereira, O.L., Stringari, D., Fabris, J., Kava-Cordeiro, V., Galli-Terasawa, L., Cunnington, J., Shivas, R.G., Groenewald, J.Z., Crous, P.W., 2011. Endophytic and pathogenic Phyllosticta species, with reference to those associated with Citrus Black Spot. Persoonia 26, 47–56. https://doi.org/10.3767/003158511X569169

Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R., Reid, G., 2010. Microbiome Profiling by Illumina Sequencing of Combinatorial Sequence-Tagged PCR Products. PLOS ONE 5, e15406. https://doi.org/10.1371/journal.pone.0015406

Goslee, S.C., Urban, D.L., 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22, 1–19.

Green, J., Bohannan, B.J.M., 2006. Spatial scaling of microbial biodiversity. Trends in Ecology & Evolution 21, 501–507. https://doi.org/10.1016/j.tree.2006.06.012

Griffith, D.M., Veech, J.A., Marsh, C.J., 2016. Cooccur: probabilistic species co-occurrence analysis in R. J Stat Softw 69, 1–17.

Hamady, M., Knight, R., 2009. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Res. https://doi.org/10.1101/gr.085464.108

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B.H., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews Microbiology 10, 497–506. https://doi.org/10.1038/nrmicro2795

Higgins, K.L., Arnold, A.E., Coley, P.D., Kursar, T.A., 2014. Communities of fungal endophytes in tropical forest grasses: highly diverse host- and habitat generalists characterized by strong spatial structure. Fungal Ecology 8, 1–11. https://doi.org/10.1016/j.funeco.2013.12.005

Hoffman, M.T., Arnold, A.E., 2010. Diverse Bacteria Inhabit Living Hyphae of Phylogenetically Diverse Fungal Endophytes. Appl. Environ. Microbiol. 76, 4063–4075. https://doi.org/10.1128/AEM.02928-09

Johnson, C.L., Versalovic, J., 2012. The Human Microbiome and Its Potential Importance to Pediatrics. Pediatrics peds.2011-2736. https://doi.org/10.1542/peds.2011-2736

Kembel, S.W., Mueller, R.C., 2014. Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. Botany 92, 303–311.

Kim, M., Singh, D., Lai-Hoe, A., Go, R., Rahim, R.A., A.n, A., Chun, J., Adams, J.M., 2012. Distinctive Phyllosphere Bacterial Communities in Tropical Trees. Microb Ecol 63, 674–681. https://doi.org/10.1007/s00248-011-9953-1

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F.S., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G.W., Hartmann, M., Kirk, P.M., Kohout, P., Larsson, E., Lindahl, B.D., Lücking, R., Martín, M.P., Matheny, P.B., Nguyen, N.H., Niskanen, T., Oja, J., Peay, K.G., Peintner, U., Peterson, M., Põldmaa, K., Saag, L., Saar, I., Schüßler, A., Scott, J.A., Senés, C., Smith, M.E., Suija, A., Taylor, D.L., Telleria, M.T., Weiss, M., Larsson, K.-H., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22, 5271–5277. https://doi.org/10.1111/mec.12481

Lau, M.K., Arnold, A.E., Johnson, N.C., 2013. Factors influencing communities of foliar fungal endophytes in riparian woody plants. Fungal Ecology 6, 365–378. https://doi.org/10.1016/j.funeco.2013.06.003

Legendre, P., Fortin, M.J., 1989. Spatial pattern and ecological analysis. Vegetatio 80, 107–138.

Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271–280.

Legendre, P., Legendre, L., 2012. Numerical Ecology, Third. ed, Developments in Environmental Modelling. Elsevier.

Lemanceau, P., Blouin, M., Muller, D., Moënne-Loccoz, Y., 2017. Let the core microbiota be functional. Trends in plant science 22, 583–595.

Liggenstoffer, A.S., Youssef, N.H., Couger, M.B., Elshahed, M.S., 2010. Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. The ISME Journal 4, 1225–1235. https://doi.org/10.1038/ismej.2010.49

Lloyd-Price, J., Abu-Ali, G., Huttenhower, C., 2016. The healthy human microbiome. Genome medicine 8, 51.

Losos, E., Leigh, E.G. (Eds.), 2004. Tropical Forest Diversity and Dynamism. University of Chicago Press.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Rio, T.G. del, Edgar, R.C., Eickhorst, T., Ley, R.E., Hugenholtz, P., Tringe, S.G., Dangl, J.L., 2012. Defining the core *Arabidopsis thaliana* root microbiome. Nature 488, 86–90. https://doi.org/10.1038/nature11237

Malloch, D.W., Pirozynski, K.A., Raven, P.H., 1980. Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (A Review). PNAS 77, 2113–2118. https://doi.org/10.1073/pnas.77.4.2113

Mantel, N., 1967. The Detection of Disease Clustering and a Generalized Regression Approach. Cancer Res 27, 209–220.

Martin, R., Gazis, R., Skaltsas, D., Chaverri, P., Hibbett, D., 2015. Unexpected diversity of basidiomycetous endophytes in sapwood and leaves of *Hevea*. Mycologia 107, 284–297.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V.H., Staley, J.T., 2006. Microbial biogeography: putting microorganisms on the map. Nature Reviews Microbiology 4, 102–112. https://doi.org/10.1038/nrmicro1341

McCune, B., Grace, J.B., Urban, D.L., 2002. Analysis of ecological communities. MjM software design Gleneden Beach, OR.

McMurdie, P.J., Holmes, S., 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8, e61217. https://doi.org/10.1371/journal.pone.0061217

Meadow, J.F., Bateman, A.C., Herkert, K.M., O’Connor, T.K., Green, J.L., 2013. Significant changes in the skin microbiome mediated by the sport of roller derby. PeerJ 1, e53. https://doi.org/10.7717/peerj.53

Mejía, L.C., Kyllo, D.A., Rojas, E., Maynard, Z., Butler, A., Van Bael, S.A., Herre, E., 2007. Ecological Implications of Anti-Pathogen Effects of Tropical Fungal Endophytes and Mycorrhizae. Ecology 88, 550–558. https://doi.org/10.1890/05-1606

Mejía, L.C., Rojas, E.I., Maynard, Z., Bael, S.V., Arnold, A.E., Hebbar, P., Samuels, G.J., Robbins, N., Herre, E.A., 2008. Endophytic fungi as biocontrol agents of Theobroma cacao pathogens. Biological Control, Special Issue: Endophytes 46, 4–14. <https://doi.org/10.1016/j.biocontrol.2008.01.012>

Melotto, M., Underwood, W., He, S.Y., 2008. Role of stomata in plant innate immunity and foliar bacterial diseases. Annu. Rev. Phytopathol. 46, 101–122.

Mueller, R.C., Paula, F.S., Mirza, B.S., Rodrigues, J.L., Nüsslein, K., Bohannan, B.J., 2014. Links between plant and fungal communities across a deforestation chronosequence in the Amazon rainforest. The ISME Journal 8, 1548–1550. https://doi.org/10.1038/ismej.2013.253

Mummey, D.L., Rillig, M.C., 2008. Spatial characterization of arbuscular mycorrhizal fungal molecular diversity at the submetre scale in a temperate grassland. FEMS Microbiol Ecol 64, 260–270. https://doi.org/10.1111/j.1574-6941.2008.00475.x

Nemergut, D.R., Schmidt, S.K., Fukami, T., O’Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P., Ferrenberg, S., 2013. Patterns and Processes of Microbial Community Assembly. Microbiology and Molecular Biology Reviews 77, 342–356. https://doi.org/10.1128/MMBR.00051-12

Norros, V., Penttilä, R., Suominen, M., Ovaskainen, O., 2012. Dispersal may limit the occurrence of specialist wood decay fungi already at small spatial scales. Oikos 121, 961–974. <https://doi.org/10.1111/j.1600-0706.2012.20052.x>

Okane, I., Nakagiri, A., Ito, T., 2001. Identity of *Guignardia* sp. inhabiting ericaceous plants. Canadian Journal of Botany 79, 101–109.

Okane, I., Nakagiri, A., Ito, T., Lumyong, S., 2003. Extensive host range of an endophytic fungus, *Guignardia endophyllicola* (anamorph: *Phyllosticta capitalensis*). Mycoscience 44, 353–363. https://doi.org/10.1007/S10267-003-0128-X

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P., 2017. M. Stevens MHH, Szoecs E, Wagner H (2017). vegan: Community Ecology Package. R package version 2.4-3.

Oldroyd, G.E.D., 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nature Reviews Microbiology 11, 252–263. <https://doi.org/10.1038/nrmicro2990>

Oses, R., Valenzuela, S., Freer, J., Sanfuentes, E., Rodriguez, J., 2008. Fungal endophytes in xylem of healthy Chilean trees and their possible role in early wood decay. Fungal Divers 33, 77–86.

Parfitt, D., Hunt, J., Dockrell, D., Rogers, H.J., Boddy, L., 2010. Do all trees carry the seeds of their own destruction? PCR reveals numerous wood decay fungi latently present in sapwood of a wide range of angiosperm trees. Fungal Ecology 3, 338–346. https://doi.org/10.1016/j.funeco.2010.02.001

Peay, K.G., Garbelotto, M., Bruns, T.D., 2010. Evidence of dispersal limitation in soil microorganisms: Isolation reduces species richness on mycorrhizal tree islands. Ecology 91, 3631–3640. https://doi.org/10.1890/09-2237.1

Peršoh, D., 2013. Factors shaping community structure of endophytic fungi–evidence from the Pinus-Viscum-system. Fungal Diversity 60, 55–69. https://doi.org/10.1007/s13225-013-0225-x

Porras-Alfaro, A., Bayman, P., 2011. Hidden Fungi, Emergent Properties: Endophytes and Microbiomes. Annual Review of Phytopathology 49, 291–315. <https://doi.org/10.1146/annurev-phyto-080508-081831>

Promputtha, I., Lumyong, S., Dhanasekaran, V., McKenzie, E.H.C., Hyde, K.D., Jeewon, R., 2007. A Phylogenetic Evaluation of Whether Endophytes Become Saprotrophs at Host Senescence. Microb Ecol 53, 579–590. https://doi.org/10.1007/s00248-006-9117-x

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rodrigues, J.L.M., Pellizari, V.H., Mueller, R., Baek, K., Jesus, E. da C., Paula, F.S., Mirza, B., Hamaoui, G.S., Tsai, S.M., Feigl, B., Tiedje, J.M., Bohannan, B.J.M., Nüsslein, K., 2013. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. PNAS 110, 988–993. https://doi.org/10.1073/pnas.1220608110

Rodriguez, R., Redman, R., 2008. More than 400 million years of evolution and some plants still can’t make it on their own: plant stress tolerance via fungal symbiosis. J Exp Bot 59, 1109–1114. https://doi.org/10.1093/jxb/erm342

Rodriguez, R.J., White Jr, J.F., Arnold, A.E., Redman, R.S., 2009. Fungal endophytes: diversity and functional roles. New Phytologist 182, 314–330. https://doi.org/10.1111/j.1469-8137.2009.02773.x

Rojas, E.I., Rehner, S.A., Samuels, G.J., Bael, S.A.V., Herre, E.A., Cannon, P., Chen, R., Pang, J., Wang, R., Zhang, Y., Peng, Y.-Q., Sha, T., 2010. Colletotrichum gloeosporioides s.l. associated with Theobroma cacao and other plants in Panamá: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. Mycologia 102, 1318–1338. https://doi.org/10.3852/09-244

Rosenberg, E., Sharon, G., Atad, I., Zilber-Rosenberg, I., 2010. The evolution of animals and plants via symbiosis with microorganisms. Environmental Microbiology Reports 2, 500–506. https://doi.org/10.1111/j.1758-2229.2010.00177.x

Santamaría, J., Bayman, P., 2005. Fungal Epiphytes and Endophytes of Coffee Leaves (<Emphasis Type="Italic">Coffea arabica</Emphasis>). Microb Ecol 50, 1–8. https://doi.org/10.1007/s00248-004-0002-1

Schulz, B., Römmert, A.-K., Dammann, U., Aust, H.-J., Strack, D., 1999. The endophyte-host interaction: a balanced antagonism? Mycological Research 103, 1275–1283.

Shade, A., Handelsman, J., 2012. Beyond the Venn diagram: the hunt for a core microbiome. Environmental Microbiology 14, 4–12. https://doi.org/10.1111/j.1462-2920.2011.02585.x

Su, S., Hsieh, C., Chang-Yang, C., Lu, C., Guan, B.T., 2010. Micro-topographic differentiation of the tree species composition in a subtropical submontane rainforest in northeastern Taiwan. Taiwan Journal of Forest Science 25, 63–80.

Su, S.-H., Chang-Yang, C.H., Lu, C.L., Tsui, C.C., Lin, T.T., Lin, C.L., Chiou, W.L., Kuan, L.H., Chen, Z.S., Hsieh, C.F., 2007. Fushan subtropical forest dynamics plot: tree species characteristics and distribution patterns. Taiwan Forestry Research Institute.

Tateno, O., Hirose, D., Osono, T., Takeda, H., 2015. Beech cupules share endophytic fungi with leaves and twigs. Mycoscience 56, 252–256. https://doi.org/10.1016/j.myc.2014.07.005

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., Ruiz, L.V., Vasco-Palacios, A.M., Thu, P.Q., Suija, A., Smith, M.E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., Piepenbring, M., Phosri, C., Peterson, M., Parts, K., Pärtel, K., Otsing, E., Nouhra, E., Njouonkou, A.L., Nilsson, R.H., Morgado, L.N., Mayor, J., May, T.W., Majuakim, L., Lodge, D.J., Lee, S.S., Larsson, K.-H., Kohout, P., Hosaka, K., Hiiesalu, I., Henkel, T.W., Harend, H., Guo, L., Greslebin, A., Grelet, G., Geml, J., Gates, G., Dunstan, W., Dunk, C., Drenkhan, R., Dearnaley, J., Kesel, A.D., Dang, T., Chen, X., Buegger, F., Brearley, F.Q., Bonito, G., Anslan, S., Abell, S., Abarenkov, K., 2014. Global diversity and geography of soil fungi. Science 346, 1256688. https://doi.org/10.1126/science.1256688

Thomas, D., Bailes, G., Vandegrift, R., Roy, B.A., 2017. Understanding and mitigating some limitations of Illumina © MiSeq for environmental sequencing of fungi. bioRxiv.

Van Der Heijden, M.G.A., Horton, T.R., 2009. Socialism in soil? The importance of mycorrhizal fungal networks for facilitation in natural ecosystems. Journal of Ecology 97, 1139–1150. https://doi.org/10.1111/j.1365-2745.2009.01570.x

Vandegrift, A.W.R., 2016. Ecological Roles of Fungal Endophytes. University of Oregon.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., Dufresne, A., 2015. The importance of the microbiome of the plant holobiont. New Phytol 206, 1196–1206. https://doi.org/10.1111/nph.13312

Veech, J.A., 2013. A probabilistic model for analysing species co-occurrence. Global Ecology and Biogeography 22, 252–260. https://doi.org/10.1111/j.1466-8238.2012.00789.x

Vellend, M., 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology 85, 183–206. https://doi.org/10.1086/652373

Venkateshwaran, M., Volkening, J.D., Sussman, M.R., Ané, J.-M., 2013. Symbiosis and the social network of higher plants. Current Opinion in Plant Biology, Growth and development 16, 118–127. https://doi.org/10.1016/j.pbi.2012.11.007

Walters, W.A., Caporaso, J.G., Lauber, C.L., Berg-Lyons, D., Fierer, N., Knight, R., 2011. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. Bioinformatics 27, 1159–1161. https://doi.org/10.1093/bioinformatics/btr087

Wilson, D., 1995. Endophyte: The Evolution of a Term, and Clarification of Its Use and Definition. Oikos 73, 274–276. <https://doi.org/10.2307/3545919>

Whalley, A.J.S., 1996. The xylariaceous way of life. Mycological Research 100, 897–922. https://doi.org/10.1016/S0953-7562(96)80042-6

Woodward, C., Hansen, L., Beckwith, F., Redman, R.S., Rodriguez, R.J., 2012. Symbiogenics: an epigenetic approach to mitigating impacts of climate change on plants. HortScience 47, 699–703.

Wright, S., 1940. Breeding Structure of Populations in Relation to Speciation. The American Naturalist 74, 232–248. https://doi.org/10.1086/280891

Zimmerman, N.B., Vitousek, P.M., 2012. Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. PNAS 109, 13022–13027. <https://doi.org/10.1073/pnas.1209872109>

# Captions

Figure 1 (A). topographic map of the Fushan FDP with the four vegetation types as classified by Su et al.(2007). (B): map of the habitat type, a composite classification based on microtopographic characteristics of quadrats, defined by Su et al. (2010). The units of the coordinates and contours are in meters, with quadrats at 20x20m scale. Figures reproduced with permission from authors.

Figure 2 (A). An overview of nested-squares, logarithmic sampling scheme Vandegrift (2016). Vertices of squares are sample sites. Units are meters. (B): Perspective diagram of Fushan Forest Dynamics Plot (Su et al., 2010). Figures reproduced with permission from authors.

Figure 3. Class-level overview of taxonomic composition of wood and leaf libraries, from all trees sampled. Proportions are out of total OTUs observed in each plant organ: 2025 OTUs in wood, 1302 in leaves.

Figure 4. Non-metric multidimensional scaling diagram of endophyte communities, with all tree hosts that were sampled at least 3 times.

Figure 5. Two PCNM vectors showing patterns of variation in endophyte communities of a single host-tree species, *Helicia* *formosana*. Size and color of bubble show differences in community predicted by this spatial pattern. For some significant subset of the total endophytes species sampled, sites with large black circles contain very different fungal species assemblages than sites with large white circles. Here, both leaf and wood endophyte communities display dissimilarity between the plot at large and the southern valley. Full arrays of ecologically significant PCNM vectors are provided in Supp Figs. 1 and 2.

Figure 6. Comparison of *Helicia* *formosana* samples against a sample the southwest valley of the plot, circled in red, using Bray-Curtis dissimilarity. Dark blue points (BC=1) share no fungal species in common with the circled sample, and increase in similarity from yellow to green (BC=0).

Figure 7. Comparisons between all *H*. *formosana* points to the core fungi of the *H*. *formosana,* using Bray-Curtis dissimilarity. Dark blue points (BC=1) contain no species from this set of core fungi, and increase in similarity from yellow to green (BC=0, 100% of core fungi present).

Figure 8. Distribution of Bray-Curtis dissimilarity among sample comparisons of all hosts, and of *Helicia* *formosana* only.

# Supplementary figures

Supplemental Figure 1 . Mantel correlogram, showing standardized Mantel statistics of distance classes between Bray-Curtis dissimilarity matrix from leaf and wood endophyte communities and distances among sampling sites. Positive Mantel’s r for a given set of comparisons indicates positive autocorrelation among a comparisons of this distance. Black filled circles indicate statistically significant correlations, after correction for multiple testing.

Supplemental Figure 2. All-host leaf dbMEM analysis. All vectors shown here correlate to some degree with spatial patterns found within leaf endophyte populations. See methods section for further details on interpretation of dbMEM vectors and their interpretation.

Supplemental Figure 3. All-host wood dbMEM analysis. All vectors shown here correlate to some degree with spatial patterns found within leaf endophyte populations. See methods section for further details on interpretation of dbMEM vectors and their interpretation.

Supplemental Figure 4. Co-occurrence networks of fungal endophytes and their tree hosts. Closest (first-degree) associations were used to select fungi as candidates for the core microbiome of a host.