HITCHHIKING IN THE CANOPY: ECOLOGICAL PATTERNS OF FOREST MYCOBIOMES

by

DANIEL C. THOMAS

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2017

DISSERTATION APPROVAL PAGE

Student: Daniel C. Thomas

Title: Hitchhiking in the Canopy: Ecological Patterns of Forest Mycobiomes

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Brendan Bohannan	Chair
Bitty Roy	Advisor
George Carroll	Core Member
Jeffrey Stone	Core Member
Lucas Silva	Institutional Representative

and

Sara D. Hodges	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded September 2017.

DISSERTATION ABSTRACT

Daniel C. Thomas

Doctor of Philosophy

Department of Biology

September 2017

Title: Hitchhiking in the Canopy: Ecological Patterns of Forest Mycobiomes

The fungal microbiome, or “mycobiome” of plants is diverse and important to host health, but the fluxes of fungi among plant hosts and with the surrounding environment are poorly understood. In chapter II, we employed sterile culture techniques and spatial sampling to examine leaves as possible vectors for transfer of their endophytic fungi from the canopy to substrate on the forest floor, as predicted by the Foraging Ascomycete hypothesis. Some foliar endophytic fungal species are also present as wood-decomposing fungi on the forest floor, that transfer of mycelium across these two life history stages can occur, that endophytic life history stages are buffered from environmental conditions in comparison to wood-decomposing fungi, and that spatial linkages between the two life history stages can be observed. In another study, described in chapter III, wood and leaf wood endophytes were sampled across a 25 ha plot, to explore landscape patterns of mycobiomes, and to explore the concept of a core microbiome in aerial plant tissues. We found that core microbiomes may be observed in a real ecological setting, but that the concept of core must be carefully defined and that some level of buffering from disturbance may be necessary to allow core microbiomes to assemble. In chapter IV, we return to examine some of the assumptions and implications of the Foraging Ascomycete hypothesis, with an agent-based model. We model the conditions under which dispersal through falling leaves may represent a fitness-enhancing dispersal strategy for fungi, and that deforestation as is currently underway throughout the world may have impacts on fungi that rely upon a canopy-inhabiting life stage for dispersal. In chapter V, some challenges associated with environmental

sampling of microbes using illumina© MiSeq sequences are critically examined. We find that biases introduced by random sampling at various stages of environmental DNA extraction and illumina© MiSeq sequencing are not well corrected by currently accepted bioinformatic algorithms. In addition, information loss from differential extraction, PCR amplification, and sequencing success, requires that users of MiSeq read libraries to interpret read abundances carefully.

This dissertation includes previously published and unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Daniel C. Thomas

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
The College of William and Mary, Williamsburg, Virginia

DEGREES AWARDED:

Bachelor of Science, 2003, Biology, William and Mary

AREAS OF SPECIAL INTEREST:

Plant-Fungal Interactions
Microbial Ecology

PROFESSIONAL EXPERIENCE:

Biological Science technician (Botany), various land management agencies, 2004-2012

Graduate teaching fellow, Department of Biology, University of Oregon, Eugene, 2012-2017

GRANTS, AWARDS, AND HONORS:

Backus Graduate Award, Mycological Society of America, 2017

Freeman Rowe Scholarship, Cascade Mycological Society, 2016

Graduate Award ,Sonoma County Mycological Society, 2016

Forest Fungal Ecology Research Award, Mycological Society of America, 2015

Grant Recipient for Canopy Studies, IDEA WILD, 2014

Scholarship, Oregon Mycological Society, 2013

Eastern Asia Pacific Summer Institute Fellowship, National Science Foundation 2013

William R. Sistrom Memorial Scholarship for Microbial Research, University of Oregon, 2013

Harvard Travellers Grant, Harvard Travellers Club, 2011.

Certificate of Merit, USDA Forest Service, 2010-2011

STAR Award, USGS Corvallis, 2009

PUBLICATIONS:

Thomas, D., Vandegrift, A., Ludden, A., Carroll, G.C., Roy, B.A. (2015). Spatial ecology of the fungal genus Xylaria in a tropical cloud forest. Biotropica, 48.3, 381-393.

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Bitty Roy, who was willing to take a gamble on me, who guided me, and tolerated my stubbornness. I thank Dr. George Carroll, whose knowledge seems boundless and whose love for fungi is contagious. I thank Dr. Roo Vandegrift for being a partner in crime on our mycological adventures and for showing me the ropes. I thank Dr. Brendan Bohannan and the microbe brews crew for the conversations that kept me thinking. I thank Graham Bailes and Tobias Policha for the making the office a place to laugh. I thank Dr. Jeff Stone for his skepticism, sarcasm, and encouragement. I thank my partner Eva, for her patience and grace with her moody mycologist.

My research was supported by numerous small grants from: The Mycological Society of America, the Harvard Travellers Club, the IDEA WILD foundation, and scholarships from the department of Biology from the University of Oregon. Our work in Taiwan was funded by an NSF EAPSI grant, NSF program 13-593.

DEDICATION

For Rosita.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

I chose for my dissertation to examine landscape patterns of endophytic fungi in forests. Prior to my entrance into graduate school I worked for ten years as a field biological technician, wherein I hiked many miles of backcountry conducting botanical surveys for governmental agencies. I came to graduate school with this landscape-level outlook of plant ecology, and with an interest in the ecological mysteries of fungi.

The study of endophytic fungi has a long and rich tradition of ecological study that fit well with my interests. Years before the current increased interest in plant microbiome research, fungal endophyte and epiphyte researchers were investigating microbiomes of plants (Rayner 1948, Bernstein 1977, Carroll 1978b). The new wave of laboratory and mathematical tools for studying microbes has given a surge to the once highly-specialized study of endophytes, placing it at the forefront of ecology, making this an exciting time to be asking questions about the role of fungi in forests.

Here I and mentors/colleagues employed two methods for investigating the ecology of forest endophytes: spatially explicit sampling of endophytes in forests at two scales (Chapters II and III), and creation of an agent-based model for one of our main working hypotheses on fungal ecology in chapter IV. Chapter V reflects on some problems associated with the common laboratory and bioinformatic techniques involved in environmental sequencing of fungi, and presents how one might deal with these challenges.

Introductory and conclusion chapters (chapters I and VI) were written solely by Daniel Thomas. Chapter II was co-authored with Roo Vandegrift, Bitty Roy, George Carroll, and Ashley Ludden. Chapter III was co-authored with Roo Vandegrift, Monica

Hsieh, Yu-Ming Ju, and Bitty Roy. Chapter IV was co-authored with Roo Vandegrift, George Carroll, and Bitty Roy. Chapter V was co-authored with Roo Vandegrift, Graham Bailes, and Bitty Roy

Bridge to Chapter II: Spatial Ecology of the Fungal Genus *Xylaria* in a Tropical Cloud Forest

The Foraging ascomycete hypothesis was put forth by George Carroll (1999). It proposes that some endophytic fungi may utilize leaves as vectors to augment dispersal. We tested this concept in the cloud forest of Northern Ecuador, using parallel sampling of the canopy and the forest floor for endophytes and sporocarps of the genus *Xylaria*. Fungi saturate the soil, the air, and the tissues of trees and organisms of the forest. Our work with the FA was an insight into one path of flux of microfungi across these boundaries, through the important and overlooked vectors of leaves. Work on the Foraging ascomycete hypothesis was extremely rewarding for me. It is the most material contribution to the natural history of fungi in which I have been a part.

CHAPTER II

SPATIAL ECOLOGY OF THE FUNGAL GENUS *XYLARIA* IN A TROPICAL CLOUD FOREST

From Thomas, D.*, Vandegrift, R.*, Ludden, A., Carroll, G. C., & Roy, B. A. 2016. Spatial ecology of the fungal genus *Xylaria* in a tropical cloud forest. Biotropica.
DOI: 10.1111/btp.12273
*D. Thomas and R. Vandegrift contributed equally to this work.

Contributions

D. THOMAS AND R. VANDEGRIFT CONTRIBUTED EQUALLY TO THIS WORK; both did field work and lab work, conceptual work, statistical work, and co-wrote the paper. A. Ludden did much of the molecular work. G. C. Carroll contributed to conceptual design and theoretical grounding. B. A. Roy contributed to design work, and contributed reagents/materials/analysis tools. All authors reviewed drafts of the paper.

Introduction

MUCH OF THE REASON FOR RECENT INTEREST IN THE PLANT MICROBIOME is economic, as awareness grows that the plant microbiome is vital to plant health (Carroll 1988, Berendsen *et al.* 2012, Berlec 2012, Chaparro *et al.* 2012), and may be important in mitigating effects of disease and climate change on human food plants (Köberl *et al.* 2011, Woodward *et al.* 2012). Fungal endophytes, an important component of the plant microbiome, are receiving particular attention (Porras-Alfaro & Bayman 2011, Jones 2013). Fungal endophytes are defined functionally, as those fungi found within living, healthy plant tissues; they make their living by not harming their host enough to induce a defensive

reaction (Clay 1990, Rudgers *et al.* 2009). Since their discovery, they have been found to be both ubiquitous and incredibly diverse in plants of all ecosystems (Arnold & Lutzoni 2007, Porras-Alfaro & Bayman 2011).

While numerous benefits to fitness for host-plant partners in the endophytic symbiosis have been observed, and many more proposed (Rodriguez *et al.* 2009), benefits for the fungal partners remain something more of a mystery. To date, the majority of endophyte research has been on temperate-zone clavicipitaceous endophytes of grasses, which often affect herbivory and host physiology, and are thus both ecologically and economically important (Clay & Schardl 2002, Schardl *et al.* 2004, Saikkonen *et al.* 2006). These fungi infect their hosts systematically and are passed on directly to their host-plant's offspring (Clay 1988). The fitness of these fungi increases with increased health and survival of their plant host. On the other hand, many non-clavicipitaceous fungal endophytes are very closely related to known plant pathogens (Carroll 1988, Freeman & Rodriguez 1993), and are well armed with energetically expensive arrays of enzymes for digestion of plant-tissues (Carroll & Petrini 1983, Schulz *et al.* 1999). Some endophytes have been observed to be latent pathogens or saprotrophs, waiting for host-plant weakness or death to be the first to colonize and digest host tissues (Chapela & Boddy 1988, Osono 2006, Promputtha *et al.* 2007, Promputtha *et al.* 2010), an obvious fitness benefit for the fungi involved.

However, many fungal endophytes neither vertically transmit to host-plant offspring, nor act as latent pathogens or saprotrophs of host tissues (Lodge 1997). The benefit of endophytism, if any, for these fungi remains unknown. Endophytism appears on the surface to be detrimental to fitness because these fungi undergo an extended period with reduced metabolic rate (Stone *et al.* 2004), and reduced or non-existent rates of sexual reproduction.

How then could the endophyte life-history strategy, which is observed in hundreds of species of fungi, and every major lineage of non-lichenized Pezizomycotina, possibly be adaptive? There are many potential benefits of endophytism to the fungal partner: the period of quiescence, or reduced metabolic rate (Stone *et al.* 2004), may

allow for persistence in the environment. The host plant potentially provides a stable carbon source, and the host may provide protection from environmental pressures such as desiccation (Chaves *et al.* 2002) and harmful UV radiation (Krauss *et al.* 1997). Endophytism may also play a role in dispersal, as we examine here

Much discussion has taken place in recent years over questions of microbial dispersal (Green *et al.* 2004, Green & Bohannan 2006, Martiny *et al.* 2006, Hanson *et al.* 2012). Dispersal is defined as any transport of propagules, individuals, or gametes that creates gene flow within or between populations (Ronce 2007, Clobert *et al.* 2012). Historically, microorganisms were thought to be functionally unlimited in their ability to disperse over the planet (Becking 1934, Fenchel & Finlay 2004). Despite this, many recent studies of microbes have uncovered evidence for dispersal limitation, or the inability of a strain or species to access and successfully establish itself in otherwise suitable habitat (Roy 2001, Telford *et al.* 2006, Green & Bohannan 2006, Grubisha et al. 2007, Peay *et al.* 2010, Galante *et al.* 2011). Such dispersal limitation may function to constrain the geographic ranges of some species, or the range of gene flow within or between local populations of a given species; indeed, such constraints on gene flow between populations are theorized as a major driver of speciation over evolutionary time scales (Clobert *et al.* 2012). There is evidence that at least some decomposer fungi are dispersal limited, even at local scales (Norros *et al.* 2012). Dispersal limitation may reduce fitness of an organism relative to competitors (Hurtt & Pacala 1995), suggesting that fungi may be under selective pressure to increase dispersal at both local and regional scales.

Dispersal involves successful transport and successful establishment of propagules (Hanson *et al.* 2012, Peay *et al.* 2012, Clobert *et al.* 2012). An endophytic life stage may enhance both of these processes: senescent leaves fall farther than the vast majority of spores are predicted to travel unassisted (Roper *et al.* 2010, Galante *et al.* 2011), carrying with them mycelium, avoiding the uncertainty inherent in the germination phase of growth from spores. In evergreen forests, leaves generally fall asynchronously, which provides low propagule density over relatively long periods of

time (in tropical cloud forests, leaves live 12 mo to >5 years; Bruijnzeel & Veneklaas 1998, Reich *et al.* 1991), in contrast to spore dispersal from a fruiting body, which provides high propagule density over relatively short periods of time (<1 year; Rogers 1979, Whalley 1996). Leaves may enhance colonization rates, by creating a sheltered microclimate favorable to inoculation. Additionally, living leaves may provide refugia for endophytic fungi, where fungi can wait out difficult conditions at low metabolic cost, benefiting from the protection afforded by the leaf tissue (Stone 1987, Schulz & Boyle 2005). The idea of endophytism as a secondary life-history strategy for decomposer fungi to span (i.e., disperse across) scarcity of primary substrates and challenging environmental conditions in both time and space is known as the Foraging Ascomycete (FA) hypothesis (Carroll 1999).

Here, we attempt to critically examine the FA hypothesis in a cloud forest ecosystem, using the genus *Xylaria* Hill ex Schrank (Xylariaceae, Ascomycota) as an example of typical endophytic fungi that may utilize a FA strategy (Fig. 2.1). Members of this genus are important saprotrophs, found primarily on decomposing dead wood—and, rarely, on leaves and fruits—on the forest floor (Whalley 1996, Lodge 1997, Rogers 2000). *Xylaria* are visible during sexual sporulation, forming relatively large, macroscopic stromata, or "fruiting" structures (Bayman *et al.* 1998, Davis & Shaw 2008). *Xylaria* are common in virtually every study that has ever been done on endophytes, especially in tropical ecosystems (see Davis *et al.* (2003) for an extensive list). We focus here on a common endophyte genus to avoid the problem of being swamped in the overwhelming diversity of fungal endophytes in the tropics (Arnold *et al.* 2000, Arnold & Lutzoni 2007). These two life stages in *Xylaria,* leaf endophyte and wood decomposer, have been observed within single, tightly defined clades (Okane *et al.* 2008). Additionally, *Xylaria* grow readily in culture, making them ideally suited for study in laboratory conditions. (Whalley 1996, Bayman *et al.* 1998).

Following the FA hypothesis, we hypothesized that (1) distributions of wood-decomposing *Xylaria* should be spatially coupled to the distributions of those same *Xylaria* in the endophytic life stage. To test this hypothesis, we used a spatially explicit

sampling scheme: we looked for spatial clustering not attributable to environmental gradients or biotic interactions, but indicative of dispersal linkage between life stages. This is in opposition to Beckman's hypothesis that microbes are unlimited in their dispersal abilities; if this is the case, *Xylaria* in both life stages should be distributed randomly and independent of each other, save for the selective impacts of the environment. Additionally, (2) if *Xylaria* endophytes display a FA life history strategy, we would expect endophytic host generalism in the tropics, as host selectivity would interfere with dispersal in systems where most available hosts are present in low densities (May 1991). The FA hypothesis also leads to the hypothesis (3) that endophytes will be released from environmental constraints relative to their corresponding decomposers. Using ITS rDNA barcode sequence comparisons (Gardes & Bruns 1993, Schoch *et al.* 2012), we matched decomposer *Xylaria* with endophytes in leaves from the canopy, and compared habitat characteristics of both. Lastly, we expect (4) the FA strategy to be a specialized survival/dispersal mechanism utilized by a subset of fungi within the genus *Xylaria*. Given the diversity of the genus, we expect variation in species' niches to modulate the selective advantage of endophytism.

Methods

FIELD.—All field work described was performed at Reserva Los Cedros, a private, protected forest preserve in the western slope of the Andes, in northwestern Ecuador (00°18′31.0″ N, 78°46′44.6″ W), at 1200 m asl. The reserve lies within the Andean Chocó bioregion, one of the most biodiverse habitats on the planet (Gentry 1992). The reserve protects approximately 6800 hectares of forest, approximately 80 percent of which is primary, premontane tropical wet and cloud forest. The Reserve also shares a border with the 305,000 hectare government-protected Cotocachi-Cayapas Ecological Reserve. Rainfall averages 2,903±186 mm per year (Policha 2014). Humidity is typically high (~100%), and daily temperatures at the site range from 15ºC to 25ºC (Policha 2014).

Seasonal variation in climate is minimal. Our sampling occurred during the early part of the wettest season, in January 2012, when fungal fruiting was presumed to be highest.

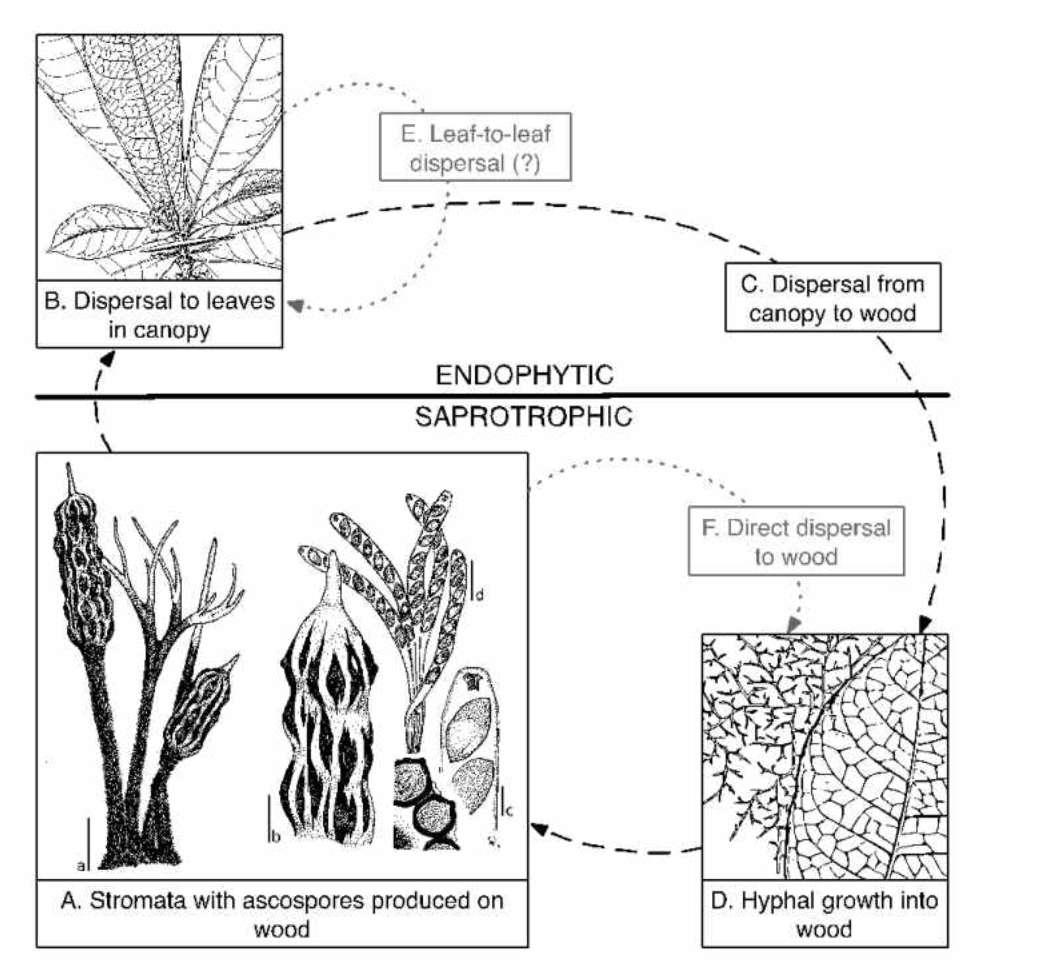


FIGURE 2.1. Schematic of proposed *Xylaria* life cycle, illustrating the Foraging Ascomycete hypothesis. Stromata (A) are produced on suitable substrate (generally wood); *Xylaria apiculata* Cooke, one of five *Xylaria* species present in both endophytic and decomposer life stages in this study, is illustrated as a typical example of the genus (scale bars: a = 2 mm; b = 1 mm (including stromatal section); c = 10 µm; d = 50 µm). The fungus disperses into the canopy (B) where it initiates endophyte infection; we presume ascospores to be the predominant mechanism of dispersal. When leaves are shed from the canopy (C), they take their endophytes with them; entire leaves may become dispersal vectors. The fungus grows from shed leaves into suitable substrate (D), and the cycle continues. Not explicitly considered in this study are other potential courses of dispersal (in grey): there may be leaf-to-leaf dispersal in the canopy (E), which would maintain endophyte infection even in the absence of sexual reproduction on the forest floor. We find no evidence for this in the literature, however, and expect it to be rare or non-existent. Direct dispersal of ascospores to suitable substrate (F) is undoubtedly a

common means of dispersal in this genus. While an interesting and important mechanism, we do not explicitly examine direct dispersal; this study focuses on elucidated the role of endophytism in the dispersal ecology of *Xylaria*. Panel B re-drawn from J. Seboth (1881).

We sampled within a previously established, "permanent" tree monitoring plot (Peck *et al.* 2008). Sampling occurred in primary forest at 1300 m, on the banks of a perennial stream and the surrounding area. The sampling area consisted of 120 individual points, spaced 10 m apart in the east-west direction and 5 m apart in the north-south direction. At each point, the two lowest leaves of the nearest tree or tree-like plant were collected for culturing of endophytes, as well as additional material for host-identification, if necessary. All xylarioid stromata within a 1.2 m radius of the point were collected from the forest floor and any aerial substrate within reach.

Previous environmental data for the plot were inaccessible, so stream mapping and individual point data were recollected later, in March 2014. Site characteristics in the plot are expected to change slowly (Policha, 2014). Slope by clinometer, canopy cover by densitometer, and aspect were measured for each point. Our sampling area was small (~0.5 ha) and is presumed to be homogeneous in soil quality and precipitation regime (Policha, 2014).

SAMPLE PROCESSING.—Leaves were washed gently in a basin of water (~30 s) to remove epiphyllous debris. Endophytes were recovered from two 2-mm diameter discs taken from each leaf using a Harris® micropunch sampling tool, for a total of 480 individual leaf discs. Discs were surface sterilized by immersion in 70 percent ethanol for 1 min, 5 percent sodium hypochlorite for 2 min, then rinsed thoroughly in sterile water and placed on water agar (2% agar) petri dishes. Fungi were individually isolated onto MEA plates (2% malt extract, 0.1% yeast extract, acidified to pH 4) as they grew out from the discs of leaf tissue (methods modified from Okane *et al.* 2008). Water agar plates with leaf discs were examined daily for a period of 9 weeks, with new isolations made as needed.

All culture work was done in a portable sterile laminar flow hood constructed using a Dayton® Blower (model MG1104058171010), $^1/_4$ inch Plexiglass®, and a Hepa-sep® filter (model STD12-12-05PEADC50). Power was supplied by a micro-hydrological power plant installed at Reserva Los Cedros.

Cultures were grown on MEA until sufficient hyphae were present for DNA extraction. Under laminar flow, all aerial mycelium were harvested, and then pressed into a Whatman FTA® card with the aid of a standard claw hammer (Dentinger *et al.* 2010). Stromata were sampled by removing outer carbonaceous layers using a flame-sterilized scalpel, and preservation of inner tissues in Whatman FTA® cards.

ENDOPHYTE TRANSFER EXPERIMENT.—In April 2014, we also collected leaves from a randomly selected tree within the plot (*Nectandra lineatifolia* (Ruiz & Pav.) Mez) for an experiment to examine the transmission of endophytic Xylariaceae to woody substrates. Eight 2-cm sections were cut from each of twelve leaves, surface sterilized as described above, and placed on sterile (twice-autoclaved) white birch tongue depressors (Puritan, Guilford, Maine, USA) as a standardized angiospermous woody substrate. Four sections from the same leaf were placed on each tongue depressor. These were incubated at room temperature in EtOH-sterilized Ziploc storage boxes (with an open container of sterilized water to maintain humidity) at the field station for 6 weeks, after which time the leaf segments were removed, the tongue depressors were air-dried in open, downward-facing, sterile plastic zipper bags, in which they were then transported back to the United States.

In our lab in Oregon, we started initial cultures from the first three tongue depressors in early August, 2014. We split each tongue depressor into three pieces lengthwise and extracted the middle piece; this was split into 12 equal pieces (~4 mm^2 each), each of which was plated onto water agar for fungal isolation, and incubated indefinitely. Subcultures were made on MEA as described above; cultures were identified to genus by a combination of morphology and DNA sequence.

DNA EXTRACTION.—Lab protocols followed Dentinger *et al.* (2010). Samples were excised from the Whatman FTA cards using a 2 mm punch tool and sterilized cutting mat. The punch tool was flame sterilized between uses, and its sterility was confirmed with extraction and PCR tests of DNA from sterile filter paper segments cut by the tool between each use.

Sigma Extract-N-Amp™ Plant PCR Kit reagents were used for extraction from Whatman© FTA cards. With each sample disc, 25 µL of Extraction reagent was added to each well and incubated for 10 minutes at 95˚C (using an Applied Biosystems© Vereti© model thermal cycler). After incubation, 25 µL of Dilution reagent was added to halt further extraction.

PCR AMPLIFICATION.—Template DNA was diluted. Generally, 1:19 dilutions worked best, though optimal dilution ranged from 1:1-1:99. DNA amplification was carried out using the fungal-specific ITS1F (5'-CTTGGTCATTTAGAGGAAGTAA-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC3-') primer sets (Gardes & Bruns 1993). DNA amplification was conducted in a standard 96-well plate with 10-uL reaction volumes (2 µL of template, 5 µL of Sigma Aldrich Jumpstart™ Taq Readymix™, 2.2 µL sterile water, 0.4 µL 25 mM $MgCl_2$, and 0.2 µL of each primer).

PCR amplification was done with an Applied Biosystems© Vereti© model thermal cycler with the following parameters: initial denaturation at 95˚C for 2 min, five cycles of denaturation at 95˚C for 30 s, annealing at 60˚C for 30 s, and extension at 72˚C for 1 min; followed by 25 cycles of denaturation of 95˚C for 30 s, annealing at 55˚C for 30 s, and extension at 72˚C for 1 min; a final extension at 72˚C for 10 min and a final step of indefinite duration at 4˚C.

DNA SEQUENCING AND ANALYSIS.—PCR products were visualized on a 1 percent agarose gel. Before sequencing, all successful PCR reactions were cleaned by adding 0.4 volumes of a master mix containing 10 percent FastAP© thermosensitive shrimp alkaline phosphatase (Thermo Scientific©) and 1 percent exonuclease I solution (New England

Biolabs©) to the PCR product, and incubation for 15 min at 37˚C followed by 15 min at 85˚C. Samples were then frozen until shipping for sequencing at Functional Biosciences, Inc (Madison, WI, U.S.A.) on ABI 3730xl instruments using Big Dye V3.1.

Forward and reverse sequences were aligned and curated in Geneious v6.0.3 (Biomatters, Auckland, New Zealand). Sequences were grouped into 97 percent similarity clusters using UClust as implemented in MacQIIME v1.7.0 with default settings. Specimens were identified morphologically with the help of Dr. Yu-Ming Ju (Academia Sinica, Taipei, Taiwan, ROC), and sequences were named via confirmed morphological identification wherever possible. In nearly all cases, 97 percent was an adequate cut-off to delineate previously defined morphological species. In one case two species occurred within a grouping (*Xylaria schweinitzii* Berk. & M.A. Curtis and *Xylaria ophiopoda* Sacc.). A maximum likelihood tree was constructed using the PhyML plugin in Geneious, and the two major branches of the tree corresponded perfectly to the two morphological species. Species groupings were adjusted to accommodate splitting that cluster. Some *Xylaria* species were unable to be identified morphologically due to immaturity or poor condition of specimens. When not in a cluster with identifiable specimens, these were assigned a species identifier, but no name. Finally, a species occurrence matrix was built for all species of *Xylaria*, both endophytes and decomposers.

STATISTICAL METHODS.—Data were analyzed using R Statistical Software, version 3.1.0 (R Core Team 2014), including the *sp* (Pebesma & Bivand 2005), *bipartite* (Dormann *et al.* 2008), and *vegan* packages (Oksanen *et al.* 2013). All scripts are publicly available online (Thomas *et al.* 2014).

Estimates of xylariaceous species richness within our plot were estimated using Chao2 and Jacknife1 estimators (Burnham & Overton 1978, Chao 1984, Colwell & Coddington 1994). Sampling effort was visualized with species accumulation curves constructed using the vegan package in R.

Spatial clustering of endophyte and decomposer *Xylaria* life stages of each species was analyzed using nearest neighbor analysis (Clark & Evans 1954) with

randomization (Fortin & Dale 2005), using a customized script in R (Thomas *et al.* 2014). Four spatial relationships were examined: clustering of (1) stromata around stromata, (2) endophytes around endophytes, (3) endophytes around stromata, and (4) stromata around endophytes. For some taxa, not all stages were present; in these cases the subset of possible comparisons was performed.

Spatial clustering of fungal observations around a stream that dissected the plot were also analyzed using custom scripts in R. When all life stages were present, combined life stages (all fungi), stromata alone, and endophytes alone were examined.

The nearest-neighbor with randomization statistic we employ here is not often utilized in ecology (but see Clark and Evans (1954) and Dixon (1994) for related usages). In each case, a nearest neighbor test statistic was generated using the average of distances of up to five (Liu 2001) nearest neighbor observations from each point, for all observations of a species and life stage. A test-statistic distribution was generated for each species using 20,000 randomly generated sampling areas with the same number of both endophytes and decomposer fungi as the actual sampling area. In each rank of nearest neighbor, or "distance class", the observed mean nearest neighbor distance ($\bar{d}_o$) was compared to the randomly generated distribution of expected mean nearest neighbor distances ($\bar{d}_e$), and the proportion of $\bar{d}_e$ values lower than the observed were taken as the probability that a given species was spatially under-dispersed significantly more than as predicted by a completely spatially random null model (i.e., that the distance between points is less than that expected by chance; this is often called "clustering" or "clumping"). P-values were considered significant at $P = 0.05$ or below; all nearest neighbor distances are reported in meters.

If the real distance to the nearest neighbor is less than the randomly generated distance to the nearest neighbor more than 95 percent of the time ($P < 0.05$), we take this to mean that the points are significantly clustered. In other words, it is more likely that observations of these species will occur in proximity to other observations of the same species than expected by chance. In the absence of environmental gradients controlling this spatial structuring within a life stage, we take this as evidence of spatial dependence:

in the case of life stages clustering to themselves (e.g., endophytes around endophytes), this is likely a signal of "true" or "autogenic" autocorrelation (Fortin & Dale 2005), or the tendency of neutral processes to cause organisms to cluster in space and time. In the case of different life stages clustering together (e.g., endophytes around decomposers), we take this as evidence that dispersal is occurring between these different life stages. Tests for spatial correlation ("autocorrelation") of environmental variables were conducted using a Mantel correlogram of environmental dissimilarity of plots against a physical distance matrix. Testing for community turnover, or decay of similarity in *Xylaria* species composition among plots with distance, was done using a Mantel correlogram of *Xylaria* species composition distance matrix against a physical distance matrix of all plots sampled (Fortin & Dale 2005).

In addition to determining if clusters are non-random, the nearest-neighbor metric we employ here allows us to examine the direction of clustering *between* life stages—that is, we can compare the distance to nearest stromata from an endophyte, or vice versa. When determining whether there is clustering of the two life stages, two *P*-values are obtained: one for stromata clustering around endophytes, and one for endophytes clustering around stromata.

We examined host-preference by endophytes using two methods: (1) we used chi-squared goodness-of-fit tests of **host preference** by our most common *Xylaria* endophyte (*X. adscendens*) and **endophyte preference** in the most common host tree (*Faramea* aff. *oblongifolia* Standl.); (2) we used bipartite network analysis to examine strength of interactions between host-plants and endophytes.

In the goodness-of-fit analysis of host-preference, the null hypothesis was that infection depended only on host commonness, and was generated from the respective ratios of species of all host trees from our plot that were found to host any xylariaceous endophyte. Reciprocally, the null hypothesis for endophyte preference was that the most common host tree would be infected by xylariaceous endophytes in roughly the same frequency that these endophytes were collected from all hosts in the plot. These hypothesized ratios were then compared to the observed ratios of host trees from which

Xylaria adscendens (Fr.) Fr. was isolated and the frequencies of endophyte species observed solely in *Faramea* aff. *oblogifolia*, using a chi-squared goodness-of-fit test with Monte Carlo simulation (from the base R *stats* package).

Network analysis followed Ikeda *et al.* (2014). Using the *bipartite* package in R (Dormann 2008) species interaction matrices were constructed and a network-wide H_2' value (Blüthgen 2006) was calculated to characterize the level of preference ("specialization") among host-plants and endophytes. These results were then compared to a null model of network assembly (Vásquez *et al.* 2007), with 10000 randomization cycles.

Tests for grouping of species by habitat characteristics—slope, canopy, distance-to-water, and aspect (separated into component northern and eastern exposures)—were done using Permutational Multiple Analysis of Variance (PerMANOVA), with the *adonis* function in *vegan* package in R. These data were visualized with non-metric multidimensional scaling (NMDS). Differences among the above characteristics for all sites containing a *Xylaria* observation were summarized in an environmental distance matrix as input for the *metaMDS* function in the *vegan* package in R (Oksanen *et al.* 2013); points were then categorized by the species of *Xylaria* observed. The *metaMDS* considers multiple possible solutions using Procrustes analysis and employs Wisconsin double standardization to reduce Kruskal stress in ordination. We considered solutions with stresses below 0.15 to be informative. Linear models of differences in habitat, used for weighting relative importance of habitat variables, were also constructed using the *adonis* function.

Results

Endophytes were isolated from 38 tree species in 19 different families, as well as a species of large fern and several large herbaceous plants when no woody hosts were present within the sampling plot. From the 480 total leaf segments, 720 unique cultures

were isolated; no leaf segment yielded zero fungi. Of the endophyte isolates, 104 (14.4%) were in the Xylariaceae (19 species in *Xylaria, Hypoxylon, Nemania,* and *Annulohypoxylon*). We collected stromata in two genera of Xylariaceae, *Xylaria* and *Kretzschmaria,* from 79 (65.8%) of the points within the plot. We found 36 species of *Xylaria*, 31 of which were found to only occur as fruiting bodies, and five of which were found as both stromata and endophytes. All five species of *Xylaria* found as endophytes were also found as fruiting bodies; there were no endophytic *Xylaria* not also recovered as stromata (Table 2.1). *Xylaria* leaf endophyte species were found to be a subset of wood decomposer species: all *Xylaria* endophyte species were also recovered as decomposer species. There were species-specific differences in the frequencies of occurrence of the leaf endophyte and decomposer (stromatal) life stages (Table 2.1): that is, frequency of one life stage does not predict frequency of the other; they are specific to particular species.

Chao2 and Jackknife1 species richness estimators predicted 52.33 (SE = 11.7) and 49.9 (SE = 4.2) *Xylaria* decomposer (stromatal) species, and 5.00 (SE = 0) and 8.0 (SE = 1.7) *Xylaria* endophyte species. This is in agreement with species accumulation curves of our sampling effort indicating that we sampled nearly completely for culturable endophyte species but that decomposer species remain to be discovered within the plot (Fig. 2.2).

Five species of *Xylaria* were found both in the leaves and as decomposers. Of these, two species demonstrated non-random clumping of differing life stages (i.e., endophyte-stage fungi were found to clump around decomposer-stage fungi, or *vice versa*): *X.* aff. *curta* ($\bar{d}_o$ (1) = 18.10, $\bar{d}_e$ (1) = 43.90 ± 17.48, P = 0.048) and *X. fissilis_1* ($\bar{d}_o$ (2) = 13.83, $\bar{d}_e$ (2) = 19.94 ± 3.84, P = 0.036) (Table 2.2; Fig. 2.3). For these five species, significant clumping within a life stage was only observed for endophytic *X. adscendens* ($\bar{d}_o$ (2) = 11.91, $\bar{d}_e$ (2) = 13.43 ± 0.89, P = 0.044).

Of the five *Xylaria* species exhibiting both decomposer and endophytic life stages, three species in the decomposer life stage appear to be closely clustering around the stream present in our sampling area (Fig. 2.4): *X.* aff. *curta* ($\bar{d}_o$ (2) = 29.67, $\bar{d}_e$ (2) =

52.75 ± 12.81, $P = 0.016$), *X. atrosphaerica* ($\bar{d}_o$ (2,3) = 21.11 and 33.17, $\bar{d}_e$ (2,3) = 34.23 ± 8.01 and 47.30 ± 9.91, $P = 0.007$ and 0.048), and *X. apiculata* ($\bar{d}_o$ (1,2,3) = 6.85, 12.74, 18.08, $\bar{d}_e$ (1,2,3) = 13.49 ± 3.50, 21.62 ± 3.99, 28.17 ± 4.54, $P = 0.006$, 0.002, 0.001). None of the species in the endophytic life stage were clustered around water.

Among the 36 species of *Xylaria* detected as decomposers, significant clustering of stromata to stromata was observed in two species (*X. multiplex* and *X. ophiopoda*). Significant clustering of stromata around streams was observed in eight species (*X.* aff. *curta, X. cuneata, X. apiculata_1, X. subtorulosa, X. multiplex, X.* sp. *13, X. enterogena,* and *X. atrosphaerica*).

Spatial correlation of environmental variables was significant only at distances below 15 m, and variance explained was extremely low (Mantel's r = 0.06, $R^2 = 0.004$, $P < 0.05$). *Xylaria* species composition was not found to be significantly autocorrelated on the scale of this study (Mantel's r = 0.01, $R^2 < 0.001$, $P = 0.394$).

TABLE 2.1. List of all *Xylaria* species recovered and the number of points in the study area (out of 120) from which each species was recovered in each life stage. Distinct ITS clusters in otherwise indistinguishable taxa are indicated by an underscore followed by a clade number on the specific epithet.

Taxa	Points with **Stromata**	Points with **Endophytes**
Xylaria adscendens (Fr.) Fr.	3	26
Xylaria anisopleura (Mont.) Fr.	3	
Xylaria apiculata_1 Cooke	9	1
Xylaria apiculata_2 Cooke	1	
Xylaria atrosphaerica (Cooke & Mass.) Callan & Rogers	4	1
Xylaria aff. *comosa* (Mont.) Fr.	5	
Xylaria cristata Speg.	1	
Xylaria cuneata Lloyd	4	
Xylaria curta_1 Fr.	1	
Xylaria curta_2 Fr.	1	

TABLE 2.1. (continued)		
Xylaria aff. *curta* Fr.	2	1
Xylaria fissilis_1 Ces.	11	5
Xylaria enterogena Mont.	11	
Xylaria fissilis_2 Ces.	2	
Xylaria globosa (Pers.) Mont.	5	
Xylaria meliacearum Læssøe	3	
Xylaria multiplex (Kunze) Fr.	3	
Xylaria ophiopoda Sacc.	5	
Xylaria schweinitzii Berk. & M.A. Curtis	16	
Xylaria scruposa_1 (Fr.) Fr.	12	
Xylaria scruposa_2 (Fr.) Fr.	4	
Xylaria subtorulosa Speg.	2	
Xylaria telfairii (Berk.) Sacc.	7	
Xylaria xanthinovelutina (Mont.) Fr.	2	
Xylaria sp. 01	1	
Xylaria sp. 02	1	
Xylaria sp. 03	1	
Xylaria sp. 05	1	
Xylaria sp. 06	1	
Xylaria sp. 07	1	
Xylaria sp. 08	1	
Xylaria sp. 10	1	
Xylaria sp. 11	1	
Xylaria sp. 12	2	
Xylaria sp. 13	2	

TABLE 2.1. (continued)		
Xylaria sp. nov. 2	1	

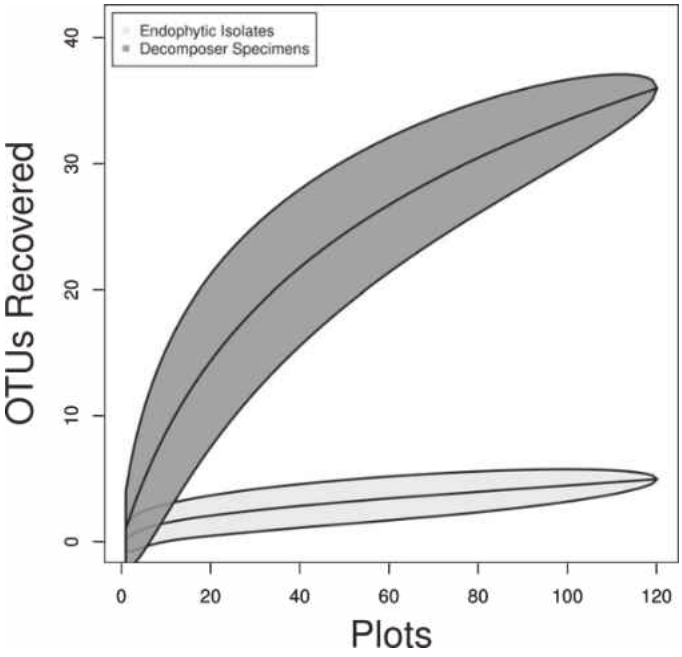


FIGURE 2.2. Species accumulation/sampling effort curve of both decomposer stromata collected on the forest floor and endophytes cultured from leaves; shaded areas are 95% confidence intervals.

Habitat preferences were not found to be significantly different among the five *Xylaria* species when we examined combined life stages (PerMANOVA, $F_{4, 58} = 1.57$, $R^2 = 0.10$, $P = 0.112$). However, when examined separately, decomposer *Xylari*a may show species-specific habitat preferences (PerMANOVA, $F_{4, 24} = 1.84$, $R^2 = 0.23$, $P = 0.07$); endophytic *Xylaria* do not ($F_{4, 29} = 0.45$, $R^2 = 0.06$, $P = 0.94$). In decomposer fungi, differences among habitats were defined most strongly by proximity to water (PerMANOVA, $F_{1, 23} = 112.42$, $R^2 = 0.44$, $P = 0.001$), followed by slope ($F_{1, 23} = 31.36$, $R^2 = 0.12$, $P = 0.001$), canopy cover ($F_{1, 23} = 20.61$, $R^2 = 0.08$, $P = 0.001$), and aspect, in its components of northern and eastern exposure ($F_{1, 23} = 11.84$, $R^2 = 0.05$, $P = 0.001$ and $F_{1, 23} = 6.20$, $R^2 = 0.02$, $P = 0.006$, respectively).

TABLE 2.2. Nearest-Neighbor analysis of spatial clusters in five species of *Xylaria*. Values shown are the observed mean nearest neighbor distance ($\bar{d}_o$), the expected mean nearest neighbor distance ($\bar{d}_e$) from a Monte Carlo simulation null model assuming complete spatial randomness (CSR), the standard deviation around the expected mean nearest neighbor distance (s_e), and the *P* values, calculated as the proportion of simulations where $\bar{d}_e < \bar{d}_o$. Bold indicates $P < 0.05$; italics indicate $0.05 < P < 0.10$; dashes indicate insufficient sample size to conduct the analysis at a given neighbor class.

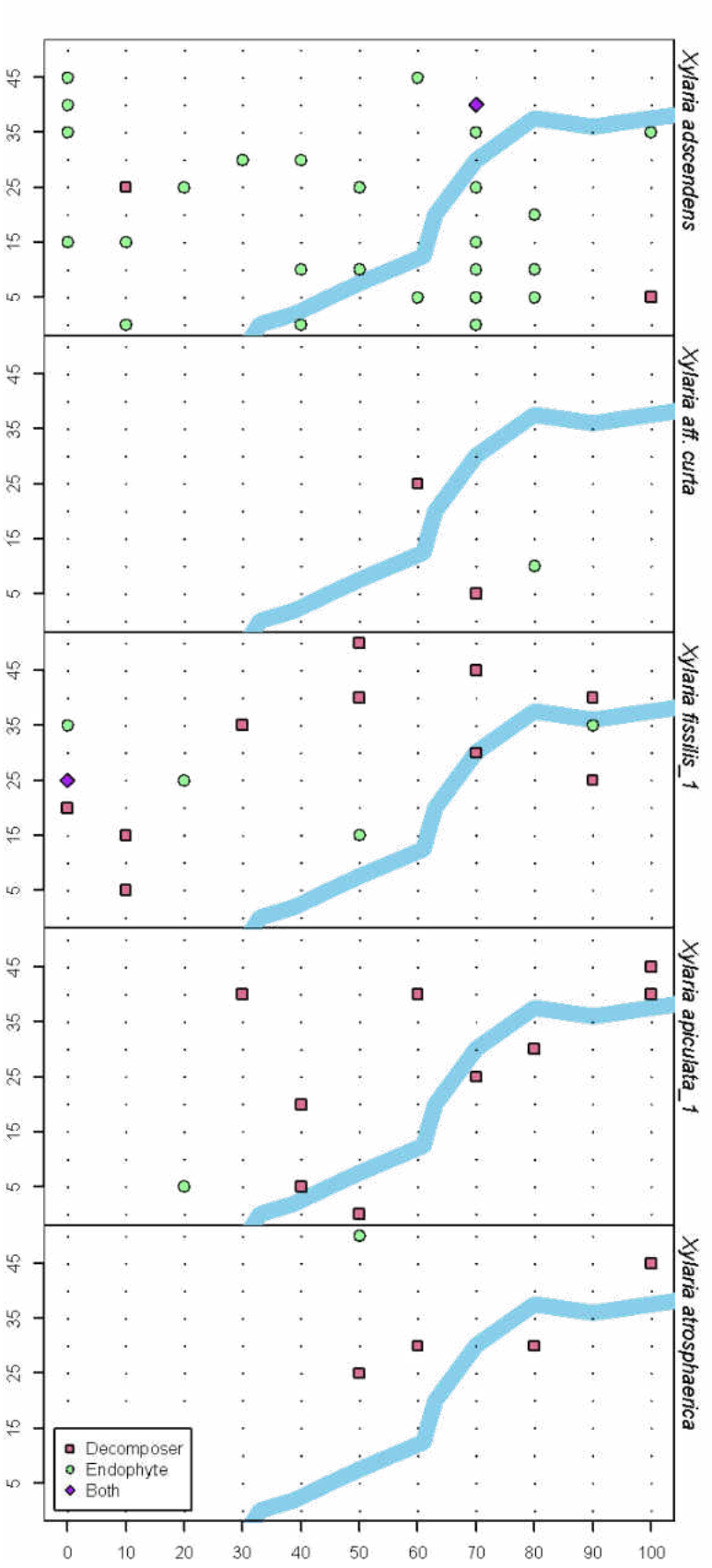
		Stromata around **Endophytes**					**Endophytes** around **Stromata**			
Taxa	Neighb or class	$\bar{d}_o$	$\bar{d}_e$	s_e	*P*		$\bar{d}_o$	$\bar{d}_e$	s_e	*P*
Xylaria aff. *curta*	1	11.2	31.5	18.2	0.152		18.1	43.9	17.5	**0.048**
	2	25.0	56.3	21.5	*0.068*		—	—	—	—
Xylaria apiculata_1	1	20.0	13.9	8.5	0.794		51.2	43.7	10.5	0.771
	2	25.0	22.3	9.3	0.705		—	—	—	—
	3	30.4	29.4	10.4	0.612		—	—	—	—
	4	36.4	36.0	11.6	0.609		—	—	—	—
	5	53.2	42.8	13.0	0.819		—	—	—	—
Xylaria fissilis_1	1	10.8	12.3	3.5	0.354		15.8	19.1	4.3	0.210
	2	13.8	19.9	3.8	**0.036**		28.4	31.5	5.3	0.285
	3	21.1	26.1	4.3	0.108		41.2	43.0	6.0	0.406
	4	26.1	31.7	4.8	0.102		58.0	55.5	8.3	0.634
	5	32.4	37.1	5.3	0.178		76.6	69.6	8.8	0.771

Table 2.2. (continued).

Xylaria adscendens	1	22.1	25.3	5.5	0.302		10.0	7.2	3.0	0.836
	2	46.2	43.4	6.2	0.715		11.9	12.4	2.7	0.494
	3	72.9	62.7	9.6	0.839		16.8	15.9	3.1	0.653
	4	—	—	—	—		19.7	19.0	3.3	0.627
	5	—	—	—	—		23.5	21.9	3.5	0.703
Xylaria atrosphaerica	1	22.4	21.8	12.9	0.639		33.4	43.9	13.4	0.230
	2	25.0	36.3	15.5	0.277		—	—	—	—
	3	36.1	50.6	17.8	0.236		—	—	—	—
	4	50.3	66.8	18.8	0.221		—	—	—	—

We found no evidence for host preference by endophytes from the family Xylariaceae. Host trees for the most common endophyte, *Xylaria adscendens,* did not vary from general abundances of host trees within the total plot (χ^2, 10000 replicates, (N = 10) = 2.45, P = 0.74*).* Relative abundances of endophytes recovered from the most common host, *Faramea* aff. *oblongifolia,* did not show a significant difference in endophyte abundances within the entire plot*,* (χ^2, 10000 replicates, (N = 26) = 19.80, P = 0.86). Network specialization did not exceed levels expected by chance alone given abundances of endophytes and host-plants (H_2' = 0.261, mean randomized H_2' = 0.290, 10000 cycles, P = 0.62; Fig. 2.5).

FIGURE 2.3 (next page). Maps of the five species of *Xylaria* displaying both endophyte and decomposer life stages. All collection points are marked; the stream is indicated with a blue line. Scale in meters.


Xylaria adscendens
Xylaria aff. curta
Xylaria fissilis_1
Xylaria apiculata_1
Xylaria atrosphaerica
Decomposer
Endophyte
Both
0
10
20
30
40
50
60
70
80
90
100
5
15
25
35
45

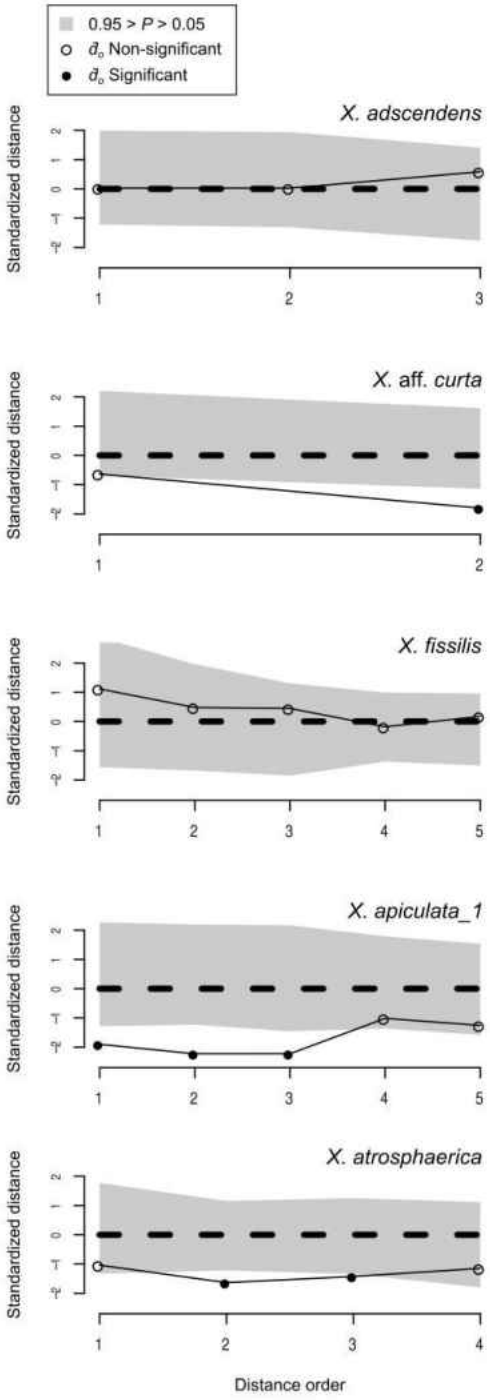


FIGURE 2.4. Graphical representation of the results of nearest neighbor Monte Carlo-type simulations for clustering of stromata around the stream. For each species, the standardized mean distance to nearest neighboring point along the stream ($\bar{d}_o$) for all available distance classes is plotted. The dashed line represents the mean distance to points along the stream of the permutations on complete spatial randomness ($\bar{d}_e$), standardized to zero for all distance classes; the y-axis units represent deviation from the permutational mean (s_e). Open points are non-significant ($P > 0.05$); closed points are significant ($P < 0.05$); the grey area represents the region where $0.95 > P > 0.05$.

ENDOPHYTE TRANSFER EXPERIMENT.—We isolated *Xylaria* from 8 of 12 segments from one of three sampled tongue depressors (22% of segments). By the sixth month, the *Xylaria* had established competitive dominance in these tongue depressor segments, and was observed to initiate fruiting in 7 of the 8 segments from which it was isolated; all stromatal primordia displayed classic *Nodulisporium* anamorphs. Unfortunately, we have not been able to obtain usable ITS sequence for these isolates, presumably due to co-extraction of PCR inhibiting fungal cell wall polysaccharides.

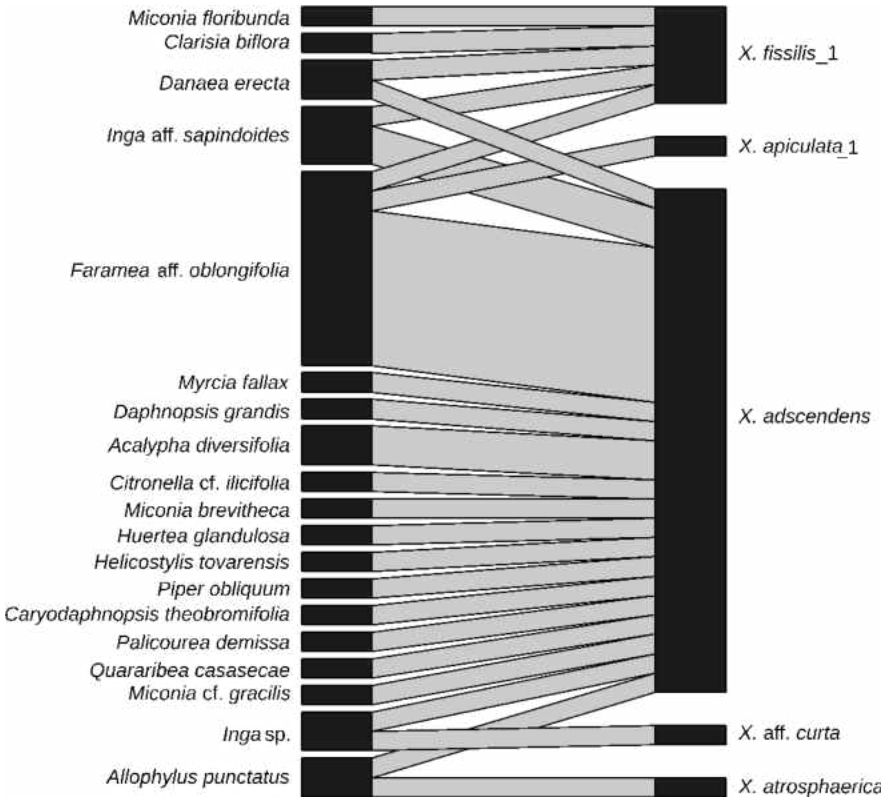


FIGURE 2.5. Bipartite network visualization of *Xylaria* endophytes (right) and plant-hosts (left). Widths of links are scaled to number of points at which endophytes were isolated from hosts.

Discussion

The Foraging Ascomycete hypothesis challenges two classical assumptions about fungal dispersal: first, that fungi are unlimited in their dispersal abilities (Becking 1934, Fenchel & Finlay 2004), and second, that sexual spores of decomposers are the sole major source of dispersal for these fungi (Norros et al. 2012, Bayman et al. 1998, Malloch & Blackwell 1992). Following these assumptions, endophytism has been supposed by some to be an accidental "dead end" infection of living plants (Bayman *et al.* 1998). The FA hypothesis proposes that for some fungi, endophytism is not an accidental "dead end", but an important mechanism of fungal dispersal—an adaptation for bridging temporal or spatial scarcity of primary substrates. Under this model, a host-plant acts as a reservoir of mycelium, distributing fungi across the range of leaf-fall.

As such, the FA hypothesis yields several testable predictions: (1) A measurable spatial linkage between endophyte and decomposer life stages for fungi utilizing a FA strategy, wherein stromata serve as sources of endophytic infection (in addition to being sources of direct dispersal) but represent relatively short "bursts" in time, while areas of endophytic infection serve as slower, more "trickling" dispersal centers. (2) A prediction of endophytic host generalism in diverse tropical forests, as strong host preference would interfere with dispersal abilities in systems where the density of any one host species is usually quite low (May 1991). This prediction may not hold in systems where strong dominant hosts are available, as in many temperate forests. (3) The FA hypothesis leads to a prediction that endophytes will be less constrained by environmental conditions than their corresponding decomposers. And, (4) we predict the FA strategy to be a specialized survival/dispersal mechanism utilized by a subset of fungi. Variation in niche or preferred habitat would modulate the selective advantage of endophytism. Thus, we predict some species in a group to be more adapted to endophytism than others.

We found significant clustering between life stages in two of the five species of *Xylaria* with both life stages, *Xylaria* aff. *curta* Fr. and *Xylaria fissilis* Ces. This suggests

spatial linkage of life stages, consistent with prediction (1). It is worth noting that the genetic marker used to link endophytic and decomposer life stages, ITS, has certain limitations. This marker is composed largely of two highly variable introns, and as such is excellent for species identification where reference sequences are available, but is not appropriate for phylogenetic approaches to clustering (Schoch et al. 2012), and is not useful for determining relatedness of individuals within a taxon. As we expect that meiotically produced ascospores are the source of endophytic infection, markers that allow the determination of relatedness between isolates, in addition to the identity of isolates, may complement ITS in future studies. Additionally, the utilization of next-generation sequencing techniques in the elucidation of endophytic communities will allow much greater depth of sampling, regardless of locus selected. Such depth of sampling will be particularly useful in further examination of the environmental constraints and host specificity of fungi suspected of utilizing a FA life history strategy.

Demonstrating the possibility of transfer from endophytic to a decomposer life stage, we have observed endophytic strains of *Nemania serpens* (Xylariaceae)—close relative of *Xylaria* (Hsieh *et al.* 2010)—from conifer needles to colonize dead *Acer macrophyllum* wood in laboratory conditions (G. C. Carroll, unpub. data). Here we explicitly tested the ability of endophytic members of the Xylariaceae to successfully transfer from leaves at our Ecuadorian site to dead woody substrates in laboratory conditions. This test conclusively demonstrates the link between endophytic and saprotrophic *Xylaria,* showing that endophytic isolates can colonize dead woody substrates from within leaves. These observations are contrary to the predictions of Bayman *et al.* (1998), who hypothesized that *Xylaria* endophytes are one-way "dead ends"—purely a sink for dispersal.

Consistent with prediction (2), we did not detect host preference by xylariaceous endophytes. However, the power of our study to detect host preferences may be limited due to the large number of hosts with few samples. Our culture and sampling efforts, though quite extensive, were insufficient to populate multivariate community analyses of host-associated xylariaceous communities (see, for example, Veresoglou & Rillig 2014).

Culture-based studies may be particularly disadvantaged when dealing with questions of endophyte host specificity because of culture bias and other limitations of culture-based studies, such as sampling depth (species accumulation curves generally saturate at impractical levels of effort per leaf) (Arnold *et al.* 2000, Arnold & Herre 2003, Lau *et al.* 2013). Some culture-based studies have addressed culture bias through the use of specialized extracts of host-plants in growth medium (Arnold *et al.* 2000, Arnold & Herre 2003, Lau *et al.* 2013), or through direct PCR/cloning methods (Higgins *et al.* 2011), but these approaches are very labor intensive in experiments involving more than a few species or hosts. We chose to work with *Xylaria* species, in particular, because they typically culture readily both from spores and from leaves as endophytes (Bayman *et al.* 1998), reducing potential culture bias. In a direct comparison of direct PCR (using cloning) versus culturing, Higgins *et. al.* (2011), reporting at the order level, found that Xylariales were somewhat more common in cultures (48%) versus clones (38.9%), but that they were common in both.

When examining questions of host specificity, endophytes are probably best analyzed as multivariate communities within hosts, or as networks of host/endophyte co-occurrences (Peršoh 2013, Higgins *et al.* 2014, Ikeda *et al.* 2014). In future efforts, culture-independent, high-throughput meta-barcode sequencing techniques combined with whole community analysis of endophytes will more adequately address host-endophyte affinities (see, for example, Peršoh 2013).

Nonetheless, our results are in agreement with many studies that indicate that most non-clavicipitaceous tropical foliar endophytes, and especially *Xylaria,* are host generalists (Bayman *et al.* 1998, Cannon & Simmons 2002, Suryanarayanan *et al.* 2002, Arnold & Lutzoni 2007, Higgins *et al.* 2011), and are supportive of the idea that plant-associated fungi in hyper-diverse regions of the tropics will tend towards host generalism (May 1991). Some have suggested that endophyte communities should be regionally unique, due to dispersal limitation (Higgins *et al.* 2014, Vaz *et al.* 2014), and that endophytes of individual plants are predicted as much by location as by host affinities. Higgins *et al.* (2011, 2014), for example, found that tropical forest grass

endophyte communities are more similar to leaves of nearby woody plants than those of distant grasses.

We found that endophytes are released from environmental constraints as compared to corresponding decomposers, as expected from prediction (3). Decomposers exhibited sensitivity to environmental variables that was not observed in endophytes, particularly to proximity of water. This is not surprising, as moisture is important for spore germination and decomposition by most free-living fungi (Moore 1986, Eveling *et al.* 1990, Gange *et al.* 2007). Indeed, it has been speculated that the evolutionary origins of the Xylariaceae are linked to adaptation for water conservation (Rogers 1979, 2000). Our findings, that *Xylaria* are found fruiting in closer proximity to water sources than expected by chance, seem to indicate a strong role of water use in the ecological and evolutionary constraints for the genus. Endophytic fungi, however, exist in the highly buffered environment of the internal tissues of their host-plants; it is predictable that environmental conditions would have a less direct effect on their distributions. We see this in our spatial clustering analysis, where endophytes are not constrained by proximity to the stream (Fig. 2.3). The unconstrained endophytic life stage may be a way that these fungi can bridge spatial and temporal gaps in suitable habitat; this is the core of the FA hypothesis, and our results here are consistent with this.

Lastly, in agreement with prediction (4), in our study all endophytic species of *Xylaria* were also recovered as decomposers from rotting wood on the forest floor. The reverse was not true; many decomposers were found only as stromata and were not detected as endophytes. Our diversity estimators and sampling effort curves indicate that we recovered most of the culturable *Xylaria* species from the leaves, but that decomposer *Xylaria* were undersampled. Okane *et al.* (2008) suggest that there may be Xylariaceae that exist solely as endophytes, but did not undertake concurrent systematic stromata collection to verify this. It is clear from our study that there are species-specific differences in the frequencies at which *Xylaria* displaying both life stages were found in the endophytic and saprotrophic phases (Table 2.1), supporting the notion that there are

dispersal or habitat differences among species. Our results suggest that endophytism is a specialist strategy for some members of the genus *Xylaria*.

We observed probable dispersal linkage in the form of spatial clustering of fungi. We also observed release from moisture limitation by two decomposer fungi through endophytism, suggesting that the endophytic life stage may be serving as a method to span dry habitats or persist during times of low moisture. We also directly observed the ability of endophytic *Xylaria* to colonize available woody substrates and initiate stromata formation. Finally, we found no evidence for host preference in endophytic *Xylaria* species. The limitations of a single observational study must be acknowledged: it remains to be seen if similar trends will be observed in some endophytic fungi of temperate zones or outside of montane cloud forests in the tropics. Nevertheless, we find these results to be consistent with the predictions of the Foraging Ascomycete Hypothesis, and a successful first step into the investigation of this intriguing and ecologically important hypothesis.

Bridge to Chapter III: Rolling a mycobiome down a hill.

Following our work in Ecuador, my colleague and coauthor Roo Vandegrift wanted to look for the signature of the Foraging Ascomycete hypothesis on a larger spatial scale, sample more deeply using next-generation sequencing, still focusing on the fungal family Xylariaceae. I also wanted to draw back to sample on a larger spatial scale, and explore deeper sampling possible with illumina© MiSeq platform sequencing. However, I was excited for the chance to characterize the endophyte mycobiomes of forests, unconstrained by a particular fungal taxon. To me, one amazing achievement of the current family of microbial ecology methods is the ability to sketch, however crudely, a portrait of the microbial landscape at a large scale. For me, this approach is a continuation of some the original work on foliar endophytes (Carroll 1978a), and has enormous practical implications.

Roo and I got our chance to do both. In summer of 2013 we were each awarded an East Asian Pacific Summer Institute grant (EAPSI) from the NSF to work with Dr. Yu-Ming Ju, one of the preeminent experts of Xylariaceae in the world, in Dr. Ju's home country of Taiwan. There Roo conducted another comparison of xylariaceous leaf endophytes and decomposers, and I conducted a general survey of spatial patterns of leaf and wood endophytes. What follows in chapter III is a characterization of the mycobiome of a forest landscape in northern Taiwan, and to explore some ecological concepts of microbiomes, namely the existence and patterns of "core" microbiomes.

CHAPTER III

ROLLING A MYCOBIOME DOWN A HILL: ENDOPHYTES IN THE TAIWANESE CLOUD FOREST

D. Thomas, R. Vandegrift, H. M. Hsieh, Y. M. Ju, B. A. Roy

Contributions:

D. Thomas and R. Vandegrift did field work. H. M. Hsieh contributed to laboratory preparation of samples and lab resources. Y. M. Ju provided materials and lab space, and did taxonomic identification. B. A. Roy provided materials and lab space and did conceptual/experimental design work. D. Thomas conducted lab work, wrote the paper and conducted the statistical analysis.

Abstract

Fungal endophytes of plants are ubiquitous and important to host plant health. Despite their ecological importance, landscape-level patterns of microbial communities in plant hosts are not well-characterized. Fungal wood-inhabiting and foliar endophyte communities from multiple tree hosts were sampled at multiple spatial scales across a 25 ha subtropical research 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. Of the variance that could be explained, host tree species was the most important driver of mycobiome community-composition. Within a single tree species, "core" mycobiomes were characterized using cooccurrence analysis. These core groups of endophytes in leaves and wood show divergent spatial patterns. For wood endophytes, a more consistent, "minimal" 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.

Introduction

Microbial community assembly and geographic patterns in microbes remain poorly understood, despite nearly a century of discussion (Baas-becking 1934 as cited in De Wit 2006, Martiny 2006, Green and Bohannan 2006, Peay 2010, Hanson 2012, Nemergut 2013). Host-associated microbes present additional complexity in modeling microbial community assembly, and raise questions concerning fidelity of host-microbe interactions. Rich microbial communities appear to be associated with all large, eukaryotic organisms (Rosenburg 2010, Hoffman 2010). Plant-fungal symbioses are important to plant and fungal fitness (Malloch 1980, Stukenbrock 2008, Vandenkoornhuyse 2015) and are at least as ancient as vascular plants (Redecker 2000, Krings 2007). Fungal endophytes, or fungi that live internally in plant tissues without incurring disease symptoms (Wilson 1995), have been shown to be widespread and important to plant health (Arnold 2003, Mejia 2008, Rodriguez 2009, Porras-Alfaro 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 (Santamaria 2005, Lundberg 2012, Bodenhausen 2013), at least partly due to host-microbe preferences (Schulz 1999, Oldroyd 2013, Venkateshwaran 2013). Plant organs have been shown to host distinct communities of endophytes (Bodenhausen 2013, Peršoh 2013, Tateno 2014, Edwards 2015). Endophyte communities are also influenced by environmental conditions (Carroll 1978, Arnold 2003, Zimmerman 2012), in spite of presumed buffering from environmental stresses by host tissues. Fungal communities are subject to spatial processes such as dispersal limitation (Peay 2010, Higgins 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.

The potential importance of microbes in adding ecological functions to their hosts (Rodriguez 2009, Johnson 2012, Woodward 2012) has led some to suggest that multicellular organisms may host *core microbiomes* (Hamady 2009, Shade 2011, Vandenkoornhuyse 2015), which are subsets of important and consistent microbial partners. Initial explorations of plant core microbiomes have been highly controlled (Lundberg 2012, Edwards 2015). Studies of plant-associated microbiomes in natural settings have rarely been framed in terms of core microbiomes (Kim 2011, Zimmerman 2012, Bodenhausen 2013, Higgins 2014, Kembel 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 2009) may be useful when carefully applied to long-studied symbioses such as ruminant gut communities (Liggenstoffer 2010) or mycorrhizal relationships (Malloch 1980, van der Heijden 2009). This definition may not always serve for describing the numerous and labyrinthine microbe-host interactions that exist outside of laboratory settings. However, other definitions of core microbiomes exist that may be more useful for ecologically modeling microbiomes (Hamady 2009).

Here we acknowledged that plant hosts exert strong influence on community membership of their endophytic compartment. 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 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. To test this, we compared community composition and ecological drivers between wood and leaf fungal endophytes in multiple species of plant host, to identify instances of differential response by microbial communities from host to environmental changes or spatial constraints. We mapped

spatial patterns in the most strongly associated endophytic fungi of a single host species, to examine patterns of turnover in a putative core microbiome.

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 (Losos & Leigh 2004) Forest Dynamics Plot (FDP) (Su 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, gramminoids, and shrubs. Vegetative communities can be broadly categorized into four community types described by dominant tree species combinations (Fig. 3.1). 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 central hilltop adjoins lowland habitat with perennial streams along its eastern and southern bases, and mid-elevation upland habitat to the north. Perennial streams join and exit the FDP through a steep valley in the southwest of the plot (Fig. 3.2). The complex topography of Fushan has been summarized by classification of each 20 m x 20 m quadrat of the FDP into one of seven habitat types, based on aspect, slope, convexity, and elevation (Fig. 3.1), which are found to influence vegetative communities (Su 2010). Soil at Fushan FDP are generally acidic, with low fertility and organic carbon content. Soils are relatively young (inceptisols) due to erosion on steep slopes and flooding disturbances in lowland habitat. High leaching and erosion cause lower nutrient levels to occur in the central hilltop. See Su et al. (2007) for more details.

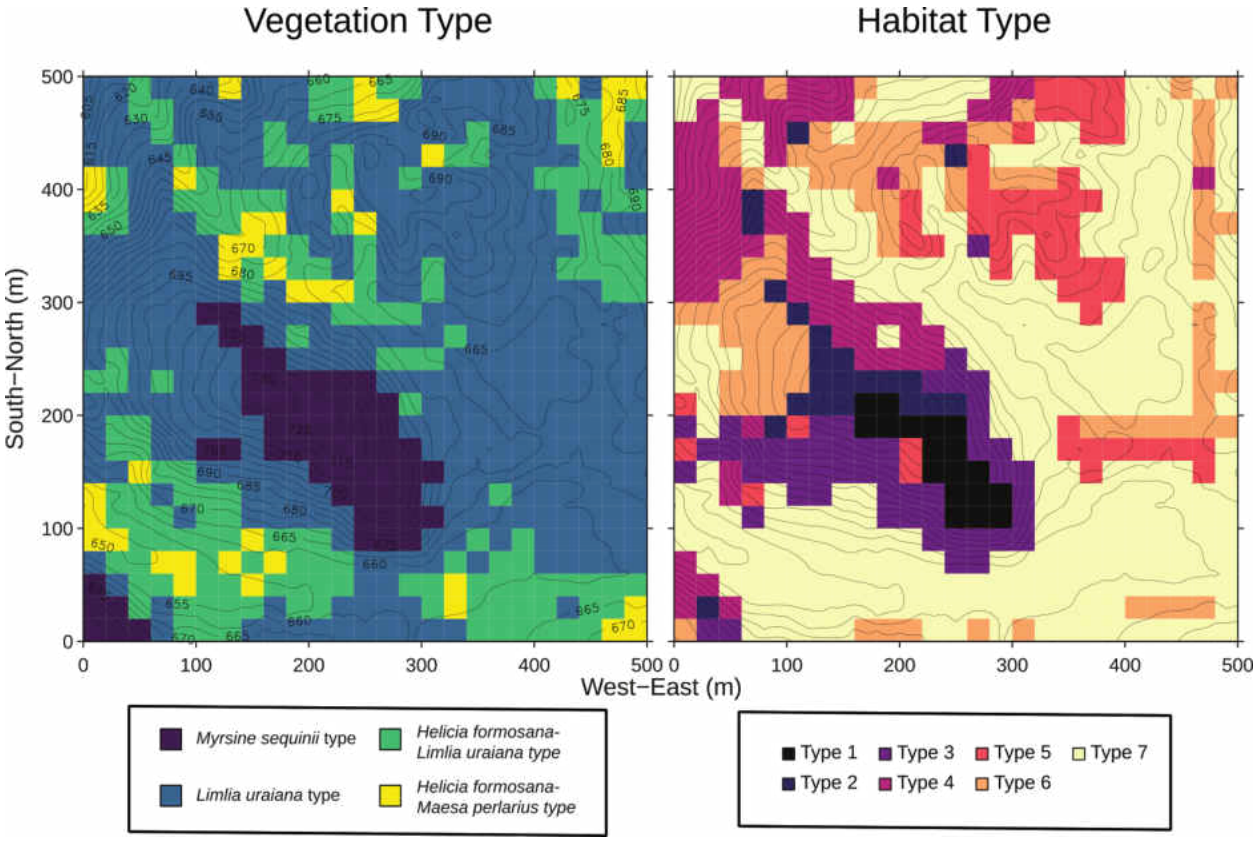


Figure 3.1. Left: topographic map of the Fushan FDP with the four vegetation types as classified by Su et al. (2007) Right: map of the habitat type, a composite classification based on microtopographic characteristics of quadrats, defined by Su et al. (2010). The units of the coordinates and contours are in meters, with quadrats at 20x20m scale. Figures reproduced with permission from authors. Click here for a higher resolution image.

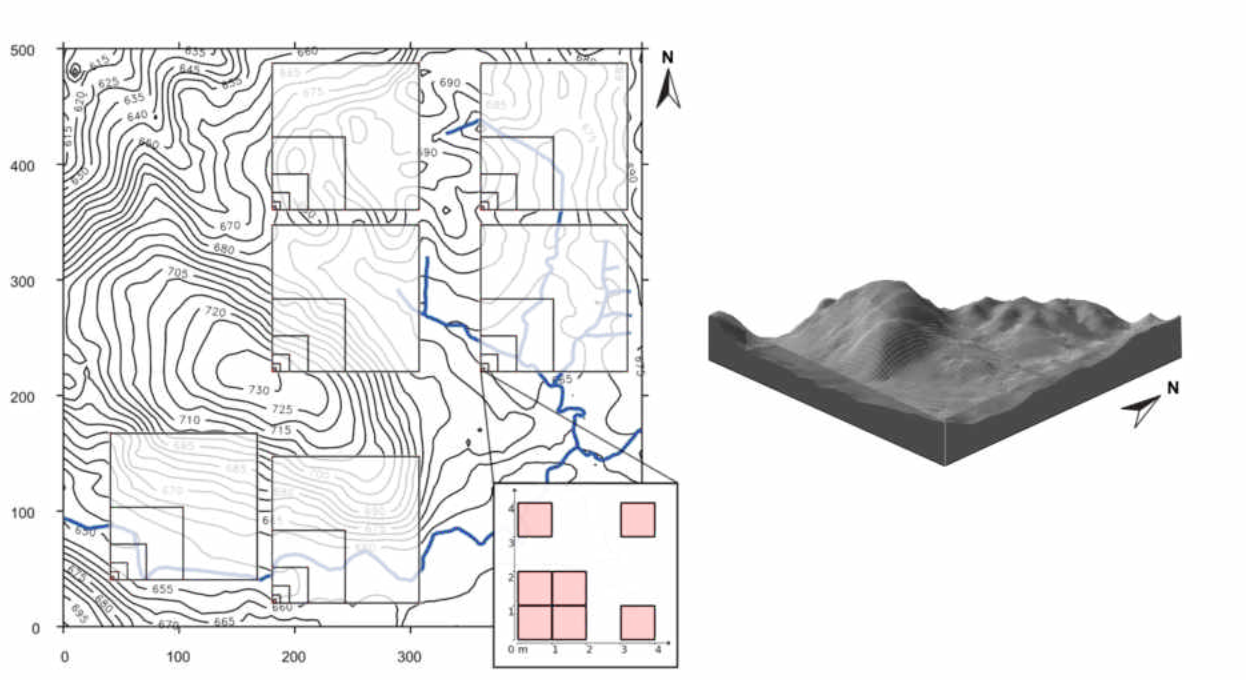


Figure 3.2. Left: An overview of nested-squares, logarithmic sampling scheme Vandegrift (2016). Vertices of squares are sample sites. Units are meters. Right: Perspective diagram of Fushan Forest Dynamics Plot (Su 2010). Figures reproduced with permission from authors. Click here for a higher resolution image.

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 2013) (Fig. 2.2). Each sub-plot contained sampled points at 1, 2, 4, 8...128 meter distances from the southwest origin of the subplot, resulting in 25 trees sampled per subplot. Sampling of each set of subplot 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. Six out of nine sets of nested squares were sampled, due to time constraints, resulting in 150 total leaf and shoot samples, though fungal endophyte DNA was not successfully amplified amplified from all samples, see below.

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 3m collapsible pole pruner. Identification of host-tree was supplied by survey data from ongoing ecological research at Fushan FDP (Su 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

Preparation and sequencing of Illumina libraries for leaves and wood were undertaken separately, with differing protocols. Protocols for leaf fungal endophyte amplicon library preparations are given in Vandegrift (2016). Protocols for wood endophytes are given in detail in chapter V. Briefly, all leaves were washed and surface-sterilized, and woody stem material was debarked with a sterile scalpel and phloem and sapwood were harvested. Fungal endophyte DNA was extracted from both in separate library preparations and ITS region 1 was amplified using a fungal-specific primer set with illumina© tagged, barcoded primers. Positive, "mock community" controls were included in the wood-endophyte library, and pure-water negative controls were included in both libraries. Samples were multiplexed and sequenced in separate illumina© Mi-Seq sequencer runs.

Bioinformatics

Details of the bioinformatics pipeline are explained in chapter V. Full scripts available in supplementary information (available here and here). Briefly, general 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, so to maximize comparability, the reads from both libraries were combined as early as possible in the bioinformatics pipeline, following merging of paired

ends. Variance stabilization of combined wood and leaf reads was done using using the *DESeq2* package in R (Love 2014, McMurdie 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.

Statistical methods

Overview

Ecological patterns of the entire fungal community of leaves and wood of all hosts were examined first. Analyses then were focused on patterns in the mycobiome of the single, most commonly-sampled host tree, *Helicia formosana*. Finally, host-fungus coccurrence patterns were used to define a core mycobiome that was also examined for ecological patterns (Fig. 3.3). Statistical analysis was conducted in R Statistical Software, version 3.3.1 (R core team 2016), with the *vegan* (Oksanen 2017), *phyloseq* (McMurdie 2013), *cooccur* (Griffith 2016), *igraph* (Csardi 2006) and *ecodist* (Goslee 2007) packages. Where required, all endophyte community comparisons were conducted using Bray-Curtis dissimilarity index (Bray 1957, McCune 2002).

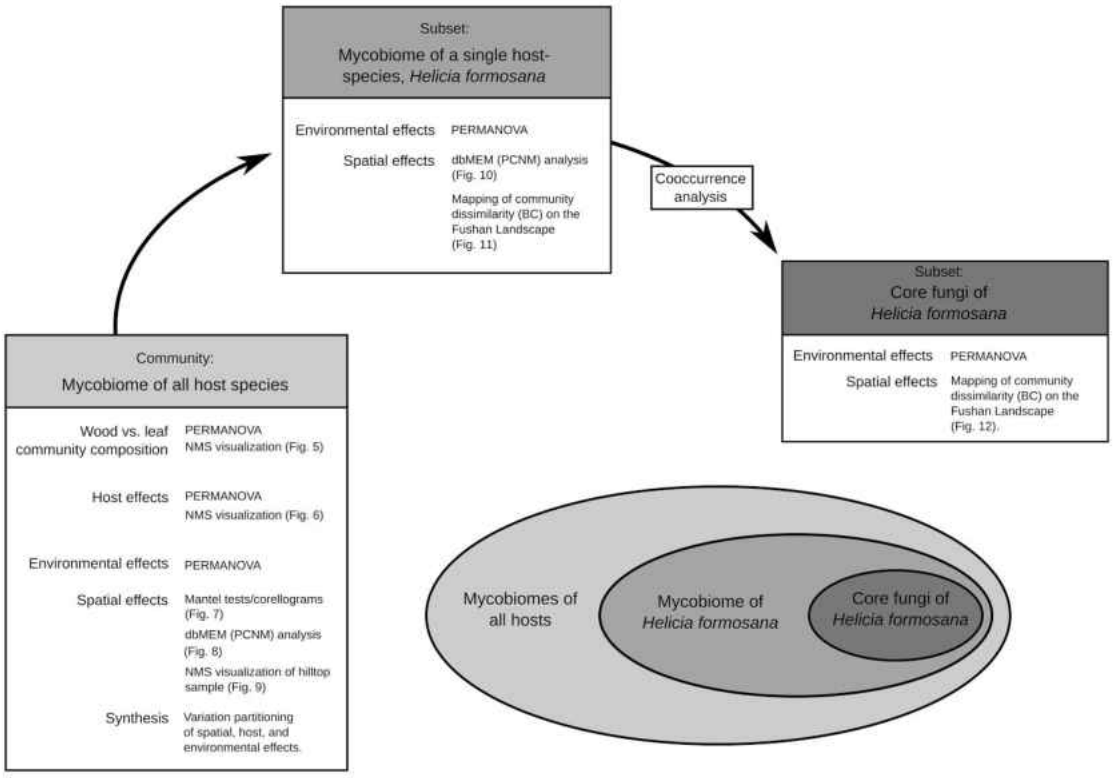


Figure 3.3. An overview of statistical methods. Analyses begin with broadscale ecological patterns of all wood and leaf samples, then subset to a single host tree species *H. formosana*, and lastly to the patterns of members of core mycobiome of *H. formosana* as defined by cooccurrence patterns. Click here for a higher resolution image.

Mycobiome of all hosts

Dissimilarity of leaf and wood endophyte communities were modeled and visualized using non-parametric multivariate analysis of variance (NPMANOVA or PERMANOVA) (Anderson 2001), and non-metric multidimensional scaling (NMS). Comparisons between leaf and wood libraries were constrained to only shared OTUs, those that were detected at least once in both leaf and wood tissue, to reduce bias from separate library preparations. Following this, all analyses were for wood and leaf endophyes were conducted separately, in parallel. Effects of host and environmental variables of vegetative community and topography (Fig. 1) on endophyte communities

were also modeled individually using PERMANOVA, and results were visualized with NMS when significant.

Spatial trends in endophyte communities were first explored using Mantel tests (Mantel 1967, Legendre 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 2011, Legendre 2012), endophyte community matrices were Hellinger-transformed (Legendre 2001), and "regressed" using Redundancy analysis (RDA) (Legendre 2012, Buttigieg 2014) against all eigenvecters ("PCNM vectors") resulting from dbMEM analysis. Stepwise model selection was then used to filter the biologically important eigenvectors (Oksanen 2017). The remaining eigenvectors were then inspected visually, and used as independent variables in linear-like models of variation partitioning (see below). Ecological patterns of interest detected in spatial analysis were also visualized by mapping Bray-Curtis distance of all wood or leaf samples from a single point of interest (indicated by PCNM vectors), in NMS ordinations.

Overall patterns of dissimilarity among in our endophyte communities were examined using variation partitioning (Peres-neto 2006, Borcard 2011, Gavilanez 2012, Buttigieg 2014). Variation partitioning attempts to explain patterns of dissimilarity among rows of a response matrix among several explanatory matrices, through comparisons of RDA (or other direct-gradient analysis) models created from all possible combinations of explanatory matrices. Here relative effects of host, environmental, and spatial variables on wood and leaf communities were tested as predictors of endophyte community dissimilarity.

Mycobiome of a single host, *Helicia formosana*

To examine ecological patterns of mycobiomes without variation resulting from host tree species, the fungal endophytes of a single host tree, *Helicia formosana* Lour. & Hemsl, were examined. This was the host tree for which the most samples (leaves, n=31; wood n=22) were available. Environmental effects on endophyte community were tested with PERMANOVA models of *H. formosana* wood and leaf endophytes against the environmental variables of vegetation class and topography. Spatial patterns were tested by constructing biologically informative PCNM vectors as above, using the subsetted matrix of sites where samples were from *H. formosana* trees. To further visualize, a single sample of interest indicated by the PCNM vectors was used as a center of comparison for all other samples. 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, cooccurrence analysis was conducted on the all-host, all-endophyte species-using a pairwise, probabilistic model (Veech 2013). Core mycobiomes of hosts were defined as the subset of fungi that showed strong cooccurrence 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. Focusing on one host, the results were 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 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.

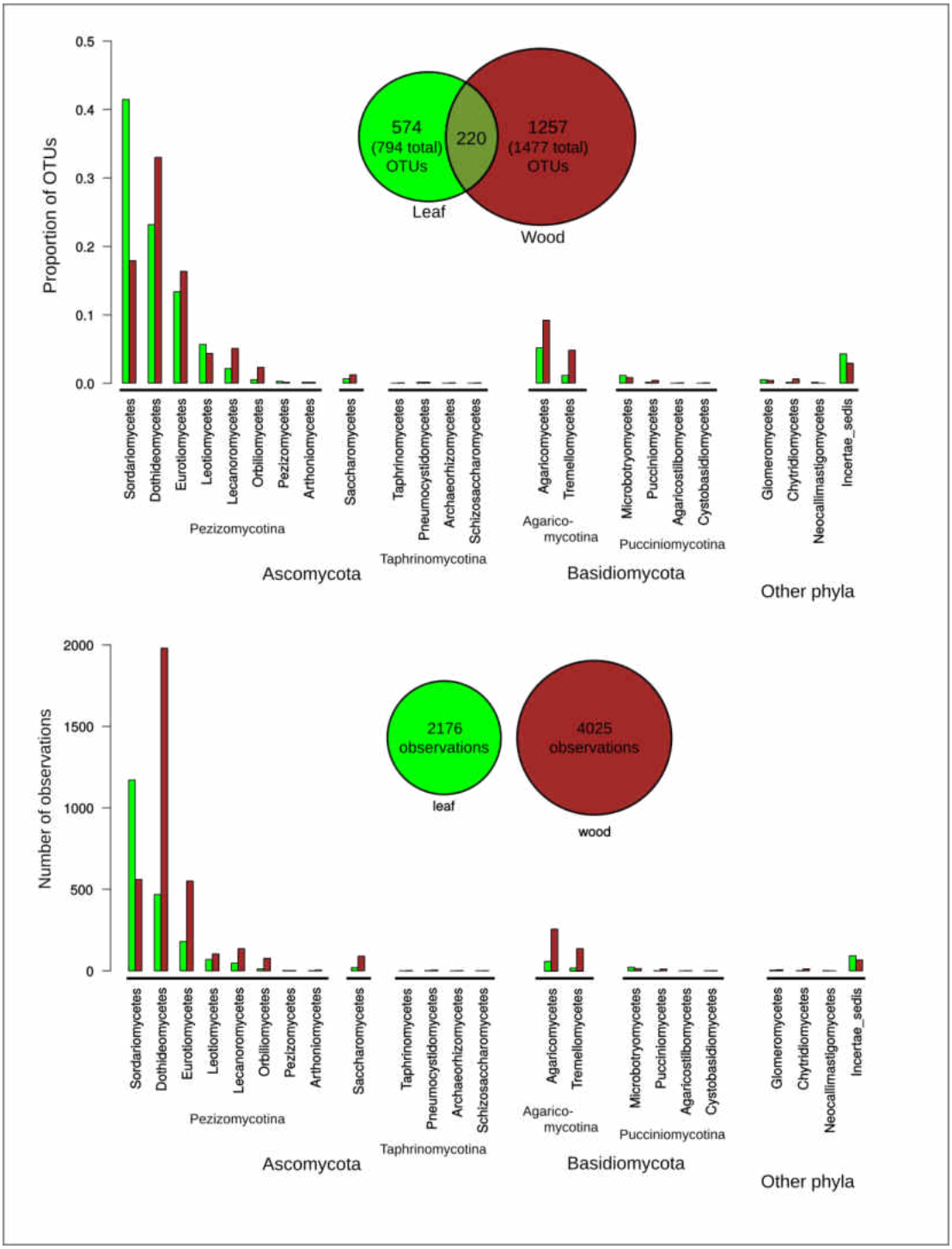
Results

Mycobiome of all hosts:

Endophyte community composition, wood vs. leaves:

After variance-stabilization, the wood endophyte library contained 1477 OTUs and the leaf library contained 794 OTUs. They shared 220 mutually-detected OTUs. (Fig. 3.4) Both leaf and wood samples were dominated by Ascomycota (91% of OTUs in leaves, 83% in wood), but a larger percentage of wood OTUs matched to Basidiomycota (15% of OTUs in wood, compared to 8% of reads in leaves). This larger percentage of Basidiomycetes was due mostly to a larger diversity of Agaricomycetes and Tremellomycetes present in the wood (Fig. 3.4). Within Ascomycota, both leaf and wood samples contained high percentages of Sordariomycetes, Dothideomycetes, and Eurotiomycetes. Dothideomycetes were present in higher relative diversity in the wood (32% of all OTUs) than in leaf samples (23% of all OTUs). The opposite was true for Sordariomycetes, which were 41% of leaf endophyte OTUs, compared to 18% of wood OTUs . As noted above, all ecological analyses were transformed to incidence data, so that the basic ecological unit for all following analyses was an non-zero observation of an OTU in a sample after cutoffs were subtracted, regardless of read abundance. Trends in numbers of observations parallel patterns in OTU diversity (Fig. 3.4); if a class of fungi contained a large diversity of OTUs, it also tended to be observed often throughout the study site.

Figure 3.4. (next page). Overview of taxonomic composition of wood and leaf libraries. Top: total numbers of unique OTUs described for each class of Fungi. Bottom: total number of observations of each class. Observations, or presence of a fungal OTU in a sample regardless of read abundance, were the unit of interest for all following analyses, rather than read abundances. Click here for a higher resolution image.


Proportion of OTUs
0.5
0.4
0.3
0.2
0.1
0.0
574
(794 total)
OTUs
220
1257
(1477 total)
OTUs
Leaf
Wood
Sordariomycetes
Dothideomycetes
Eurotiomycetes
Leotiomycetes
Lecanoromycetes
Orbiliomycetes
Pezizomycetes
Arthoniomycetes
Saccharomycetes
Taphrinomycetes
Pneumocystidomycetes
Archaeorhizomycetes
Schizosaccharomycetes
Agaricomycetes
Tremellomycetes
Microbotryomycetes
Pucciniomycetes
Agaricostilbomycetes
Cystobasidiomycetes
Glomeromycetes
Chytridiomycetes
Neocallimastigomycetes
Incertae_sedis
Pezizomycotina
Taphrinomycotina
Agarico-
mycotina
Pucciniomycotina
Ascomycota
Basidiomycota
Other phyla
Number of observations
2000
1500
1000
500
0
2176
observations
leaf
4025
observations
wood
Sordariomycetes
Dothideomycetes
Eurotiomycetes
Leotiomycetes
Lecanoromycetes
Orbiliomycetes
Pezizomycetes
Arthoniomycetes
Saccharomycetes
Taphrinomycetes
Pneumocystidomycetes
Archaeorhizomycetes
Schizosaccharomycetes
Agaricomycetes
Tremellomycetes
Microbotryomycetes
Pucciniomycetes
Agaricostilbomycetes
Cystobasidiomycetes
Glomeromycetes
Chytridiomycetes
Neocallimastigomycetes
Incertae_sedis
Pezizomycotina
Taphrinomycotina
Agarico-
mycotina
Pucciniomycotina
Ascomycota
Basidiomycota
Other phyla

Leaf and wood endophyte communities are distinct, even when analyses are constrained to only species present in both Illumina libraries (PERMANOVA, $F(1, 206) = 34.5$, $p < 0.01$, $R2 = 0.14$, permutations = 10000) (Fig. 3.5).

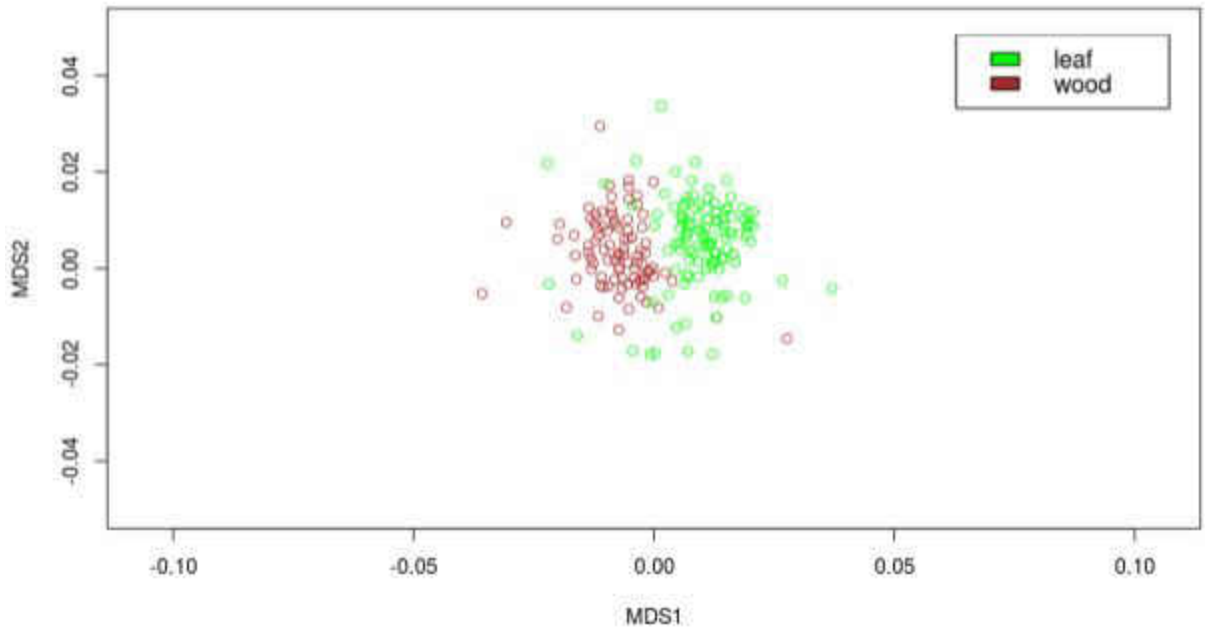


Figure 3.5. Non-metric multidimensional scaling diagram, comparing leaf and wood endophytes of all host trees, using shared species only. Plot has been scaled in to maximize visibility, two far outliers have been removed. To see entire NMS with outliers, and for a higher resolution image, click here.

Host effects on endophyte community composition:

Host species is the strongest single predictor of similarity within both leaf (PERMANOVA, $F(33, 89) = 2.1$, $p < 0.01$, $R^2 = 0.44$, permutations = 10000) and wood endophyte communities (PERMANOVA, $F(29,61) = 1.48$, $p < 0.01$, $R^2 = 0.41$, permutations = 10000) (Fig. 3.6).

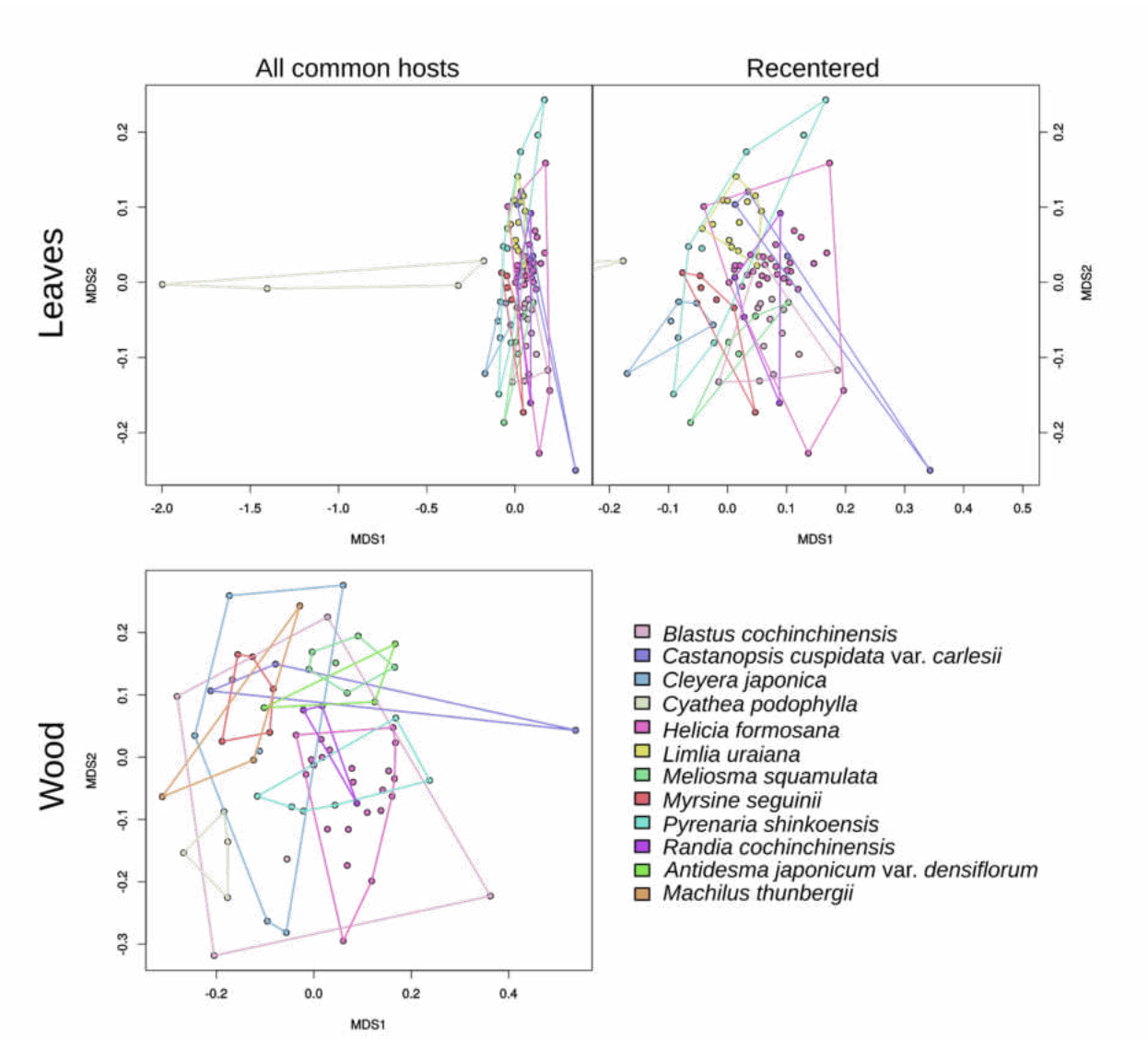


Figure 3.6. Non-metric multidimensional scaling diagram of endophyte communities, with all tree hosts that were sampled at least 3 times. Leaf plot has been recentered to maximize visibility in upper right, excluding the very unique communities of *Cythea japonica*. Click here for a higher resolution image.

Environmental effects on endophyte community composition:

Taken alone, composite environmental variables are predictors of similarity in both wood endophyte communities (surrounding above-ground vegetative community: (PERMANOVA, $F(3,87) = 1.5$, $p < 0.01$, $R^2 = 0.05$, permutations = 10000), micro-topographic conditions (PERMANOVA, $F(6,84) = 1.28$, $p < 0.01$, $R^2 = 0.08$,

permutations = 10000), and also in leaf endophyte communities (surrounding above-ground vegetative community: (PERMANOVA, $F(3, 119) = 2.19$, $p < 0.01$, $R^2 = .05$, permutations = 10000), micro-topographic conditions (PERMANOVA, $F(6, 116) = 1.31$, $p < 0.01$, $R^2 = 0.06$, permutations = 10000).

Spatial patterns all-host mycobiomes

Mantel tests:

Wood endophyte community displayed a weak pattern of community-turnover/distance-decay over the entire study area (Mantel's $r = 0.07$, $p = 0.031$) (Fig. 3.7). Leaf communities displayed no global distance decay relationship (Mantel's $r = -0.01$, $p = 0.67$), but displayed local negative autocorrelation in comparisons of samples approximately 200 meters apart (Mantel correlogram, Mantel's $r = -0.10$, $p < 0.05$) (Fig. 3.7) indicating that some portion of these samples at this distance apart contained communities more similar than expected under a null model of complete spatial randomness.

dbMEM analyses

Our sampling scheme yielded 5 biologically significant PCNM vectors for leaf samples, explaining 6.6% of endophyte community variation (Redundancy analysis, constrained inertia = 0.06, Unconstrained inertia = 0.89, $F(5,117) = 1.65$, $P < 0.01$). Three of five of these PCNM vectors can be considered part of a general north-south pattern that can be combined/detrended as such, and the smallest scale PCNM is probably indicative of endogenous autocorrelation (Borcard 2011) . The remaining PCNM vector centers on the hill of the Fushan FDP (Fig. 3.8), and correlates strongly

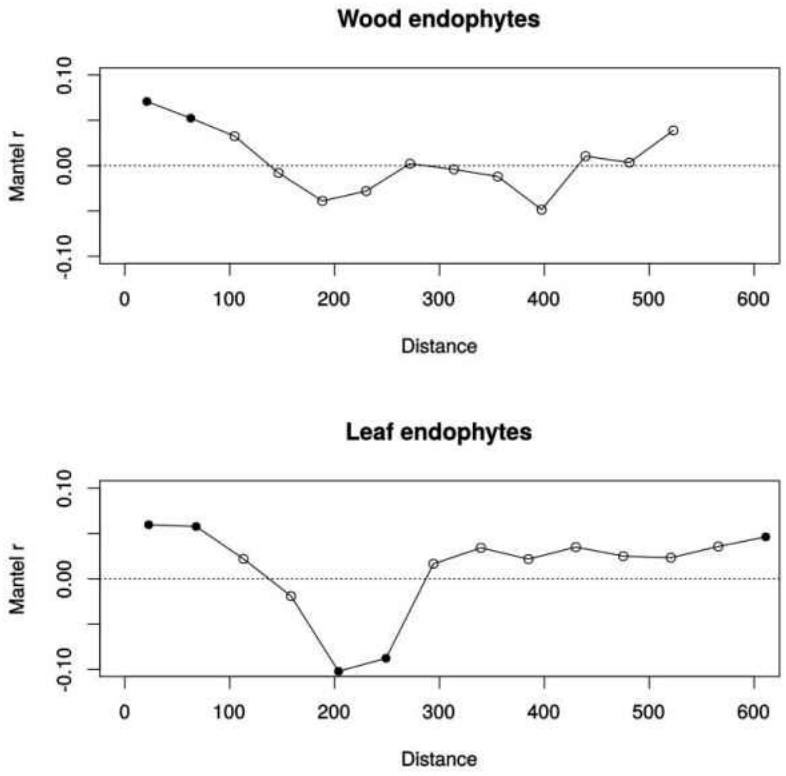


Figure 3.7. Mantel correlograms of spatial correlation of community dissimilarity of endophyte community. Distance units are meters. Black dots indicate statistical significance. Wood endophytes show weak global distance decay trends. Leaf endophytes do not display global distance decay but have a strong signal of local negative autocorrelation at comparisons around 200 m. Click here for a higher resolution image.

with environmental variables of topography and vegetative community (Linear model/multiple regression, adj-R^2=0.64, $F(9,113)=25.65$, $p < 0.01$), highlighting this point as important focal point for further comparisons. For leaves, this hilltop point is consistently central in all stable NMS solutions of similarity among all-host comparisons (Fig. 3.9), and community dissimilarity from this hilltop point is a predictor of dissimilarity among all points ((PERMANOVA, $F(1,121) = 8.6$, $p < 0.01$, $R^2 = 0.067$, permutations = 10000).

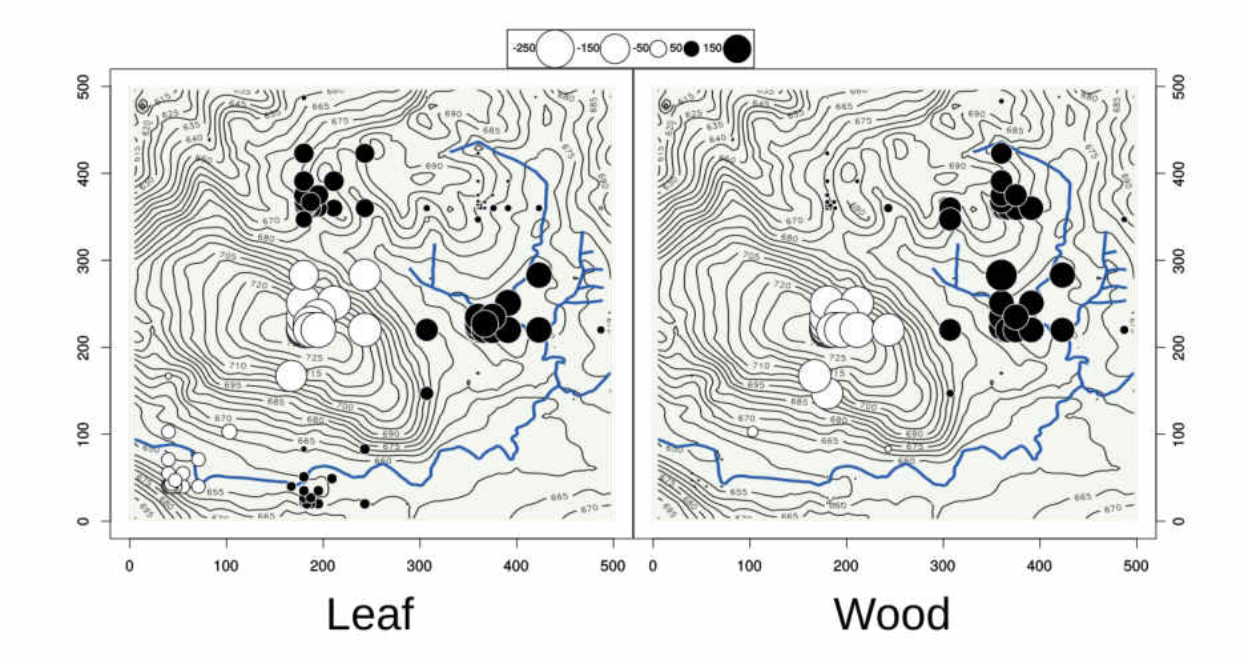


Figure 3.8. Two PCNM vectors showing patterns of variation of all-host endophyte communities of leaf and wood, plotted over a map of Fushan FDP. Both leaf and wood endophyte communities showed some response to the central hill of the plot. Click here for a higher resolution image.

From wood endophyte samples, 4 biologically significant PCNM vectors were described, explaining 6% of variation (Redundancy analysis, constrained inertia = 0.06, Unconstrained inertia = 0.89, $F(5,117) = 1.65$, $P < 0.01$). One PCNM correlates strongly with topographical variables (Linear model/multiple regression, adj-R2=.78, $F(9,81)=36.39$, $p < 0.01$) and is also centered on the hilltop (Fig. 3.8). Two of the remaining PCNMs for wood probably represent fine-scale endogenous autocorrelation and the final is not explained well by available variables or visual inspection.

Variation partitioning

Most of the variation found among samples in our endophyte communities was unexplained. In wood, host effects explain 5% of total community variation (Redundancy analysis, tested with permutational ANOVA, $F(29,54) = 1.20$, $P = 0.001$).

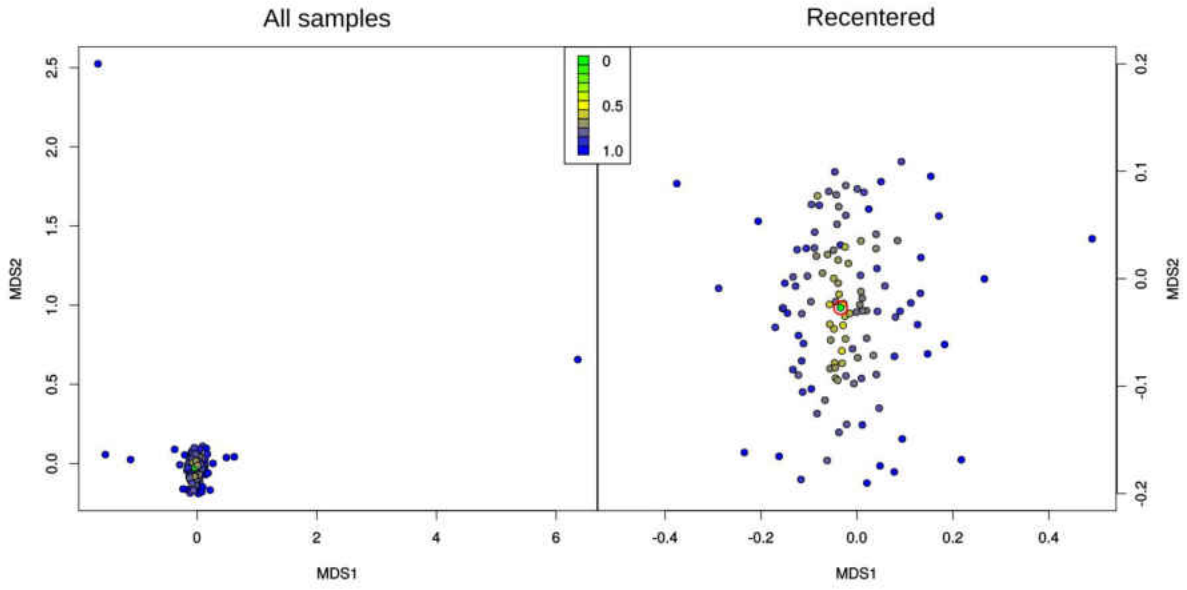


Figure 3.9. Non-metric multidimensional scaling diagram of leaf endophyte communities. Color indicates community dissimilarity (Bray-Curtis), from a single sample on the central hill of the plot. Dark blue points (BC=1) share no fungal species in common with the hilltop sample, and increase in similarity from yellow to green (BC=0). Leaf plot has been recentered to maximize visibility right, losing 4 samples. Hilltop sample is circled in red on the right. Click here for a higher resolution image.

Spatial patterns from wood endophytes were not independent of host spatial patterns (Redundancy analysis, tested with permutational ANOVA, $F(4,54) = 1.09$, $P = 0.195$). Environmental variables (microtopography and vegetative community) were not observed to explain changes in wood endophyte community directly (0% inertia explained).

Explained variation in leaf endophyte community is also mostly correlated with host effects (10% out of 11% explained; Redundancy analysis, tested with permutational ANOVA, $F(9,107) = 2.34$, $P = 0.001$). Independent of host, an additional 1% of leaf endophyte community variation is explained by spatial patterns (Redundancy analysis, tested with permutational ANOVA, $F(5,107) = 1.25$, $P = 0.001$). Environmental variables were also not observed to independently explain changes in leaf endophyte community

(0% inertia explained).

Mycobiome of a single host, Helicia formosana

Environmental variables were not found to directly explain any variance in community of *H. formosana* endophytes, for leaves (PERMANOVA, permutations=10000. Topography: $F(4,26) = 0.80$, $p = 0.89$, $R^2 = 0.11$. Vegetative community: $F(3,27) = 1.13$, $p = 0.24$, $R^2 = 0.11$), or wood (Topography: $F(4,17) = 1.03$, $p = 0.31$, $R^2 = 0.20$, permutations. Vegetative community: $F(3,18) = 1.07$, $p = 0.23$, $R^2 = 0.15$). Leaf and wood endophyte community each yielded one biologically significant PCNM vector (RDA, leaves: constrained inertia = 0.044, Unconstrained inertia = 0.72, $F(1,29) = 1.78$, $P < 0.01$. RDA, wood: constrained inertia = 0.052, Unconstrained inertia = 0.75, $F(1,20) = 1.38$, $P < 0.01$). These PCNMs both display a pattern of dissimilarity centered on the southwest valley (Fig. 3.10). Centering the Bray-Curtis comparisons on this region shows that leaf samples in this region share fungal OTUs (Fig. 3.11).

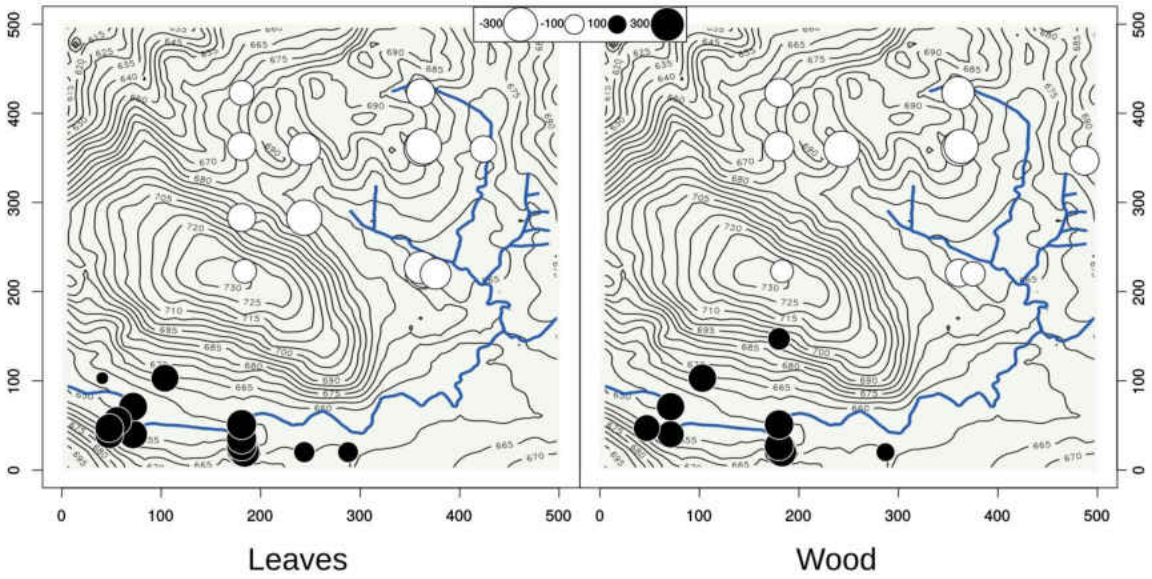


Figure 3.10. Two PCNM vectors showing patterns of variation of single host-tree, *Helicia formosana*, endophyte communities of leaf and wood, plotted over a map of Fushan FDP. Both leaf and wood endophyte communities display dissimilarity between the plot at large and the southern valley. Click here for a higher resolution image.

Cooccurrence analysis:

8 out of 774 possible fungal OTUs showed patterns of cooccurrence with *Helicia formosana* in leaf tissue, and 10 out of 1477 possible taxa from wood tissues (Table 3.1). These fungi were considered members of the *H. formosana* core mycobiome for further analysis.

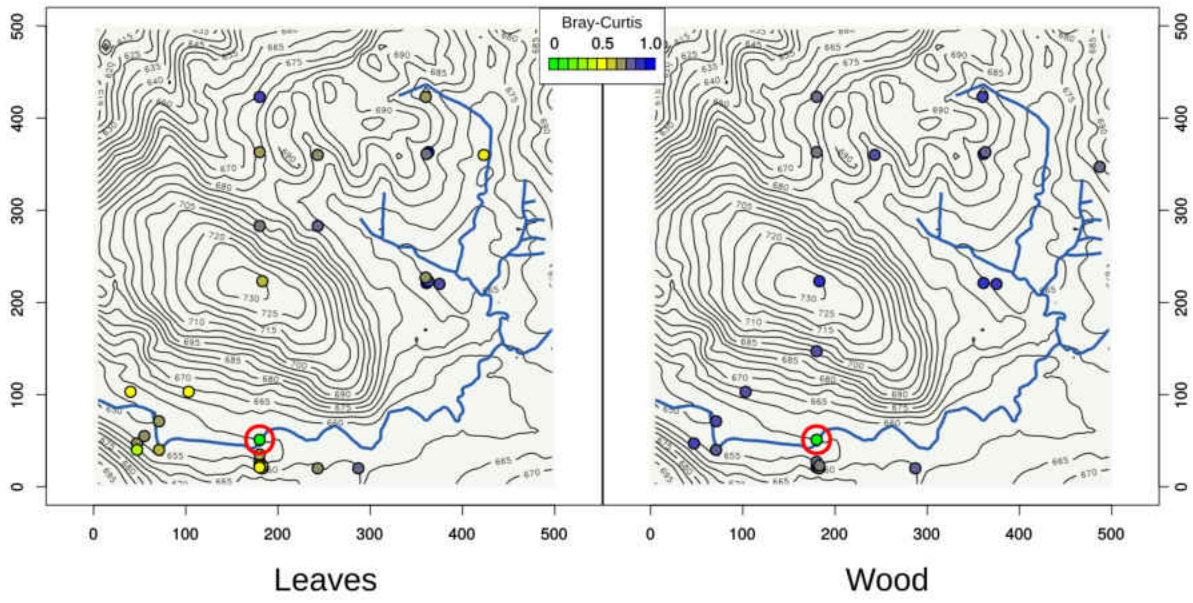


Figure 3.11. Map of Bray-Curtis dissimilarity values over the Fushan FDP, resulting from comparisons between red circled point all other *Helicia formosana* samples. Dark blue points (BC=1) share no fungal species in common with the circled sample, and increase in similarity from yellow to green (BC=0). Click here for a higher resolution image.

Table 3.1. Core mycobiome of *Helicia formosana*, defined by cooccurrence patterns. Click here for a higher resolution image.

Leaves

Taxa	Kingdom	Phylum	Subphylum	Class	Order	Family	Genus
OTU65:128Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Pleosporales		
OTU75:21Leaf	Fungi	Ascomycota	Pezizomycotina	Sordariomycetes	Diaporthales		
OTU226:101Leaf	Fungi	Ascomycota	Pezizomycotina	Orbiliomycetes	Orbiliales	Orbiliaceae	Arthrobotrys
OTU136:130Leaf	Fungi	Basidiomycota	Pucciniomycotina				
OTU221:103Leaf	Fungi	Ascomycota	Pezizomycotina	Eurotiomycetes	Eurotiales	Trichocomaceae	Aspergillus
OTU59:96Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Botryosphaeriales	Phyllosticaceae	Phyllosticta
OTU133:116Leaf	Fungi	Zoopagomycota					
OTU5434:100Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Botryosphaeriales	Phyllosticaceae	Phyllosticta

Wood

Taxa	Kingdom	Phylum	Subphylum	Class	Order	Family	Genus
OTU284:131Leaf	Fungi	Ascomycota	Pezizomycotina	Eurotiomycetes	Chaetothyriales		
OTU1126:199wood	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Capnodiales		
OTU256:38Leaf	Fungi	Basidiomycota	Agaricomycotina	Tremellomycetes	Tremellales		
OTU841:237wood	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Capnodiales		
OTU68:55Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes			
OTU65:128Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Pleosporales		
OTU81:81Leaf	Fungi	Ascomycota	Pezizomycotina	Sordariomycetes	Xylariales	Amphisphaeriaceae	Pestalotiopsis
OTU216:90Leaf	Fungi	Ascomycota	Pezizomycotina	Eurotiomycetes	Verrucariales		
OTU235:131Leaf	Fungi	Ascomycota	Pezizomycotina	Dothideomycetes	Capnodiales	Capnodiaceae	
OTU113:194wood	Fungi	Ascomycota	Pezizomycotina	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	

Core fungi of *Helicia formosana*

No direct relationship between topographic and vegetative community and was detected in either leaf endophytes (PERMANOVA, permutations=10000. Topography: $F(4,24) = 1.30$, $p = 0.26$, $R^2 = 0.18$. Vegetative community: $F(3,25) = 0.57$, $p = 0.79$, $R^2 = 0.06$), or wood endophytes (PERMANOVA, permutations=10000. Topography: $F(4,17) = 1.05$, $p = 0.35$, $R^2 = 0.19$. Vegetative community: $F(3,18) = 0.86$, $p = 0.53$, $R^2 = 0.13$). Visual inspection of spatial patterns show that leaves within the southern valley of the plot contained relatively high proportions of core fungi (Fig. 3.12). Wood contained high proportions of core fungi consistently throughout the plot (Fig. 3.12). In leaves, presence or absence of just these core species in *H. formosana* leaf fungal communities is a partial predictor of entire fungal community structure (PERMANOVA, $F(1, 29) = 3.38$, $p < 0.01$, $R^2 = .10$, permutations = 10000), and for wood endophyte community structure (PERMANOVA, $F(1, 20) = 1.29$, $p = 0.047$, $R^2 = 0.06$, permutations = 10000).

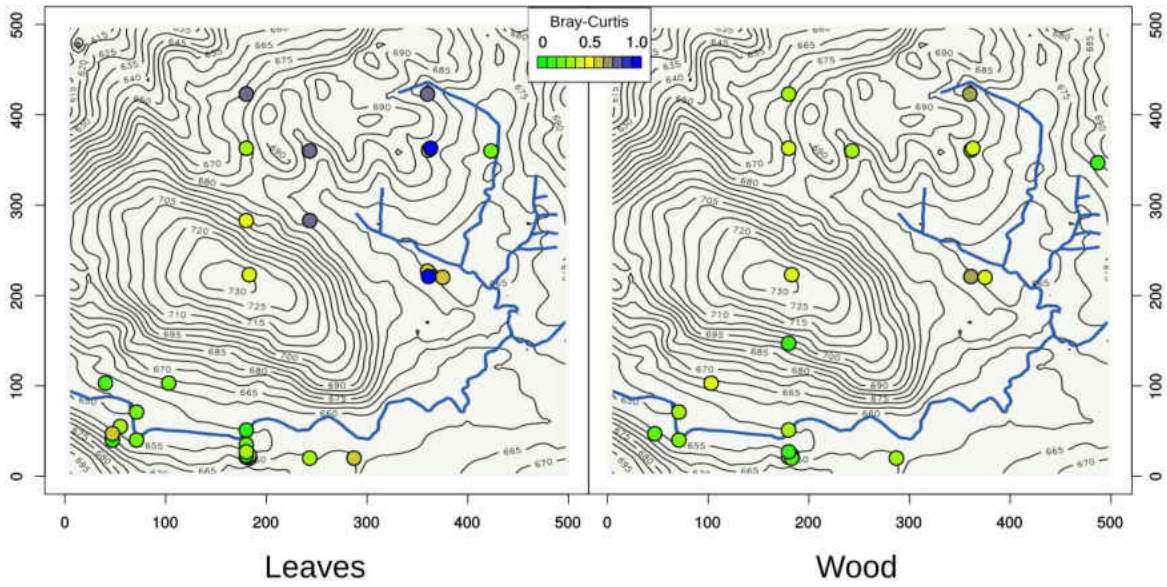


Figure 3.12. Map of Bray-Curtis dissimilarity values over the Fushan FDP, resulting from comparisons between all *H. formosana* points and the core fungi of the *H. formosana*. Dark blue points (BC=1) contain no species from this set of core fungi, and increase in similarity from yellow to green (BC=0, 100% of core fungi present). Click here for a higher resolution image.

Summary comparison

The above analysis compared patterns of community dissimilarity at several levels (Fig. 3.13, Table 3.2). Wood and leaf endophytes of all host-trees showed an identical, high mean Bray-Curtis dissimilarity among all samples (all-host leaf endophyte mean BC=0.9, sd=0.10; wood endophyte mean BC=0.9, sd=0.07). samples are more similar to one another when considering only one host species, *Helicia formosana* (leaf mean BC=0.78, sd =0.12; wood mean BC = 0.81, sd=0.07). This variation can then be partitioned into two groups: Non-core fungi from these hosts show a similar, high level of dissimilarity among samples (leaf mean BC=0.86, sd =0.11 ; wood mean BC = 0.86, sd=0.06). As expected, core fungi assemblages from *Helicia* samples have a lower mean BC (leaf mean BC=0.50, sd =0.27; wood mean BC = 0.40, sd=0.17). Leaf core fungi are more dynamic than wood, showing a higher mean Bray-Curtis dissimilarity and greater variance.

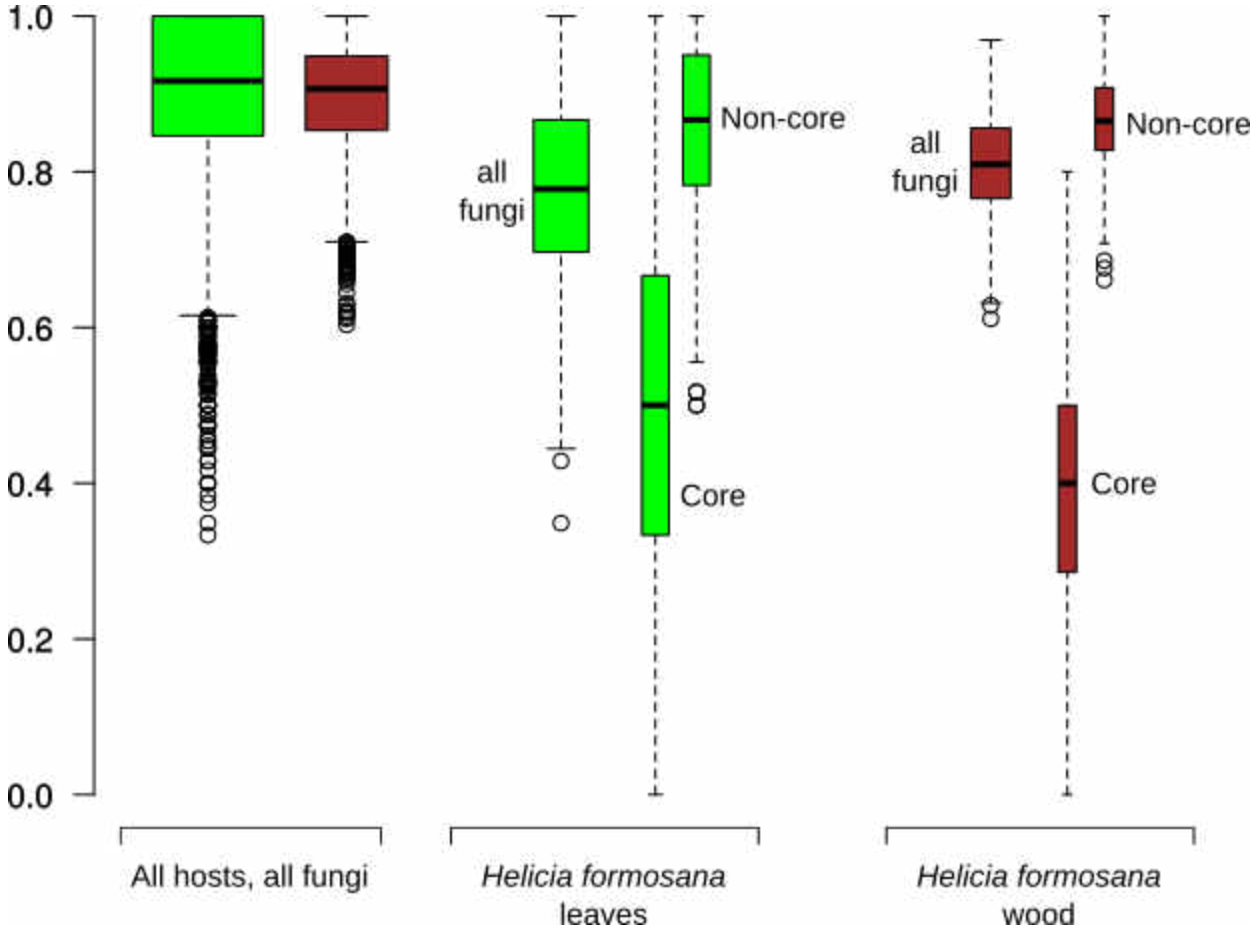


Figure 3.13. Distribution of Bray-Curtis dissimilarity among sample comparisons of all hosts, and of *Helicia formosana* only. Click here for a higher resolution image.

Organ	All hosts, all endophytes	*Helicia*, all endophytes	*Helicia*, non-core	*Helicia*, core-fungi
leaf	0.90 (+/- 0.10)	0.78 (+/-0.12)	0.86 (+/-0.11)	0.50 (+/-0.27)
wood	0.90(+/-0.70)	0.81(+/-0.07)	0.86 (+/-0.06)	0.40 (+/-0.17)

Table 3.2. Summary mean and standard deviation of Bray-Curtis dissimilarity among sample comparison of all hosts and of *Helicia formosana* only.

Discussion

The fungal mycobiomes of trees at Fushan FDP are highly variable, and we uncovered only a small part of the reason for their enormous variability. When all host trees are compared, the average dissimilarity between any two trees is extremely high, (Fig. 3.13, Table 3.2). Samples become somewhat more similar on average when constrained to a single host, for wood and leaves, a result of the strong effects of host (Fig. 3.6) . But we do see an assemblage of fungi, 8 species in leaves and 10 in wood, "the core" that behave differently. Removing these fungi from consideration brings the mycobiome of their host, *H. formosana*, nearly back to background levels of dissimilarity among samples of the entire study, indicating that these few species are the ones through which host effects are manifested (Fig. 3.13).

These two sets of core fungi show differing spatial patterns (Fig. 3.12). In leaves, these core fungi are most consistently present in the southern valley, and are often completely missing in other areas of the study. In wood, they are more "loyal", and coexist more reliably with *H. formosana* throughout the plot. This may perhaps be due to the high rate of turnover in leaves, which are flushed mostly sterile (Arnold 2003), and are shed within 1 to several years, in contrast with the longer lifespan of woody tissues. Applying terminology proposed by Hamady and Knight (2009), core woody endophytes here may be best described by the "minimal" core model: they are few in number among a large and highly variable microbiome, but are consistently present throughout the study. In contrast, leaf endophytes may be described better by "gradient" or "subpopulation" core models, where a core group of associated microbes may establish with a particular host, but whose presence is highly conditional on space and environment.

Among the endophytes of all hosts, the central hill of FDP was important. We observed in leaves a homogenizing spatial effect with a radius of ~200 m, centered around the hill of the FDP (Fig. 3.7, 3.8). The hill of the Fushan plot was central point in

the community dissimilarity space of all the samples (Fig. 3.9). This is surprising, because the hilltop is a very distinct environment from the surrounding lowlands (Fig. 3.1), which have more in common with each other than with the hilltop. We were limited here by our coarse environmental data in the arguments that can be made for neutral spatial effects versus environmental filters as major predictors. However, this suggests that neutral effects may have been at work: the hilltop may be acting as a dispersal obstacle among the lowland areas, causing local structuring of microbial communities, especially the sheltered southwestern valley, and acting also as a common crossroads through which more widely dispersed microbes must pass. Being the exposed, high point of an area frequently subject to hurricanes, this hilltop may also be a local source of microbial species that are wind-dispersed. Conversely, where we see the most stable cooccurrence relationships are to be found in the relatively sheltered southwestern valley of the FDP.

The presence of a core group of microbes in a host can be seen as a kind of stabilization or structuring of a portion of a host's microbiome, as a result of interactions among hosts and select microbes. Extensive dispersal and disturbance can disrupt the effects of species interactions and beta diversity/local structure in communities and gene pools (Wright 1940, Cadotte 2006, Vellend 2010). We see that a single, relatively small land feature, a hill representing an 80m elevation gain, can alter the distribution of microbes of a landscape, disrupting seemingly strong microbe-host affinities. However, when defining core microbiomes, it may be important to consider the different organs of hosts as very different refugia for microbes: here the more stable environment of woody tissues appeared to host a more consistent assemblage of fungi. Similarly, the leaves *Helicia formosana* trees in the more sheltered southwestern valley held more consistent microbial communities than in more exposed areas of the plot. We conclude that even the strongest biological interactions between microbe and host can be disrupted by neutral processes or environmental changes. This implies that for a consistent core microbiome to develop, either local habitat or host must provide some measure of

stability through time and space for local community structuring of microbes to occur.

Bridge to Chapter IV: Agent-based model of the Foraging Ascomycete hypothesis

Metabarcode surveys such as the ones employed in chapter III have greatly improved our ability to pick up ecological signals in noisy systems like environmental microbial samples. For instance, in this chapter we detected the spatial trends of 18 species of fungi out of thousands that may be especially important and could be the targets for cultivation and further study. Ideally, such surveys are only the first step to exploring landscape patterns of microbes, to be followed by experimental tests of hypotheses generated by the patterns observed in these surveys.

However, In microbial ecological research, questions are often of scales that make experimental manipulations infeasible, and true replication cannot be accomplished even if repeated efforts were possible. Researchers dealing with large- or medium-scale questions often must simply move to the next biogeography study, in the hopes that it will inform their hypotheses. The wealth of information coming from these observational studies, combined with minimal ability to conduct experimental tests, pushes the field of microbial ecology deeper into the dilemma described by Lawton (1999), wherein ecologists spend too much time too close to the fireplace guessing at general laws of flames.

I do not have a solution for this dilemma. But in the work that follows, I found some intellectual satisfaction with combination of carefully targeted biogeographic studies (chapters II and III), and the integration of concepts and data from these environmental sampling efforts, as first principles and parameters of an agent-based model (ABM).

Bottom-up simulation models, such as ABMs, have many benefits beyond powers of prediction, which are often not even intended to predict in the strict sense (Epstein 2009). In one sense, to code a functioning ABM is to construct a precise and detailed formulation of a hypothesis. This process of expressing hypotheses in the form of functioning computer code is a type of intellectual honesty check - can we successfully simulate the natural processes in which we are interested? If we cannot, have we misunderstood our system? Successful recreation of patterns seen in nature using an ABM lends weight to a hypothesis, clarifies communication of ideas, and allows exploration of the logical outcomes from these principles (Grimm 2005). It can also promote the generation of new hypotheses and questions. Though far from a substitute for experimental manipulations, the construction of simulation models is a step towards more rigorous inquiry for large scale ecological questions.

CHAPTER IV

AN AGENT-BASED MODEL OF THE FORAGING ASCOMYCETE HYPOTHESIS.

Daniel Thomas, Roo Vandegrift, George Carroll, Bitty Roy

Contributions:

D. Thomas coded all scripts, ran all simulations and analysis of results, and wrote the paper. R. Vandegrift, B. A. Roy, and G. C. Carroll contributed conceptually to the ecological theory.

Abstract

Plant-fungal interactions are of paramount importance. Building useful ecological models of plant-fungal interactions is challenging, due to the complexity of habitat, varying definitions of biological basic units of interest, various spatial scales of dispersal, and non-linear, emergent properties of plant-fungal systems. Here we show that the bottom-up approach of agent-based models is useful for exploring the ecology of fungi. We constructed an agent-based model of the Foraging Ascomycete hypothesis, which proposes that some fungi maintain an endophytic life stage to enhance dispersal and bridge gaps in substrate in space and time. We characterized the general conditions in which dispersal through leaves may be worth the metabolic and fitness costs of endophytism. We also modeled possible effects of deforestation on leaf endophytes, highlighting how agent-based models can be useful for asking questions about changing ecosystems. In the competition simulations, leaf- borne fungal dispersal allowed fungi with lower dispersal capabilities to compete effectively against fungi whose spores were dispersed at much greater distances and concentrations. However, this benefit was reduced or lost without sufficient retention of endophyte infection in the canopy, or with deforestation.

Introduction

Plant-fungal symbioses are ancient (Stukenbrock 2008, Redecker 2000), ubiquitous, and important (Vandenkoornhuyse 2015). All large organisms are observed to host complex microbiomes (Rosenburg 2010), and plants are no exception, with both epiphytic and endophytic fungi and bacteria present on and within all tissues (Rodriguez 2009, Rosenbleuth 2006). These symbionts are known to be extremely diverse (Arnold 2000, Arnold 2007) and some are important to plant health (Mejia 2008, Arnold 2003a, Porras-Alfaro 2011, Rodriguez 2009).

Extensive literature has explored benefits conferred to plant hosts by endophytic fungi and bacteria. However, benefits conferred to the endophytic microbes themselves are not as well explored. In particular, the reduced reproductive activity, and costly array of unique metabolites produced by fungal endophytes to maintain the endophytic phase (Carroll 1983, Kusari 2012) make the endophytic life history strategy seem like an uncertain investment, from the fungal perspective. Nevertheless, diverse fungi are observed that can both decompose wood or litter, and exist as an endophyte in a different, living host (Lodge 1997), suggesting there are benefits to the life history strategy.

Carroll (1999) proposed that some endophyte-competent fungi may utilize an endophytic phase to increase dispersal, a concept known as the Foraging Ascomycete hypothesis (Fig. 4.1). Thomas and Vandegrift et al. (2015) expanded this concept, proposing that some endophytic fungi utilize the endophytic phase to bridge spatial and temporal gaps in substrates and suitable environmental conditions. We use the term *viaphytic* ("via," road; "phyte," plant) recently proposed by Nelson (2016), to describe endophytic fungi that are observed to transfer from endophytic infections to woody substrates. This term is distinct from "endophyte competence" (Hardoim 2008), which denotes the ability of a microorganism to endophytically infect a host, but does not inform the ability to disperse beyond this endophytic state.

Endophytes provide important services to forests, through their hosts. Some endophyte species have been shown to increase increase drought tolerance or even disease resistance, often in the form of local "adaptation" (Rodriguez 2009, Giaque 2013). Presumably, many more important plant symbionts remain to be discovered (Gazis 2012, Suryanarayanan 2009). Even if not directly contributing to plant fitness, the community at large of plant-associated, commensal microbes may be important to preventing disease (Herre 2007). Microbial partners to plants may become more important in the current context of climate-change associated stresses (Woodward 2012). Land use changes have been shown to induce changes in microbial population dynamics (Arnold 2003b, Rodrigues 2013). In the status quo of rapid change, models of effects of environmental change of plant-microbial communities are increasingly pertinent.

Attempts to model microbiome community assembly and dynamics are in their infancy (Nemergut 2013). The astounding diversity of microbiomes, the complexity of real-world environmental systems, and the particular difficulties of quantifying fungal individuals, all appear to have stunted the development of robust and useful ecological models for fungi. Here an Agent-Based Model (ABM) approach (Grimm 2005) is employed to examine the Foraging Ascomycete hypothesis, as set of competition "experiments" among viaphytic and non-viaphytic fungi. Agent-based models take a "bottom-up" approach to understanding systems of many interacting actors, often including an explicitly spatial and stochastic behaviours that can prove difficult to realistically model with traditional population and community ecology mathematical models. For these reasons, ABMs may prove increasingly useful in future ecological modeling of fungi and other microbes.

The Foraging Ascomycete agent-based-model is presented here using the standard 'ODD' (Overview, Design concepts, and Details) protocol for describing agent-based models (Grimm 2006, Grimm 2010). Following this several sets of simulations are

reviewed, which explore the theoretical benefits and limits of viaphytism as part of a fungal life-history strategy. In addition, several simple scenarios of deforestation are simulated, to highlight the potential for ABMs to help in the understanding of microbial ecology in the context of current environmental challenges.

Figure 4.1. Visualization of the Foraging Ascomycete hypothesis, also known as "viaphytism". Leaves are infected endophytically by spores, then act as dispersal vectors of fungi to new substrates.

Methods

Methods I. ODD protocol

Purpose

The purpose of this model was to explore the feasibility of Viaphytism (Carroll 1999, Thomas and Vandgrift 2015, Nelson 2017), as part of a fungal life history and dispersal strategy. An ABM approach was used to explore the possible advantages to fitness and

dispersal conferred by endophytism in fungi, by enacting competition-type scenarios among fungi with and without endophyte-competence.

Entities, state variables, and scales

Three agent types were placed on a spatial grid: trees, fungi, and woody debris.

Tree-agents represent individual adult trees with diameter-at-breast height greater than 10 cm. State variables of trees included position, leaf dispersal ability, state of endophyte infection (positive or not), and rate of endophyte loss. Leaf dispersal ability is a positive integer, where larger values represented longer-range and more plentiful leaf deposition (see submodels). State of endophyte infection denoted whether a tree carries the endophyte stage of an endophyte-competent fungus in its leaves. Successful infection from fungal spores changed a tree-agent's infection state to positive. Infections could be lost, and this loss was controlled by the endophyte-loss state variable, a number between 0 and 1, representing the probability that an infection was lost at each timestep.

A fungus agent represented a mycelium, resulting from a single reproductive event, either a spore- or leaf-vectored inoculation of wood. State variables of fungi included: position, spore dispersal ability, stored energy (biomass), and endophyte-competence. Like leaf dispersal in trees, spore dispersal ability was a positive integer, with larger values representing longer-range and more plentiful spore deposition across the landscape when sporulation occurs (see submodels). Energy was representative of biomass and potential energy gain from decomposition of woody debris. Sufficient energy stores allowed for a sporulation event. Endophyte-competence denoted the ability of a fungus to reside as an endophyte in leaves after infection of leaves from fungal spores. In terms of the model, "endophyte competence" indicated whether a fungus can change the endophyte infection status of a tree during a sporulation event, and then disperse through leaves, or viaphytism.

Woody debris agents represented the biomass deposited on the forest floor from the canopy. State variables of woody debris were position and stored energy (biomass). New wood were given a starting amount of energy, and this wood biomass was converted incrementally to fungal biomass if fungi were present in the cell.

Grid cells were not given attributes, except for the agents they held, and their location, in the form of x and y coordinates. For all the scenarios examined, the grid spanned one square hectare (100m by 100m), wherein each grid cell represented one square meter. The grid was toroidal, and agents of all types could occur at all grids, though fungi did not persist for long periods without woody debris also present because of energy constraints.

Model-wide, environmental state variables included the rate of deposition of new woody debris, number and spatial clustering parameters. Trees could be removed at any time to simulate to effects of deforestation.

Process overview and scheduling

Time steps began with the placement of new woody debris on the landscape. Following this, agents were chosen randomly to act, regardless of type. See Figure 4.2 for a summary schematic of model processes for one time step.

Fungus agents began time steps with a test of their biomass (energy) reserves. If energy was high enough, sporulation occurred, possibly instantiating new fungus agents on woody debris. If the sporulating fungus was endophyte-competent, the spores could also change the endophyte infection status of trees on the landscape to positive. Sporulation resulted in a loss of energy for the parent fungus agent. Following this, fungus agents decomposed the woody debris available in their grid cell, resulting in a gain of energy for each fungus agent present and a loss of stored energy in the woody

debris. If the woody debris at a grid cell had died, fungus agents continued to respire, subtracting from their energy each turn until they have energy < 1, upon which they died.

Tree agents began by dropping leaves. If a tree had a positive endophyte infection state, these leaves dispersed to the landscape and could inoculate woody debris, instantiating a new fungus agent. Trees could also be removed from the landscape, which if requested occurred at the very beginning of a step, before deposition of woody debris.

Woody debris agents were placed at the beginning of each time step, in multiple random locations at the start of each step. The exact number of woody debris agents laid down each step was random. The initial energy in each varied and the total energy represented by all the new woody debris approximately equalled the New Wood Energy state variable set by model user. Woody debris agents then tested their biomass (energy) state variable: when energy < 1, the agent is removed from the landscape.

After all agents present on the landscape acted, data collection took place, and the time step was complete. Under model default settings, each time step was intended to represent an ecologically relevant period of approximately 3 months.

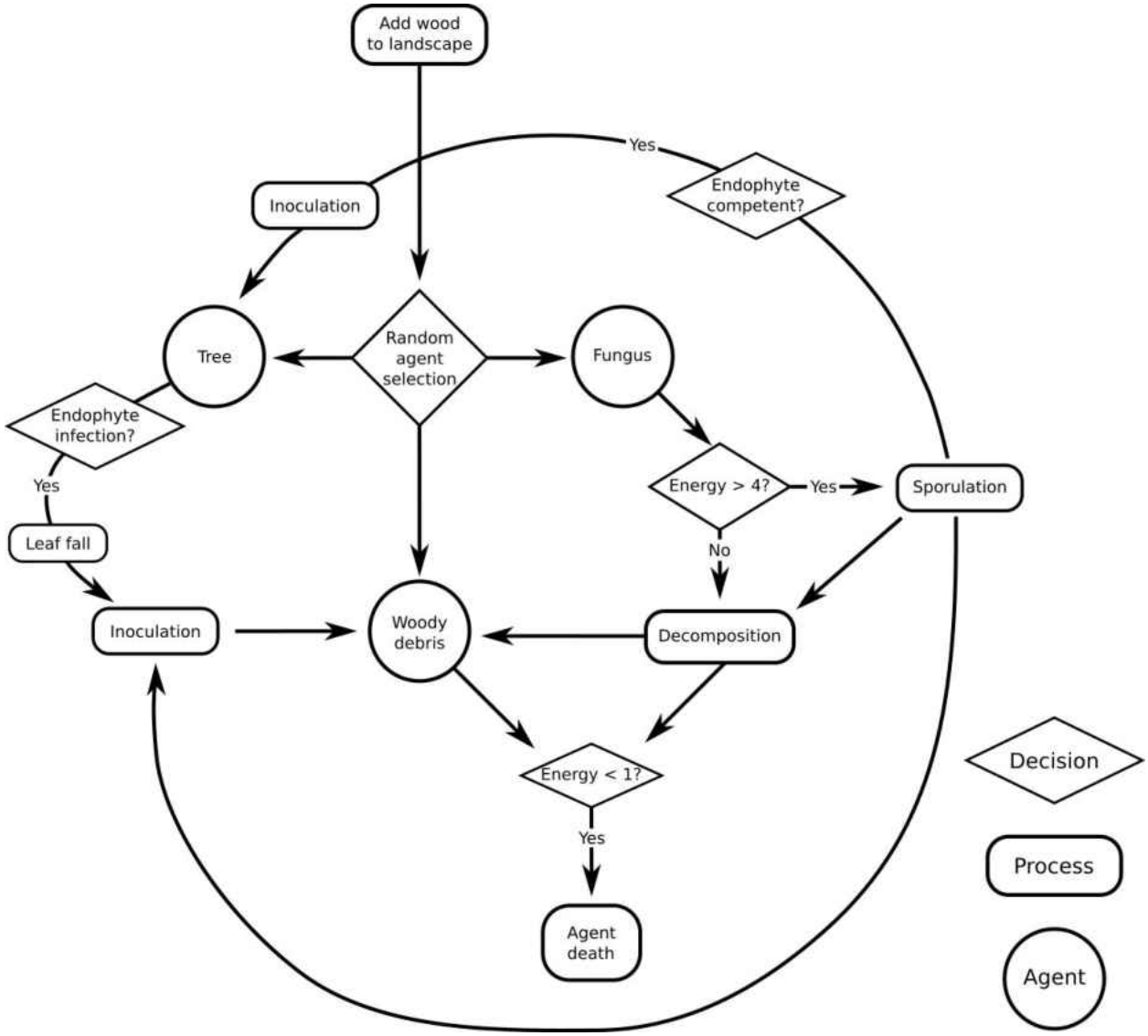


Figure 4.2. Schematic of processes possible during one timestep of the model. For a higher resolution image click here.

Design concepts

Basic principles

This ABM was primarily a model of dispersal and competition among fungi. Patterns of spore dispersal at various scales were measured in Galante et al (2011), Norris et al. (2012), Peay et al (2012), and others. These studies showed that the negative exponential family of functions can be parameterized to fit abundances and probabilities of spore-dispersal in nature. Leaf fall has also been shown to be well described by

exponential decay functions (Ferrari and Sugita 1996). These well-established patterns of dispersal served as first principles in this model, guiding the behavior of both tree and fungus agents.

However, the purpose of the model was to explore the hypothesis that some fungi utilize an endophytic life stage to enhance dispersal and to persist on the landscape during times of scarcity, intense competition, or environmental stress (Carroll 1999, Thomas and Vandegrift 2015). This viaphyte life history strategy, where some fungi alternate endophytic and free-living phases, was a basic principle of the model, and the focus of the simulations presented below.

Emergence

Emergent properties of interest were: (1) emergence of endophytism is a beneficial life history strategy despite its costs, (2) differential responses of viaphytic fungi to changes in substrate (woody debris), as compared to non-viaphytic fungi, (3) changes in abundances of endophyte-competent fungi that result from changes among spatial relationships of trees, including deforestation, that weren't well modeled as a function of simple abundances of trees.

Adaptation, Objectives, Learning, and Prediction

Fungus agents were intended to seek reproductive success, which can be measured either by number of substrates occupied or sporulation events. However, fungus agents were not given the ability to modify their behaviors to increase fitness. As such, they did not take any measure of success, memory of past events, or predictions of future conditions, into account during their actions.

Sensing

Fungus agents' decisions were based primarily on internal sensing of biomass (stored energy) to decide when to initiate sporulation and external sensing of distance woody debris and trees when sporulating, to determine the probability of infection.

Inoculation of woody debris by endophyte infected trees also sensed the distance to woody debris to calculate probabilities of infection.

Interaction

Interactions among fungi were indirectly competitive, mediated through wood debris agents. Woody debris agents were consumed by fungal agents as a source of energy, and the presence of existing fungus agents associated with a woody debris agent reduced the likelihood of establishing new fungus agents on a woody debris agent.

Stochasticity

Several stochastic processes were used in the model to emulate the variable environment of forest ecosystems. Amount of wood deposition per step, number of successes in sporulation/inoculation, initial placement of trees and woody debris, and methods of tree selection in deforestation all involved stochastic selections of agents and locations. These are described in the sub models.

Observation

At the end of each step in the model the following were recorded: total numbers of fungus agents, woody debris agents occupied by fungus agents of both endophyte-competent and non-competent fungi, total sporulation events by both types of fungi, percent of trees infected by endophytes, and for deforestation scenarios, total number of trees on the landscape.

Initilization

Model default density of ~600 trees in a 1 ha plot were intended to approximate that of wet tropical forests (Crowther et al 2015). Initial conditions of the model were intended to emulate a recent small disturbance in a forest landscape, where a larger than usual amount of uncolonized woody debris has been randomly deposited. Unless otherwise specified, all model runs began with one fungus agent of each type, randomly

associated with a woody debris agent. These initial fungus agents were assumed to have established themselves and began the model with a starting energy sufficient to sporulate 2 or three times. Endophytism in the model could be disabled, allowing competition experiments between two non-viaphytic fungi. Dispersal coefficients were assigned to both types of fungi, and to trees for dropping leaves, though this last setting is typically held at a default value from leaf fall data (see submodels). Default initial Woody debris agents had a total biomass/energy of 30 (this could be changed by the user). Rate of new woody debris deposition on the landscape could also be set prior to initialization, though this was typically held a default value found to allow aggressive, non-viaphytic fungi to persist on the landscape. Initialization states were intended vary among model runs, to explore the benefits and limits of a viaphyte-style life history strategy.

Input

Deforestation scenarios required time-series input data, in the form of timing, intensity, and spatial nature of tree removal. Otherwise the model does not require input data.

Submodels

Submodels are listed below in Figure 4.2 schematically as processes. In addition, we describe procedures for initial placement of trees, and two deforestation submodels.

Wood deposition

Wood deposition (Fig. 4.3) was given total energy budget per timestep (A), that was defined by the user/defaults before initiating a model run. To simulate the variety of sizes of woody debris that occur in forest settings, however, each new woody debris agent (W) was given variable (random) initial energy (e), taken from the iteratively smaller range of energy remaining. As agents were added, a tally of energy used ("a") was maintained. This tally "a" ultimately approximately equalled the wood deposition rate given by the user/default, and the submodel exited.

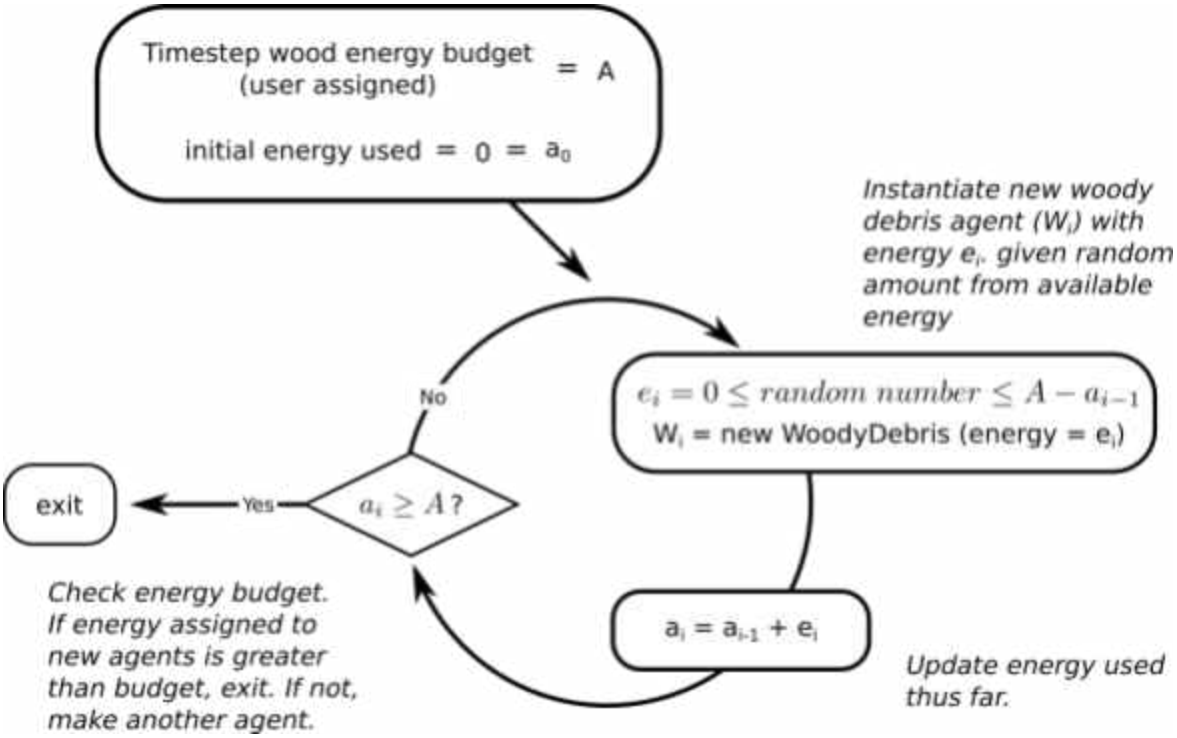


Figure 4.3. Wood deposition submodel. For a higher resolution image click here.

Sporulation and inoculation

Calculation of probability of infection of a woody debris or tree agent from spores was an exponential decay function of distance ("x") from self (fungus agent), multiplied by a dispersal ability coefficient ("D") assigned by the user (Fig. 4.4). Viaphytic and non-viaphytic fungi could have been – and usually were – assigned distinct dispersal abilities. Probability of inoculation of woody debris agent was furthered multiplied by the fraction of current, remaining energy ("E_c") over starting energy ("E_i"), to give a handicap to colonization of woody debris agents by new fungi, if the wood is already inhabited by other fungi.

Leaf fall and leaf-vectored wood inoculation

Leaf fall was treated similarly to sporulation (Fig. 4), except that it occurred at every time step, as an action of all Tree agents, without any energy budgeting. For the

purposes of this study, leaf fall for all tree agents was calibrated at d=4. The equation for determining the probability of inoculation of a Woody debris agent was identical:

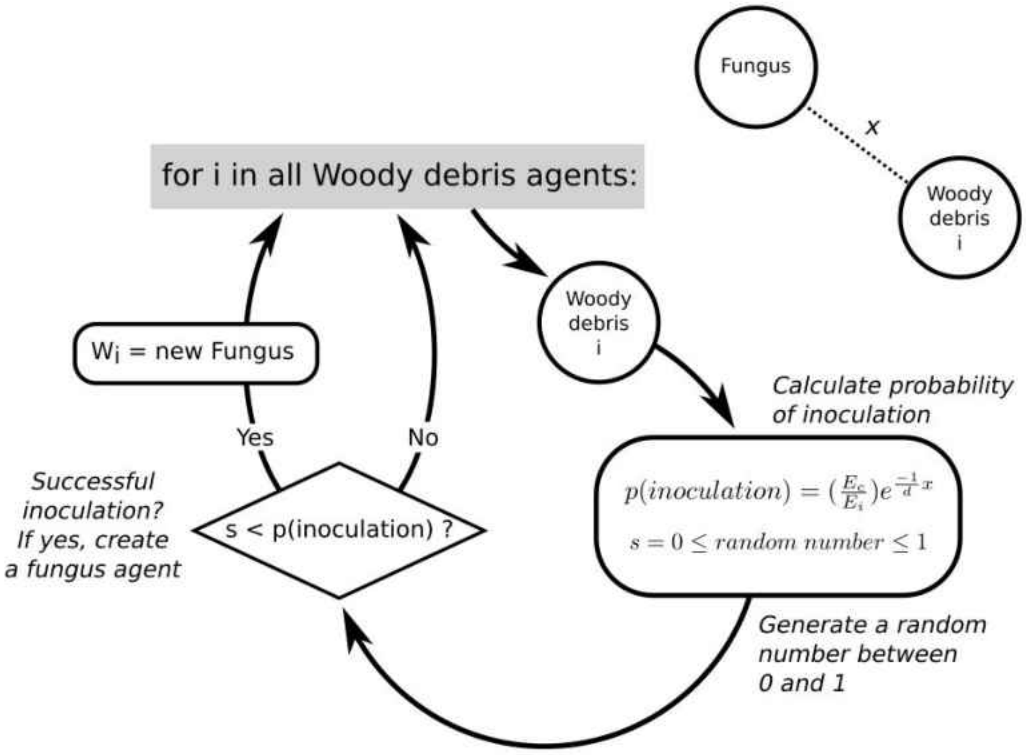
$$p(inoculation) = (\frac{E_c}{E_i})e^{\frac{-1}{d}x}$$

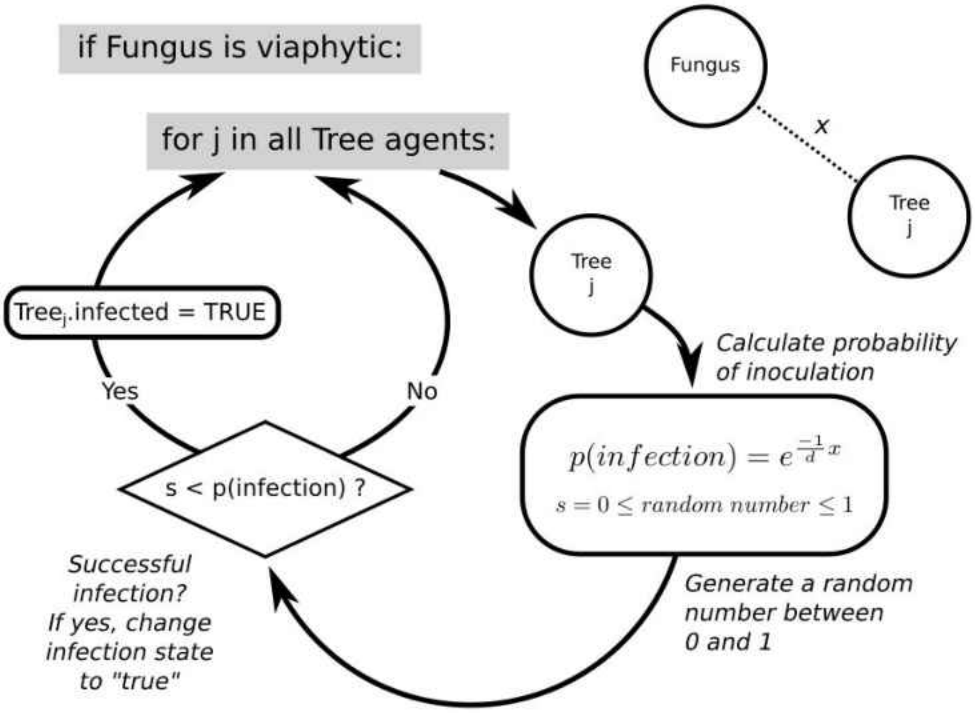
"E_c" and "E_i" are current and initial energy, "x" is the distance between Tree agent and Woody debris agent, and "d" is the dispersal ability coefficient for trees, usually held at d=4.

Decomposition

Decomposition as modeled here was a simple one-way transferral of energy from Woody debris agents to their associated Fungus agents. Every time-step, each Fungus agent on a grid cell with a Woody debris agent gained one energy, and caused the Woody debris agent to lose one energy. Thus, a cell with numerous Fungus agents would show rapid decomposition of the resident Woody debris agent, and became increasingly difficult to for new Fungus agents to access. After a Woody debris agent dropped below energy=1, it was removed from the model. Fungus agents at this empty cell then respired away stored energy at a rate of one per step until dropping below one unit of energy, followed by removal from the model, unless a new Woody debris agent randomly arrived at the cell.

Figure 4.4. (next page). Sporulation submodel, for both woody debris inoculation and endophyte infection of trees. For a higher resolution image click here.


Fungus
X
Woody debris i
for i in all Woody debris agents:
Woody debris i
Calculate probability of inoculation
$p(inoculation) = (\frac{E_c}{E_i})e^{\frac{-1}{d}x}$
$s = 0 \leq random\ number \leq 1$
Generate a random number between 0 and 1
s < p(inoculation) ?
Yes
No
W$_i$ = new Fungus
Successful inoculation? If yes, create a fungus agent


if Fungus is viaphytic:
Fungus
X
Tree j
for j in all Tree agents:
Tree j
Calculate probability of inoculation
$p(infection) = e^{\frac{-1}{d}x}$
$s = 0 \leq random\ number \leq 1$
Generate a random number between 0 and 1
s < p(infection) ?
Yes
No
Tree$_j$.infected = TRUE
Successful infection? If yes, change infection state to "true"

Tree placement

Initial tree placement on the model landscape followed a "Thomas" process (Thomas 1949), controlled by three, user-defined parameters: the poisson-process rate of parent points that become centers of tree clusters ("kappa" or κ), a secondary poisson-process rate for child points that become Tree agents ("mu" or μ) the spread (variance) of child points ("sigma" or **σ**). Default settings were intended to create approximately 600 trees per hectare (Crowther et al 2015).

Tree removal

Tree removal could be programmed into model runs at any time. Two types of tree removal were included as functions in the model, to emulate two broad categories of deforestation: (1) thinning, or selective logging, where trees are removed at +/- the same rate, throughout the landscape, interspersed among leave trees, or (2) fragmenting, where contiguous blocks of forest are removed. The first attempted to emulate the results of selective logging, often in the form of "highgrading." The second was intended to model land use conversions - homesteading, conversion to agriculture, etc. (Kettle and Koh 2014).

Thinning of trees required one argument from the user, the intensity of the thin. This number is between 0 and 1, indicating the proportion of trees to be removed, each of which are randomly, independently selected from the pool of the entire set of trees on the landscape.

Fragmentation of forest accepts two arguments, the number and radius of fragments. Fragment center locations are assigned randomly, then all trees within the user-assigned radius from each center are protected, and the remaining trees are removed from the model (Fig. 4.5).

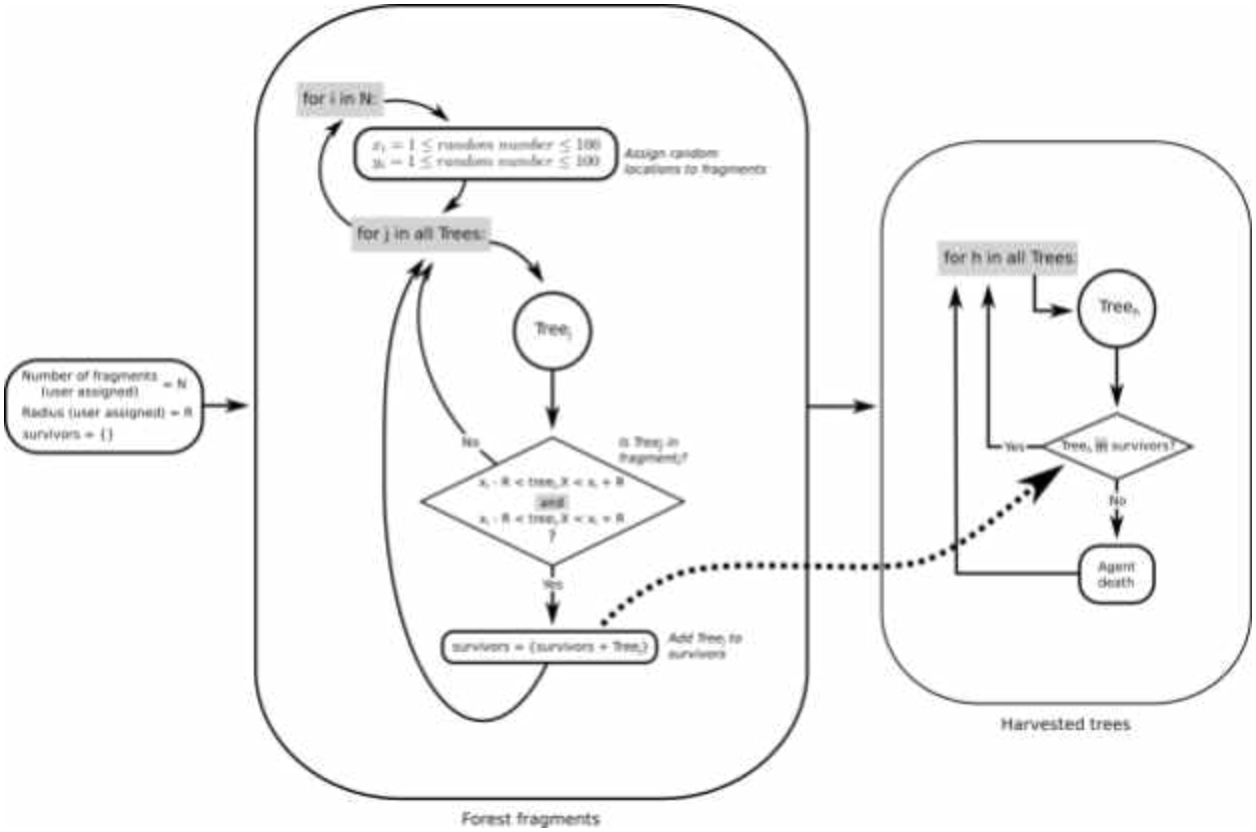


Figure 4.5. Forest fragmentation submodel. For a higher resolution image click here.

Methods II: Simulations

Basic behaviour of the model was characterized by first defining a non-viaphytic, "typical" fungus. Negative exponential models with various dispersal coefficient settings were compared with estimates of spore dispersal from empirical studies (Norros 2012, Peay 2012, Galante 2011) to characterize a well-dispersed fungus. This well-dispersed fungus was then tested on the model landscape using a parameter sweep of dispersal coefficients in the "tropical forest" of default model settings. Next, a model viaphytic fungus was defined as the lowest-dispersing viaphytic fungus that could cooccur and compete successfully with the model non-viaphytic fungus. Once defined, these model viaphyte (d=2) and non-viaphyte (d=10) fungi were used as default settings for Fungus agents in subsequent simulations to explore properties of the model.

Additional tests included:

- Sensitivity of fungus agents to initial amounts of wood on the landscape, and to regular rates of wood deposition after initialization.
- Sensitivity of viaphytic fungi to residence times of endophytic infection in host-trees.
- Sensitivity of viaphytic fungi to deforestation, using three deforestation scenarios: (1) one-time thinning of various intensities, (2) serial thinning, and (3) fragmentation of the 1 ha forest plot into 15 m-diameter clusters.

In most parameter sweeps, 100 simulations of 50 timesteps were run for each level of the variable of interest. In deforestation simulations, which were run for 100 timesteps, with "harvests" introduced at timestep 51.

Results

Model fungus agent calibrations

A non-viaphyte fungus with a dispersal coefficient of d=10 was found to fit expectations from empirical data of aggressively dispersed fungi on real landscapes and to persist reliably on model landscapes (Fig. 4.6). At lower dispersal abilities, populations often went to zero, or had not finished decomposing initial wood deposits within 50 steps. At d=10, Fungus agents were able to fully colonize initial deposits wood on the landscape, then maintain a lower, steady population where woody debris agents inoculated by fungi were approximately equal to the amount of new woody debris energy deposited on the landscape each turn.

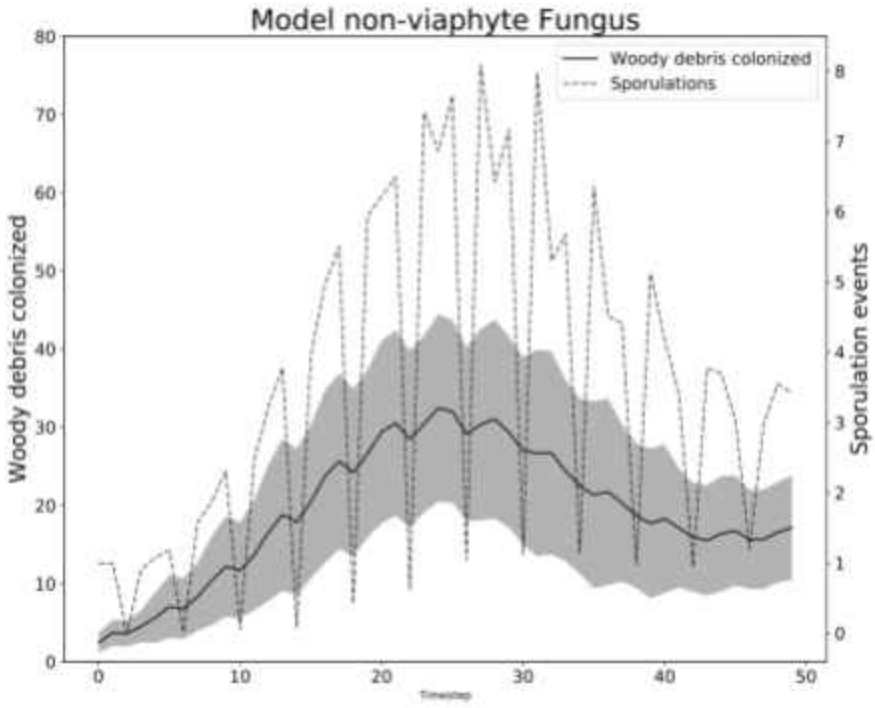


Figure 4.6. Behavior of model non-viaphytic fungi on the default model landscape. Error lines are one standard deviation from the mean. For a higher resolution image click here.

A viaphyte fungus with a dispersal coefficient of d=2 was the lowest-dispersing fungus that maintained a balanced competition with our model non-viaphytic fungus (Fig. 4.7). Below this dispersal level (d<2) for viaphytes the model non-viaphyte fungus agents clearly outcompeted the viaphytes, keeping infected trees and inoculated substrates to near zero levels. The reverse was true above (d>2) this dispersal level.

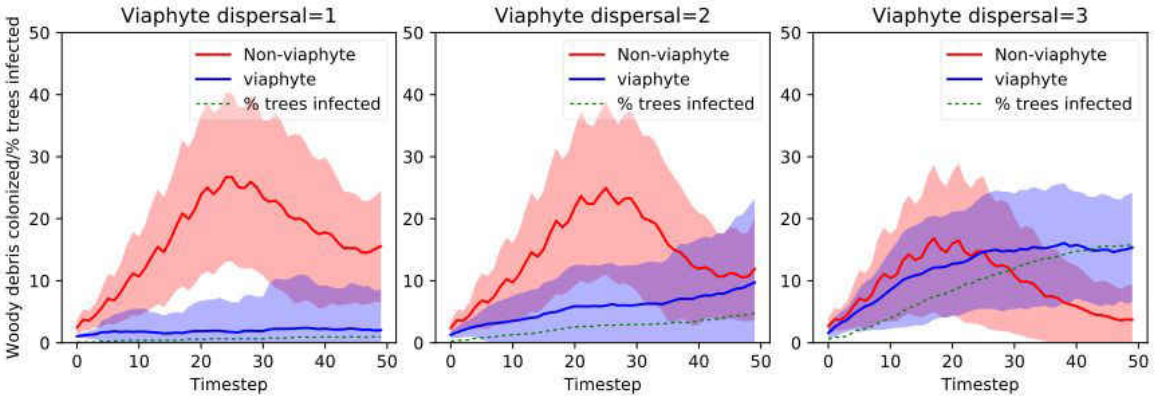


Figure 4.7. Competition of viaphytic fungi with various dispersal abilities against a model non-viaphytic fungus. For a higher resolution image click here.

Importance of initial wood deposition and subsequent rates of wood deposition

Model, non-viaphytic Fungus agents increased their populations rapidly when presented with large abundances of woody debris (Fig. 4.8). Larger initial deposits of wood on the landscape were often consumed as or more quickly than small abundances. Higher initial abundance of wood was equivalent to more continuous distributions of woody debris, with fewer gaps in substrate, making all woody debris on the landscape generally more available. These conditions allow exponential population growth of fungi and quicker consumption of wood, despite absolute wood biomass being much greater.

Subsequent, per-step wood deposition was important for this reason also (Fig. 4.9), as sufficient wood was required to sustain fungus agents in the long-term, but also to enable the initial explosive exploitation of wood on the landscape by bridging gaps between islands of existing substrate reserves on the landscape. Less-aggressively--dispersed model viaphytes responded less dramatically to abundances of woody debris on the landscape, taking longer to reach peak abundances. With viaphytic fungi, abundances was also highly influenced by the increasing number of endophytically-infected trees.

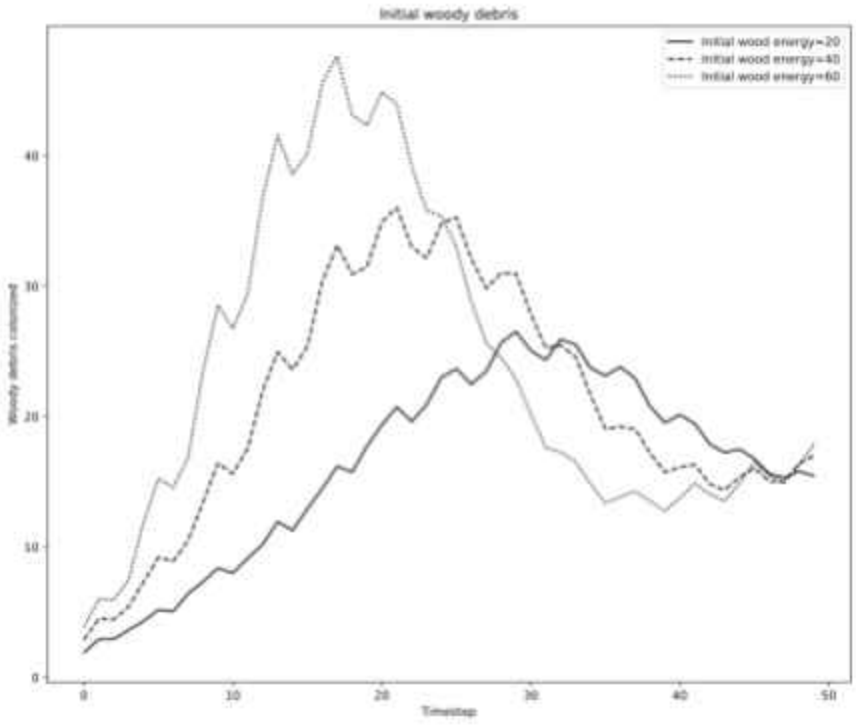


Figure 4.8. Response by fungi to varying amounts of substrate on the landscape. For a higher resolution image click here.

Importance of residence times of endophytic infection in host-trees

All benefits conferred by the endophytic phase were contingent upon a low rate of loss by trees of their endophyte infection. Under model defaults, endophyte loss greater than 5% per time-step caused loss of all competitive advantage by model viaphytes (Fig. 4.10).

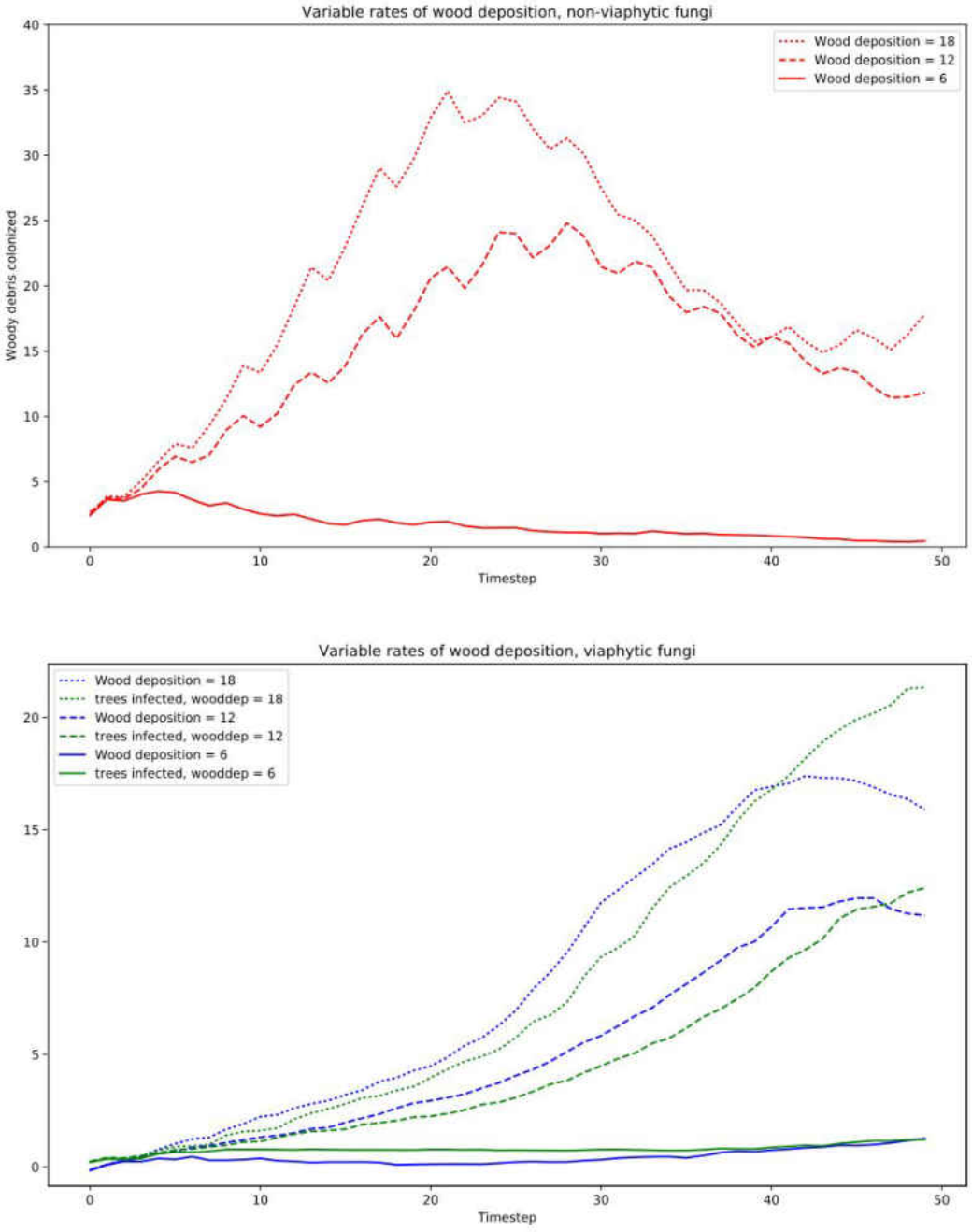


Figure 4.9. Differential response by model non-viaphyte and viaphytes to per-step wood deposition rates. For a higher resolution image click here.

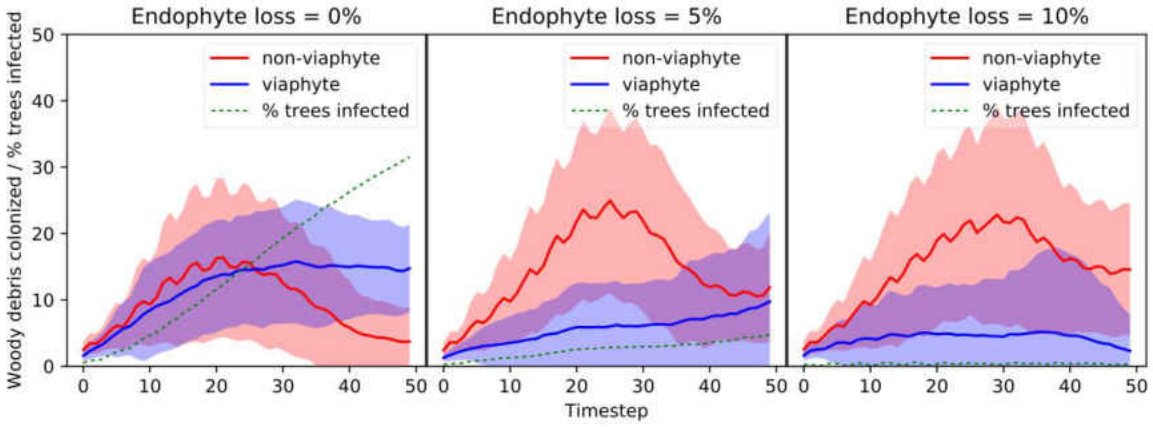


Figure 4.10. Effect of endophyte infection loss rates on viaphyte success. [For a higher resolution image click here](#).

Deforestation and its ecological consequences

Consequences of removing trees depended on the intensity, timing, and spatial arrangements of the removal of trees. Without any cutting, model viaphytes showed an increasingly stable presence on the landscape, as the reservoir of fungus in the canopy incrementally increased (Fig. 4.11a) Drastic thins (70-100%) reduced this stability (Fig. 4.11b). Lighter thins (10-30%) appeared to affect established populations of endophytes minimally. Serial thinning, or 10% removal of trees every 10 steps, beginning at step 51, had less impact on viaphyte populations than the comparable event of thinning 40% of trees at once (Fig. 4.11c). As modeled here, fragmentation of forest had similar effect to comparable thins, but endophyte populations remained stable at higher rates of removal, with viaphytic fungi recovering to competitive abundances even at 70% removal of trees (Fig. 4.11d).

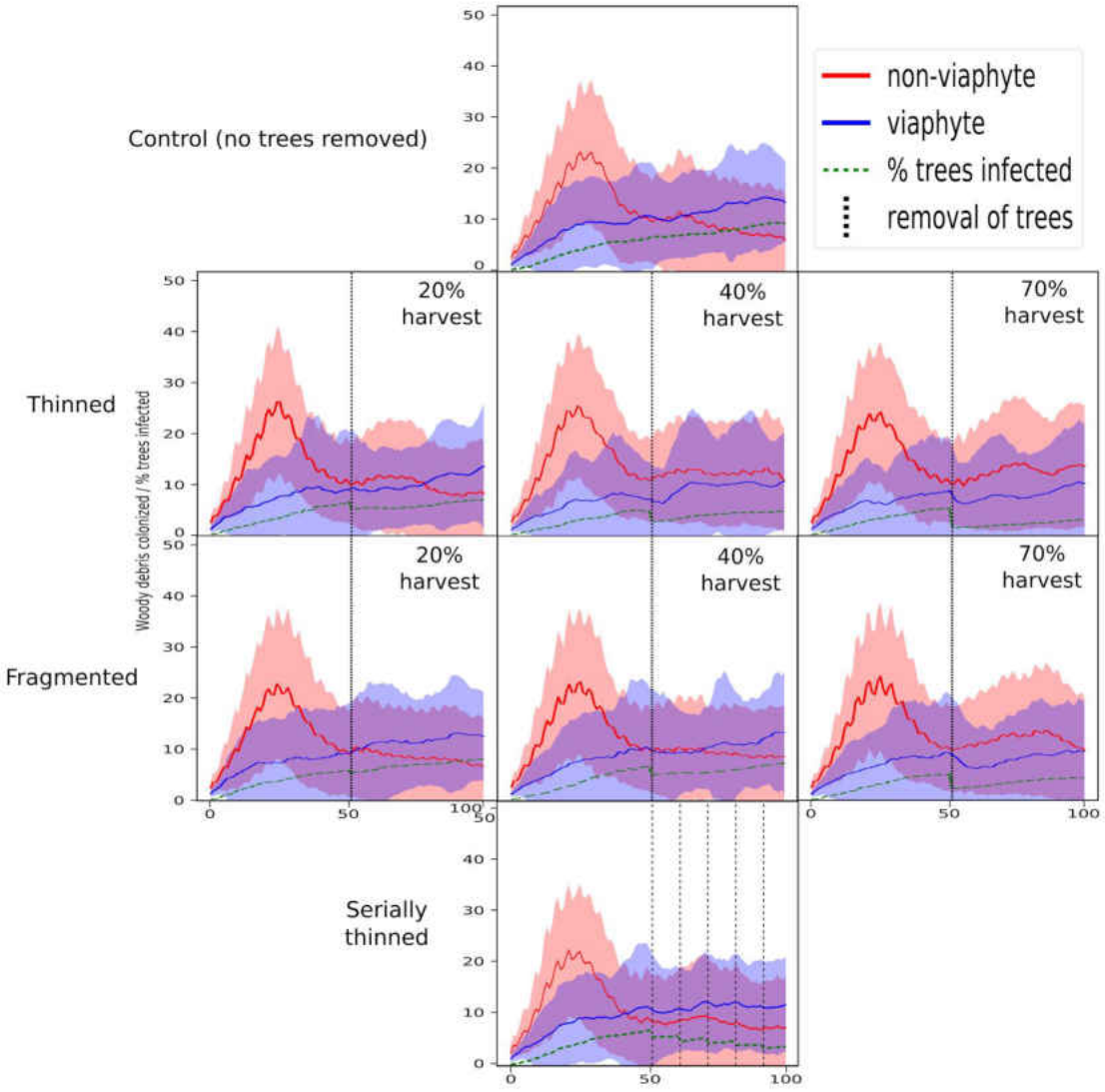


Figure 4.11. Deforestation scenarios - thinning, fragmentation, and serial thinning. For a higher resolution image click here.

Discussion

Under some conditions modeled here, the utilization of leaves as dispersal vectors and refugia in times of scarcity allowed a fungus to persist and compete on a landscape of other, far-better dispersed fungi.

Persistence of fungi on the landscape was due, in large part, to the highly spatially and temporally autocorrelated pattern of colonization that results from spore dispersal modeled as a negative logarithmic decay pattern: Fungus agents had to bridge gaps in woody debris over distance and time to persist on the landscape. Thus fungi without viaphytism could only overcome this limitation simply by increasing dispersal (or other strategies not examined in this model). This created a negative feedback, since most spores must fall locally in order for some percent to reach farther distances, meaning that once established, a fungus typically rapidly colonized and consumed all local substrate. This "boom and bust" cycle of exponential growth and collapse was risky. If new substrate were not found, local extinction was very possible. In addition, both in the model and in nature, offspring of a sporulating fungus are often vegetatively incompatible with parents (Paoletti 2016), and were, in one sense, competitors of their own kin when consuming substrate.

Viaphytic fungi, alternatively,could take refuge in - and augment dispersal with - an endophyte phase. Neither leaves or spores of these endophytes were modeled as very widely dispersed. Instead, viaphytic fungi relied on an incremental but steady increase over time on the landscape. The implications of this positive-feedback would surely be increased if wood deposition were spatially linked to the presence of canopy trees - as written the model allocated random dispersal of woody agents across the landscape, regardless of the presence of canopy. Certain types of deforestation were implicated here as more problematic than others for allowing endophytes to persist on the landscape. Small scale, regular disturbances were more sustainable in terms of endophyte populations than large single harvesting events, reducing the chance of stochastic removal of a species from the landscape and enabling populations of endophytes to regenerate into other trees from remnant trees. Host preferences are not modeled here, and would exacerbate any negative effects of selective thinning. Fragmentation as modeled here was less impactful on endophyte populations than comparable dispersed

thinning, as it left large blocks of contiguous forest to remain on much of the landscape. Fragmentation as modeled here is of small scale, and does not necessarily reflect larger scale fragmentation such as is occurring throughout much of the tropics (Kettle and Koh 2014), however.

This study and other studies (Boswell 2012) have demonstrated just some of potential uses for simulation-based models in mycology. Effort must be made to find realistic and useful modeling solutions for mycology, as understanding the bewildering diversity and complexity of ecology of fungi becomes ever more urgent in a changing world.

Bridge to Chapter V: Understanding and mitigating some limitations of Illumina© MiSeq for environmental sequencing of fungi.

The preceding agent-based model yielded interesting results that may help clarify and expand the discussion of the Foraging Ascomycete hypothesis. I feel that just as important as the actual results of the model runs are the methodological implications: I included an ABM in my dissertation work to highlight their potential usefulness to mycologists asking ecological questions. We continue the methodological theme with the final chapter, a report on two phenomena associated with next-generation sequencing read libraries: unequal sampling error and index misassignment. Both are known by industry and research communities, but in combination they have special consequences for environmental metabarcoding studies such as are used by microbial ecologists. The methods employed in Chapter III, and related methods, are the current standard protocols for fungal metabarcoding studies. When I first began to learn about illumina(c) MiSeq datasets, there was little information about these apparently ubiquitous sources of error. They still have yet to be acknowledged in their full importance, though researchers such as Jusino et al. (2016) and others are beginning to report on the full extent of the issue. Here I give my small contribution to this important, ongoing discussion.

CHAPTER V

UNDERSTANDING AND MITIGATING SOME LIMITATIONS OF ILLUMINA© MISEQ FOR ENVIRONMENTAL SEQUENCING OF FUNGI.

Dan Thomas, Roo Vandegrift, Graham Bailes, Bitty Roy

Contributions:

D. Thomas and G. Bailes conducted laboratory work. B. A. Roy contributed to conceptual discussion and provided lab space, reagents and funding. R. Vandegrift provided conceptual discussion and feedback on mansucript. D. Thomas wrote the manuscript, coded all scripts, and conducted analyses.

Abstract

ITS-amplicon using illumina miseq sequencing platform are the current standard tool for fungal ecology studies. Here we report on some of the particular challenges experienced while creating and using a ribosomal RNA gene amplicon library for an ecological study. Two significant complications were encountered. First, artificial differences in read abundances among OTUs were observed, apparently resulting from PCR bias at two stages: PCR amplification of genomic DNA with ITS-region illumina-sequence-adapted-primers, and during Illumina sequencing. These differential read abundances were only partially corrected by a common variance-stabilization method. Second, tag-switching, or shifting of amplicons to the incorrect sample indices, occurred at high levels in positive controls. An example of a bioinformatic method to estimate the rate of tag switching is shown, some recommendations on the use of positive controls and primer choice are given, and one approach to reducing potential false positives resulting from these technological biases is presented.

Key words: fungi, index-switching, ITS, PCR bias, OTU splitting, positive controls, tag-hopping, tag-switching, variance-stabilization

Introduction

ITS or 16s amplicon libraries sequenced with Illumina © MiSeq sequencing technology are the current standard tool for bacterial and fungal ecology studies. The power of next generation sequencing technologies like MiSeq, however, are balanced by their limits and biases, fueling a lively discussion on their proper implementation in microbial ecology (Pinto 2012, Lindahl 2013, Persoh 2013, McMurdie 2014 Tedersoo 2015, Nugyen 2015 and 2016, Taylor 2016). Our study continues the discussion of some of the issues surrounding metabarcoding methods, with a focus on (1) the difficulties of ecological interpretation of read abundances from Illumina-sequencer results and (2) misassignment of sample indexes to reads, also known as "index-switching," "tag-hopping", or "tag-switching."

Read abundances resulting from next generation sequencing studies with multiple samples and multiple biological units of interest ("OTUs") are an example of a multinomial, "roll-of-dice" sampling process at both levels (Anders 2010, McMurdie 2014). Differences in read abundances among samples or among OTUs within samples may represent real biological differences, but they must first be tested and adjusted for the natural differences that occur when "dice are rolled." Here we observe that initial PCR amplification of the ITS region of environmental samples of genomic DNA and Illumina-platform sequencing of the resulting libraries may both introduce this family of errors into distributions of read abundances. The variability of read abundances from next generation studies are probably most effectively modeled with negative binomial distributions (Anders 2010). Failure to adequately correct for these sources of variation could result in read distributions that give the impression of ecological patterns, such as

species abundance distributions as predicted by neutral models (Baldridge 2016).

Another source of bias in next-generation sequence studies is the erroneous assignment of sample identity to a read, or tag-switching. The mechanisms for this error are, to date, poorly explained and seem to vary with platform (Sinha 2017, Carlson 2012). Prescriptions for mitigating the effects of misassignment are various (Nugyen 2015 and 2016, Carlson 2012, Kong 2017).

Here we report on some of the particular challenges that result from these two sources of error, and their interaction. These were experienced while creating and using a ribosomal RNA gene amplicon library for a fungal ecology study (see chapter III). Synthetic mock communities are recommended as an alternative to standard mock communities (Jusino 2016). For studies using standard mock communities, a simple method is given: observe the abundance of OTUs from mock communities in negative controls, to estimate the potential levels of index-switching. Minimum abundances for observations of OTUs can then be chosen as a balance of removing as many tag-switching events as possible, while retaining as much ecological signal as possible. Additional discussion is given to some hazards and limitations of illumina MiSeq sequence data.

Methods

The following protocols were part of an ecological study (see chapter III), examining landscape level patterns of leaf and wood endophytes. Leaf and wood libraries were prepared separately, and the data presented here is from the wood endophyte library. This library included positive and negative controls and 91 ecological samples.

Wood endophyte sample preparation

Wood was debarked and phloem and sapwood was collected using tools that were ethanol- and flame-sterilized between cuts. 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 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 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 $MgCl_2$ (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 $MgCl_2$ (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 10^6 sequences, of which approximately 5.5 x 10^6 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

General bioinformatics protocols followed the USEARCH/UPARSE pipeline version 8.1 (Edgar 2013) wherever possible. Full scripts available in supplementary information (available here and here). Libraries of leaf and wood fungal endophyte DNA were prepared separately, so to maximize comparability, the reads from both libraries were combined as early as possible in the bioinformatics pipeline, following merging of paired ends, quality filtering, and chimera checks of each library individually.

FASTX toolkit software was used to visualize quality and trim unpaired read ends. After

removing low quality end base-calls from each direction, paired ends were merged using the USEARCH algorithm ("fastq_mergepairs" command). Quality filtering of merged reads was implemented using the USEARCH algorithm ("fastq_filter" command) with an Expected Error approach. Primer sequences were removed from all sequences. Small numbers of reads containing "floating" primer sequences, forward and reverse primer

Table 5.1. Taxa used in mock community (MC) positive control.

Proposed Taxa	MC#	Phylum	Sub-phylum	Class
Phaeocryptopus gaeumannii	MC1	Ascomycota	Pezizomycotina	Dothideomycetes
Rhabdocline parkerii	MC2	Ascomycota	Pezizomycotina	Leotiomycetes
Atractiellales sp.	MC3	Basidiomycota	Pucciniomycotina	Atractiellomycetes
Diaporthe sp.	MC4	Ascomycota	Pezizomycotina	Sordariomycetes
Plectania milleri	MC5	Ascomycota	Pezizomycotina	Pezizomycetes
Trametes versicolor	MC6	Basidiomycota	Agaricomycotina	Agaricomycetes
Lewia infectoria (Alternaria infectoria)	MC7	Ascomycota	Pezizomycotina	Dothideomycetes
Ramularia eucalypti	MC8	Ascomycota	Pezizomycotina	Dothideomycetes
Dothideomycetes sp.	MC9	Ascomycota	Pezizomycotina	Dothideomycetes
Coniocheta sp.	MC10	Ascomycota	Pezizomycotina	Sordariomycetes
Penicillium sp.	MC11	Ascomycota	Pezizomycotina	Eurotiomycetes
Penicillium nodusitatum	MC12	Ascomycota	Pezizomycotina	Eurotiomycetes
Hormonema sp.	MC13	Ascomycota	Pezizomycotina	Dothideomycetes
Ophiostoma sp.	MC14	Ascomycota	Pezizomycotina	Sordariomycetes
Phyllosticta sp.	MC15	Ascomycota	Pezizomycotina	Sordariomycetes
Xylaria nigripes	MC16	Ascomycota	Pezizomycotina	Sordariomycetes
Shizosaccharomyces pombe	MC17	Ascomycota	Taphrinomycotina	Schizosaccharomycetes
Saccharomyces cerevisiae	MC18	Ascomycota	Saccharomycotina	Saccharomycetes
Nemania serpens	MC19	Ascomycota	Pezizomycotina	Sordariomycetes
Psilocybe cyanescens	MC20	Basidiomycota	Agaricomycotina	Agaricomycetes
Epichloe sylvaticum	MC22	Ascomycota	Pezizomycotina	Sordariomycetes
Mucor racemosus	MC23	Mucoromycota	Mucoromycotina	(Mucorales)
Hypoxylon notatum	MC24	Ascomycota	Pezizomycotina	Sordariomycetes

sequences in central regions (Balint 2014), were presumed erroneous and removed using custom scripts. First chimera checks were conducted using the UCHIME algorithm ("uchime_ref" command) using the UNITE vers. 7.0 ITS1 reference database formatted for UCHIME. Leaf and wood libraries were concatenated at this point, and all reads were trimmed to ITS1 region only, using locations verified by the ITSx software (Bengtsson-Palme 2013). A 95% similarity radius according to UCLUST similarity

algorithms in the ITS1 region was used to define OTUs. This radius was shown in our positive controls to cause less artificial splitting of fungal species (Results, Fig. 6), and while not noticeably causing artificial lumping of positive control species within the same same genus. Assignment of taxonomy to OTUs was accomplished using a modified version of the UNITE vers. 7.0 database (Kõljalg 2013): all accessions in this database not identified to at least class-level were removed. This was done to avoid the possibility that other highly probable matches with more complete taxonomic information would be ignored during taxonomic assignments. Biome-format tables were constructed with *usearch* algorithms of the *usearch_global* program, which also allowed for inclusion of taxonomy information. Site metadata was added using the *biom-format* package (McDonald 2012). Some reformatting of taxonomic metadata of usearch-generated biom tables was required for parsing in downstream analyses. Variance stabilization of read-abundances was conducted using the *DESeq2* package in R (Love 2014, McMurdie 2013) after removal of controls.

Fungal species intentionally placed into positive control samples were distinguished from contaminants by querying with BLAST algorithm (Altschul 1990) the sequences found in our illumina library control samples against a database of Sanger-generated sequences of the 23 intended members of our mock-community. High confidence matches were assumed to be original, intentional members of the mock community, and the remaining sequences to be contaminants. Similarly, patterns of tag-switching were examined by querying all sequences from negative controls against this mock-community database. As most of the mock-community species were not common lab contaminants, and as care was taken during preparation of the mock community to avoid cross-contamination of Illumina libraries with DNA from positive controls before amplification of all samples with illumina-tagged primers, the presence of members of mock-community species was interpreted as tag-switching of reads from positive control to negative control indices.

Artificial OTU splitting of mock-community fungal species was observed even at a 95%

similarity radius for OTU formation. Due to the possible biases of PCR, OTU splitting, and tag-switching (Fig. 5.3), high minimum cutoffs were applied to all observations use in further analyses (see chapter III). 60 reads, or 1.0×10^{-5} of total wood endophyte library size, were subtracted from all observations of OTUs in each sample, and observations with less than 1 read were adjusted to zero. As potential results of contamination from tag switching, all observations of positive control fungi were removed from non-control ("ecological") samples. After minimum abundance cutoffs and removal of any observations of mock community members in the study, 15.5% of total reads were lost, and 80.4% of observations were lost.

Results:

Positive controls:

Positive controls recovered 22 of 23 species included in our mock-community. One species included in our mock communities, *Schizosaccharomyces pombe,* was not detected. A rank abundance plot of positive-control read abundances displayed a negative binomial (geometric)-type distribution, typical of amplicon libraries (McMurdie 2014). ITS-only positive controls displayed less dramatic differences among read abundances of OTUs, though large differences were still observed (Fig. 5.1, 5.2).

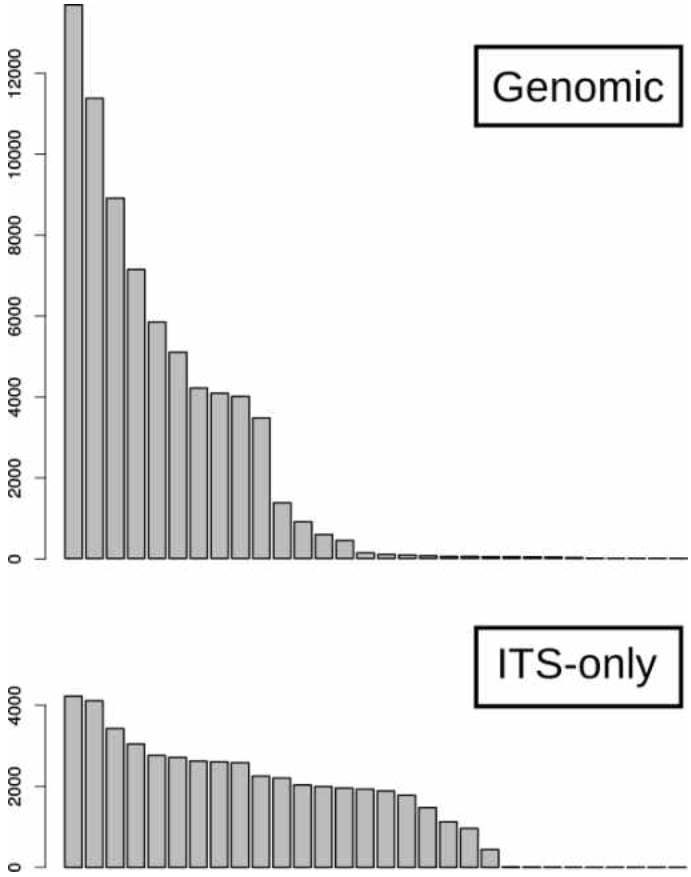


Figure 5.1. Ranked read distribution of genomic and ITS-only positive controls, by OTU. Singletons are removed after 30 OTUs. Click here for a higher resolution image.

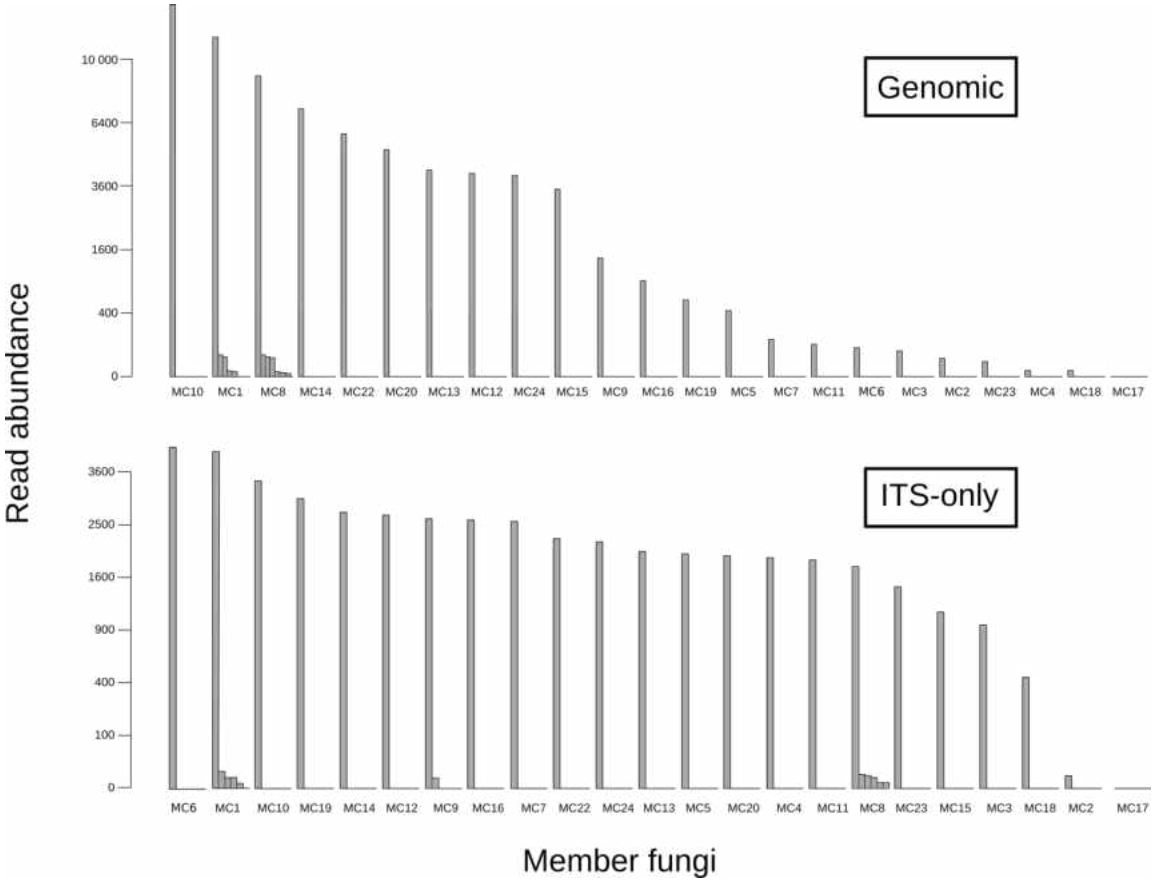


Figure 5.2. Ranked read distribution of abundances of genomic and ITS-only positive controls, including OTU splitting of mock-community members. Vertical axis is square-root transformed. Contaminants have been removed. Click here for a higher resolution image.

Variance stabilization:

Transformation by *DESeq*2 algorithms adjusted total read levels to more equal proportions among all samples (Fig. 5.4), and reduced the scale of artificial differences from PCR bias among read abundances of OTUs within our positive controls (Fig. 5.5), and therefore also presumably in ecological samples (Fig. 5.6). Despite this, read differences of one order of magnitude were found among our genomic mock-community samples and even after variance stabilization (Fig 5.5).

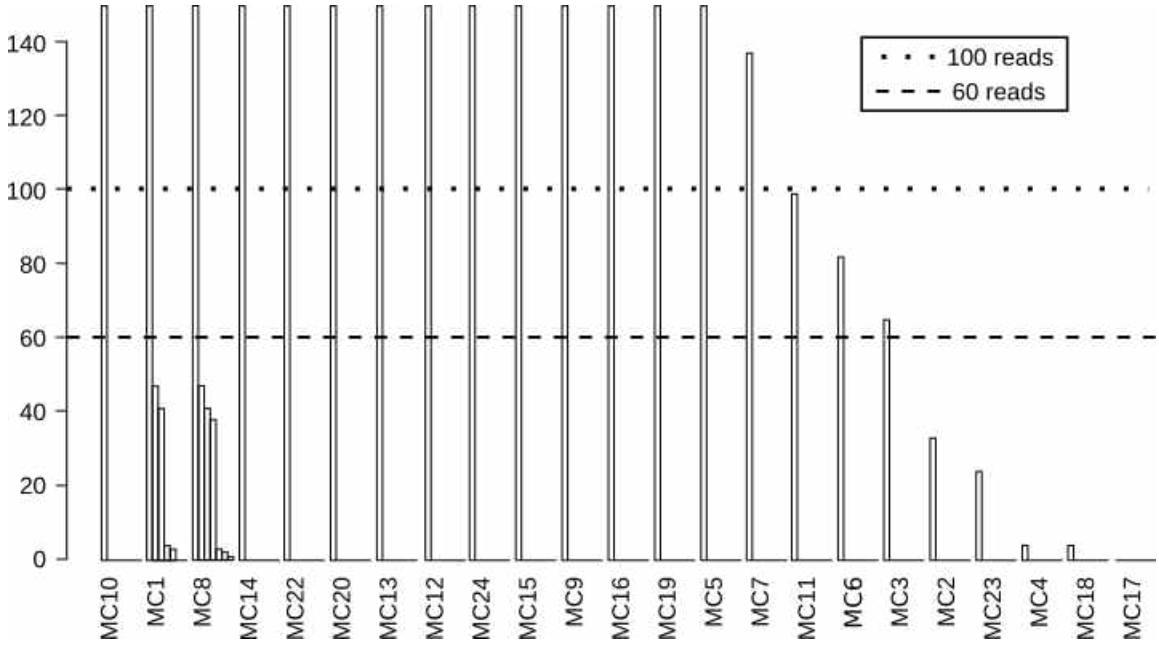


Figure 5.3. Truncated ranked read distribution of abundances of genomic positive control, including OTU splitting of mock-community members. The 100-read line represents the level around which tag-switching errors were observed to occur in the study (see Fig. 5.7), the 60-read line represents the abundance which authors of the subsequent ecological study chose as a minimum abundance cut-off for observations. Click here for a higher resolution image.

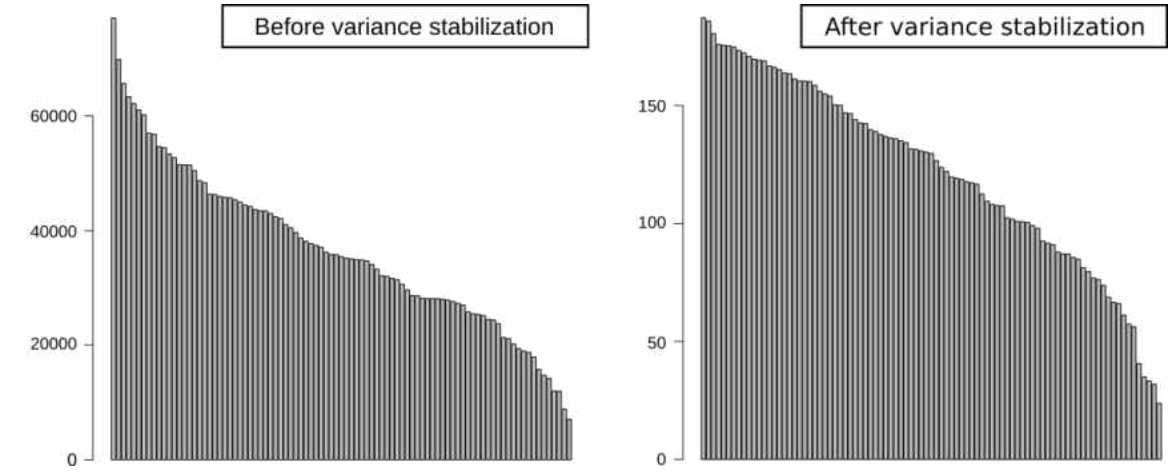


Figure 5.4. Ranked distribution of read abundances per sample for entire wood endophyte library, before and after variance stabilization using *deseq2* algorithms. Click here for a higher resolution image.

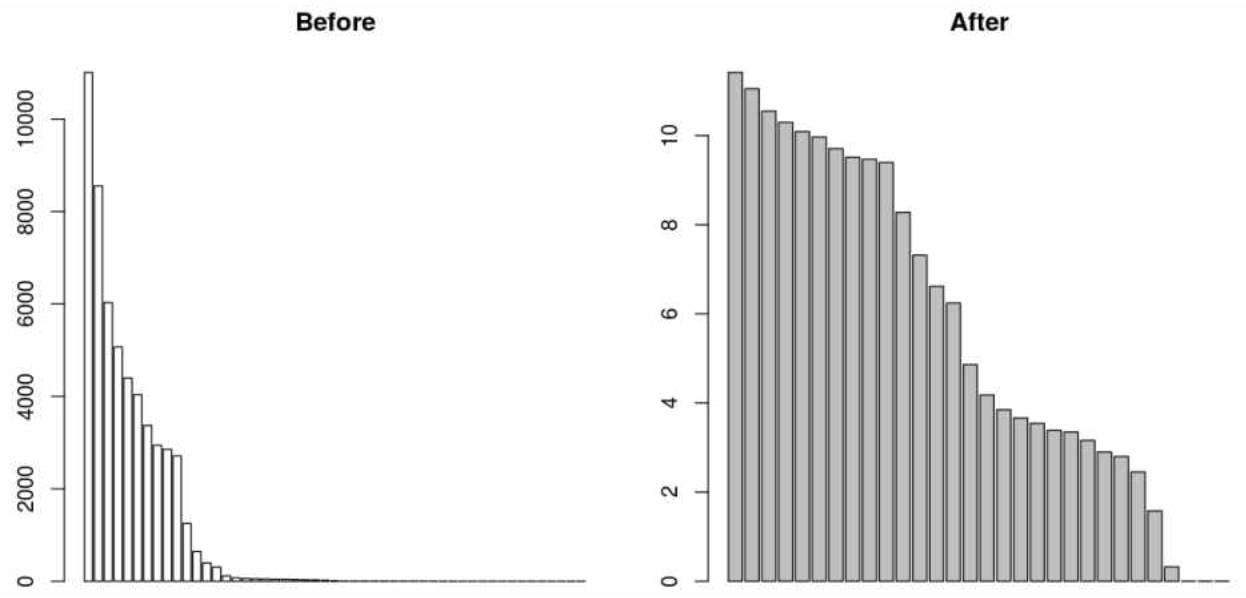


Figure 5.5. Ranked distribution of read abundances per OTU for genomic positive control, before and after variance stabilization using *deseq2* algorithms. Click here for a higher resolution image.

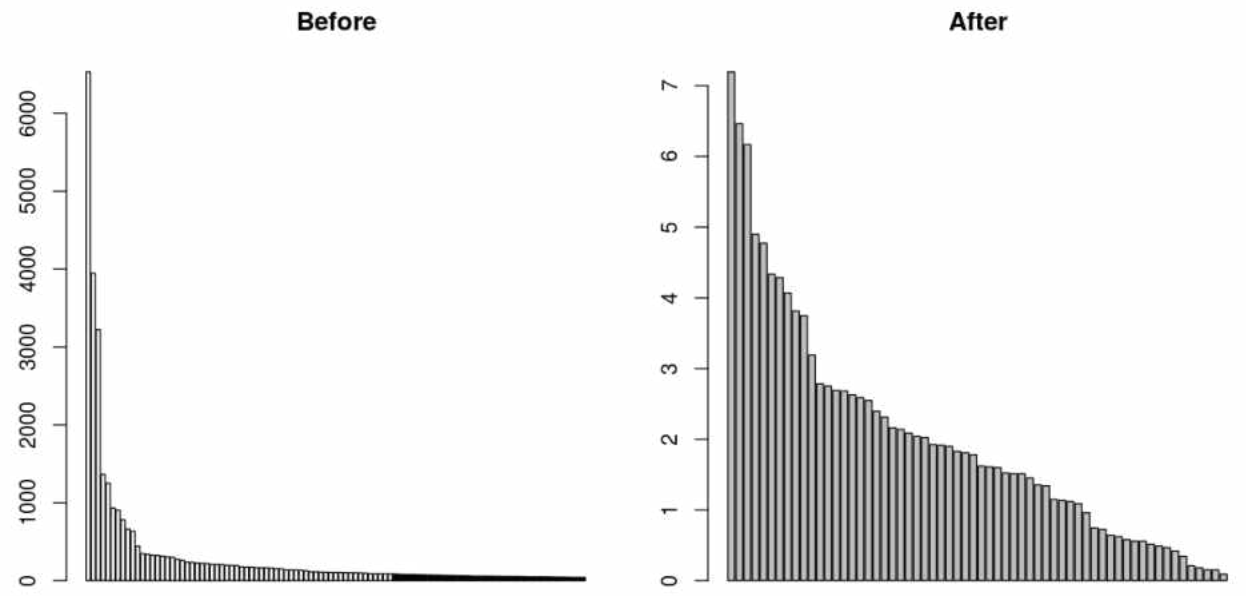


Figure 5.6. Ranked distribution of read abundances per OTU for one randomly selected sample from an the subsequent ecological study, before and after variance stabilization using *deseq2* algorithms. Blackened OTUs (left) represent OTUs that are removed by variance stabilization. Click here for a higher resolution image.

Negative controls:

Using an OTU similarity radius of 95%, pure water control contained 54 OTUs, with abundances of individual OTU observations up to 544 reads. 13 of these OTUs present in negative controls matched with high confidence to intended, original members of our positive controls (Fig. 5.7).

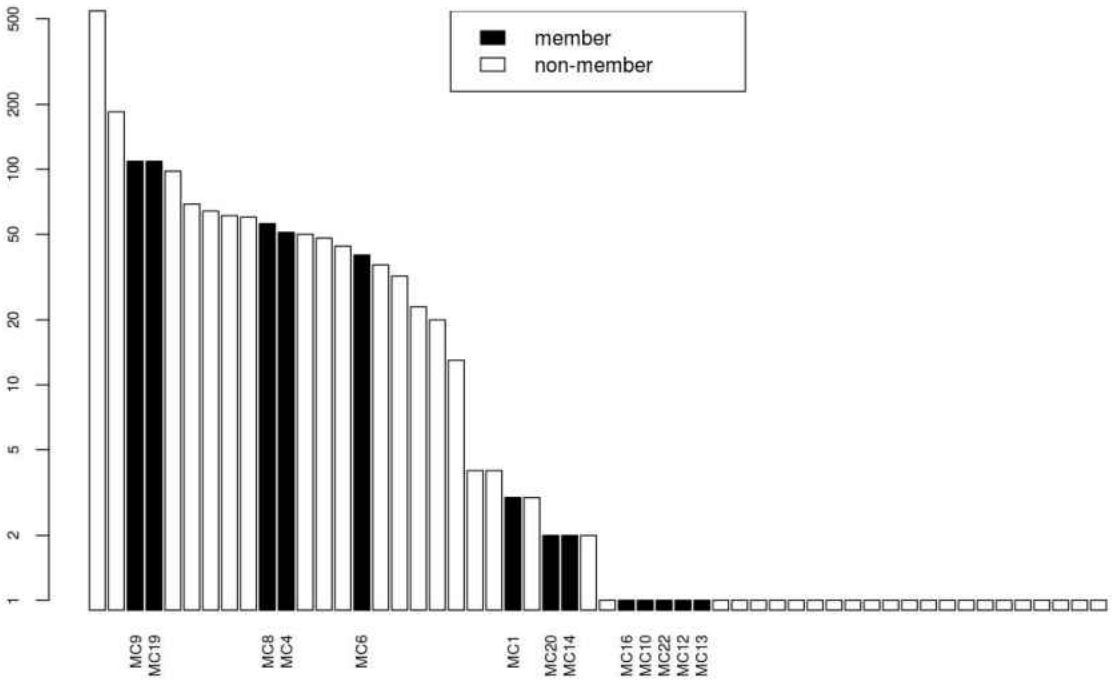


Figure 5.7. Ranked read distribution of OTUs from a pure water negative control. Black bars indicate OTUs that are also members of positive control, indicating probable misassignment of reads. Click here for a higher resolution image.

Discussion

The approximately negative binomial curve of genomic controls, compared to the less

dramatic differences in abundances shown by our ITS-only positive controls (Fig. 5.5), suggests that much of the potential bias within this study, and potentially amplicon sequencing studies in general, may originate in the first PCR step. ITS regions of organisms must be "found" amid the other regions of many genomes of the thousands of organisms present in any environmental sample of DNA. Initial conditions that allow easier discovery, such as larger ITS copy numbers (Schoch 2012) or ease of DNA extraction (Fredericks 2005), may be very important in determining which organisms' barcode regions are ultimately represented in amplicon sequence libraries. Even after adjusting for these differences among samples through negative-binomial variance stabilization, large artificial read abundances within our positive controls remained.

Negative controls showed high levels of mislabeled sequences probably originating from our positive control mock-community. These patterns of tag switching indicate that mock-community DNA disproportionately affected other samples in the study, probably due to relatively high concentrations of the simpler mock community DNA as compared to diverse ecological samples.

To utilize the data presented here in downstream ecological analyses (see chapter III), we chose to use a presence/absence transformation of data to correct for artificial differences in read abundance. This method results in the elevation of low-abundance observations of OTUs to equal importance with higher abundance observations in downstream statistical analyses. This was deemed appropriate given that even after variance stabilization, large artificial differences in abundances remained in the positive control. This elevation of importance for low-abundance OTUs can be problematic, as low-abundance observations and rarely observed OTUs are more likely to be spurious (Huse 2010, Brown 2015): here OTU splitting and tag-switching caused low-abundance, erroneous OTUs that were present as more than singletons (Fig. 5.3). Mock-community samples were used to estimate levels of tag-switching and generate minimum read-abundances for observations, as well as gauge appropriate levels of similarity for definition of OTUs, and

to gain insight into levels of PCR amplification biases. If tag-switching is shown to be common in an amplicon study, and presence/absence transformation is used to correct for artificial differences in read abundances, higher minimum cutoffs per observation than traditional removal of singleton OTUs are appropriate to avoid the elevation of spurious OTUs to the same importance as real OTUs. This of course results in a large loss of information about the presence of rare organisms (Brown 2015) (Fig. 5.6), but here strong ecological signals remained even after strict minimum read abundance cutoffs (see chapter III).

Illumina© has stated that index misassignment occurs at low levels and is likely due to contamination from free, unligated adapter/primer oligonucleotides (Illumina 2017). In the present study however, free primers should have been largely removed at two stages in library preparation: cleaning of all PCR products with Magbind beads and size selection of fragments larger than 250 BP. Despite this, erroneous, tag-switched reads appear at levels equal to the rate at which of many "real" OTUs appeared in the positive control (Fig. 5.3, 5.7)) . Thus dismissing the issue of tag switching due of the relatively low rate of index misassignment misses the mark for microbial metabarcoding studies: the researcher is confronted with balancing the loss of real ecological information against the need to remove possible incidences of index swap that could create false positives. In this study, after bioinformatic processing, 7 of 22 species present as OTUs in the mock community were present below the 100 base pair abundance of tag-switched observations in negative controls (Fig. 5.3).

The issue confusion of real OTU observations with those resulting from tag switching is compounded if species of interest are included in the positive control. Contrary to the recommendations of Kong (2017), researchers should also take care to avoid the use of species of biological interest as members of their mock communities, as observations of these species outside of positive control samples may then be called into question as potential relics of tag-switching. In the present study, significant levels of tag-switching

were evident, especially in the members of the positive community itself (Fig. 5.7), so all OTUs identified as original intentional members in the positive controls were removed entirely from the wood endophyte library for downstream analyses. This study highlights the need for strict cutoffs and careful implementation of positive controls, and a framework for estimating rates of tag switching from these. The most promising toolset for estimating rates of tag-switching is completely synthetic positive controls proposed by Jusino et al. (2016), which are mock-community constructed from fungal-like ITS-region oligonucleotides that do not represent any organisms in nature.

CHAPTER VI

CONCLUSION

The field of microbial ecology has accelerated with the rise of next-generation sequencing technologies. In nearly every ecosystem imaginable, microbes are being sampled, including fungi. This has created a bit of culture shock in mycology. Until recently, mycology was a field characterized by careful phylogenetic analyses, exhaustively detailed species descriptions, laborious culture studies, and traditional cellular biology and genetical studies. It was a field that had to carefully and painstakingly check all of its ecological findings against the reality that most fungi in natural settings could not be detected (O'Brien 2005). All ecological mycology studies were (and probably still are) underrepresenting true diversity and abundance of fungi. Admirably and understandably, mycologists have been somewhat reluctant to take on the role of ecologists.

However, If no other discipline will, mycology in particular will force real introspection by ecology of its own contradictions. This is because fungi defy any single ecological definition of the individual, and operate on all spatial and temporal scales available to life. They vary in size from single cells to perhaps the largest organism on earth (Smith 1993). Their diversity is probably exceeded only by prokaryotes (Mindes 2011), their ecological functions are numerous and vital, with which they tie together the fitness of unrelated species (Gorzelak 2015). And they and other microbes seem to be literally everywhere (Cuadros-Orellan 2013, Yahr 2016). Massive species matrices, rather than interactions among a few species, have become the new basic unit of many ecological analyses. For these reasons fungi teach us to realize how poorly we ecologists were perceiving the biological world, with tools that were problematic but tenable with macro-organisms, such as the simplification of interactions into trophic levels, crude

categorization of spatial patterns in “regional” and “local” scales, black-boxing of microbes, and an over-reliance on simplistic ecological, evolutionary and mathematical models. Ecologists are now faced with a world saturated and shimmering with innumerable species, varying on all spatial and temporal scales, and uncertain, shifting niche spaces. The scientific and mathematical language we create to describe this beautiful complexity is one of the big tasks ahead for the imaginations of researchers.

I hope that the research here of my and my collaborators’ research has contributed in some small way to the way forward for ecological mycology. Our contributions presented here have been of two types: (1) empirical, in the form of data on fungal dispersal and the behavior of forest microbiomes on a landscape level and (2) methodological, with technical suggestions on the current techniques on metabarcoding, and with suggestion of use of simulation models to increase rigor in largely observational projects where experimental manipulation may not be feasible.

REFERENCES 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*(3), 403 – 410.

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10), R106.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46.

Arnold, A., Maynard, Z., Gilbert, G., Coley, P., & Kursar, T. (2000). Are tropical fungal endophytes hyperdiverse? *Ecology Letters*, *3*(4), 267–274.

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*(3), 388–398.

Arnold, A. E., & Lutzoni, F. (2007). Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? *Ecology*, *88*(3), 541–549.

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. *Proceedings of the National Academy of Sciences*, *100*(26), 15649–15654.

Baas-Becking, L. G. M. (1934). *Geobiologie; of inleiding tot de milieukunde*. WP Van Stockum & Zoon NV The Hague, the Netherlands.

Baldridge, E., Harris, D. J., Xiao, X., & White, E. P. (2016). An extensive comparison of species-abundance distribution models. *PeerJ*, *4*, e2823.

Bálint, M., Schmidt, P.-A., Sharma, R., Thines, M., & Schmitt, I. (2014). An illumina metabarcoding pipeline for fungi. *Ecology and Evolution*, *4*(13), 2642–2653.

Bayman, P., Angulo-Sandoval, P., Báez-ortiz, Z., & Lodge, D. J. (1998). Distribution and dispersal of *Xylaria* endophytes in two tree species in Puerto Rico. *Mycological Research*, *102*(8), 944 – 948.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y. J. K., Sanli, K., Eriksson, K. M., Vik, U., Veldre, V., & Nilsson, R. H. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution*, *4*(10), 914–919.

Berendsen, R. L., Pieterse, C. M., & Bakker, P. A. (2012). The rhizosphere microbiome and plant health. *Trends in Plant Science*, *17*(8), 478 – 486.

Berlec, A. (2012). Novel techniques and findings in the study of plant microbiota: Search for plant probiotics. *Plant Science*, *193*, 96 – 102.

Bernstein, M. E., & Carroll, G. C. (1977). Internal fungi in old-growth Douglas-fir foliage. *Canadian Journal of Botany*, *55*(6), 644–653.

Blaalid, R., Kumar, S., Nilsson, R. H., Abarenkov, K., Kirk, P. M., & Kauserud, H. (2013). ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources*, *13*(2), 218–224.

Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecology*, *6*(1), 9.

Bodenhausen, N., Horton, M. W., & Bergelson, J. (2013). Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana. PLOS ONE*, *8*(2), 1–9.

Borcard, D., Gillet, F., & Legendre, P. (2011). *Numerical Ecology with R*. Springer.

Boswell, G. P., & Davidson, F. A. (2012). Modelling hyphal networks. *Fungal Biology Reviews*, *26*(1), 30 – 38. Hyphal networks: mechanisms, modelling and ecology.

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349.

Brown, S. P., Veach, A. M., Rigdon-Huss, A. R., Grond, K., Lickteig, S. K., Lothamer, K., Oliver, A. K., & Jumpponen, A. (2015). Scraping the bottom of the barrel: are rare high throughput sequences artifacts? *Fungal Ecology*, *13*, 221 – 225.

Bruijnzeel, L. A., & Veneklaas, E. J. (1998). Climatic conditions and tropical montane forest productivity: The fog has not lifted yet. *Ecology*, *79*(1), 3–9.

Burnham, K. P., & Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika*, *65*(3), 625–633.

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*(3), 543–550.

Cannon, P. F., & Simmons, C. M. (2002). Diversity and host preference of leaf endophytic fungi in the iwokrama forest reserve, guyana. *Mycologia*, *94*(2), 210–220.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., et al. (2010). Qiime allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), 335–336.

Carlsen, T., Aas, A. B., Lindner, D., Vrålstad, T., Schumacher, T., & Kauserud, H. (2012). Don't make a mista(g)ke: is tag switching an overlooked source of error in amplicon pyrosequencing studies? *Fungal Ecology*, *5*(6), 747 – 749.

Carroll, G. (1988). Fungal endophytes in stems and leaves: From latent pathogen to mutualistic symbiont. *Ecology*, *69*(1), 2–9.

Carroll, G. (1999). The foraging ascomycete. In *16th International Botanical Congress, Abstracts*, vol. 309.

Carroll, G., & Petrini, O. (1983). Patterns of substrate utilization by some fungal endophytes from coniferous foliage. *Mycologia*, *75*(1), 53–63.

Carroll, G. C., & Carroll, F. E. (1978). Studies on the incidence of coniferous needle endophytes in the pacific northwest. *Canadian Journal of Botany*, *56*(24), 3034–3043.

Chambers, J. Q., Higuchi, N., Schimel, J. P., Ferreira, L. V., & Melack, J. M. (2000). Decomposition and carbon cycling of dead trees in tropical forests of the central Amazon. *Oecologia*, *122*(3), 380–388.

Chambers, J. Q., Schimel, J. P., & Nobre, A. D. (2001). Respiration from coarse wood litter in central amazon forests. *Biogeochemistry*, *52*(2), 115–131.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, *11*(4), 265–270.

Chaparro, J. M., Sheflin, A. M., Manter, D. K., & Vivanco, J. M. (2012). Manipulating the soil microbiome to increase soil health and plant fertility. *Biology and Fertility of Soils*, *48*(5), 489–499.

Chapela, I. H., & Boddy, L. (1988). Fungal colonization of attached beech branches. *New Phytologist*, *110*(1), 47–57.

Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P. P., Osório, M. L., Carvalho, I., Faria, T., & Pinheiro, C. (2002). How plants cope with water stress in the field? photosynthesis and growth. *Annals of Botany*, *89*(7), 907–916.

Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, *35*(4), 445–453.

Clay, K., & C. Schardl (2002). Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *The American Naturalist*, *160*(S4), S99–S127. PMID: 18707456.

Clay, K. (1988). Fungal endophytes of grasses: a defensive mutualism between plants and fungi. *Ecology*, *69*(1), 10–16.

Clay, K. (1990). Fungal endophytes of grasses. *Annual Review of Ecology and Systematics*, *21*(1), 275–297.

Clinton, P., Buchanan, P., Wilkie, J., Smaill, S., & Kimberley, M. (2009). Decomposition of nothofagus wood in vitro and nutrient mobilization by fungi. *Canadian Journal of Forest Research*, *39*(11), 2193–2202.

Clobert, J. (2012). *Dispersal ecology and evolution*. Oxford University Press.

Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions: Biological Sciences*, *345*(1311), 101–118.

Crowther, T. W., Glick, H., Covey, K., Bettigole, C., Maynard, D., Thomas, S., Smith, J., Hintler, G., Duguid, M., Amatulli, G., et al. (2015). Mapping tree density at a global scale. *Nature*, *525*(7568), 201–205.

Cuadros-Orellana, S., Leite, L. R., Smith, A., Medeiros, J. D., Badotti, F., Fonseca, P. L., Vaz, A. B., Oliveira, G., & Góes-Neto, A. (2013). Assessment of fungal diversity in the environment using metagenomics: a decade in review. *Fungal Genomics & Biology*, *3*(2), 1.

Davis, E. C., Franklin, J. B., Shaw, A. J., & Vilgalys, R. (2003). Endophytic xylaria (xylariaceae) among liverworts and angiosperms: phylogenetics, distribution, and symbiosis. *American Journal of Botany*, *90*(11), 1661–1667.

Davis, E. C., & Shaw, A. J. (2008). Biogeographic and phylogenetic patterns in diversity of liverwort-associated endophytes. *American Journal of Botany*, *95*(8), 914–924.

De Wit, R., & Bouvier, T. (2006). "Everything is everywhere, but, the environment selects"; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*, *8*(4), 755–758.

Dentinger, B. T. M., Margaritescu, S., & Moncalvo, J.-M. (2010). Rapid and reliable high-throughput methods of DNA extraction for use in barcoding and molecular systematics of mushrooms. *Molecular Ecology Resources*, *10*(4), 628–633.

Dixon, P. (1994). Testing spatial segregation using a nearest-neighbor contingency table. *Ecology*, *75*(7), 1940–1948.

Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite package: analysing ecological networks. *Interaction*, *1*, 0–2413793.

Edgar, R. C. (2013). UPARSE: highly accurate otu sequences from microbial amplicon reads. *Nature methods*, *10*(10), 996–998.

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. *Proceedings of the National Academy of Sciences*, *112*(8), E911–E920.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, *11*(4), 12.

Eveling, D., Wilson, R., Gillespie, E., & Bataillé, A. (1990). Environmental effects on sporocarp counts over fourteen years in a forest area. *Mycological Research*, *94*(7), 998 – 1002.

Fenchel, T., & Finlay, B. J. (2004). The ubiquity of small species: Patterns of local and global diversity. *BioScience*, *54*(8), 777–784.

Ferrari, J. B., & Sugita, S. (1996). A spatially explicit model of leaf litter fall in hemlock-hardwood forests. *Canadian Journal of Forest Research*, *26*(11), 1905–1913.

Fredricks, D. N., Smith, C., & Meier, A. (2005). Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *Journal of Clinical Microbiology*, *43*(10), 5122–5128.

Freeman, S., & Rodriguez, R. J. (1993). Genetic conversion of a fungal plant pathogen to a nonpathogenic, endophytic mutualist. *Science*, *260*(5104), 75–79.

Gabor, C., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Galante, T. E., Horton, T. R., & Swaney, D. P. (2011). 95% of basidiospores fall within 1 m of the cap: a field-and modeling-based study. *Mycologia*, *103*(6), 1175–1183.

Gange, A., Gange, E., Sparks, T., & Boddy, L. (2007). Rapid and recent changes in fungal fruiting patterns. *Science*, *316*(5821), 71–71.

Gardes, M., & Bruns, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology*, *2*(2), 113–118.

Gavilanez, M. M., & Stevens, R. D. (2013). Role of environmental, historical and spatial processes in the structure of neotropical primate communities: contrasting taxonomic and phylogenetic perspectives. *Global Ecology and Biogeography*, *22*(5), 607–619.

Gentry, A. H. (1992). Tropical forest biodiversity: distributional patterns and their conservational significance. *Oikos*, (pp. 19–28).

Gorzelak, M. A., Asay, A. K., Pickles, B. J., & Simard, S. W. (2015). Inter-plant communication through mycorrhizal networks mediates complex adaptive behaviour in plant communities. *AoB PLANTS*, *7*, plv050.

Goslee, S. C., Urban, D. L., et al. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*(7), 1–19.

Green, J., & Bohannan, B. J. (2006). Spatial scaling of microbial biodiversity. *Trends in Ecology & Evolution*, *21*(9), 501 – 507.

Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., et al. (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, *432*(7018), 747.

Griffith, D. M., Veech, J. A., Marsh, C. J., et al. (2016). cooccur: Probabilistic species co-occurrence analysis in r. *J Stat Softw*, *69*(2), 1–17.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Peér, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø¸, R., Visser, U., & DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1), 115 – 126.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The odd protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760 —2768.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, *310*(5750), 987–991.

Grubisha, L. C., Bergemann, S. E., & Bruns, T. D. (2007). Host islands within the California Northern Channel islands create fine-scale genetic structure in two sympatric species of the symbiotic ectomycorrhizal fungus *Rhizopogon*. *Molecular Ecology*, *16*(9), 1811–1822.

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research*, *19*(7), 1141–1152.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506.

Hardoim, P. R., van Overbeek, L. S., & van Elsas, J. D. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology*, *16*(10), 463 – 471.

Herre, E. A., Mejía, L. C., Kyllo, D. A., Rojas, E., Maynard, Z., Butler, A., & Van Bael, S. A. (2007). Ecological implications of anti-pathogen effects of tropical fungal endophytes and mycorrhizae. *Ecology*, *88*(3), 550–558.

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.

Higgins, K. L., Coley, P. D., Kursar, T. A., & Arnold, A. E. (2011). Culturing and direct PCR suggest prevalent host generalism among diverse fungal endophytes of tropical forest grasses. *Mycologia*, *103*(2), 247–260.

Hoffman, M. T., & Arnold, A. E. (2010). Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Applied and environmental microbiology*, *76*(12), 4063–4075.

Hsieh, H.-M., Lin, C.-R., Fang, M.-J., Rogers, J. D., Fournier, J., Lechat, C., & Ju, Y.-M. (2010). Phylogenetic status of *Xylaria* subgenus *Pseudoxylaria* among taxa of the subfamily xylarioideae (Xylariaceae) and phylogeny of the taxa involved in the subfamily. *Molecular Phylogenetics and Evolution*, *54*(3), 957–969.

Hurtt, G. C., & Pacala, S. W. (1995). The consequences of recruitment limitation: reconciling chance, history and competitive differences between plants. *Journal of theoretical biology*, *176*(1), 1–12.

Huse, S. M., Welch, D. M., Morrison, H. G., & Sogin, M. L. (2010). Ironing out the wrinkles in the rare biosphere through improved otu clustering. *Environmental Microbiology*, *12*(7), 1889–1898.

Ikeda, A., Matsuoka, S., Masuya, H., Mori, A. S., Hirose, D., & Osono, T. (2014). Comparison of the diversity, composition, and host recurrence of xylariaceous endophytes in subtropical, cool temperate, and subboreal regions in japan. *Population ecology*, *56*(2), 289–300.

illumina, i. (2017). Effects of index misassignment on multiplexing and downstream analysis.

Johnson, C. L., & Versalovic, J. (2012). The human microbiome and its potential importance to pediatrics. *Pediatrics*, *129*(5), 950–960.

Jusino, M. A., Palmer, J. M., Banik, M. T., Edgar, R. C., & Lindner, D. L. (2016). Everything is not everywhere: Designing a synthetic fungal ITS mock community for ngs of environmental samples. In *Abstracts MSA Annual Meeting 2016*, vol. 130, (p. 53).

Kettle, C. J., & Koh, L. P. (2014). *Global forest fragmentation*. CABI.

Kim, M., Singh, D., Lai-Hoe, A., Go, R., Abdul Rahim, R., A.N., A., Chun, J., & Adams, J. M. (2012). Distinctive phyllosphere bacterial communities in tropical trees. *Microbial Ecology*, *63*(3), 674–681.

Köberl, M., Müller, H., Ramadan, E. M., & Berg, G. (2011). Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. *PLoS One*, *6*(9), e24452.

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., Poldmaa, 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. *Molecular Ecology*, *22*(21), 5271–5277.

Kong, H. H., Andersson, B., Clavel, T., Common, J. E., Jackson, S. A., Olson, N. D., Segre, J. A., & Traidl-Hoffmann, C. (2017). Performing skin microbiome research: a method to the madness. *Journal of Investigative Dermatology*.

Krauss, P., Markstädter, C., & Riederer, M. (1997). Attenuation of uv radiation by plant cuticles from woody species. *Plant, Cell & Environment*, *20*(8), 1079–1085.

Krings, M., Taylor, T. N., Hass, H., Kerp, H., Dotzler, N., & Hermsen, E. J. (2007). Fungal endophytes in a 400-million-yr-old land plant: infection pathways, spatial distribution, and host responses. *New Phytologist*, *174*(3), 648–657.

Lau, M. K., Arnold, A. E., & Johnson, N. C. (2013). Factors influencing communities of foliar fungal endophytes in riparian woody plants. *Fungal Ecology*, *6*(5), 365–378.

Lawton, J. H. (1999). Are there general laws in ecology? *Oikos*, *84*(2), 177–192.

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280.

Legendre, P., & Legendre, L. F. (2012). *Numerical ecology*, vol. 24. Elsevier.

Leigh, E. G. (1975). Structure and climate in tropical rain forest. *Annual Review of Ecology and Systematics*, *6*, 67–86.

Liggenstoffer, A. S., Youssef, N. H., Couger, M., & 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*(10), 1225–1235.

Lindahl, B. D., Nilsson, R. H., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjøller, R., Kõljalg, U., Pennanen, T., Rosendahl, S., Stenlid, J., & Kauserud, H. (2013). Fungal community analysis by high-throughput sequencing of amplified markers —a user's guide. *New Phytologist*, *199*(1), 288–299.

Liu, C. (2001). A comparison of five distance-based methods for spatial pattern analysis. *Journal of Vegetation Science*, *12*(3), 411–416.

Lodge, D. J. (1997). Factors related to diversity of decomposer fungi in tropical forests. *Biodiversity & Conservation*, *6*(5), 681–688.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, *15*(12), 550.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Del Rio, T. G., et al. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature*, *488*(7409), 86–90.

Malloch, D., & Blackwell, M. (1992). Dispersal of fungal diaspores. *The fungal community: its organization and role in the ecosystem*, (pp. 147–171).

Malloch, D., Pirozynski, K., & Raven, P. (1980). Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (a review). *Proceedings of the National Academy of Sciences*, *77*(4), 2113–2118.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2 Part 1), 209–220.

Cadotte, M. W. (2006). Dispersal and species diversity: A meta-analysis. *The American Naturalist*, *167*(6), 913–924. PMID: 16649154.

Marie-Josée Fortin, F. M. D. (2005). Spatial analysis. a guide for ecologists.

Martiny, J. B. H., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., et al. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102–112.

May, R. M. (1991). A fondness for fungi. *Nature*, *352*(6335), 475–476.

McDonald, D., Clemente, J. C., Kuczynski, J., Rideout, J. R., Stombaugh, J., Wendel, D., Wilke, A., Huse, S., Hufnagle, J., Meyer, F., Knight, R., & Caporaso, J. G. (2012). The biological observation matrix (biom) format or: how I learned to stop worrying and love the ome-ome. *GigaScience*, *1*(1), 7.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, *8*(4), 1–11.

McMurdie, P. J., & Holmes, S. (2014). Waste not, want not: Why rarefying microbiome data is inadmissible. *PLOS Computational Biology*, *10*(4), 1–12.

McWilliam, A. L. C., Roberts, J. M., Cabral, O. M. R., Leitao, M. V. B. R., de Costa, A. C. L., Maitelli, G. T., & Zamparoni, C. A. G. P. (1993). Leaf area index and above-ground biomass of terra firme rain forest and adjacent clearings in Amazonia. *Functional Ecology*, *7*(3), 310–317.

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*, *46*(1), 4 – 14. Special Issue: Endophytes.

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J. H. M., Piceno, Y. M., DeSantis, T. Z., Andersen, G. L., Bakker, P. A. H. M., & Raaijmakers, J. M. (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science*, *332*(6033), 1097–1100.

Moore, A. M. (1986). Temperature and moisture dependence of decomposition rates of hardwood and coniferous leaf litter. *Soil Biology and Biochemistry*, *18*(4), 427–435.

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*(3), 342–356.

Nguyen, N. H., Smith, D., Peay, K., & Kennedy, P. (2015). Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytologist*, *205*(4), 1389–1393. 2014-17526.

Nguyen, N. H., Williams, L. J., Vincent, J. B., Stefanski, A., Cavender-Bares, J., Messier, C., Paquette, A., Gravel, D., Reich, P. B., & Kennedy, P. G. (2016). Ectomycorrhizal fungal diversity and saprotrophic fungal diversity are linked to different tree community attributes in a field-based tree experiment. *Molecular Ecology*, *25*(16), 4032–4046.

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*(6), 961–974.

O'Brien, H. E., Parrent, J. L., Jackson, J. A., Moncalvo, J.-M., & Vilgalys, R. (2005). Fungal community analysis by large-scale sequencing of environmental samples. *Applied and Environmental Microbiology*, *71*(9), 5544–5550.

Okane, I., Srikitikulchai, P., Toyama, K., Læssøe, T., Sivichai, S., Hywel-Jones, N., Nakagiri, A., Potacharoen, W., & Suzuki, K.-i. (2008). Study of endophytic Xylariaceae in thailand: diversity and taxonomy inferred from rDNA sequence analyses with saprobes forming fruit bodies in the field. *Mycoscience*, *49*(6), 359.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2017). *vegan: Community Ecology Package*. R package version 2.4-2.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., Simpson, G. L., Solymos, P., Stevens, M., & Wagner, H. (2013). *vegan: Community Ecology Package. R package version 2.0-10. 2013*. R package version 2.4-2.

Oldroyd, G. E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, *11*(4), 252–263.

Osono, T. (2006). Role of phyllosphere fungi of forest trees in the development of decomposer fungal communities and decomposition processes of leaf litter. *Canadian Journal of Microbiology*, *52*(8), 701–716.

Paoletti, M. (2016). Vegetative incompatibility in fungi: From recognition to cell death, whatever does the trick. *Fungal Biology Reviews*, *30*(4), 152–162.

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*(12), 3631–3640.

Peay, K. g., Schubert, M. G., Nguyen, N. H., & Bruns, T. D. (2012). Measuring ectomycorrhizal fungal dispersal: macroecological patterns driven by microscopic propagules. *Molecular Ecology*, *21*(16), 4122–4136.

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in r. *R news*, *5*(2), 9–13.

Peck, M., Tirira, D., Mariscal, A., & Paredes, K. (2008). Developing a sustainable network for primates in Ecuador (primenet). Darwin Initiative, final report with notes.

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, *87*(10), 2614–2625.

Peršoh, D. (2013). Factors shaping community structure of endophytic fungi–evidence from the *Pinus-Viscum*-system. *Fungal Diversity*, *60*(1), 55–69.

Pinto, A. J., & Raskin, L. (2012). PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLOS ONE*, *7*(8), 1–16.

Policha, T. (2014). *Pollination biology of the mushroom-mimicking orchid genus* Dracula. Ph.D. thesis, University of Oregon.

Porras-Alfaro, A., & Bayman, P. (2011). Hidden fungi, emergent properties: Endophytes and microbiomes. *Annual Review of Phytopathology*, *49*(1), 291–315. PMID: 19400639.

Promputtha, I., Hyde, K. D., McKenzie, E. H., Peberdy, J. F., & Lumyong, S. (2010). Can leaf degrading enzymes provide evidence that endophytic fungi becoming saprobes? *Fungal Diversity*, *41*(1), 89–99.

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. *Microbial Ecology*, *53*(4), 579–590.

Rayner, R. (1948). Latent infection in *Cofffea arabica* l. *Nature*, *161*(4085), 245–246.

Redecker, D., Kodner, R., & Graham, L. E. (2000). Glomalean fungi from the Ordovician. *Science*, *289*(5486), 1920–1921.

Reich, P., Uhl, C., Walters, M., & Ellsworth, D. (1991). Leaf lifespan as a determinant of leaf structure and function among 23 amazonian tree species. *Oecologia*, *86*(1), 16–24.

Rodrigues, J. L. M., Pellizari, V. H., Mueller, R., Baek, K., Jesus, E. d. 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. *Proceedings of the National Academy of Sciences*, *110*(3), 988–993.

Rodriguez, R. J., White Jr, J. F., Arnold, A. E., & Redman, R. S. (2009). Fungal endophytes: diversity and functional roles. *New Phytologist*, *182*(2), 314–330.

Rogers, J. D. (1979). The Xylariaceae: systematic, biological and evolutionary aspects. *Mycologia*, *71*(1), 1–42.

Rogers, J. D. (2000). Benefactors' lecture thoughts and musings on tropical Xylariaceae. *Mycological Research*, *104*(12), 1412–1420.

Ronce, O. (2007). How does it feel to be like a rolling stone? ten questions about dispersal evolution. *Annu. Rev. Ecol. Evol. Syst.*, *38*, 231–253.

Roper, M., Seminara, A., Bandi, M., Cobb, A., Dillard, H. R., & Pringle, A. (2010). Dispersal of fungal spores on a cooperatively generated wind. *Proceedings of the National Academy of Sciences*, *107*(41), 17474–17479.

Rosenberg, E., Sharon, G., Atad, I., & Zilber-Rosenberg, I. (2010). The evolution of animals and plants via symbiosis with microorganisms. *Environmental Microbiology Reports*, *2*(4), 500–506.

Rosenblueth, M., & Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Molecular plant-microbe interactions*, *19*(8), 827–837.

Roy, B. (2001). Patterns of association between crucifers and their flower-mimic pathogens: host jumps are more common than coevolution or cospeciation. *Evolution*, *55*(1), 41–53.

Rudgers, J. A., Afkhami, M. E., Rúa, M. A., Davitt, A. J., Hammer, S., & Huguet, V. M. (2009). A fungus among us: broad patterns of endophyte distribution in the grasses. *Ecology*, *90*(6), 1531–1539.

Saikkonen, K., Lehtonen, P., Helander, M., Koricheva, J., & Faeth, S. H. (2006). Model systems in ecology: dissecting the endophyte–grass literature. *Trends in plant science*, *11*(9), 428–433.

Santamaría, J., & Bayman, P. (2005). Fungal epiphytes and endophytes of coffee leaves (coffea arabica). *Microbial Ecology*, *50*(1), 1–8.

Schardl, C. L., Leuchtmann, A., & Spiering, M. J. (2004). Symbioses of grasses with seedborne fungal endophytes. *Annu. Rev. Plant Biol.*, *55*, 315–340.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., & Consortium, F. B. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Sciences*, *109*(16), 6241–6246.

Schulz, B., & Boyle, C. (2005). The endophytic continuum. *Mycological research*, *109*(6), 661–686.

Schulz, B., Römmert, A.-K., Dammann, U., Aust, H.-J., & Strack, D. (1999). The endophyte-host interaction: a balanced antagonism? *Mycological Research*, *103*(10), 1275 – 1283.

Seboth, J. (1881). *Faramea tinguana*. *Martius C., Eichler AG, and I. Urban. Flora Brasiliensis*, *6*(5).

Shashkova, T., Popenko, A., Tyakht, A., Peskov, K., Kosinsky, Y., Bogolubsky, L., Raigorodskii, A., Ischenko, D., Alexeev, D., & Govorun, V. (2016). Agent based modeling of human gut microbiome interactions and perturbations. *PloS one*, *11*(2), e0148386.

Sinha, R., Stanley, G., Gulati, G. S., Ezran, C., Travaglini, K. J., Wei, E., Chan, C. K. F., Nabhan, A. N., Su, T., Morganti, R. M., Conley, S. D., Chaib, H., Red-Horse, K., Longaker, M. T., Snyder, M. P., Krasnow, M. A., & Weissman, I. L. (2017). Index switching causes "spreading-of-signal" among multiplexed samples in illumina HiSeq 4000 DNA sequencing. *bioRxiv*.

Smith, M. L., Bruhn, J. N., & Anderson, J. B. (1992). The fungus *Armillaria bulbosa* is among the largest and oldest living organisms. *Nature*, *356*(6368), 428.

Stone, J. K. (1987). Initiation and development of latent infections by *Rhabdocline parkeri* on Douglas-fir. *Canadian Journal of Botany*, *65*(12), 2614–2621.

Stone, J. K., Polishook, J. D., White, J. F., et al. (2004). Endophytic fungi. *Biodiversity of Fungi. Elsevier Academic Press, Burlington*, (pp. 241–270).

Stukenbrock, E. H., & McDonald, B. A. (2008). The origins of plant pathogens in agro-ecosystems. *Annu. Rev. Phytopathol.*, *46*, 75–100.

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

Su, S.-H., Hsieh, C.-F., Chang-Yang, C.-H., Lu, C.-L., Guan, B. T., et al. (2010). Micro-topographic differentiation of the tree species composition in a subtropical submontane rainforest in northeastern taiwan. *Taiwan J For Sci*, *25*(1), 63–80.

Suryanarayanan, T., Murali, T., & Venkatesan, G. (2002). Occurrence and distribution of fungal endophytes in tropical forests across a rainfall gradient. *Canadian Journal of Botany*, *80*(8), 818–826.

Suryanarayanan, T., Thirunavukkarasu, N., Govindarajulu, M., Sasse, F., Jansen, R., & Murali, T. (2009). Fungal endophytes and bioprospecting. *Fungal Biology Reviews*, *23*(1), 9 – 19.

Tateno, O., Hirose, D., Osono, T., & Takeda, H. (2015). Beech cupules share endophytic fungi with leaves and twigs. *Mycoscience*, *56*(3), 252 – 256.

Taylor, D. L., Walters, W. A., Lennon, N. J., Bochicchio, J., Krohn, A., Caporaso, J. G., & Pennanen, T. (2016). Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for illumina amplicon sequencing. *Applied and Environmental Microbiology*, *82*(24), 7217–7226.

Team, R. C. (2014). R: A language and environment for statistical computing. R foundation for statistical computing. Vienna, austria; 2014.

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, H., Hildebrand, F., Bork, P., & Abarenkov, K. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys*, *10*, 1–43.

Telford, R. J., Vandvik, V., & Birks, H. J. B. (2006). Dispersal limitations matter for microbial morphospecies. *Science*, *312*(5776), 1015–1015.

Thomas, D., Vandegrift, R., Carroll, G., & Roy, B. (2014). Code: Spatial ecology of the genus Xylaria in a tropical cloud forest. https://doi.org/10.6084/m9.figshare.1224345.v1

Thomas, D. C., Vandegrift, R., Ludden, A., Carroll, G. C., & Roy, B. A. (2016). Spatial ecology of the fungal genus *Xylaria* in a tropical cloud forest. *Biotropica*, *48*(3), 381–393.

Thomas, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika*, *36*(1/2), 18–25.

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*(6), 1139–1150.

Vandegrift, A. W. R. (2016). *Ecological Roles of Fungal Endophytes*. Ph.D. thesis, 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 Phytologist*, *206*(4), 1196–1206. 2014-18496.

Vaz, A., Fontenla, S., Rocha, F. S., Brandao, L. R., Vieira, M. L., De Garcia, V., Aristóteles, G.-n., & Carlos, A. (2014). Fungal endophyte b-diversity associated with Myrtaceae species in an Andean patagonian forest (Argentina) and an atlantic forest (Brazil). *Fungal Ecology*.

Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, *116*(7), 1120–1127.

Veech, J. A. (2013). A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, *22*(2), 252–260.

Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. PMID: 20565040.

Veneklaas, E. J. (1991). Litterfall and nutrient fluxes in two montane tropical rain forests, Colombia. *Journal of Tropical Ecology*, *7*(3), 319–336.

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*, *16*(1), 118 – 127. Growth and development.

Veresoglou, S. D., & Rillig, M. C. (2014). Do closely related plants host similar arbuscular mycorrhizal fungal communities? a meta-analysis. *Plant and soil*, *377*(1-2), 395–406.

W., K. S., & C., M. R. (2014). Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. *Botany*, *92*(4), 303–311.

Whalley, A. (1996). The Xylariaceous way of life. *Mycological research*, *100*(8), 897–922.

Wilson, D. (1995). Endophyte: the evolution of a term, and clarification of its use and definition. *Oikos*, (pp. 274–276).

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*(6), 699–703.

Wright, S. (1940). Breeding structure of populations in relation to speciation. *The American Naturalist*, *74*(752), 232–248.

Yahr, R., Schoch, C. L., & Dentinger, B. T. (2016). Scaling up discovery of hidden diversity in fungi: impacts of barcoding approaches. *Phil. Trans. R. Soc. B*, *371*(1702), 20150336.

Zimmerman, N. B., & Vitousek, P. M. (2012). Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. *Proceedings of the National Academy of Sciences*, *109*(32), 13022–13027.