UNIVERSITY OF CALIFORNIA

SANTA CRUZ

THE SPILLOVER OF FUNGAL PATHOGENS FROM LOCAL PLANT SPECIES TO NOVEL NON-NATIVE SPECIES IN A CALIFORNIA COASTAL GRASSLAND

A thesis submitted in partial satisfaction
of the requirements for the degree of

MASTER OF ARTS

in

ECOLOGY AND EVOLUTIONARY BIOLOGY

by

Zackery Richard Carmean Shearin

December 2021

The Thesis of Zackery Richard Carmean Shearin
is approved:

_____
Professor Ingrid M. Parker

_____
Professor Gregory S. Gilbert

_____
Professor Rachel Meyer

_____
Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

The spillover of fungal pathogens from local plant species to novel non-native species in a California coastal grassland by Zackery Richard Carmean Shearin

Abstract:

The rare species advantage postulates that dominant plant species experience greater disease pressure, as their abundance can support the spread of specialized pathogens. Yet most pathogens are not host specific – pathogen host ranges frequently encompass multiple closely related plant species, resulting in increased sharing of pathogens between closely phylogenetically related community members. We examined the role host phylogenetic rarity plays in pathogen spread in a diverse plant community by inserting novel hosts (non-native plant species newly introduced into a novel habitat) into a coastal grassland and identifying the patterns of pathogen spillover from resident plant species onto the novel hosts. We identified fungi present in leaf tissues of novel hosts via metagenomic sequencing of the fungal internal transcribed spacer (ITS) region to generate operational taxonomic units (OTUs). We found that 161/342 (47.1%) OTUs identified were singletons. Excluding singletons, the median number of novel plant species colonized by a given OTU was 4, and the mean was 11.5, while 19/181 (10.5%) of non-singleton OTUs were found on 14 (50%) or more novel hosts. The proportion of broad generalist and rare fungal taxa on novel hosts were similar to several other systems, including other grasslands, temperate and boreal forests, tropical rainforests, wetlands, and deserts. We did not find a relationship between phylogenetic relatedness of a novel host to the surrounding vegetation and the fungal species richness on that host. Weighting phylogenetic distance by local plant

species abundance did not strengthen the relationship between phylogenetic distance and fungal richness on novel hosts.

Acknowledgements:

This research was funded by a grant from the National Science Foundation DEB-1655896 to Ingrid M. Parker and Gregory S. Gilbert at the University of California, Santa Cruz. I thank undergraduate students Elizabeth Davis, Diana Tataru, Taelor Whittington, Joia Fishman, Natalie Gonzalez, Hector Borjas, Elizabeth Martin, Nate Blackmore, and Lexi Necarsulmer, and family members Anusha Radhakrishnan and Eli Gilbert for assisting with greenhouse work, field work, and sample collection. I also thank Mat Lutz and Natalie Gonzalez for assisting with sample processing, and Miranda Melen for creating Figure 3. A special thank you goes to the Parker-Gilbert Laboratory Manager, Asa Conover, for handling the majority of the molecular work and data cleanup. I would also like to thank the Glen Millhauser lab at UCSC for assisting with lyophilization and Beth Shapiro at UCSC for helping us design our PCR and sequencing protocols. Thank you to the members of the Parker-Gilbert lab group for providing feedback at multiple points during this experiment's execution and write up. I would like to thank my committee members, especially Rachel Meyer for agreeing to work with me on such an accelerated schedule and for their flexibility. Finally, I would like to give special thanks to my advisors, Ingrid Parker and Greg Gilbert, who have never wavered in their support for me and my graduation despite an inordinate number of personal struggles on my end. I would like to thank them for their assistance with developing my ideas, expertise in both writing and statistics, assistance with

understanding and developing the interpretations of the results of this research, and emotional support.

Introduction:

Two interrelated phenomena help describe the spread of a pathogen within a plant community – pathogen host ranges and plant community structure, including both numerical and phylogenetic components (Gilbert et al., 2012; Parker et al., 2015; Termorshuizen, 2016). Traditional approaches to understanding pathogen spread in plant communities frame the interaction in terms of host abundance – increased host density increases the probability of disease spread, resulting in higher disease prevalence on dominant host species (Bagchi et al., 2010; Bell et al., 2006). The role of host abundance on disease spread through complex plant communities is perhaps best understood in relation to the rare species advantage, which postulates that dominant plant species experience greater disease pressure because their abundance can support the spread of specialized pathogens. This provides an opportunity for rarer, less susceptible species to proliferate due to decreased competition with the dominant plant species (Connell et al., 1984; Metz et al., 2010). As a result, the plant community maintains a higher level of diversity, which in turn reduces the probability of disease spread overall by interspersing the hosts susceptible to any given pathogen among non-susceptible hosts, a phenomenon known as the dilution effect (Johnson et al., 2013).

However, the quantity of disease that develops may not be a simple function of the density of a focal host, because pathogens are often able to infect and spread across

multiple host species (Barrett et al., 2009). If a pathogen has evolved the ability to infect a host, it is more likely to be capable of also infecting a closely related host than it is a more distantly related one – a concept known as phylogenetic signal in host range (Gilbert & Webb, 2007, Parker et al., 2015). As a result, there is always a degree of pathogen spillover between closely related hosts. Phylogenetic signal in host range suggests that the dilution effect – that is, the protective effect of biodiversity – is influenced by the phylogenetic diversity of the plant community. A community with many closely related species would be unlikely to benefit from the dilution effect. It also suggests that the host range of pathogens has a strong influence on whether community diversity will mediate the impact of disease. Despite these broad-reaching implications, however, few studies have looked at the interaction between pathogen host range and the phylogenetic structure of the plant community (Gilbert & Parker, 2010; Parker et al., 2015).

The ever-increasing movement of plant material by humans and geographic range expansion caused by climate change create many opportunities for evolutionarily novel encounters between plant hosts and pathogens (Anderson et al., 2004; Knops et al., 1999; Thomson et al., 2010). Elucidating the underlying patterns of pathogen host range can inform our understanding of key processes in invasion ecology. Two major hypotheses that describe the interaction between plant invaders and pathogens are the Enemy Release Hypothesis and the Biotic Resistance Hypothesis (Termorshuizen, 2016). The Enemy Release Hypothesis postulates that when invaders leave their native range, they also leave behind the coevolved pathogens that ordinarily would have

regulated their spread and abundance (Blumenthal, 2006; Liu & Stiling, 2006; Mitchell & Power, 2003). Enemy release is likely to be most important in systems where the dominant pathogens in the native range are highly specialized (Parker & Gilbert, 2007).

In contrast, the Biotic Resistance hypothesis states that pathogens that spill over onto introduced host species have the potential to regulate invasions (Byun & Lee, 2017; Flory & Clay, 2013; Knevel et al., 2004). Introduced plant species may be infected by pathogens through spillover from plants in the surrounding resident community. Pathogens that colonize introduced species are expected to either be broad generalists that infect a wof host plants, or given the documented phylogenetic signal in host ranges, to be more specialized pathogens that have evolved traits needed to infect closely related neighbors (Barrett et al., 2009; Barrett & Heil, 2012; Philibert et al., 2011). Introduced species that are phylogenetically distant from the local community would be expected to only be colonized by highly generalist pathogens; novel species with abundant close relatives in the local community should have a greater probability of also being colonized by more selective polyphagous pathogens (Bufford et al., 2016). Testing these expectations requires describing the structure of host specificity among pathogens that colonize novel hosts and assessing whether the richness of the fungal assemblage on novel hosts is a function of the abundance of close relatives in the local community.

Given the effects of phylogenetic distance between novel species and local neighbors on pathogen spillover, there are several specific expectations. First, phylogenetically isolated novel hosts are expected to have lower taxon richness of

fungal colonists than novel hosts with abundant close relatives as neighbors. Second, novel hosts are expected to be colonized by a mixture of more specialized and more generalist fungi; fungi with broad host ranges are likely to be found on a random set of novel hosts, whereas fungi with narrow host ranges would be more likely associated with novel hosts that have abundant closely related neighbors in the surrounding community. Together, these patterns suggest that there should be a relationship between the richness of fungi found on a novel host and the phylogenetic distance to local species. Accounting for additional effects of the local abundance of each host species on spillover (i.e., density dependence) should strengthen the relationship between fungal r and phylogenetic distance to the local community.

We explored the interacting roles of pathogen host range and phylogenetic rarity by measuring which pathogens of resident species colonize novel hosts introduced into a Pacific coastal grassland. Specifically, we asked:

- What is the relative frequency of generalist pathogens and more host-selective pathogens in the community of fungi colonizing novel hosts?
- Can we predict spillover of fungal species onto focal hosts based on composition of plant species in the surrounding community?
    - Does phylogenetic rarity predict fungal species richness?
    - How does the local abundance of closely related neighbors contribute to fungal species richness?

Methods:

Study site description. Our study area consisted of ten 20-m diameter circular sites (314 m^2) located in the Great Meadow on the University of California Santa Cruz (UCSC) campus on the central coast of California, USA (36° 59' 18.09" N, 122° 3' 31.29" W). This 30-ha grassland experiences a Mediterranean climate, with dry summers and winter rains. Previous work by the Parker-Gilbert research group identified 43 vascular plant species in these sites with a lognormal distribution of relative abundance. Introduced Eurasian annual grasses dominate the sites, with species such as *Bromus diandrus* and *Avena barbata* found in high abundance throughout. Plots sampled within the meadow are the same used by Parker et al. (2015), save four (plots 1, 4, 6, and 10), which were either disturbed or experienced woody plant encroachment. Plots were randomly placed based on GPS location and ranged from 44-260m apart at their centers (Table 1).

Introduction of novel hosts. Within each of the plots we placed randomized arrays consisting of 28 plant species that, to our knowledge, have never been found in the state of California, either in the wild or grown as horticultural species (Table 2). We selected herbaceous species based on their phylogenetic relationships to locally important clades (*Poaceae*, *Asteraceae*, *Fabaceae*) to obtain a wide range of phylogenetic rarity relative to the local species (Figures 1 and 2). 14 different plant families were represented in the final selection. We ordered seeds from a range of geographically diverse native plant nurseries to avoid horticultural species or plants bred for disease resistance.

Seeds were sown in the UCSC greenhouses and transplanted into 3.8 cm X 14 cm Cone-tainers™ (Stuwe & Sons, Inc., Corvallis, OR) after germination and placed into RL98 Cone-tainer™ trays (Figure 3). The randomized arrays were placed in the field, in triplicate, with spaces on all four sides of each Cone-tainer™ to minimize light competition and to expose all novel hosts more uniformly to the same biotic and abiotic conditions. These arrays were bottom watered with cotton string wicks and tubs of water and placed in a line from the center of the plots to the Northwest. They were surrounded by a cage of thin deer netting to discourage large mammal herbivory. Before placing plants into the meadow, all leaves and cotyledons with any sign of necrosis were removed. Control samples of each novel host were taken from plants that were not placed in the field at this stage. The novel hosts were left in the field for 51 days, 28 days longer than the previous experiment, facilitated by favorable weather conditions (Parker et al., 2015). Novel host species were not permitted to flower or produce seed to eliminate any biosafety risk created by this project.

Harvesting and Sample Preparation. Novel host arrays were removed from the field on 28 April 2018. Ten samples from the leaves of each plant were collected using a standard 6-mm diameter hole punch; for plants with narrow leaves, an approximately similar surface area was excised using scissors. If a plant presented too little leaf tissue to obtain ten samples, additional samples were taken from the other two plants of the same plant species from the same site, so that the total number of samples per species per site was 30. Samples were pooled by species within a site and surface sterilized to eliminate epifoliar fungi by submerging them in 70% ethanol for one minute, then 0.5%

sodium hypochlorite for one minute, and finally rinsed in sterile deionized water. Samples were stored in 1-mL microcentrifuge tubes at -80 C until lyophilization and DNA extraction were performed.

Local species abundance. We quantified the relative abundance of all extant vascular plant species in each of the plots. For each plot, we quantified percent cover of eight, 50 cm X 100 cm rectangular quadrats, four each from randomly selected points on a North-South and East-West transect. We used a visual estimation of cover by dividing each quadrat into five 20 cm X 50 cm subsections and two observers independently quantified percent cover for each subsection. The two independent observations were averaged to obtain the quadrat percent cover. This method was verified by Parker et al. (2015) and found to have high repeatability. Rare species were quantified by estimating the area of coverage at the plot scale (314 m^2). Rare species were defined as having been found in three or fewer quadrats within a given plot.

DNA extraction and library preparation. The surface-sterilized samples stored frozen in microcentrifuge tubes were lyophilized in 600 mL lyophilization flasks containing approximately 60 microcentrifuge tubes each at 200 mbar for 48 hours and stored at room temperature.

Prior to DNA extraction, leaf material from samples with >20 mg of lyophilized biomass were partitioned into two subsamples. Subsamples were separately weighed and independently processed for DNA extraction. Extractions were performed using Quick-DNA Fungal/Bacterial Miniprep Kits (Zymo Research, Tustin, CA; Cat. No.

D6005) following manufacturer's protocol (including addition of 0.5% (v/v) beta-mercaptoethanol to the Genomic Lysis Buffer) with one modification: input leaf tissue was soaked overnight at room temperature in the ZR BashingBead Lysis Tube with 750 µl of BashingBead Buffer prior to bead-beating. Bead-beating was performed for 10 minutes with a BioSpec Mini-Beadbeater-16. Samples were visually inspected for homogeneity and bead-beat for an additional 5 minutes if intact leaf tissue remained.

Libraries of fungal internal transcribed spacer (ITS) rRNA gene amplicons were prepared using a two-step PCR process. Fungal ITS amplification (PCR1) was performed using forward primer ITS-1F_KYO2 (5'-TAGAGGAAGTAAAAGTCGTAA-3', Toju et al. 2012) and reverse primer ITS-4 (5'-TCCTCCGCTTATTGATATGC-3', White et al. 1990), which was modified to include a 5' 13-bp stub sequence (5'-GCTCTTCCGATCT-3'). The stub sequence was added to provide a priming site for PCR2. PCR2 primers added 8-bp dual-matched molecular identification tags (MID tags) to both ends of each amplicon. MID tags were unique for each of the samples per library, which included several negative (blank) and positive controls. Positive controls were created using a mock community of non-biological synthetic ITS sequences as described by Palmer et al. (2018).

25 µl PCR reactions were carried out in 96-well plates. The two DNA extracts for each plant species were pooled prior to PCR1. PCR1 consisted of 11 µl template DNA, 0.75 µl of each primer (10 µM), and 12.5 µl of KAPA HiFi HotStart ReadyMix (2X) (Kapa Biosystems, Wilmington, MA). PCR1 conditions were: 95°C for 3 min, 29 cycles of 98°C for 30 s, 60°C for 15 s, 72°C for 50 s, and a final 1 min at 72°C.

PCR1 products were visualized and DNA-quantified by gel electrophoresis. 5 µl of each PCR1 product was run on a 1% agarose-TBE gel at 120V for 45 minutes. Gels were imaged using an iBright CL750 Imaging System (Invitrogen, Carslbad, CA). Images were processed for DNA quantification using iBright Analysis Software (Thermo Fisher Scientific, Waltham, MA). A standard curve was constructed using a DNA dilution series to relate lane brightness to known DNA concentration, measured using Qubit 1X dsDNA HS Assay (Invitrogen, Carslbad, CA). Gel image brightness values for each PCR1 product were then used to calculate DNA concentration. Using these concentration values, PCR1 products were diluted to 0.4 ng/µl in nuclease-free water.

PCR2 reactions consisted of 10.5 µl diluted PCR1 product, 2 µl of combined forward and reverse indexing primers (each at 5 µM), and 12.5 µl of KAPA HiFi HotStart ReadyMix (2X) (Kapa Biosystems, Wilmington, MA). PCR2 conditions were: 95°C for 3 min, 8 cycles of 98°C for 30 s, 56°C for 15 s, 72°C for 50 s, and a final 1 min at 72°C.

PCR2 products were visualized and DNA-quantified as described above. Samples that failed to produce visible bands after both rounds of PCR were re-amplified following the same 2-step PCR protocol, except with 35 cycles for PCR1 instead of 29. Successfully amplified PCR2 products were pooled to construct a library, and cleaned with magnetic KAPA Pure Beads (Kapa Biosystems, Wilmington, MA) using a 0.6X beads-to-sample ratio. Each library was loaded onto one PacBio Sequel II SMRT Cell by diffusion loading and sequenced with a movie length of 15 hours in

circular consensus sequencing (CCS) mode. Creation of SMRTbell sequencing libraries and PacBio sequencing was performed by Brigham Young University's DNA Sequencing Center.

Analyses:

Statistics were done using R 4.1.0 (R Core Team, 2021). For all measures, "global" refers to a calculated value based on all samples in the study regardless of site or species.

Bioinformatics. Circular consensus sequences (CCS) were generated and demultiplexed using PacBio's SMRT Link v8.0 software. The "DADA2" package was used in R to denoise CCS reads and to assign taxonomy using the UNITE database general FASTA release for all eukaryotes (Callahan et al. 2016, Abarenkov et al. 2020). Reads were clustered into operational taxonomic units (OTUs) using the "tip_glom" function from the "phyloseq" package in R.

Phylogenetic tree generation. A dated phylogenetic tree of the 44 local plant species as well as the 28 novel hosts was generated using the "V.PhyloMaker" (https://github.com/jinyizju/V.PhyloMaker) package in R according to APG4 taxonomy (The Angiosperm Phylogeny Group et al., 2016; Figure 1). The *Rumex* clade, containing three species, was manually adjusted as it was not fully resolved (Grant et al., 2020). Species names were updated using the "Plantminer" function from "Taxize" to match the most current names given in The Plant List vs 1.1 (http://www.theplantlist.org/).

Quantifying biodiversity and phylogenetic distance measures. Species richness (S), Shannon diversity index (H), and Pielou's evenness index (J) were calculated for the local species in each plot as well as for the entire field site using the package "vegan" in R (Table 1). The pairwise phylogenetic distance (time of independent evolution in millions of years - Myr) between any two plant species was calculated using the "cophenetic" function from "picante", and the distribution of phylogenetic distances between novel hosts and the local plant community was visualized (Figure 2). Nearest taxon distance (NTD) is defined as the pairwise distance to the nearest taxon neighbor within the local plant community for each novel host. The 10^{th} quantile phylogenetic distance (PD10) is the pairwise distance of the local plant species that sits at the 10^{th} quantile if the local plant species are sorted by pairwise distance to the focal species (Parker et al., 2015; Gilbert and Parker, *In Review*). One can consider PD10 as representing the influence of multiple close relatives. The mean pairwise distance (meanPD) was calculated as a simple mean of all the pairwise distances of the local community to the focal host.

Rarefaction. As the fungal communities have not previously been studied in these host plants or at our site, we created sample-based rarefaction curves to evaluate how well the design adequately captures the variety of fungi present in the hosts, and to estimate the total richness of the community.

Species accumulation curves of observed fungal taxa were generated for each of the 28 host species, drawing randomly from the sites they were sampled from and repeating for 100 permutations. A separate rarefaction curve was generated by

randomly selecting one host species from each of the 10 random sites in random order (1000 permutations to calculate median and 95% CI) to represent the null model, which assumes that each of the novel host species to be equivalent fungal hosts.

Fungal abundance and host range. The distribution of fungal richness per host was explored as the number of OTUs associated with each host combined across all ten sites. Singletons, fungi found on only a single sample, were visualized separately from non-singletons.

Phylogenetic distance weighted by local species abundance. The weighted PD10 was calculated by sorting all the neighboring vegetation species by phylogenetic distance from the focal novel host, and then calculating the cumulative relative abundances from the closest to most distant relatives using the function "weighted.mean" in R. The phylogenetic distance that corresponded to the 10^{th} quantile in cumulative relative abundance is the weighted PD10 value. The mean weighted mean PD was calculated as the arithmetic mean of the phylogenetic distances from novel hosts to each of the neighboring plant species, weighted by the relative abundance of the neighboring plant species.

Results:

Fungal community analysis. Of 280 novel hosts placed in plots, 240 survived the field experiment and generated viable sequence data. The lowest survival rate was 40%, whereas 11 had a 100% survival rate (Table 2). Most mortality was apparently due to insufficient watering or herbivory. The phylogenetic distance between novel hosts and

species in the local community varied widely (Figure 2), with the NTD ranging from 0.54 to 222.85 Myr (mean = 89.55 ± 60.54, global = 62.47), PD10 ranging from 39.06 to 263.37 Myr (mean = 134.62 ± 73.16, global = 126.61), and meanPD ranging from 212.19 to 263.56 Myr (mean = 229.21 ± 13.01, global = 231.67).

Our sampling appears to have captured the majority of fungal taxa for some novel hosts, but for other hosts, the species accumulation curves have not yet saturated (Figure 4). One novel host curve was significantly above the null curve – *Pseudognaphalium obtusifolium*. Chao estimators for the number of fungal OTUs expected on individual host species ranged from 16.2 to 317.8 (mean = 111.1 ± 69.1), with a global estimate of 576.4 across all 28 host species (Table 2). Based on the Chao estimators, we captured 59.3% of the "true" richness of the fungal community on novel hosts, and for individual hosts that percent ranged from 6.3% (*Corydalis curvisiliqua*) to 67.9% (*Ipomopsis rubra*).

Of the 342 OTUs identified, 161 (47.1%) were singletons. (Figure 5). Excluding singleton OTUs, four fungi were limited to a single host species, while three OTUs, two species in the yeast genus *Malassezia* and an unidentified fungus in the *Pleosporaceae*, were found on all 28 host species (Table 3). The number of OTUs found on individual novel host samples varied widely, from 12 to 124 (mean = 44.2 ± 24.2) (Figure 6).

The majority (90%) of OTUs were found in 12 or fewer samples and 8 or fewer novel hosts. The mean number of samples per OTU was 6.56 ± 16.84 and the median

was 2. The mean number of novel hosts per OTU was 3.62 ± 4.86 and the median was 2. Excluding singletons, 90% of OTUs were found in 25 or fewer samples and 14 or fewer host species. The mean number of samples per OTU excluding singletons was 11.51 ± 22.02 and the median was 4. The mean number of novel hosts per OTU excluding singletons was 5.94 ± 5.77 and the median was 4. OTU641, *Malassezia* sp., was found on all 28 novel hosts and in 152 (43.2%) samples, the most cosmopolitan fungus in our study (Table 3). Histograms of number of taxa per host and per sample follow the classic hollow curve, with most hosts and samples having few fungal OTUs and a few having very many (Figure 5).

Local community analysis. Surveys of the local vegetation within each of the 10 plots found percent cover ranging from 55.0% to 90.8% (mean = 73.7 ± 11.1%). Three grasses (family *Poaceae*) were the most abundant species in the meadow vegetation: *Bromus diandrus* was the most abundant local species, with a mean percent cover across all sites of 18.86 ± 4.72%, followed by *Avena barbata* (mean = 17.55 ± 7.61%) and *Brachypodium distachyon* (mean = 7.63 ± 6.25%). The majority of local species (31, or 70.5%) had a percent cover of less than 1%. There were 44 plant species across all 10 plots (Global local species richness, S), with individual plots ranging from 17 to 27 (mean = 22.7 ± 3.02) (Table 1). The global Shannon diversity index (H) was 2.34 (1.50 - 2.47, mean = 1.91 ± 0.31), and the global Pielou's evenness index (J) was 0.62 (0.49 - 0.75, mean = 0.61 ± 0.090).

Phylogenetic analyses. There was no relationship between the number of fungal taxa on a host and its phylogenetic distance from the local community (Figure 7). Linear

regressions indicated no statistically significant relationship for any of the three measures of phylogenetic isolation: NTD ($p = 0.48$, $R^2 = 0.019$, $N = 26$), PD10 ($p = 0.38$, $R^2 = 0.029$, $N = 26$), or meanPD ($p = 0.48$, $R^2 = 0.019$, $N = 26$). Weighting phylogenetic distance by local host species abundance did not result in a statistically significant trend (Figure 8). Linear regressions showed no significant relationship for PD10 weighted by local plant species abundance ($p = 0.63$, $R^2 = 0.0089$, $N = 26$), nor for meanPD weighted by mean local plant species abundance ($p = 0.37$, $R^2 = 0.031$, $N = 26$). Similarly, regressions run with singletons excluded or by using Chao estimators indicated no statistically significant relationships.

Phylogenetic distance from individual local species in the family *Asteraceae* do show a statistically significant relationship with the number of fungal taxa found on novel hosts (Figure 9). The pairwise phylogenetic distance between novel hosts and *Baccharis pilularis*, a prominent *Asteraceae* in the local plant community, accounts for 25.4% of the observed variance in OTU richness in novel hosts ($p = 0.006$, $R^2 = 0.254$, $N = 26$, $F = 8.872$, $y = -0.02628x + 13.5696$).

The richness of fungal OTUs on plant samples was shaped more strongly by host species than location in the meadow. A two-way ANOVA of the number of OTUs as a function of novel host species and site showed statistically significant variation in the number of OTUs among host species ($p < 0.001$, $F = 6.50$, $N = 27$) but not among site ($p = 0.79$, $F = 0.61$, $N = 9$). Figure 6 visualizes the variation in the number of OTUs for each novel host.

Discussion:

Fungal community structure. Novel hosts showed a broad range of OTU richness, and the completeness of our sampling varied across the hosts (Figure 4). Several species, such as *Lupinus perennis*, *Rumex altissimus*, and *Chamaecrista fasiculata* show a nearly flat species accumulation curve, reaching an asymptote after only a few samples. By contrast, a few species, namely *Pseudognaphalium obtusifolium*, *Liatris spicata*, and *Phaecelia congesta*, are steeply curved and do not reach an asymptote. This appears to be supported by Chao estimates, as novel hosts with nearly asymptotic curves have some of the lowest estimated "true richness", whereas the three steeply curved novel hosts mentioned have the largest estimates (Table 2). It is possible that the unusually high richness of the latter three species is a result of their physiology – their leaves are covered in long, sticky trichomes that may be effective at capturing and retaining fungal spores from the environment. This dense protective barrier may also make surface sterilization procedures less effective, so the richness of the fungal communities in their tissues may be lower than observed. Overall, most of the species were fairly well sampled – adequate to address our questions about factors influencing fungal colonization across hosts.

A small proportion of the fungal taxa we found were broadly host-generalist. Of the 342 OTUs identified, 19 (5.6%) were found on at least half of the 28 novel hosts, distributed across multiple families (Table 3, Figure 5). Despite differences in both plant and fungal richness, this broadly falls within the range (5-10%) found in other grasslands (Toju et al., 2019; Whitaker et al., 2018, Kendig et al., 2020), temperate and

boreal forests (Arnold & Lutzoni, 2007; Bálint et al., 2014; Christian et al., 2016), deserts (Massimo et al., 2015), wetlands (Clay et al., 2016), and tropical forests (Christian et al., 2017). Interestingly, the proportion of broad generalists seems to remain in the same range regardless of sampling depth or whether culturing or metagenomic methods are utilized. Abundant, broad generalists play an important role in plant microbiomes and plant community structure (Facelli et al., 2018). Their relative abundance and impact on a large number of plants make them an important player in structuring and regulating plant communities through disease.

In total, 47.1% of observed fungal taxa were singletons (Figure 5B). As with broad generalists, the relative proportion of singletons remains consistent across biomes and methodologies (± 10% of our result). Understanding the prevalence and biological impact of rare taxa is another important piece of understanding fungal assemblages in plant microbiomes (Põlme et al., 2017; Gil-Martínez et al., 2018). There are several challenges to interpreting and comparing these data. Foundational work on this subject utilized culturing methods, which are inherently incapable of identifying the majority of fungi (Wu et al., 2019), so historical data on the fungal microbiomes of plants is of limited usability; though metagenomic methodologies largely circumvent this issue, greatly improving our understanding of the true diversity of microbiomes in plants (Unterseher et al., 2010; Delmont et al., 2011). Another challenge is to interpret the biological significance of singletons – they could be specialized fungi that only infect one or a few hosts, or they could simply be rare (Lim et al., 2012). Determining how rare fungal taxa, especially host-specific taxa, impact disease pressure and plant

community structure is an important challenge for disease ecologists – their impact may be smaller than broad generalists, but they can regulate plant communities by exerting disease pressure on specific hosts. Understanding which taxa are host-specific and what impact they have on their hosts would require inoculation experiments, which require a large investment and can only include taxa that are culturable. Increasing the sample size of a metagenomic microbiome survey like this would also help resolve the question for some taxa. Future work could also utilize simulations to estimate the probability of different levels of specialization, given the frequency with which individual OTUs were found (Ferrer & Gilbert, 2003).

Prominent fungal community members and their potential role. Yeasts are less likely than filamentous fungi to grow on culture media commonly used for plant microbiome studies (Di Menna, 1957; Abdillah et al., 2020) so our understanding of the prevalence and ecological importance of yeasts are underdeveloped; though research on yeasts in plants have been ongoing for decades (Starmer et al., 1978; Starmer et al., 2003). Of the 19 broad generalists identified in our study, eight are yeasts, including two of the three taxa that were present on all 28 novel hosts – *Malassezia* sp. and *Malassezia globosa* (Table 3). Yeasts belonging to the genus *Malassezia* are in a sister clade to plant pathogenic smuts and *Exobasidium* (Wang et al., 2014), and are known as part of the normal human skin microbiome and opportunistic human pathogens (Findley et al., 2013). With metagenomic assessments of microbiomes of plants, corals, and other hosts, *Malassezia* spp. have been identified in a variety of habitats (Amend 2014) and it is increasingly clear that they are not

exclusive to human microbiomes. Their potential role as a member of the microbial community in plants deserves additional attention (Roy et al., 2009). According to a review article by Amend (2014), *Malassezia* spp. have been associated with diseased tissue in corals. Given that they are also associated with disease in mammals, it is reasonable to assume they could play a role in plant disease as well, though whether they are a causative agent of disease or simply a symptom will need to be investigated. Obtaining axenic cultures for inoculation studies is the optimal next step forward, though culturing *Malassezia* spp. is challenging (Ashbee, 2007). Verifying that *Malassezia* spp. is actively metabolizing in plant isolates, especially in diseased tissue can be accomplished with mRNA-based studies.

The significance of phylogeny in plant disease. Previous work by Parker et al. (2015) at our site found that novel plant species with short phylogenetic distances to species in the neighboring community suffered greater disease pressure, and we hypothesized that this was because those hosts would be colonized by a greater number of fungal taxa from the local plant community. Our findings did not support that hypothesis either based just on the phylogenetic distance to neighboring plants (Figure 7) or when those distances were weighted by local plant abundance (Figure 8). An assumption of our study is that the fungi that colonize the novel hosts would primarily have spilled over from the local plants within each 20 m radius plot; it is possible that inoculum coming from reservoirs outside the plots were responsible for a large proportion of available inoculum. This would be consistent with airborne fungal communities being broadly mixed across a landscape, as indicated by Crandall et al. (2020).

Additionally, the measures of plant cover we used to estimate local plant abundance may be insufficient to represent the relative influence of local plants on novel hosts' fungal communities. The shrub *Baccharis pilularis*, for example, is atypical from most other plants in the meadow community because it is taller than any other plant at our study site, standing 2-3 meters at maturity; this may allow fungal spores originating from its leaves to travel much further than those of other plants. *B. pilularis* is also the most prominent evergreen species at our study site – the majority of local plant species are annuals or biennials. Perennials may serve as important disease reservoirs for pathogens that infect annual species (Thrall et al., 1993; Kendig et al., 2020), so *B. pilularis* may exert a disproportionately large effect on novel hosts' fungal communities. Our study site was also bordered on one side by a mixed-evergreen forest and by an organic farm on another side. If the airborne fungal communities were indeed as homogeneous as suggested by Crandall et al. (2020), fungi originating in those habitats and outside the areas we measured may have had significant influence on microbiome formation in the novel hosts. Finally, we did not differentiate between pathogens and other functional groups, so it is possible that a relationship exists between the number of pathogens on a novel host and phylogenetic distance, but it was masked by the inclusion of non-pathogens. While it is still an approach in development, classifying OTUs into functional groups is possible (Liang et al., 2019). This is a necessary step forward to understand whether pathogens are specifically sensitive to the phylogenetic relationships between host plant species.

We found a statistically significant relationship between the number of fungal OTUs found on a novel host and the pairwise distance from novel hosts to *Baccharis pilularis* (Figure 9), as well as to all other local plant species of the family *Asteraceae*. This is driven in part because two of the three species with the greatest fungal richness (*Pseudognaphalium obtusifolium* and *Liatris spicata;* Figure 6) are in the *Asteraceae*, suggesting that there may be a relationship between that family and patterns of fungal community richness at our site. In Figure 9, *P. obtusifolium* and *L. spicata* are the two uppermost and leftmost points, at around 47 Myr, and are clearly driving the regression trendline. It may be that these species have traits that allow them to support a wide variety of fungal endophytes, or it could be that the trichome-filled leaf surface of these species could make them difficult to properly surface sterilize. Resolving whether the taxa found on species like these are epifoliar or colonize host tissue by modifying and validating the efficacy of surface sterilization techniques is integral in determining whether this pattern is a result of a biological signal or methodological error. A more focused sampling of families closely related to *Asteraceae* can elucidate whether *Asteraceae* have a larger relative impact on fungal species richness on novel hosts.

Conclusion. We were able to partially answer the two main questions we posited. First, we focused on describing the structure of fungal communities that develop on novel host plants. We asked: "What is the relative frequency of generalist pathogens and more host-selective pathogens in the community of fungi colonizing novel hosts?" We can assert that the proportion of both broad generalists and rare species is similar to several other communities that have been studied. However, an important limitation is the

difficulty in untangling whether rare species are rare because they are specialists or because they are simply uncommon in the habitat. Only four OTUs were found in more than one sample but on a single host, which is suggestive of possible specialization, while we could not determine whether the other OTUs found on only a single host were specialized or not. Accepting those caveats (and excluding singletons), we found the fungal community was dominated by species that utilized a few hosts--not one, and not many.

Our second question was: "Can we predict spillover of fungal species onto focal hosts based on composition of plant species in the surrounding community?" We also asked: "Does phylogenetic rarity predict fungal species richness?", and "How does the local abundance of closely related neighbors contribute to fungal species richness?" We were unable to predict the degree of spillover of fungal species onto novel hosts as a function of phylogenetic distance on a community-wide level. Abundance of plants in the local community did not appear to contribute significantly to fungal species richness, though there may be confounding factors that were not addressed with this experiment that warrant inquiry. For example, alternative hosts in surrounding habitats (e.g. mixed evergreen forest, agricultural land, and the campus arboretum) may have donated pathogens to the novel hosts over a larger spatial scale to a degree that swamped out the effect of the local community. Novel hosts may not have been in the field long enough to be colonized by more specialized pathogens. Additionally, microbes that are incidental colonists, or epifoliar fungi that escaped surface sterilization, might obscure patterns in the community of pathogenic fungi.

We suggest several avenues for further inquiry into these questions. To separate the patterns of pathogenic fungi from incidental species, one could replicate these analyses using subsets of the data including just OTUs with known functions. To elucidate the question of host-specific fungal taxa, inoculation studies would provide definitive information on host range and the potential for spillover for culturable fungi. An increased sampling depth and simulations can resolve some of the unculturable fungi. For common, yet unculturable taxa like *Malassezia* spp., obtaining axenic cultures would be ideal, but future research can also rely on mRNA-based approaches and identifying taxa of interest predominantly in diseased tissues. Information on which fungi are associated with hosts in the surrounding plant community would help trace the fungi on novel hosts back to the source of pathogen spillover. Finally, we suggest validating surface sterilization methods combined with a more focused sampling of families closely related to *Asteraceae* to resolve whether the observed fungal richness in some *Asteraceae* spp. is the result of them having a disproportionately large impact on the fungal community in novel hosts.

Tables and Figures:

Table 1 - Pertinent measures of the local plant community by plot. Values for species richness (S), Shannon diversity index (H), and Pielou's evenness index (J) for each plot from which local species abundance data was calculated. Additionally, the mean nearest taxon distance (NTD), mean 10^{th} quantile phylogenetic distance (PD10), and mean of the mean pairwise distances (meanPD), of the novel host species to the local plant community are presented for each site in millions of years (Myr). The mean values across sites were calculated as well as global values for the entire experiment. Coordinates and elevation in meters for the center point of each plot are included. Plots with an asterisk differ from the plots used in Parker et al. (2015) and were moved due to disturbance or woody plant encroachment.

	Geographic location			**Biodiversity measures**			**Phylogenetic Measures (Myr)**		**Distance**
Plot	**Latitude**	**Longitude**	**Elevation**	**S**	**H**	**J**	**Mean NTD**	**Mean PD10**	**Mean meanPD**
1*	36.9861679	-122.059408	135.1379	25	2.02	0.63	62.28	120.66	228.04
2	36.9865279	-122.059722	141.4272	22	1.89	0.61	95.61	170.17	232.35
3	36.9868500	-122.059333	144.4752	21	1.75	0.58	90.96	119.13	232.36
4*	36.9873960	-122.059629	157.1840	23	2.15	0.69	67.56	129.29	224.24
5	36.9878110	-122.05901	153.3144	27	2.47	0.75	74.62	138.32	227.15
6*	36.9876560	-122.058453	156.1437	22	2.23	0.72	101.75	131.54	226.07
7	36.9878459	-122.057661	156.0576	17	1.74	0.62	96.84	162.33	230.65
8	36.9868829	-122.057887	146.304	22	1.49	0.48	83.80	140.57	234.23
9	36.9863770	-122.058792	140.5128	21	1.50	0.49	69.98	134.44	225.72
10*	36.9858000	-122.059011	133.0146	27	1.88	0.57	83.75	146.52	230.32
Mean				**22.7**	**1.91**	**0.61**	**82.72**	**139.30**	**229.11**
Standard Error				**3.02**	**0.31**	**0.09**	**13.64**	**16.59**	**3.34**
Global				**44**	**2.34**	**0.62**	**62.47**	**126.61**	**231.67**

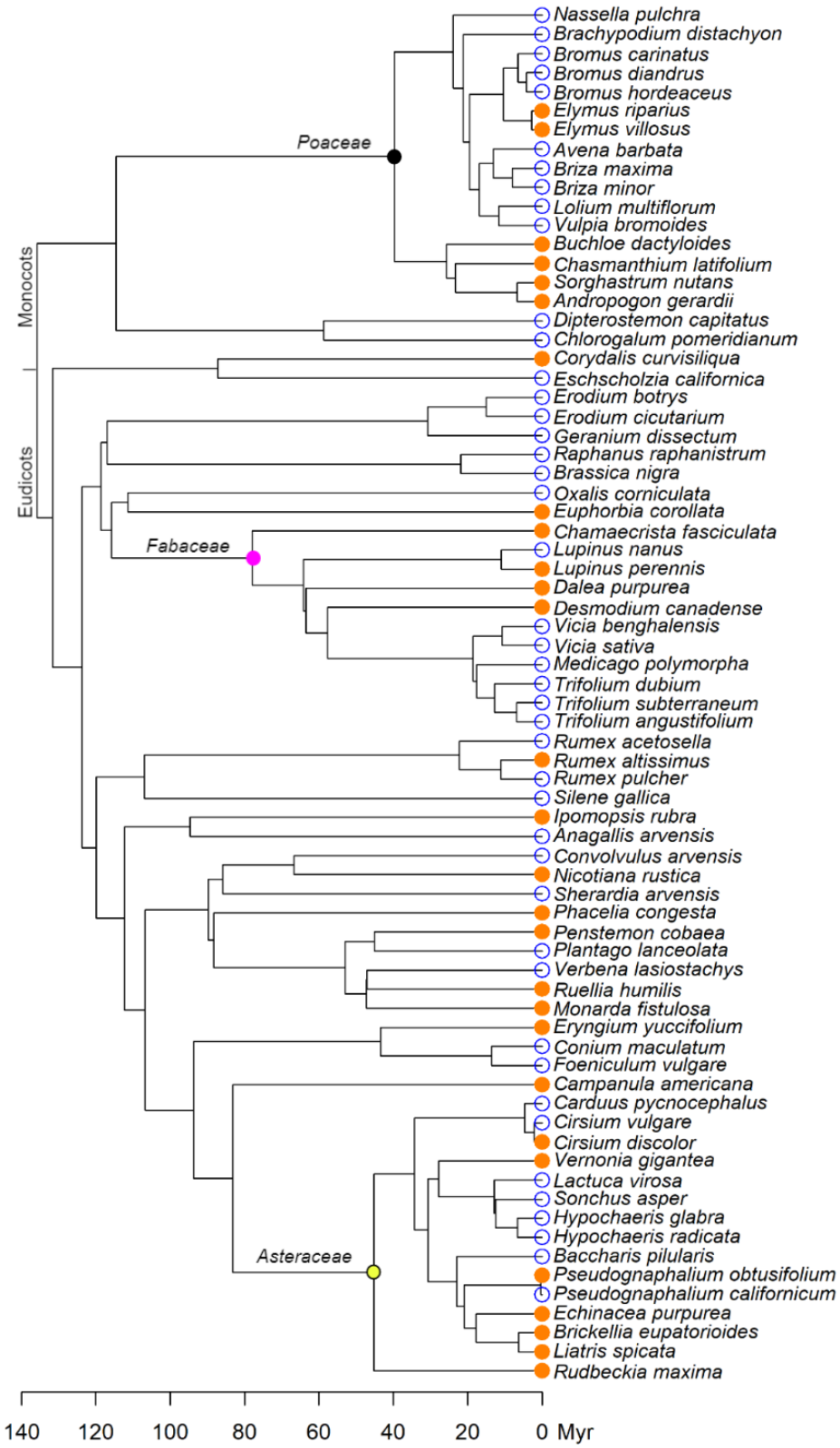
Table 2 - Novel host species selected for this experiment. 28 novel host species were selected for this experiment, representing a wide array of phylogenetic rarity in relation to the local plant community. Survival rate displays the number of sites, out of 10, that viable data was able to be retrieved from for each novel host. The OTUs column displays the number of operational taxonomic units that were identified on each novel host species, whereas the Survival (Surv) column shows the number of sites sampled (out of 10) that had sampleable hosts at the end of the study. Chao is the Chao1 estimator of the estimated "true" number of fungal taxa on each host species. Mean global phylogenetic distance of each novel host from the local plant community is displayed using three measures of phylogenetic distance and is in millions of years (Myr): nearest taxon distance (NTD), 10^{th} quantile phylogenetic distance (PD10), and mean pairwise distance (meanPD).

Species Name	Family	Surv	OTUs	Chao	NTD	PD10	meanPD
Andropogon gerardii	*Poaceae*	10	26	73.5	79.50	79.50	225.95
Brickellia eupatorioides	*Asteraceae*	10	41	69.9	41.81	61.35	213.22
Buchloe dactyloides	*Poaceae*	10	31	69.0	79.50	79.50	225.95
Campanula americana	*Campanulaceae*	4	24	71.5	166.41	166.41	232.79
Chamaecrista fasciculata	*Fabaceae*	8	21	51.3	155.93	155.93	238.29
Chasmanthium latifolium	*Poaceae*	10	33	69.1	79.50	79.50	225.95
Cirsium discolor	*Asteraceae*	9	47	117.9	4.23	68.56	212.19
Corydalis curvisiliqua	*Fumariaceae*	6	21	174	174.40	263.36	263.56
Dalea purpurea	*Fabaceae*	9	43	131	127.03	127.03	233.73
Desmodium canadense	*Fabaceae*	9	40	94.4	115.26	115.26	232.12
Echinacea purpurea	*Asteraceae*	7	27	57.6	41.81	61.35	213.22
Elymus riparius	*Poaceae*	10	53	88.4	20.88	39.06	215.79
Elymus villosus	*Poaceae*	10	45	72.1	20.88	39.06	215.79
Eryngium yuccifolium	*Apiaceae*	9	49	116.7	86.94	187.46	232.05
Euphorbia corollata	*Euphorbiaceae*	5	21	29.7	222.85	231.57	250.13
Ipomopsis rubra	*Polemoniaceae*	10	51	109.0	189.32	224.68	245.29
Liatris spicata	*Asteraceae*	7	78	192.8	41.81	61.35	213.22
Lupinus perennis	*Fabaceae*	4	12	16.2	22.01	128.43	231.50
Monarda fistulosa	*Lamiaceae*	10	61	139.0	94.54	213.48	235.84
Nicotiana rustica	*Solanaceae*	9	69	146.2	133.30	213.48	235.84
Penstemon cobaea	*Scrophulariaceae*	10	39	93.0	90.32	213.48	238.22
Phacelia congesta	*Hydrophyllaceae*	10	89	185.3	176.50	213.48	239.30
Pseudognaphalium obtusifolium	*Asteraceae*	9	124	270.3	0.55	61.35	212.28
Rudbeckia maxima	*Asteraceae*	8	49	154.0	90.52	90.52	218.99
Ruellia humilis	*Acanthaceae*	9	27	84.0	94.33	213.48	235.83
Rumex altissimus	*Polygonaceae*	10	21	34.0	22.24	239.75	241.27
Sorghastrum nutans	*Poaceae*	9	36	82.4	79.50	79.50	225.95
Vernonia gigantea	*Asteraceae*	9	59	317.8	55.66	61.35	213.50
Mean		**8.6**	**44.2**	**111.1**	**89.55**	**134.61**	**229.21**
Standard Error		**1.8**	**24.2**	**69.1**	**60.54**	**73.16**	**13.01**
Global		**240**	**342**	**576.4**			

Table 3 - Most common fungal taxa found on novel hosts. The 20 most common fungal operational taxonomic units (OTUs) ranked by the number of novel hosts they were found on (maximum of 28). The number of samples each OTU was identified in out of a maximum of 240 is also included. Taxonomy was assigned using the R package DADA2 and the UNITE database general FASTA release for all eukaryotes. Reads were clustered into operational taxonomic units (OTUs) using the "tip_glom" function from the "phyloseq" package in R.

OTU	Family	Species Name	Novel Hosts	Samples
641	*Malasseziaceae*	*Malassezia* sp.	28	152
154	*Pleosporaceae*	Undetermined	28	142
640	*Malasseziaceae*	*Malassezia globosa*	28	70
157	*Pleosporaceae*	Undetermined	27	125
968	*Filobasidiaceae*	*Filobasidium* sp.	26	107
958	*Chrysozymaceae*	*Oberwinklerozyma* sp.	23	49
1349	*Mycosphaerellaceae*	Undetermined	22	53
1016	*Mrakiaceae*	*Itersonilla* sp.	20	65
946	*Sporidiobolaceae*	*Rhodotorula* sp.	20	53
207	*Sporomiaceae*	Undetermined	20	42
1013	*Mrakiaceae*	*Udeniomyces* sp.	19	61
1032	*Sclerotiniaceae*	Undetermined	18	49
987	*Filobasidiaceae*	*Nagnishia* sp.	18	34
771	*Psathyrellaceae*	Undetermined	16	40
848	*Phaeotremellaceae*	*Gelidatrema spencermartinsiae*	16	27
267	*Botryosphaeriaceae*	Undetermined	15	25
65	*Aspergillaceae*	*Aspergillus* sp.	15	18
897	*Kondoaceae*	Undetermined	14	41
936	*Cystobasidiomycetes Incertae sedis*	*Sakaguchia* sp.	14	40
770	*Bolbitiaceae*	*Panaeolus* sp.	13	19

Figure 1 - Newick tree of the phylogenetic relationships between novel hosts and local species. Tree tips highlighted in orange are novel hosts (28), whereas tree tips with empty blue circles are resident meadow species (44). Major families with three or more representatives are labelled with colored circles indicating the common ancestor for each labelled family – *Poaceae* in black, *Fabaceae* in purple, and *Asteraceae* in yellow. Branches representing Monocots and Eudicots are also labelled. Units for the tree are in millions of years (Myr).


Nassella pulchra
Brachypodium distachyon
Bromus carinatus
Bromus diandrus
Bromus hordeaceus
Elymus riparius
Elymus villosus
Poaceae
Avena barbata
Briza maxima
Briza minor
Lolium multiflorum
Vulpia bromoides
Buchloe dactyloides
Chasmanthium latifolium
Sorghastrum nutans
Andropogon gerardii
Monocots
Dipterostemon capitatus
Chlorogalum pomeridianum
Corydalis curvisiliqua
Eschscholzia californica
Erodium botrys
Erodium cicutarium
Geranium dissectum
Raphanus raphanistrum
Brassica nigra
Eudicots
Oxalis corniculata
Euphorbia corollata
Chamaecrista fasciculata
Fabaceae
Lupinus nanus
Lupinus perennis
Dalea purpurea
Desmodium canadense
Vicia benghalensis
Vicia sativa
Medicago polymorpha
Trifolium dubium
Trifolium subterraneum
Trifolium angustifolium
Rumex acetosella
Rumex altissimus
Rumex pulcher
Silene gallica
Ipomopsis rubra
Anagallis arvensis
Convolvulus arvensis
Nicotiana rustica
Sherardia arvensis
Phacelia congesta
Penstemon cobaea
Plantago lanceolata
Verbena lasiostachys
Ruellia humilis
Monarda fistulosa
Eryngium yuccifolium
Conium maculatum
Foeniculum vulgare
Campanula americana
Carduus pycnocephalus
Cirsium vulgare
Cirsium discolor
Vernonia gigantea
Lactuca virosa
Sonchus asper
Hypochaeris glabra
Hypochaeris radicata
Asteraceae
Baccharis pilularis
Pseudognaphalium obtusifolium
Pseudognaphalium californicum
Echinacea purpurea
Brickellia eupatorioides
Liatris spicata
Rudbeckia maxima
140
120
100
80
60
40
20
0 Myr

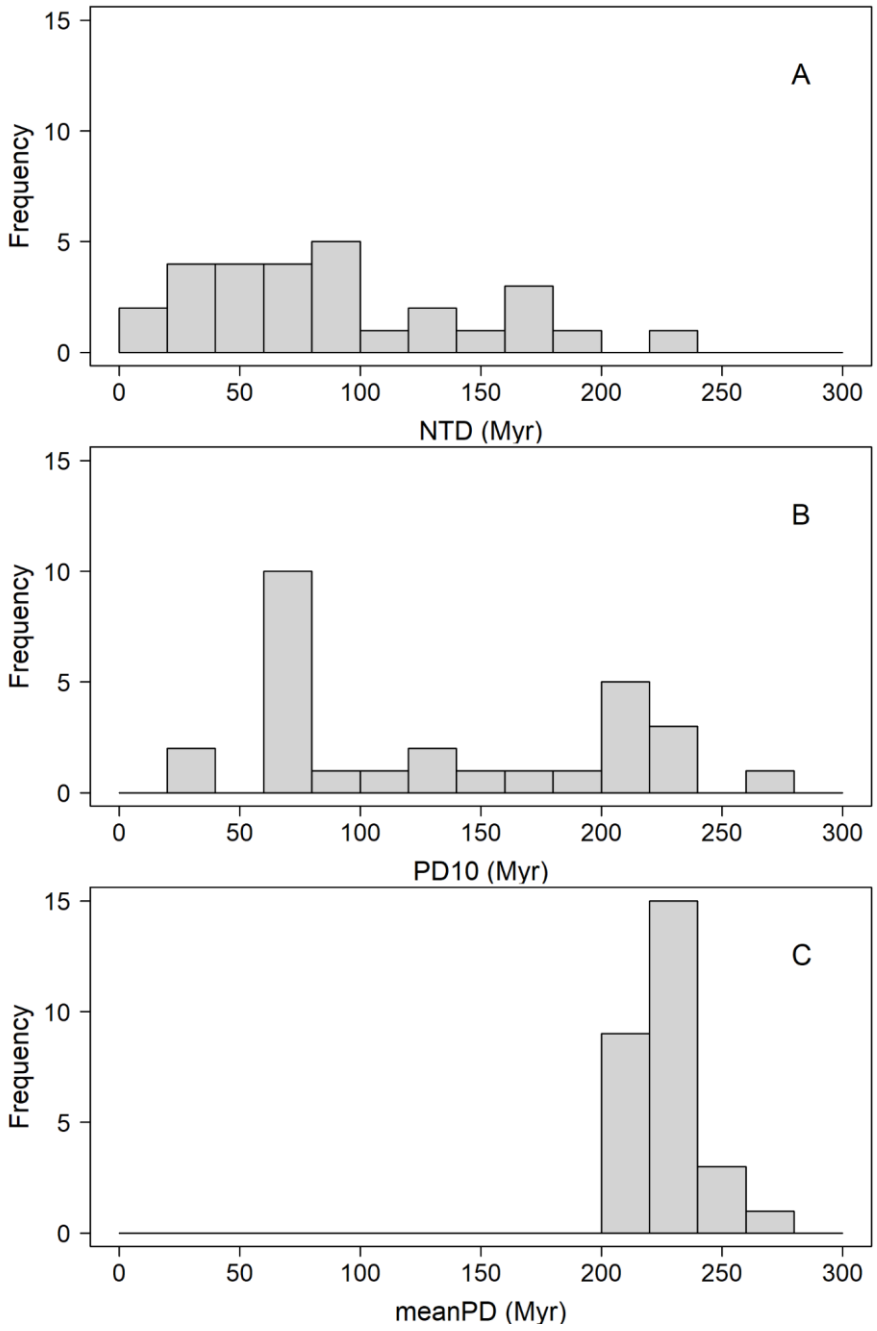


Figure 2 - Distribution of novel host species in relation to the local plant community as a function of three phylogenetic distance metrics. Phylogenetic distance metrics were calculated for each of the 28 novel hosts in relation to the local plant community shown in Table 3. A) Phylogenetic distance to the most closely related taxon (NTD), B) the 10^{th} quantile phylogenetic distance to the local plant community (PD10), and C) the mean pairwise distance from the novel host to each member of the local plant community (meanPD).

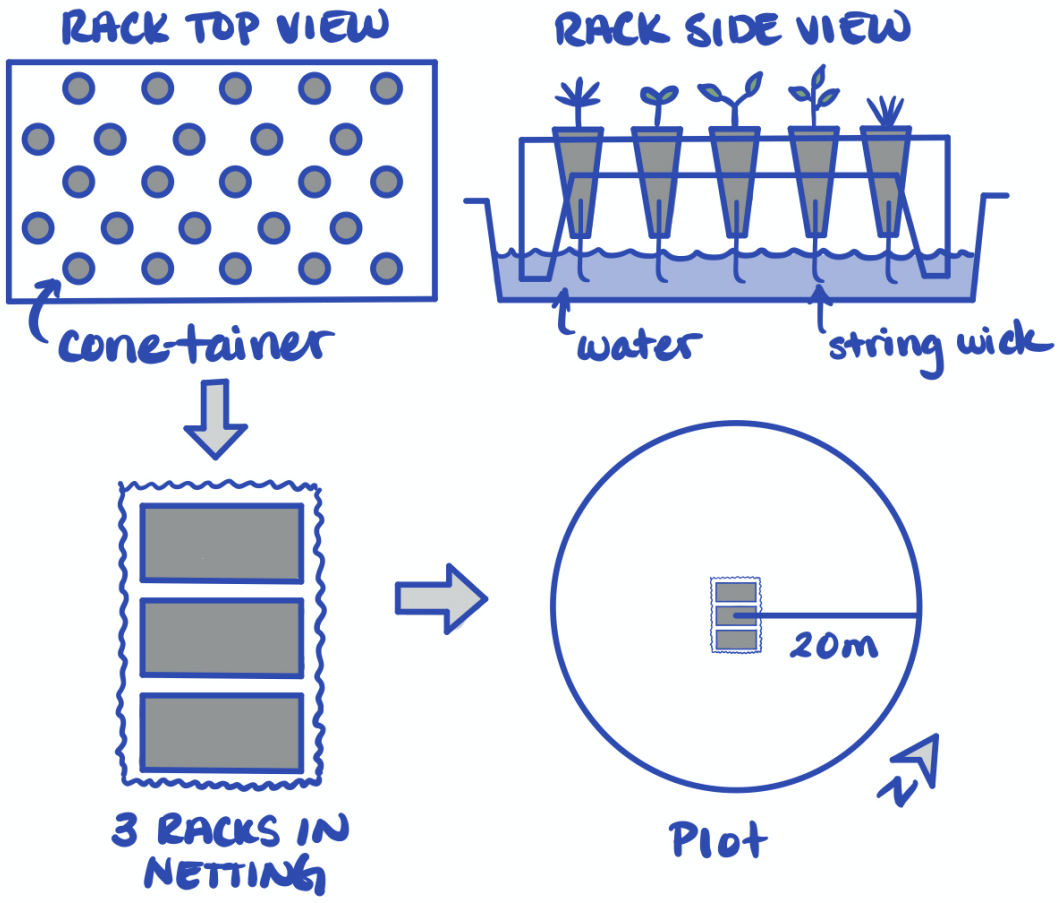


Figure 3 - Schematic representation of field experiment design. One each of the 28 novel hosts were randomly arranged in Cone-tainer™ racks with empty spaces on each cardinal side. Each Cone-tainer™ had a cotton string wick inserted approximately 3 cm into the soil from the bottom with sufficient length to reach the bottom of the watering trough the rack assembly was placed into. Racks were placed in triplicate in the center of each 20 m radius plot in a Northwest facing arrangement and enclosed in thin deer netting to prevent large mammal herbivory.

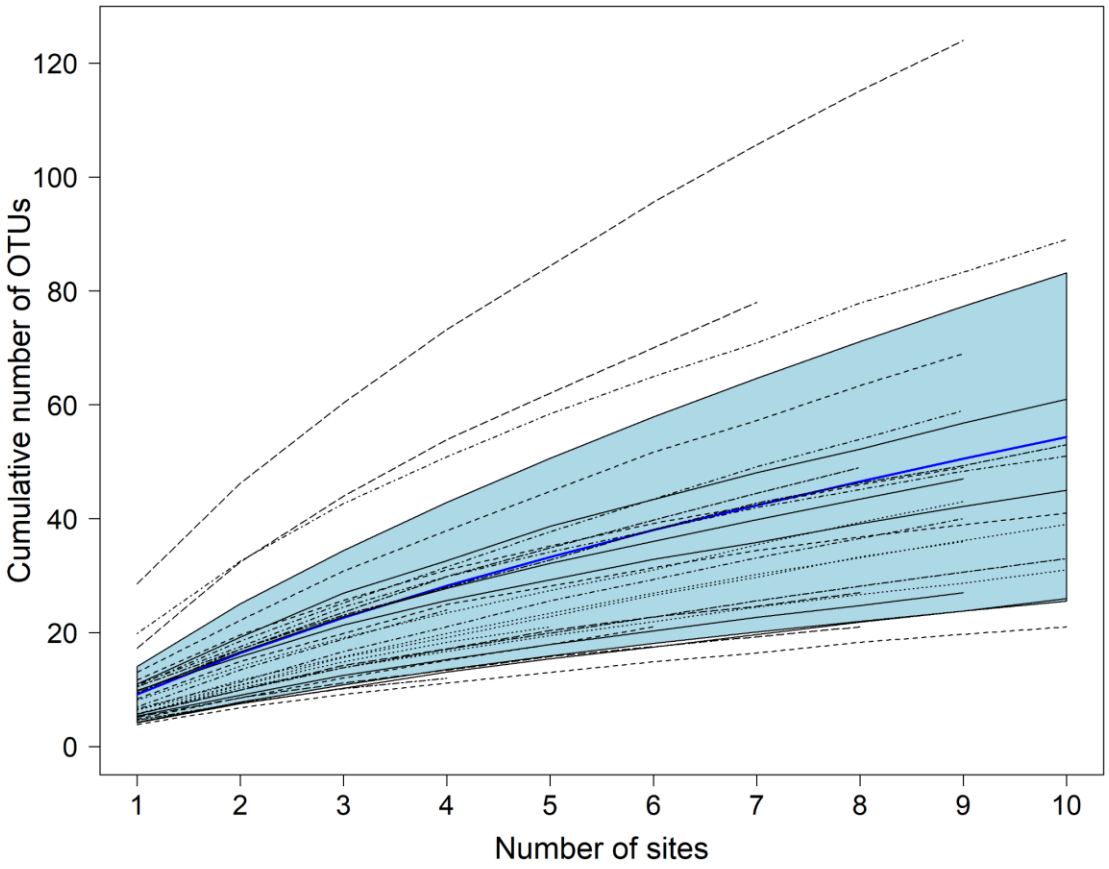


Figure 4 - Species accumulation curves for fungal communities on novel hosts. The Dashed lines represent the 28 novel hosts and were generated by randomly sampling among the sites where the novel host survived for 100 permutations. The blue enveloped curve represents the null hypothesis – that each of the novel host species are equivalent fungal hosts – and was generated by randomly selecting one host species from each of the 10 random sites in random order for 1000 permutations to generate the median and 95% CI.

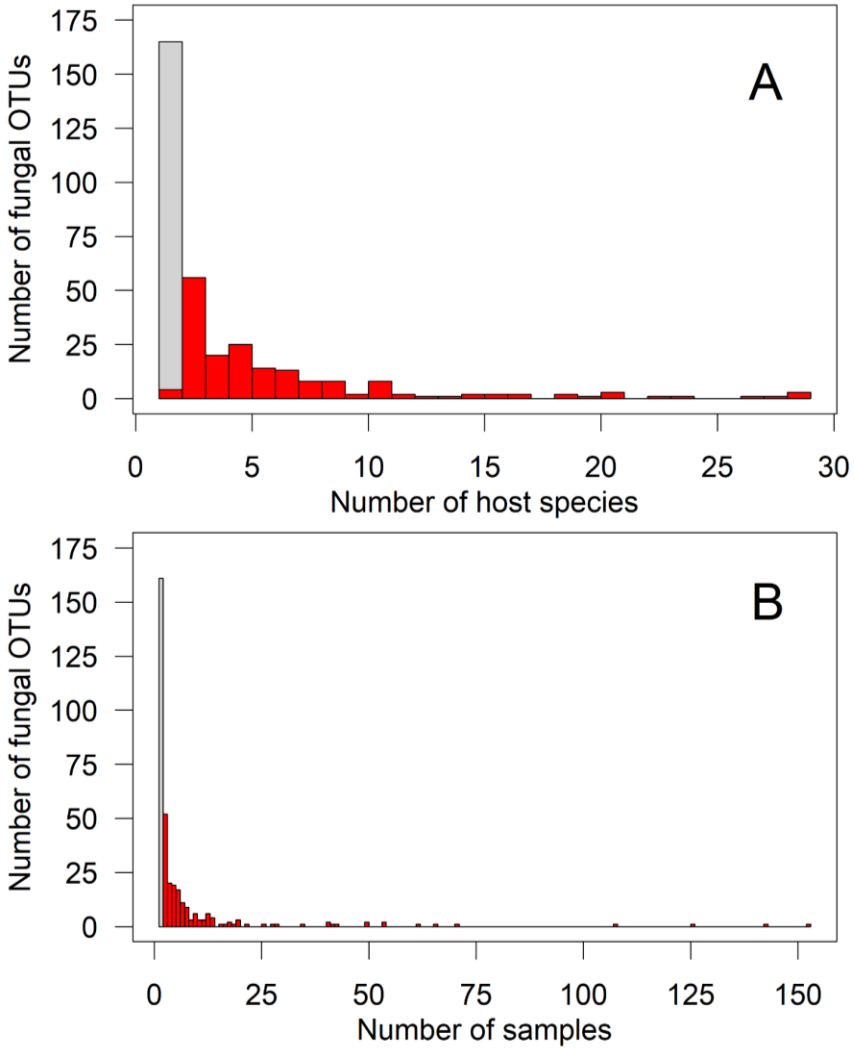


Figure 5 - Abundance and host range of fungal taxa within the fungal community found on novel hosts. The number of fungal operational taxonomic units (OTUs) by A) the number of novel host species and B) the number of samples in which they were detected. The red bars represent the fungal community excluding singletons (fungal taxa detected on only one sample), whereas the grey bar represents the singletons. The number of fungal OTUs as a function of the number of novel hosts they infect provides a visual representation of the variation in host ranges among fungi in the community, whereas fungal OTUs as a function of the number of samples provides a visual representation of the distribution of abundances of fungal taxa.

Figure 6 - Range of fungal community richness on novel hosts, arranged by phylogenetic relationships between novel hosts. Boxplots of the number of operational taxonomic units (OTUs) found on each novel host sample. The observed fungal richness ranged from 4-10 OTUs per sample. Combined values for all samples from each host species are provided in Table 2. Host species are arranged by phylogenetic relationships, and families that are represented by more than one novel host are labelled with colored circles indicating the common ancestor for each labelled family – *Poaceae* in black, *Fabaceae* in blue, and *Asteraceae* in orange. Branches representing Monocots and Eudicots are also labelled. Units for the tree are in millions of years (Myr).

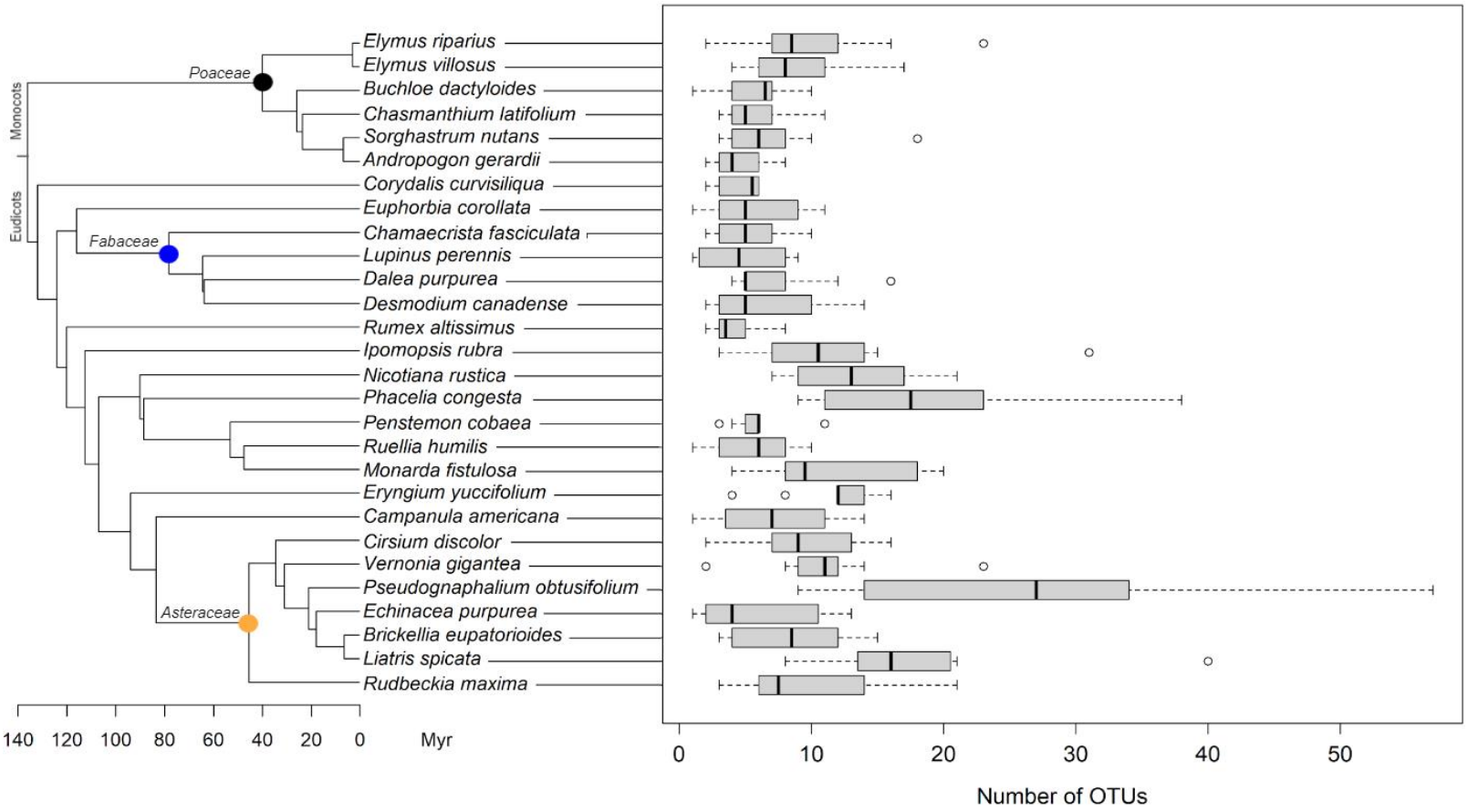
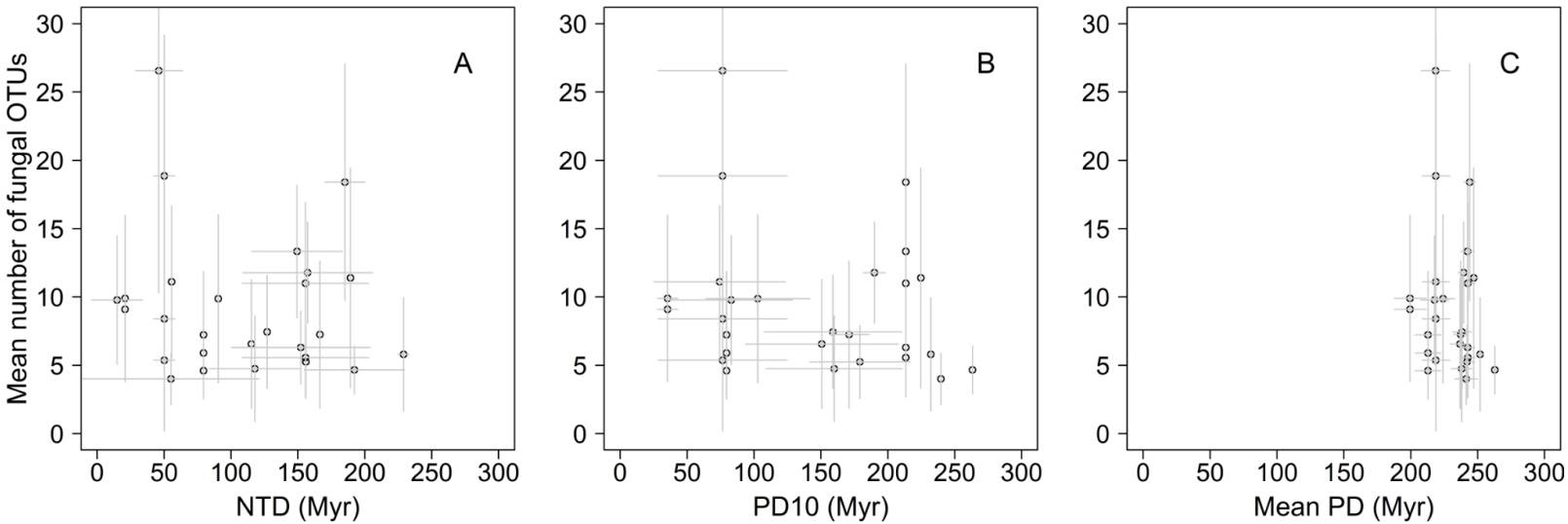

Elymus riparius
Elymus villosus
Buchloe dactyloides
Chasmanthium latifolium
Sorghastrum nutans
Andropogon gerardii
Corydalis curvisiliqua
Euphorbia corollata
Chamaecrista fasciculata
Lupinus perennis
Dalea purpurea
Desmodium canadense
Rumex altissimus
Ipomopsis rubra
Nicotiana rustica
Phacelia congesta
Penstemon cobaea
Ruellia humilis
Monarda fistulosa
Eryngium yuccifolium
Campanula americana
Cirsium discolor
Vernonia gigantea
Pseudognaphalium obtusifolium
Echinacea purpurea
Brickellia eupatorioides
Liatris spicata
Rudbeckia maxima
Poaceae
Fabaceae
Asteraceae
Monocots
Eudicots
140 120 100 80 60 40 20 0
Myr
0 10 20 30 40 50
Number of OTUs

Figure 7 - Phylogenetic isolation as a predictor of the mean number of fungal taxa found on novel hosts. The mean number of fungal operational taxonomic units (OTUs) found on novel hosts as a function of phylogenetic distance based on pairwise phylogenetic distances from each novel host to each of the species in their surrounding community. Three metrics for phylogenetic distance were used: A) nearest taxon distance (NTD), B) 10^{th} quantile phylogenetic distance (PD10), and C) mean pairwise distance (meanPD), and were measured in millions of years (Myr) of independent evolution. None of the linear regressions were statistically significant. Standard error bars for each novel host are displayed in grey.


Mean number of fungal OTUs
NTD (Myr)
PD10 (Myr)
Mean PD (Myr)
A
B
C
0
5
10
15
20
25
30
0
50
100
150
200
250
300

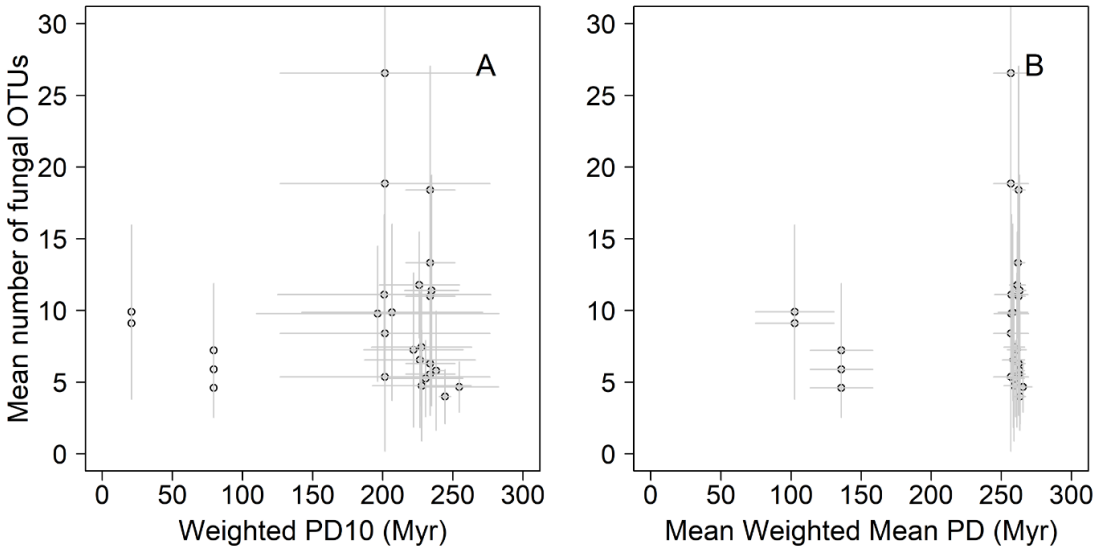


Figure 8 - Phylogenetic isolation weighted by local species abundance as a predictor of the mean number of fungal taxa found on novel hosts. The mean number of fungal operational taxonomic units (OTUs) found on novel hosts as a function of phylogenetic distance weighted by local plant species abundance, based on pairwise phylogenetic distances from each novel host to each of the species in their surrounding community. Two metrics for phylogenetic distance between novel plant hosts and the local plant community were used: A) mean weighted 10^{th} quantile phylogenetic distance (wPD10), and B) weighted mean pairwise distance (wmeanPD) and were measured in millions of years (Myr) of independent evolution. None of the linear regressions were statistically significant. Standard error bars for each novel host are displayed in grey.

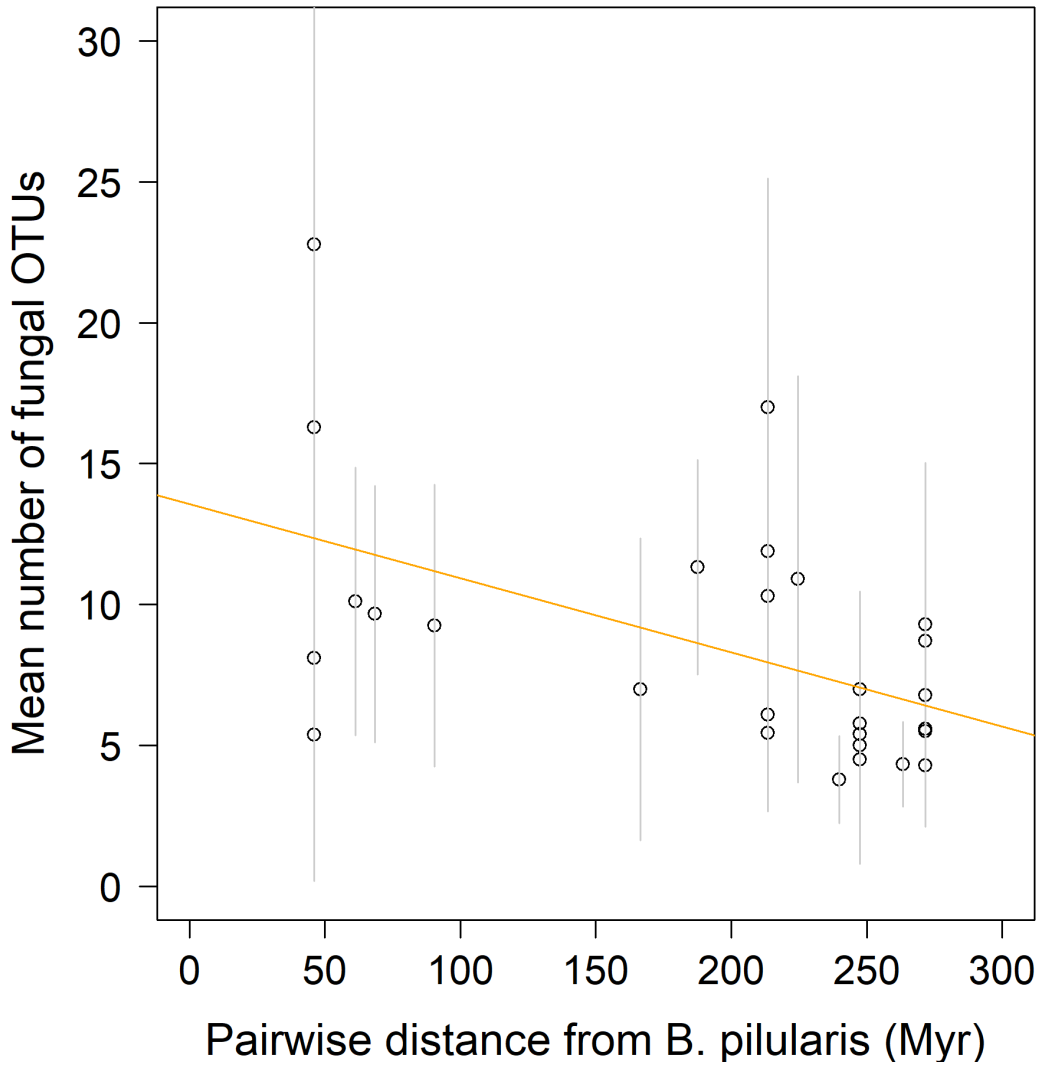


Figure 9 - Pairwise phylogenetic distance from *Baccharis pilularis* as a predictor of the mean number of fungal taxa found on novel hosts. The mean number of fungal operational taxonomic units (OTUs) found on novel hosts as a function of pairwise phylogenetic distance from the focal local plant species, *Baccharis pilularis*. The relationship is statistically significant. Standard error bars for each novel host are displayed in grey.

Appendices:

Appendix 1 - Table of the abundance of plant species in the local community. The percent cover measurements for all local plant species identified within each plot during the quadrat and rare species surveys. Mean percent cover, standard error (S.E.), and relative abundances (Rel. Abund.) were calculated for each of the 44 local species. All values are percentages.

Plant species	Plot										Mean ± S.E.	Rel. Abund.
	1	2	3	4	5	6	7	8	9	10		
Anagallis arvensis	0.978	0.256	0.008	3.284	4.756	0.563	0.978	0.031	0.303	1.413	1.257 ± 1.563	1.705
Avena barbata	19.172	12.728	12.316	22.444	7.769	11.650	25.891	32.672	13.681	17.200	17.552 ± 7.609	23.804
Baccharis pilularis	0.105	0.000	0.637	3.213	5.892	5.414	0.000	0.637	0.000	1.083	1.698 ± 2.298	2.303
Brachypodium distachyon	3.784	18.192	10.500	10.816	0.859	0.094	10.556	7.753	13.750	0.000	7.630 ± 4.755	10.349
Brassica nigra	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.146	0.000	0.015 ± 0.046	0.020
Briza maxima	0.045	0.000	0.000	0.719	14.613	5.900	0.000	0.000	0.000	0.019	2.130 ± 4.755	2.888
Briza minor	0.000	0.000	0.000	0.000	0.085	0.010	0.000	0.000	0.000	0.063	0.016 ± 031	0.021
Bromus carinatus	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.029	0.003 ± 0.009	0.004
Bromus diandrus	21.966	14.884	22.875	13.025	15.641	14.697	14.975	26.481	20.484	23.478	18.851 ± 4.719	25.565
Bromus hordeaceus	7.509	0.483	0.000	0.091	2.434	0.966	0.069	0.000	0.070	3.634	1.526 ± 2.437	2.069
Carduus pycnocephalus	0.046	0.033	0.030	0.095	4.228	0.325	0.000	0.011	0.018	0.030	0.482 ± 1.320	0.653
Chlorogalum pomeridianum	0.000	0.029	0.000	0.000	0.000	0.045	2.077	0.000	0.000	0.000	0.215 ± 1.654	0.292
Cirsium vulgare	0.000	0.000	0.000	0.000	0.034	0.000	0.000	0.000	0.000	0.000	0.003 ± 0.011	0.005
Conium maculatum	0.000	0.000	0.000	0.000	0.000	0.955	0.000	0.000	0.000	0.000	0.096 ± 0.302	0.130
Convolvulus arvensis	0.275	0.014	0.115	2.634	0.053	0.000	0.000	0.488	4.459	0.001	0.804 ± 1.516	1.090
Dipterostemmon capitalus	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	<0.001 ± <0.001	<0.001
Erodium botrys	0.015	0.392	0.018	0.006	0.216	0.000	<0.001	0.042	0.031	0.007	0.073 ± 0.129	0.098
Erodium cicutarium	0.007	0.025	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.029	0.007 ± 0.011	0.009
Eschscholzia californica	4.303	6.053	0.030	0.000	1.213	1.028	18.631	1.044	0.000	0.026	3.233 ± 5.787	4.384

Foeniculum vulgare	0.000	0.000	0.000	0.000	0.000	0.000	1.281	0.000	0.731	0.000	0.201 ± 0.443	0.273
Geranium dissectum	12.131	1.125	0.678	11.166	14.066	5.878	0.000	1.797	0.017	11.875	5.873 ± 5.824	7.965
Hypochaeris glabra	0.007	0.000	<0.001	0.000	0.734	0.029	0.000	0.000	0.013	0.001	0.078 ± 0.231	0.106
Hypochaeris radicata	0.775	0.009	0.049	2.931	2.884	0.030	0.029	0.186	0.003	0.356	0.725 ± 1.175	0.983
Lactuca virosa	0.000	0.000	0.000	<0.001	0.000	0.000	0.000	0.000	0.000	0.000	<0.001 ± <0.001	<0.001
Lolium multiflorum	1.201	0.006	0.000	0.309	0.563	11.006	0.000	1.413	0.050	3.053	1.760 ± 3.388	2.387
Lupinus nanus	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.001 ± 0.002	0.001
Medicago polymorpha	0.000	<0.001	0.006	0.000	0.000	0.000	0.000	0.001	0.000	<0.001	0.001 ± 0.002	0.001
Nassella pulchra	0.000	0.010	0.013	0.000	0.000	0.000	0.020	0.000	0.764	0.000	0.081 ± 0.240	0.109
Oxalis corniculata	0.000	0.000	0.000	0.013	0.000	0.001	0.000	0.000	0.000	0.007	0.002 ± 0.004	0.003
Plantago lanceolata	0.000	0.082	0.010	0.000	0.000	0.000	0.000	0.035	0.000	0.000	0.013 ± 0.027	0.017
Pseudognaphaliumcalif ornicum	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.001 ± 0.002	0.001
Raphanus sativus	1.231	3.963	6.572	6.956	4.159	5.716	12.295	3.906	0.049	0.496	4.532 ± 3.650	6.146
Rumex acetosella	0.046	0.125	0.038	0.161	1.144	1.053	0.000	0.122	0.053	1.628	0.437 ± 0.598	0.593
Rumex pulcher	<0.001	0.000	0.000	0.099	0.000	0.000	0.000	0.036	0.173	0.000	0.031 ± 0.059	0.042
Sherardia arvensis	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	<0.001 ± 0.001	0.001
Silene gallica	0.000	0.000	0.000	0.000	0.034	0.000	0.000	0.000	0.000	0.004	0.001 ± 0.011	0.005
Sonchus asper	0.000	0.000	0.000	0.166	0.021	0.000	0.006	0.095	0.003	0.000	0.029 ± 0.056	0.040
Trifolium angustifolium	0.002	0.017	0.010	0.000	0.000	0.000	0.000	0.000	0.010	0.001	0.004 ± 0.006	0.005
Trifolium dubium	0.150	0.000	0.000	0.000	1.131	0.022	0.000	0.000	0.00	0.088	0.139 ± 0.352	0.189
Trifolium subterraneum	<0.001	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.001 ± 0.005	0.002
Verbena lasiostachys	0.000	0.000	0.000	0.000	0.339	0.000	0.000	0.000	0.000	0.000	0.034 ± 0.107	0.046
Vicia benghalensis	2.619	4.317	7.250	0.706	0.001	0.000	0.025	0.003	0.000	0.000	1.492 ± 2.501	2.024
Vicia sativa	0.007	0.000	3.844	0.822	0.000	3.106	0.003	0.023	0.000	2.169	0.997 ± 1.485	1.353
Vulpia bromoides	2.231	1.795	0.006	0.903	7.991	0.023	0.004	0.459	0.144	3.550	0.997 ± 1.485	1.353

Appendix 2 - Map of plot locations within the study site. Satellite image of the Great Meadow on the University of California Santa Cruz campus overlaid with the positions of the center point of each plot. Plots with the "-2018" suffix are plots that were changed from Parker et al. (2015) due to disturbance or woody plant encroachment. Original positions of the changed plots are shown without the "-2018" designator. A mixed evergreen forest borders the Eastern side of the study site (left side of image), and an organic farm borders the North (bottom of image). The Pacific Ocean is approximately 1 km to the West (right side of image). The prefix "GM-" stands for Great Meadow.

References:

Abarenkov K, Zirk A, Piirmann T, et al (2020) UNITE generate FASTA release for eukaryotes. Version 04.02.2020. UNITE Community.

Abdillah A, Khelafia S, Raoult D, Bittar F, Ranque S (2020) Comparison of three skin sampling methods and two media for culturing *Malassezia* yeast. Journal of Fungi 6:350. https://doi.org/10.3390/jof6040350

Amend A (2014) From dandruff to deep-sea vents: *Malassezia*-like fungi are ecologically hyper-diverse. PLoS Pathogens 10(8) e1004227. https://doi.org/10.1371/journal.ppat.1004277

Anderson PK, Cunningham AA, Patel NG, et al (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology & Evolution 19:535–544. https://doi.org/10.1016/j.tree.2004.07.021

Arnold AE, Lutzoni F (2007) Diversity and host range of foliar fungal endophytes: Are tropical leaves biodiversity hotspots? Ecology: 88:541-549 https://doi.org/10.1890/05-1459

Ashbee HR (2007) Update on the genus *Malassezia*. Medical Mycology 45:287-303. https://doi.org/10.1080/13693780701191373

Bagchi R, Swinfield T, Gallery RE, et al (2010) Testing the Janzen-Connell mechanism: pathogens cause overcompensating density dependence in a tropical tree. Ecology Letters 13:1262–1269. https://doi.org/10.1111/j.1461-0248.2010.01520.x

Bálint M, Bartha L, O'Hara RB, et al (2014) Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. Molecular Ecology 24:235-248. https://doi.org/10.1111/mec.13018

Barrett LG, Heil M (2012) Unifying concepts and mechanisms in the specificity of plant–enemy interactions. Trends in Plant Science 17:282–292. https://doi.org/10.1016/j.tplants.2012.02.009

Barrett LG, Kniskern JM, Bodenhausen N, et al (2009) Continua of specificity and virulence in plant host–pathogen interactions: causes and consequences. New Phytologist 183:513–529. https://doi.org/10.1111/j.1469-8137.2009.02927.x

Bell T, Freckleton RP, Lewis OT (2006) Plant pathogens drive density-dependent seedling mortality in a tropical tree. Ecology Letters 9:569–574. https://doi.org/10.1111/j.1461-0248.2006.00905.x

Blumenthal DM (2006) Interactions between resource availability and enemy release in plant invasion. Ecology Letters 9:887–895. https://doi.org/10.1111/j.1461-0248.2006.00934.x

Bufford JL, Hulme PE, Sikes BA, et al (2016) Taxonomic similarity, more than contact opportunity, explains novel plant–pathogen associations between native and alien taxa. New Phytologist 212:657–667. https://doi.org/10.1111/nph.14077

Byun C, Lee EJ (2017) Ecological application of biotic resistance to control the invasion of an invasive plant, *Ageratina altissima*. Ecology and Evolution 7:2181–2192. https://doi.org/10.1002/ece3.2799

Callahan BJ, McMurdie PJ, Rosen MJ, et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13:581–583. https://doi.org/10.1038/nmeth.3869

Christian N, Sullivan C, Visser ND, Clay K (2016) Plant host and geographic location drive endophyte community composition in the face of perturbation. Microbial Ecology 72:621-632.

Christian N, Herre EA, Mejia LC, Clay K (2017) Exposure to leaf litter microbiome of healthy adults protects seedlings from pathogen damage. Proceedings of the Royal Society B: Biological Sciences 284:1858. https://doi.org/10.1098/rspb.2017.0641

Clay K, Shearin ZRC, Bourke KA, Bickford WA, Kowalski KP (2016) Diversity of fungal endophytes in non-native *Phragmites australis* in the Great Lakes. Biological Invasions 18:2703-2716. https://doi.org/10.1007/s10530-016-1137-y

Connell JH, Tracey JG, Webb LJ (1984) Compensatory recruitment, growth, and mortality as factors maintaining rain forest tree diversity. Ecological Monographs 54:141–164. https://doi.org/10.2307/1942659

Crandall SG, Saarman N, Gilbert GS (2020) Fungal spore diversity, community structure, and traits across a vegetation mosaic. Fungal Ecology 45:100920. https://doi.org/10.1016/j.funeco.2020.100920

Delmont TO, Robe P, Cecillon S, et al (2011) Accessing the soil metagenome for studies of microbial diversity. Applied and Environmental Microbiology 77:1315-1324. https://doi.org/10.1128/AEM.01526-10

Di Menna ME (1957) The isolation of yeasts from soil. Microbiology 17:678-688. https://doi.org/10.1099/00221287-17-3-678

Facelli E, McKay SF, Facelli JM, et al (2018) A soil-borne generalist pathogen regulates complex plant interactions. Plant and Soil 433:101-109. https://doi.org/10.1007/s11104-018-3828-x

Ferrer A, Gilbert GS (2003) Effect of a tree host species on fungal community composition in a tropical rain forest in Panama. Diversity and Distributions 9:455-468. https://doi.org/10.1046/j.1472-4642.2003.00039.x

Findley K, Oh J, Yang J, et al (2013) Topographic diversity of fungal and bacterial communities in human skin. Nature 498:367-370. https://doi.org/10.1038/nature12171

Flory SL, Clay K (2013) Pathogen accumulation and long-term dynamics of plant invasions. Journal of Ecology 101:607–613. https://doi.org/10.1111/1365-2745.12078

Gil-Martínez M, López-Garcia Á, Domínguez MT, et al (2018) Ectomycorrhizal fungal communities and their functional traits mediate plant-soil interactions in trace element contaminated soils. Frontiers in Plant Science 9:1682. https://doi.org/10.3389/fpls.2018.01682

Gilbert GS, Webb CO (2007) Phylogenetic signal in plant pathogen–host range. PNAS 104:4979–4983. https://doi.org/10.1073/pnas.0607968104

Gilbert GS, Parker IM (2010) Rapid evolution in a plant-pathogen interaction and the consequences for introduced host species: Pathogen evolution and novel hosts. Evolutionary Applications 3:144–156. https://doi.org/10.1111/j.1752-4571.2009.00107.x

Gilbert GS, Magarey R, Suiter K, Webb CO (2012) Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. Evolutionary Applications 5:869–878. https://doi.org/10.1111/j.1752-4571.2012.00265.x

Gilbert, GS and Parker, IM. Phylogenetic distance metrics for studies of focal species in communities: quantiles and cumulative curves. In Revision at PLoS ONE

Grant KD, Koenemann D, Mansaray J, Ahmed A, Khamar H, El Oualidi Jl, Burke, JM (2020) Phylogeny of docks and sorrels (*Rumex*, Polygonaceae) reveals plasticity of reproductive systems. bioRxiv. Preprint. 12 September 2020. https://doi.org/10.1101/2020.09.11.293118

Johnson PTJ, Preston DL, Hoverman JT, Richgels KLD (2013) Biodiversity decreases disease through predictable changes in host community competence. Nature 494:230–233. https://doi.org/10.1038/nature11883

Kendig AE, Spears ER, Dawn SC, Flory SL, Mordecai EA (2020) Native perennial and non-native annual grasses shape pathogen community composition and disease severity in a California grassland. Journal of Ecology 109:900-912. https://doi.org/10.1111/1365-2745.13515

Knevel IC, Lans T, Menting FBJ, et al (2004) Release from native root herbivores and biotic resistance by soil pathogens in a new habitat both affect the alien *Ammophila arenaria* in South Africa. Oecologia 141:502–510. https://doi.org/10.1007/s00442-004-1662-8

Knops JMH, Tilman D, Haddad NM, et al (1999) Effects of plant species richness on invasion dynamics, disease outbreaks, insect abundances and diversity. Ecology Letters 2:286–293. https://doi.org/10.1046/j.1461-0248.1999.00083.x

Liang M, Liu X, Parker IM, et al (2019) Soil microbes drive phylogenetic diversity-productivity relationships in a subtropical forest. Science Advances 5:eeax5008. https://doi.org/10.1126/sciadv.aax5088

Lim GS, Balke M, Meier R (2012) Determining species boundaries in a world full of rarity: Singletons, species delimitation methods. Systematic Biology 61:165-169. https://doi.org/10.1093/sysbio/syr030

Liu H, Stiling P (2006) Testing the enemy release hypothesis: a review and meta-analysis. Biol Invasions 8:1535–1545. https://doi.org/10.1007/s10530-005-5845-y

Massimo NC, Nandi Devan MM, Arendt KR, et al (2015) Fungal endophytes in aboveground tissues of desert plants: Infrequent in culture, but highly diverse and distinctive symbionts. Microbial Ecology 70:61-76. https://doi.org/10.1007/s00248-014-0563-6

Metz MR, Sousa WP, Valencia R (2010) Widespread density-dependent seedling mortality promotes species coexistence in a highly diverse Amazonian rain forest. Ecology 91:3675–3685. https://doi.org/10.1890/08-2323.1

Mitchell CE, Power AG (2003) Release of invasive plants from fungal and viral pathogens. Nature 421:625–627. https://doi.org/10.1038/nature01317

Palmer JM, Jusino MA, Banik MT, Lindner DL (2018) Non-biological synthetic spike-in controls and the AMPtk software pipeline improve mycobiome data. PeerJ 6:e4925. https://doi.org/10.7717/peerj.4925

Parker IM, Gilbert GS (2007) When there is no escape: The effects of natural enemies on native, invasive, and noninvasive plants. Ecology 88(5):1210-1224.

Parker IM, Saunders M, Bontrager M, et al (2015) Phylogenetic structure and host abundance drive disease pressure in communities. Nature 520:542–544. https://doi.org/10.1038/nature14372

Philibert A, Desprez-Loustau M-L, Fabre B, et al (2011) Predicting invasion success of forest pathogenic fungi from species traits. Journal of Applied Ecology 48:1381–1390. https://doi.org/10.1111/j.1365-2664.2011.02039.x

Põlme S, Bahram M, Jacquermyn H, et al (2017) Host preference and network properties in biotrophic plant-fungal associations. New Phytologist 217:1230-1239. https://doi.org/10.1111/nph.14895

Roy M, Watthana S, Stier A, et al (2009) Two mycoheterotrophic orchids from Thailand tropical dipterocarpacean forests associate with a broad diversity of ectomycorrhizal fungi. BMC Biology 7:1-17. https://doi.org/10.1186/1741-7007-7-51

Starmer WT, Phaff HJ, Miranda M, Miller MW (1978) *Pichia cactophilia*, a new species of yeast found in decaying tissue of cacti. International Journal of Systematic and Evolutionary Microbiology 28:318-325.

Starmer WT, Schmedicke RA, Lachance MA (2003) The origin of the cactus-yeast community. FEMS Yeast Research 3:441-448. https://doi.org/10.1016/S1567-1356(03)00056-4

Termorshuizen AJ (2017) Ecology of Fungal Plant Pathogens. In: The Fungal Kingdom. John Wiley & Sons, Ltd, pp 387–397.

The Angiosperm Phylogeny Group, Chase MW, Christenhusz MJM, et al (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society 181:1-20. https://doi.org/10.1111/boj.12385

Thomson LJ, Macfadyen S, Hoffmann AA (2010) Predicting the effects of climate change on natural enemies of agricultural pests. Biological Control 52:296–306. https://doi.org/10.1016/j.biocontrol.2009.01.022

Thrall PH, Biere A, Antonovics J (1993) Plant life-history and disease susceptibility - the occurrence of *Ustilago violacea* on different species within the Caryophyllaceae. Journal of Ecology 81:489-498. https://doi-org.oca.ucsc.edu/10.2307/2261527

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-Coverage ITS primers for the DNA-based identification of Ascomycetes and Basidiomycetes in environmental samples. PLoS ONE 7:e40863.

Toju H, Kurokawa H, Kenta T (2019) Factors influencing leaf- and root-associated communities of bacteria and fungi across 33 plant orders in a grassland. Frontiers in Microbiology 10:241. https://doi.org/10.3389/fmicb.2019.00241

Untersheher M, Jumpponen A, Öpik M, et al (2010) Species abundance distributions and richness estimations in fungal metagenomics - lessons learned from community ecology. Molecular Ecology 20:275-285. https://doi.org/10.1111/j.1365-294X.2010.04948.x

Wang QM, Theelen B, Groenwald M, Bai FY, Boekhout T (2014) *Moniliellomycetes* and *Malasseziomycetes*, two new classes in *Ustilaginomycotina*. Persoonia 33:41-47. https://doi.org/10.3767/003158514X682313

Whitaker BK, Reynolds HL, Clay K (2018) Foliar fungal endophyte communities are structured by environment but not host ecotype in *Panicum virgatum* (switchgrass). Ecology 99:2703-2711. https://doi.org/10.1002/ecy.2543

White TJ, Bruns T, Lee SJWT, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18:315–322

Wu B, Hussain M, Zhang W, et al (2019) Current insights into fungal species diversity and perspective on naming the environmental DNA sequences of fungi. Mycology 10:127-140. https://doi.org/10.1080/21501203.2019.1614106