Fungal pathogen diversity and composition respond rapidly to changes in plant diversity and composition

By

Haley Burrill

B.S., The University of California, Santa Cruz, 2017

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

_____

Chair: Dr. James Bever

_____

Dr. Helen Alexander

_____

Dr. Benjamin Sikes

Date Defended: 8 May 2020

The thesis committee for Haley Burrill certifies that this is the
approved version of the following thesis:

Fungal pathogen diversity and composition respond rapidly to changes in plant diversity and composition

_____

Chair: Dr. James Bever

_____

Dr. Helen Alexander

_____

Dr. Benjamin Sikes

Date Approved: 3 July 2020

Abstract

Accumulating evidence suggests that plant pathogens play a major role in maintaining plant community diversity. Accumulation of host-specific pathogens is expected to negatively impact productivity at low plant diversity (i.e., monocultures), thereby allowing establishment of other plant species. However, in high diversity plant communities, reduced abundance of individual plant species and/or an increase in microbial diversity may inhibit pathogen accumulation and therefore facilitate plant diversity maintenance. In addition to plant diversity, increased phylogenetic distance between plant community members may also affect pathogen accumulation as pathogens are likely to be shared between closely related plant species (i.e. within family or genus). For this reason, under-dispersed plant communities with relatively closely related species are likely to encounter similar risks as monocultures. In order to better understand the ecology of plant-pathogen interactions, fungal pathogen composition was analyzed from plots planted with combinations of 18 plant species from three plant families: Poaceae, Fabaceae, and Asteraceae. Plots were planted in monoculture, 2, 3, or 6 species richness mixtures; either representing multiple families (over-dispersed) or one family (under-dispersed). Soil samples were collected 4 months after plot planting from each of the plant diversity treatments. Soil DNA was extracted, amplified for the fungal ribosomal DNA region ITS2, barcoded, and sequenced. Using bioinformatics and closed reference OTU labeling, we analyzed linear models of known fungal plant pathogens and plant species richness, as well as plant community composition. We tested whether 1) fungal pathogen diversity increases with plant species richness and phylogenetic dispersion, and 2) fungal pathogen relative abundance varies between phylogenetic groups. There was a positive response of fungal plant pathogen species richness to planted species richness treatments with no effects from non-planted species

richness or from phylogenetic dispersion. There were no significant responses of total fungal richness to planted species richness, non-planted species or phylogenetic dispersion. There were significant differences in fungal pathogen community composition between proportion of plant families and species within the plots. A perMANOVA comparing fungal pathogen community response to plant treatments showed significantly different fungal pathogen communities due to the proportion of Poaceae and Fabaceae but not Asteraceae. Pathogen composition also differed in response to proportion of *Coreopsis tinctoria,* both in the planted species treatments and realized percent cover. Total fungal community composition was not significantly different between plant families, though it was significantly different with increased proportion of planted *Dalea pupurea* and realized proportion of *Andropogon gerardii.* We identify several fungal pathogen taxonomic groups that increase or decrease in response to proportion of each plant family to investigate potential pathogen associations. The rapid response of fungal pathogens to plant treatments support previous research suggesting the major role of plant pathogens in plant community composition and can mediate productivity benefits of plant diversity.

Acknowledgments

Firstly I would like to acknowledge the history of the eastern Kansas region where I did my research. Prior to colonization, the Kansa and Osage peoples were driven out of this region, massacred by colonists, and have since been displaced multiple times. The landscape we know as prairies would not exist without these communities who tended to the land before colonization. In developing this land, indigenous peoples' history and relationship to the land has been taken from them. It is essential to recognize this dark history when working on such lands in an attempt to pay respects for those who have been deeply disrespected.

A big thank you to Guangzhou Wang for teaching me molecular work and helping with bioinformatics, statistical analysis, and soil collection. Austin Yoder and Camille Delavaux were also essential in helping with molecular work and bioinformatics. Much gratitude to Kristen Mecke for her hard work in setting up and maintaining the experimental plots. Thanks to the KU Field Station for the space and funding to do this project, in addition to funding from the Nation Science Foundation. The most gratitude to my advisor, Jim Bever, for making time to help me along the way, and my other committee members Helen Alexander and Ben Sikes for supporting my pursuit of this degree. Last, but not least, this accomplishment would not have been possible without the support of my friends, family, and the Bever/Schultz lab.

Table of Contents

List of Figures

List of Tables

Introduction

Global biodiversity is rapidly reducing due to a variety of factors, dominantly driven by conversion of native ecosystems for human use such as monoculture cropland and development. In addition to species extinction, biodiversity loss can degrade ecosystem functions (Tilman *et al.* 2014). Therefore, understanding the mechanisms of biodiversity maintenance is essential to effectively conserve and restore these ecosystems. In plant communities, diversity can be mediated by microbial community composition (Vogelsang *et al.* 2006). Through measuring responses of plant productivity to differential microbial communities, plant-soil feedback (PSF) experiments have been able to influence of plant-microbe interactions in greenhouse experiments (Bever et al. 1997, Crawford *et al.* 2019). Meanwhile, field experiments demonstrate how microbes drive plant species coexistence via negative frequency dependence (Mangan *et al.* 2010). It has been suggested that plant pathogens are likely driving this negative frequency dependence through negative PSF, thereby driving plant community structure (Bever *et al.* 2015, Crawford *et al.* 2019). However, the sequence of events through which pathogens influence plant communities and vice versa has not been fully demonstrated. We investigate a poorly understood step in the feedback process by measuring the response of pathogen communities to variation in plant community composition.

Plant pathogens—microbial organisms that have antagonistic effects on plants—have a notorious history of eradicating particular host species and reshaping plant communities, as with the case of Chestnut Blight (Stephenson 1986) and sudden oak death (Rizzo & Garbelotto 2003). However, most of these extreme cases are due to the invasion of a nonnative microorganism as a result of the inherent globalist nature of travel and trade. These introduced plant pathogens can cause widespread alterations in plant communities, such as the mass-destruction of Chestnut

Blight to a dominant tree species. However, this catastrophic result is certainly most often the exception to an otherwise easily overlooked silent series of interactions in native plant communities. By suppressing dominant plants, soil-borne pathogens may help maintain diverse native plant communities (Mangan *et al.* 2010; Parker *et al.* 2015; Bever *et al.* 2015).

While recent studies have identified the importance of pathogens in maintaining diversity via negative soil feedbacks (Mangan *et al.* 2010; Bever *et al.* 2015; Crawford *et al.* 2019), evidence of plant community diversity response to pathogens is lacking. Due to the host-specificity of plant pathogens, they are likely drivers of plant community composition. These soil-borne pathogens can act as mediators in the plant community – if one plant begins to dominate, the host-specific pathogens will accumulate and attack, keeping it at a lower frequency and thereby allowing other plants to coexist (Eppinga *et al.* 2018, Bagchi *et al.* 2014). Many plant pathogens have coevolved with plant hosts that have been susceptible to their methods of attack. In this way, phylogenetically related plants tend to share pathogen associations due to their inherited response mechanisms to distress (Gilbert & Parker 2016). A plant community with only plants within the same family (under-dispersed), for example, is more likely to accumulate pathogens than a community with a wide range of phylogenetic relatedness (over-dispersed). Therefore, we hypothesize that, in over-dispersed plant communities, pathogen communities will be significantly different due to the host-specificity of soil-borne pathogens.

Fungi are key plant pathogens implicated in PSF, and whose communities may respond to changes in plant host diversity. Fungal pathogens of plants are the most widespread and destructive (Ingold & Hudson 1993), likely playing important roles in negative plant-soil feedbacks. Previous work on the relationship between fungal pathogen diversity and plant species richness, however, is limited. In a large biodiversity experiment, foliar disease severity

was found to decrease as plant species richness increased (Mitchell *et al.* 2002). A similar study looking at foliar fungal pathogen diversity and severity found a positive relationship of plant diversity with pathogen diversity and a negative relationship between plant diversity and foliar pathogen infection (Rottstock *et al.* 2014). However, foliar pathogens and soil-borne pathogens behave differently (Gilbert 2002). We do not know whether this positive relationship between foliar fungal plant pathogens and plant diversity is mirrored by belowground fungal pathogens. Since most plants have host-specific pathogen associations, we might expect higher diversity plant communities to house higher soil-borne pathogen diversity. Alternatively, with increasing plant species richness, there may be lower density of individual host species which may decrease the spread of individual pathogens by "dilution" (Collins *et al.* 2020). This dilution may decrease pathogen detection with increased plant species richness. Regardless of these responses to plant diversity, we expect that changes in plant composition will drive changes in composition of pathogens in the community because of their host-specificity.

To test these hypotheses, we designed a large experiment with plots manipulating plant species richness, phylogenetic dispersion, and plant composition. The study was conducted in a tallgrass prairie restoration on post-agricultural soil, in northeastern region of Kansas, US. This ecosystem is known to include long-lived, deep-rooted plants, which are ideal subjects to study pathogen accumulation. We varied planting of three well-studied plant families, Poaceae, Fabaceae, and Asteraceae, which are commonly represented in prairies with considerable variation of individual species within each family. For example, from this pool we were able to include early and late successional plants within each family, C3 and C4 grasses, and plants with varying above- and below-ground expanse (see *Appendix A*). We used six plants from each family, manipulating plots to contain either one, two, three, or six species; using plant species

richness as a measure of diversity. These plants were either within one family (phylogenetically under-dispersed) or more than one family (over-dispersed), so that we could examine response of fungal pathogen communities to the phylogenetic dispersion of plants. We characterized fungal pathogens using DNA sequencing of soils in plots to test the response of fungal pathogen communities to plant diversity, phylogenetic dispersion, and plant composition treatments.

Methods

Study System

This study was conducted in the floristically diverse tallgrass prairie region of North America. Plot were established in summer 2018, at the KU Field Station in Lawrence, KS, US (39.052462, -95.191656). Historically this land was tallgrass prairie, followed by cropland and pasture, today considered "post-agricultural" with predominantly cool-season nonnative grasses (Kettle *et al.* 2000). As part of the experiment setup, we tilled the resident soil and added soil made available because of a high-way widening construction from an unplowed prairie remnant near Welda, KS, (38.179600, -95.265695) approximately 100 km south of the experiment site. This provided experimental plots with an initial microbial inocula of remnant prairie microbes.

Experimental Design

A total of 240 plots (1.5 m x 1.5 m) were designed to equally represent each of the 18 plant species (6 from each of the three plant families, Poaceae, Fabaceae, and Asteraceae) within each combination of plant species richness (1, 2, 3, and 6), phylogenetic dispersion (under or over), and precipitation (50 or 150% ambient). Plots varied in plant diversity, phylogenetic dispersion and composition across 72 monoculture plots, 72 with 2 species mixtures, 48 with 3 species mixtures, and 48 with 6 species (Table 1; Figure 1; *Appendix E*). These plots represent two replicates of the same 120 plant combinations, with half set up to receive 150% water treatment (150% of annual precipitation), while the other 120 replicated plots would receive 50% water (50% of annual precipitation). However, this water treatment began after samples for this analysis were collected and therefore precipitation effects will not be considered in these analyses. Soil samples collected from these replicate future precipitation treatments were pooled

prior to analysis. 120 pooled samples were analyzed here: 36 monocultures, 36 two-species, 24 three-species, and 24 six-species. Two-species plots either contained two plant families (Poaceae and Fabaceae, Poaceae and Asteraceae, or Fabaceae and Asteraceae) to represent over-dispersion; three- and six-species plots either contained all three families (over-dispersion) or species all within one plant family (under-dispersion) (Figure 1).

Table 1 - Plant species used; 6 per family.

Poaceae	Fabaceae	Asteraceae
Schizachrium scoparium	*Amorpha canescens*	*Liatris pycnostachya*
Andropogon gerardii	*Dalea candida*	*Coreopsis tinctoria*
Koeleria macrantha	*Dalea purpureum*	*Echinacea pallida*
Elymus canadense	*Desmanthus illinoensis*	*Eupatorium altissimum*
Bouteloua gracilis	*Desmodium canadense*	*Silphium integrifolium*
Panicum virgatum	*Chamaecrista fasticulata*	*Helianthus mollis*

Figure 1 – A factorial description of plot design. Number of plant species denotes the plant species richness treatments (monoculture, 2, 3, or 6 species). Phylogenetically underdispersed combinations of plants are all within one plant family (e.g. 2 grasses, or 3 legumes, or 6 asters). Phylogenetically overdispersed combinations of plants are from more than one plant family (e.g. 1 grass and 1 aster, or 1 species from each family, or 2 species from each family). Monocultures are inherently underdispersed. The sets describe the combinations in which each of the 18 plant species are represented once (i.e. within sets species are randomly chosen without replacement from the pool

of 18 plant species). There are 18 monocultures, two sets of 9 underdispersed 2-plant plots, two sets of 9 overdispersed 2-plant plots, two set of 6 underdispersed 3-plant plots, two sets of 6 overdispersed 3-plant plots, three sets of underdispersed 6-plant plots, and three sets of overdispersed 6-plant plots. This approach equally represents each plant species in each richness treatment by phylogenetic dispersion combination.

Experiment Details

Prairie seedlings from the three most common plant families (Poaceae, Fabaceae, Asteraceae) were planted in May 2018. A total of 18 species were used (*Table 1*). Seeds were purchased from producers located near eastern Kansas: Hamilton Native Outpost, Stock Seed, Missouri Wildflowers, and Prairie Moon. Each plot was seeded with each plant species being equally represented by pre-weighing 100 seeds per species and evenly dividing the final mix of species by weight, resulting in 1800 seeds per blend. Resident soil microbes were augmented with soil microbes from native prairie in two ways. First, we added 3.81 cm of soil from an unplowed native prairie soil from Welda, KS, which was then tilled into the resident soil to a depth of 15.24 cm. In addition, 18 seedlings previously inoculated with native Welda soil were planted into each plot. Seeds were sowed into flats with autoclaved sterile potting soil and placed in cold-moist stratification for 4 weeks prior to germination. When large enough, the seedlings were transplanted into Stuewe and Sons groove tubes (GT51D) with 98 mL of Welda soil and grown in a greenhouse for 5 weeks prior to being planted into a hexagonal array within each plot.

Plant data

Weedy plants—those that were not the planted/seeded species—were removed early summer of 2018. Only one round of weeding was accomplished, removing all forbs but leaving any grasses below 3 inches. Percent cover of plants was collected August 2018, 3 months after planting. A 1 m x 1 m quadrat was placed in the center of each plot, approximately 0.5 m in towards the center from the south side. Percent cover values were estimated percentages of cover per species in the plot, allowing totals to exceed 100%. Identified species were included on a

species list (*Appendix B*) and unknowns were flagged for later identification. Those that were too small to identify were labeled as "unkforb" for unknown forbs or "unkgrass" for unknown grass. Percent cover data was converted into realized proportion of plants within each plant family, so that Asteraceae, Poaceae, and Fabaceae could be compared to the planted proportions. Each plant family percent cover value was divided by the percent cover of all plants within each plot, to calculate the proportion of plants in the plot within each plant family.

Soil Collection

In September 2018, approximately 4 months after planting, soil was collected from the 240 plots. A total of two 20 cm soil cores were taken from each plot, added to a sample bag, and then paired plots of matched plant composition were pooled across the future rainfall treatments (i.e., 1,21; 2,22; 3,23; etc.; Figure 2). Coring devices were rinsed of dirt in a water bucket, then sterilized in 10% bleach bucket between plots. Immediately following soil collection, samples were kept on ice, then later transferred to a -20 °C freezer within 5 hours (Delavaux *et al.* 2020). Homogenized samples were thawed to sieve out roots and, following the Qiagen DNeasy PowerSoil kit, 0.25 g of the remaining soil was weighed for DNA extraction. The extraction kit utilizes a series of buffers and ethanol to isolate DNA from other soil particulates.
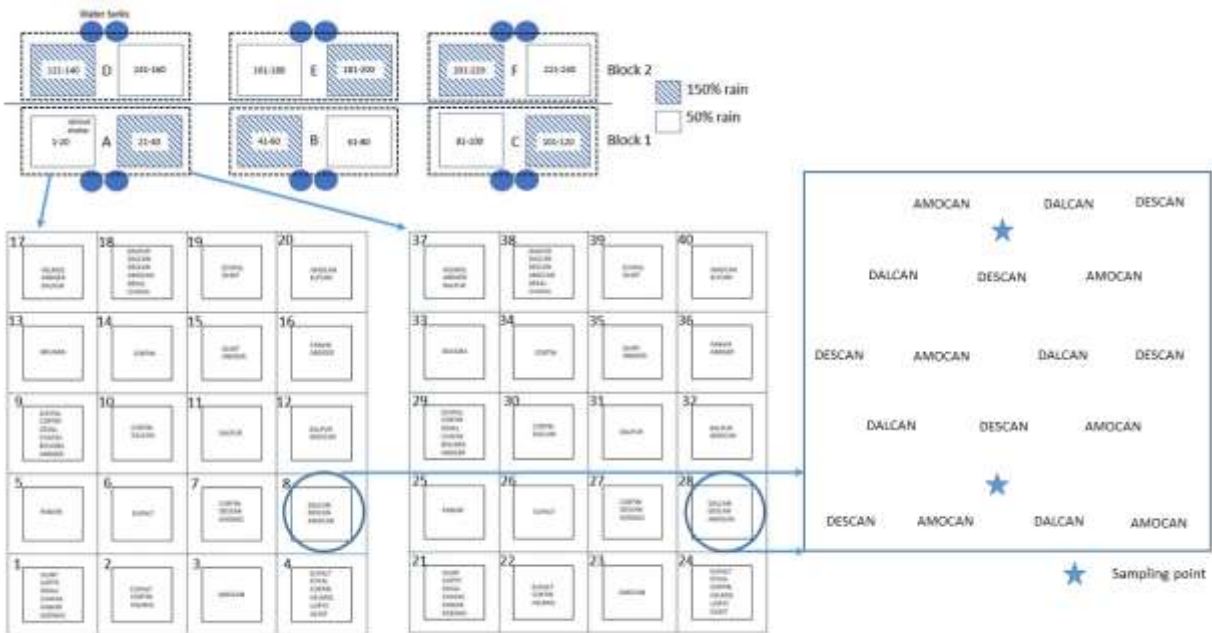
Figure 2 – Experimental plot layout, with Blocks 1 & 2 separated by a road. Subblocks (A-F) have replicate plant design (as seen in zoom of plots 1-20 and 21-40), as each of these receive rainfall manipulations (50% or 150%) prior to this data collection. Final zoom view shows soil sampling points within each plot, arrows denoting the pooling of replicate plots.

DNA Sequencing

Using extracted DNA, we performed next-generation sequencing to target the internal transcribed spacer (ITS2) region with primers of fungal forward ITS4 and reverse fITS7 (Ihrmark *et al.* 2012). We did two rounds of PCR, the first (PCR1) used a mixture of 1 μL DNA, 0.5 μL forward and 0.5 μL reverse primers, 12.5 μL Phusion buffer, and 10.5 μL sterilized DI water per sample. Our fungal PCR cycle was 94°C for 5 minutes, then 35 cycles at 94°C for 30 seconds, 57°C for 30 seconds, and 72°C for 30 seconds, followed by a final extension step of 72°C for 7 minutes and then kept at 4°C until retrieved from the thermocycler. Following PCR1, we used the NucleoMag 96 PCR kit, binding DNA to metal beads that attract it towards a magnetic plate, using a series of buffers to pipet non-DNA particles out of the samples. The

second PCR attached 5 µL of cleaned DNA from PCR1, 2.5 µL of each Illumina barcode primer, 25 µL Phusion, and 10 µL of DI water. The barcoding PCR program was as follows: 98℃ for 30 seconds, then 10 cycles of 98℃ for 10 seconds, 55℃ for 30 seconds, and 72℃ for 30 seconds, a final 72℃ cycle for 5 minutes, and then 4℃ until. The barcoded amplicons were cleaned again using the same microbead kit, and DNA concentrations measured using a Qubit. Products from all samples were pooled at equimolar concentrations with final pool at 4uM. The pool was size and concentration verified on an Agilent TapeStation (Santa Clara, CA, USA), then sequenced on an Illumina MiSeq (San Diego, CA, USA) with 300bp, paired-end chemistry at the KU Genome Sequencing Core.

Bioinformatics

We processed sequencing data through a bioinformatic pipeline to form operational taxonomic units (OTUs) by clustering fungal amplicons and match OTUs to a known fungal database to help identify fungal pathogens. The entire bioinformatics pipeline is available in *Appendix C*. Sequencing data were already demultiplexed (e.g. split by sample) by the GSC. We first trimmed primers using cutadapt (Martin 2011), then used dada2 as implemented in QIIME2 (Bolyen *et al.* 2019) to remove low quality reads and truncate them at 260 base pairs. Sequences were clustered into OTUs with a dynamic threshold using the open reference OTU algorithm (Abarenkov *et al.* 2020) and with the UNITE taxonomy 2.19 database to 97% matching (Nilsson *et al.* 2019). To remove potentially spurious reads (e.g. PCR or sequencing artefacts), the dataset was filtered to remove reads that occurred less than 5 times across the entire datasets. Finally, we converted the reads to an OTU table using QIIME2 (Bolyen *et al.* 2019). We uploaded the OTU table to the reference database FUNGuild (Nguyen *et al.* 2016) to identify known functional groups. Of the 7272 OTUs, 1904 were matched in the FUNGuild, with 1394 either "probable" or

"highly probable", and remaining (e.g. possible) excluded as per Nguyen *et al.* recommendations (2016). The 254 putative fungal plant pathogens (about 3% of the original OTUs) were the main focus for downstream analyses of fungal pathogen responses to planting treatments, but we also analyzed data on all fungal OTUs for comparison.

Statistical Analysis (*Appendix D*)

To assess planting effects on fungal diversity, we performed regression analyses of fungal pathogen diversity and total fungal diversity (both H') with planted species richness using non-planted species richness as a covariate. To calculate total fungal diversity and fungal pathogen diversity, we used the diversity function in vegan, with the default Shannon-Wiener index (Oksanen *et al.* 2019).

We compared similarity of total fungal or pathogen only communities among planting treatments to test how they impacted community structure. First, fungal OTU tables were rarefied to the same sequencing depth, using the sample with the lowest number of reads (54) in the R package vegan (Oksanen *et al.* 2019). The dissimilarity matrix was then calculated using Bray-Curtis method and differences among plant treatments analyzed using the adonis2 function in vegan with default parameter of 1000 permutations (Oksanen *et al.* 2019).

Using dissimilarity matrices of all fungi and fungal pathogens, we performed perMANOVA tests to assess the response in variance of fungal community composition explained by block (for spatial non-independence), plant diversity, phylogenetic dispersion, proportion of weedy plants (calculated from species cover data), proportion of planted proportion of Poaceae, Fabaceae, Asteraceae, as well as proportions of the 18 planted species. In order to assess whether these patterns were influenced by non-planted species or the success of any one species, we performed the same perMANOVA analyses using actual plant cover data as fixed

factors, represented as realized proportion of Poaceae, Fabaceae, Asteraceae, as well as proportions of each of the 18 plant species.

To assess relative abundance of fungal pathogens within taxonomic groups we used usearch10 (see *Appendix E;* Edgar 2010) in R to assign fungal OTUs to phyla, family, genus, and some to species. We use the relative abundances of fungal families and particularly abundant genera in linear models with proportion of each plant family (planted species) per plot and report fungal pathogen groups with significant increases and decreases. Here we focus on fungal genera, since very few OTUs could be identified to species. In order to assess whether planted species non-planted species have an effect, we performed the same analysis using actual plant cover data.

Results

Fungal pathogen diversity

Fungal pathogen diversity increased with increased plant species richness (Figure 3, $F_{3,116}=1.878$, $p=0.02$) with no effects from non-planted species ($F_{3,116}=1.878$, $p=0.4$) or from phylogenetic dispersion ($F_{3,116}=1.878$, $p=0.4$). Total fungal diversity had no significant response to increased plant species richness (Figure 3, $F_{3,116}=0.1853$, $p=0.6$), non-planted species ($F_{3,116}=0.1853$, $p=0.8$) or phylogenetic dispersion ($F_{3,116}=0.1853$, $p=0.8$).

Fungal pathogen composition

There were significant differences in fungal pathogen community composition between plant families and species. Fungal pathogen communities were significantly different due to spatial block ($p=0.004$), planted proportion of Fabaceae ($p=0.001$), planted species richness ($p=0.01$), proportion of weeds ($p=0.04$), and marginally with planted proportion of Poaceae ($p=0.08$), but not with planted proportion Asteraceae ($p=0.3$), or phylogenetic dispersion ($p=0.9$). Fungal pathogen communities were also significantly affected by the proportion of *S. scoparium* ($p=0.03$), *D. canadense* ($p=0.01$), and *C. tinctoria* ($p=0.04$), as well as the interaction of planted species richness with proportion of Asteraceae and weeds ($p=0.001$). Using realized proportion as predictors, fungal pathogen communities were significantly affected by spatial block ($p=0.004$), planted species richness ($p=0.02$), the realized proportion of Poaceae ($p=0.002$), the realized proportion of Fabaceae ($p=0.001$), but not the realized proportion of Asteraceae (0.2) or phylogenetic dispersion ($p=0.9$). Fungal pathogen communities were also significantly affected by realized proportion of *D. purpurea* ($p=0.05$), *C. tinctoria* ($p=0.03$), and plant species richness interacting with the proportion of each of the three plant families ($p=0.001$).

There were also significant differences in total fungal community composition between plant families and species. Total fungal communities were significantly different due to spatial block (p=0.001), but not phylogenetic dispersion (p=0.9) or planted species richness (p=0.4), the proportion of Poaceae (p=0.2), the proportion of Fabaceae (p=0.7), the proportion of Asteraceae (p=0.3), or the proportion of weeds (p=0.7). Total fungal communities were significantly affected by planted proportion of *D. purpurea* (p=0.05), as well as the interaction of planted species richness with proportion of Fabaceae and weeds (p=0.003). Focusing on realized plant composition, total fungal communities were significantly affected by spatial block (p=0.001), but not phylogenetic dispersion (p=0.9), planted species richness (p=0.4), the planted proportions of Poaceae (p=0.2), Fabaceae (p=0.09), or Asteraceae (p=0.5). Total fungal communities were significantly different between realized proportion of *A. gerardii* (p=0.05), the interaction of planted species richness with realized proportion of Fabaceae (p=0.05), and the interaction of planted species richness with realized proportion of Asteraceae (p=0.004).

In analyses of changes in the relative abundance of pathogen OTUs with proportion of planted plant family proportions, we find evidence of several fungal families responding to particular plant families (Table 2). The relative abundance of fungal pathogens within Hyponectriaceae (Figure 4) significantly increased with proportion of Fabaceae ($F_{1,118}$=14.93, p=0.0002), but decreased with proportion of Poaceae ($F_{1,118}$=4.039, p=0.05) and proportion of Asteraceae ($F_{1,118}$=2.862, p=0.09). The relative abundance of fungal pathogens within Mycospaerellaceae (Figure 5) significantly increased with proportion of Fabaceae ($F_{1,118}$=7.308, p=0.008), but decreased with proportion of Poaceae ($F_{1,118}$=3.561, p=0.06) with no response driven by the proportion of Asteraceae ($F_{1,118}$=0.628, p=0.4). The relative abundance of fungal pathogens within Pleosporaceae (Figure 6) significantly decreased with proportion of Fabaceae

($F_{1,118}$=7.105, p=0.009), but increased with proportion of Poaceae ($F_{1,118}$=3.671, p=0.06) with no

response driven by Asteraceae ($F_{1,118}$=0.495, p=0.5). We also find evidence of several fungal

genera responding to particular plant families (Table 3). The relative abundance of fungal

pathogens within *Monographella* (Figure 8) significantly increased with proportion of Fabaceae

($F_{1,118}$=14.93, p=0.0002), but decreased with proportion Poaceae ($F_{1,118}$=4.039, p=0.05) and

proportion of Asteraceae ($F_{1,118}$=2.862, p=0.09). The relative abundance of fungal pathogens

within *Erysiphe* (Figure 7) significantly increased with proportion of Fabaceae ($F_{1,118}$=9.315,

p=0.003), but decreased with proportion of Asteraceae ($F_{1,118}$=2.795, p=0.09) with no effect

driven by proportion Poaceae ($F_{1,118}$=1.724, p=0.2). The relative abundance of fungal pathogens

within *Cercospora* (Figure 9) significantly increased with proportion of Fabaceae ($F_{1,118}$=8.158,

p=0.005), but decreased with proportion of Poaceae ($F_{1,118}$=4.081, p=0.05) with no effect driven

by proportion Asteraceae ($F_{1,118}$=0.647, p=0.4).
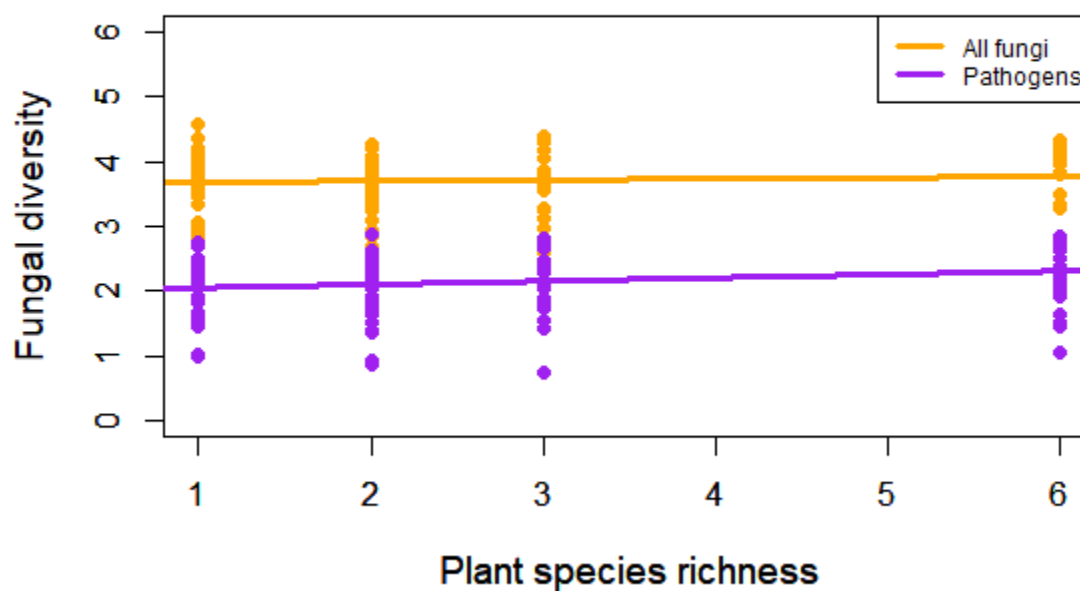
Figure 3 – General linear model of total fungal diversity (p=0.6) and fungal pathogen diversity (p=0.02) in response to planted species richness, with non-planted species as a covariate.

Table 2 – Linear model results of fungal pathogen OTU relative abundance by family significant to marginally significant increases (+) and decreases (-) with proportion of plant family.

Fungal family	Proportion plant family	$F_{1,118}$	p-value	Incr/Decr
Hyponectriaceae	Fabaceae	14.93	0.0002	+
	Poaceae	4.039	0.05	-
	Asteraceae	2.862	0.09	-
Mycosphaerellaceae	Fabaceae	7.308	0.008	+
	Poaceae	3.561	0.06	-
Pleosporaceae	Fabaceae	7.105	0.009	-
	Poaceae	3.671	0.06	+

Figure 4 – General linear model of relative abundance of fungal pathogens within Hyponectriaceae increasing with proportion of Fabaceae ($F_{1,118}$=14.93, p=0.0002) and decreasing with proportion of Poaceae ($F_{1,118}$=4.039, p=0.05) and Asteraceae ($F_{1,118}$=2.862, p=0.09).

Figure 5 – General linear model of relative abundance of fungal pathogens within Mycosphaerellaceae increasing with proportion of Fabaceae ($F_{1,118}$=7.38, p=0.008) and decreasing with proportion of Poaceae ($F_{1,118}$=3.561, p=0.06).

Figure 6 – General linear model of relative abundance of fungal pathogens within Pleosporaceae decreasing with proportion of Fabaceae ($F_{1,118}$=7.105, p=0.009) and increasing with proportion of Poaceae ($F_{1,118}$=3.671, p=0.06).

Table 3 – Linear model results of fungal pathogen OTU relative abundance by genera significant to marginally significant increases (+) and decreases (-) with proportion of plant family.

Fungal genus	Plant family	F1,118	p-value	Incr/Decr
Monographella	Fabaceae	14.93	0.0002	+
	Poaceae	4.039	0.05	-
	Asteraceae	2.862	0.09	-
Erysiphe	Fabaceae	9.315	0.003	+
	Asteraceae	2.796	0.09	-
Cercospora	Fabaceae	8.158	0.005	+
	Poaceae	4.081	0.05	-

Figure 7 – General linear model of relative abundance of fungal pathogens within *Monographella* increasing with proportion of Fabaceae ($F_{1,118}=14.93$, *p*=0.0002) and decreasing with proportion of Poaceae ($F_{1,118}=4.039$, *p*=0.05) and Asteraceae ($F_{1,118}=2.862$, *p*=0.09).
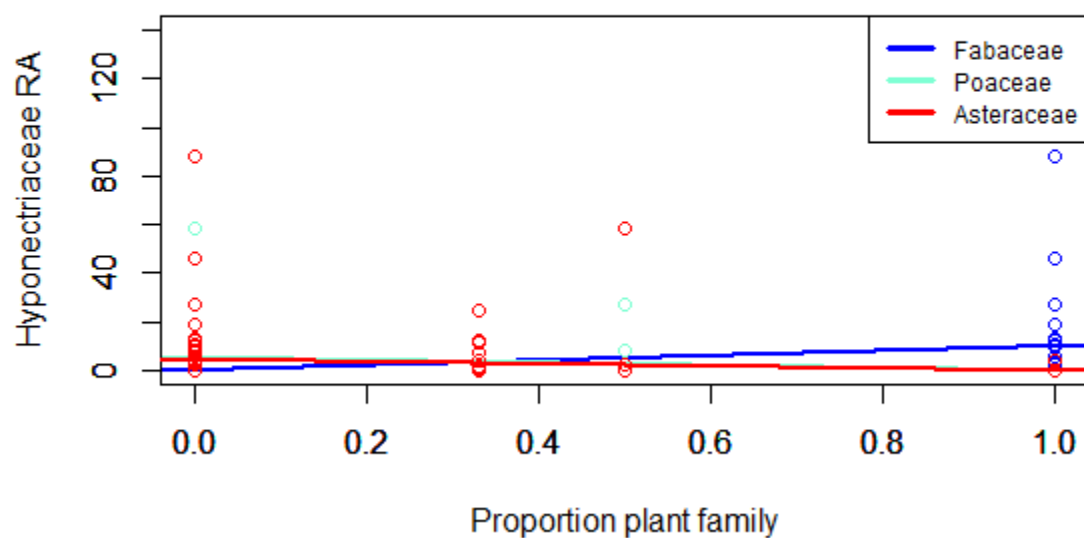

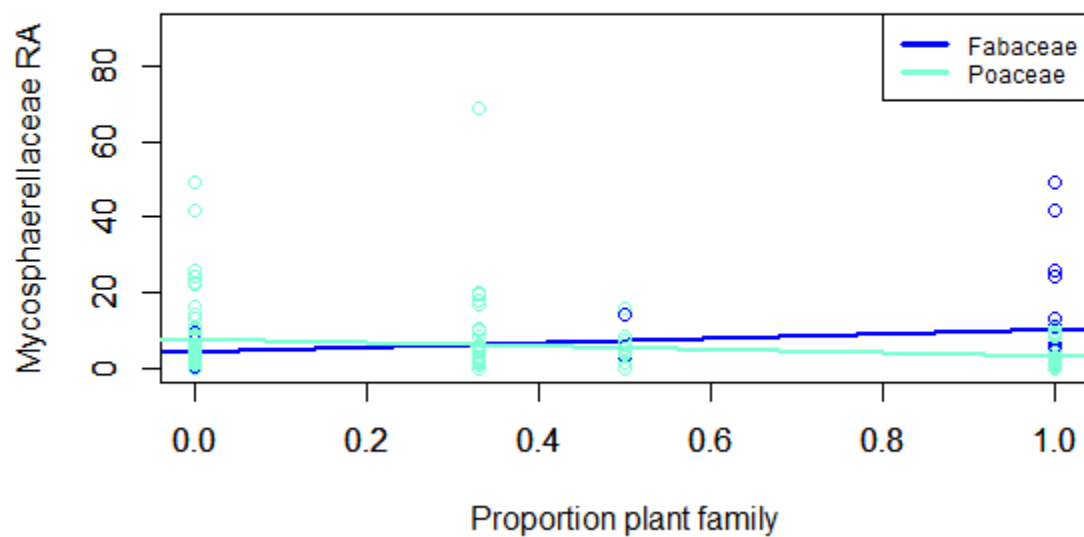Figure 8 – General linear model of relative abundance of fungal pathogens within *Erysiphe* increasing with proportion of Fabaceae ($F_{1,118}=9.315$, *p*=0.003) and decreasing with proportion of Asteraceae ($F_{1,118}=2.796$, *p*=0.09).
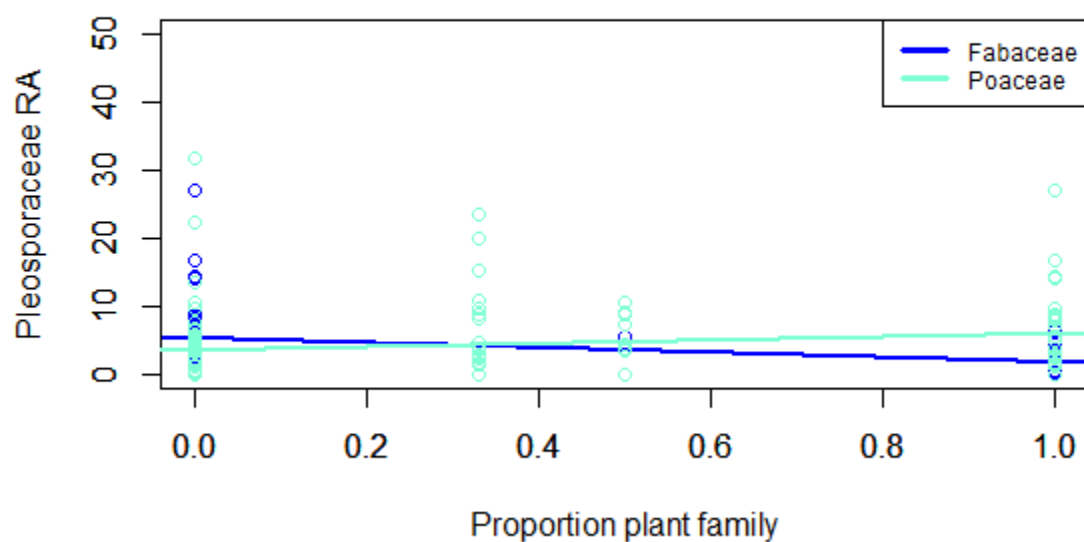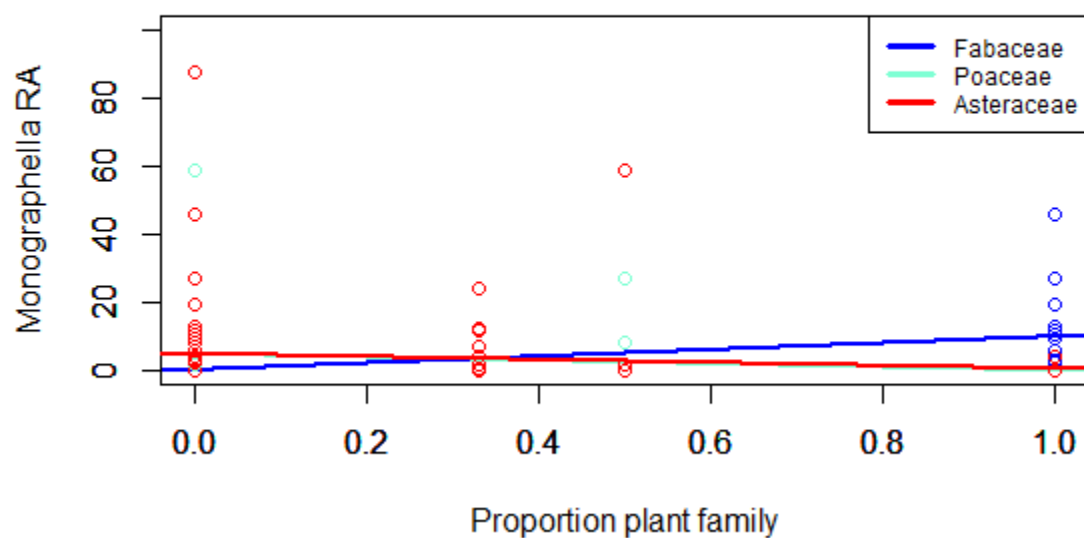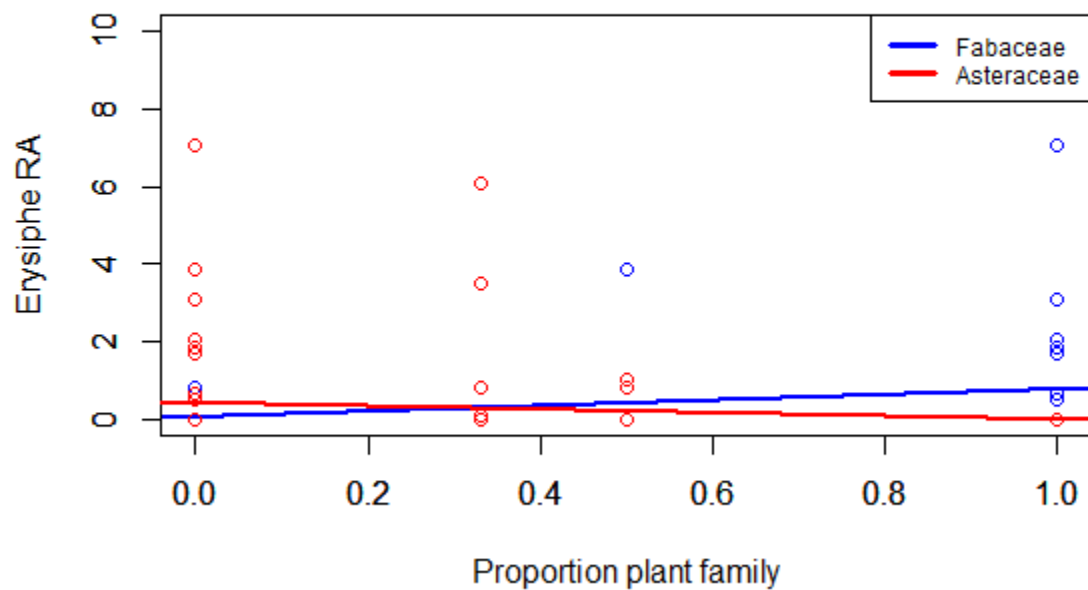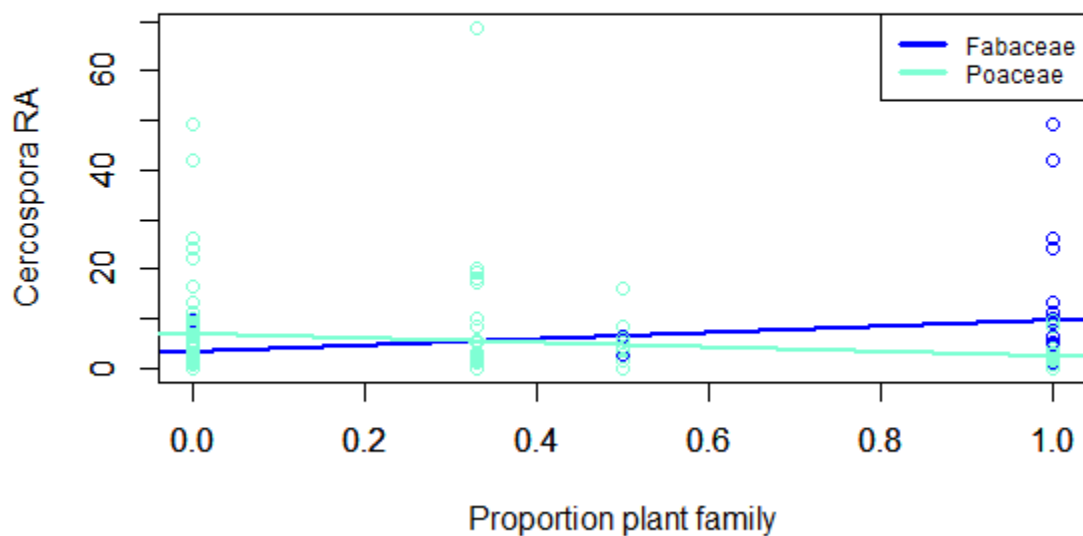
Figure 7 – General linear model of relative abundance of fungal pathogens within *Cercospora* increasing with proportion of Fabaceae ($F_{1,118}=8.158$, *p*=0.005) and decreasing with proportion of Poaceae ($F_{1,118}=4.081$, *p*=0.05).

Discussion

The results of this experiment show fungal pathogen diversity increases with plant species richness, while total fungal diversity had no significant response. The presence and amount of non-planted species in a plot was not associated with either fungal or fungal pathogen diversity. Previous studies have demonstrated the negative relationship between pathogen load and plant species richness (Mitchell *et al.* 2002, Hantsch *et al.* 2013) which demonstrates how lower plant diversity can cause pathogen accumulation. Our findings complement these studies by identifying that fungal pathogen diversity rapidly increased in response to plant species richness manipulations. Although we might have expected plant pathogens to respond to plant phylogenetic dispersion, due to shared pathogens of phylogenetically closely related plant species, our findings did not support that expectation. However, fungal pathogen composition differed in response to changes in proportion of plant family. In this way, the response of individual plant families is likely stronger than the overarching phylogenetic dispersion of the plot.

We report rapid responses of fungal pathogen composition to plant families as well as individual plant species. In 3 months, distinct fungal pathogen communities formed in plots with Fabaceae and Poaceae, but not Asteraceae. These findings were consistent in models using planted species only, as well as realized cover data. We did not see these differences in similar analyses of total fungal community composition. These rapid shifts of fungal pathogens in response to the proportion of plant family align with the expectations that host-specific pathogens are shared within plant family and accumulate where the family is more abundant (Gilbert & Parker 2016). Investigations of plant pathogen host ranges have found increasing disease pressure on closely related communities (Gilbert & Webb 2007; Parker *et al.* 2015), and

our observations of compositional differences are consistent with fungal pathogen accumulation generating similar effects in fitness of plants in plots dominated by single plant families.

Differences in the relative abundance of specific fungal pathogens among these treatments may indicate their association with specific plant families. For example, fungi in the families Hyponectriaceae and Mycosphaerellaceae increased with Fabaceae, but decreased with Poaceae. Similarly, pathogens within the genera *Monographella, Erysiphe,* and *Cercospora* all increased in plots with a greater proportion of Fabaceae. Fungal pathogen communities differed based on individual plants as well, including with greater dominance by the legumes *D. canadense* and *D. purpurea*, as well as the forb *C. tinctoria*. However, these results only represent the matched fungal OTU sequences, some matched to species, but most only matched to genus or family. Despite this drawbacks of fungal sequencing data analysis, our findings provide evidence of rapid responses of fungal pathogen communities to the plant community composition manipulations.

By demonstrating rapid host-specific changes in pathogen structure, our results provides complementary evidence of pathogens role in facilitating plant species coexistence. The differential relative abundances of fungal pathogen OTUs we see with plant taxonomic groups is likely early detection of pathogen specificity (Gilbert & Parker 2016). As pathogen differentiation with host composition is a critical assumption of feedback models and of pathogen driven coexistence of plant species (Bever *et al.* 1997), our results are consistent with pathogens generating negative feedbacks between phylogenetically distant plant species. This expectation of negative feedback increasing with plant phylogenetic distance was supported by a recent meta-analysis (Crawford *et al.* 2019), and together, these results suggest that pathogen dynamics will increase the phylogenetic richness of plant communities.

The increase of pathogen diversity with increasing plant species richness supports the dilution effect hypothesis (Keesing *et al.* 2015). This hypothesis suggests that, in low plant species richness, host plants are at greater abundance and therefore more susceptible to disease. With higher plant species richness, host plants are relatively less abundant and therefore less likely to be infected. Above-ground disease severity has been found to decrease with increased plant diversity (Mitchell *et al.* 2002). As thoroughly described in Collins *et al.* (2020), the dilution effect caused by higher abundance of non-host plant species reduces pathogen driven negative feedbacks on host plants. Since relative abundance of each host is lower, so is the possibility of encountering its associated pathogen, and therefore we would expect decreased infection. Our results are consistent with pathogen dilution contributing to the productivity benefits of increasing plant species and phylogenetic richness (Maron *et al.* 2011, Schnitzer *et al.* 2011, Wang *et al.* 2019).

Only three months after planting we found that pathogen diversity and composition had changed significantly in response to plant community composition manipulations. Such a rapid response is remarkable given expectations of a time lag of microbial community differentiation (Bever *et al.* 1997, 2012). Nevertheless, only a small proportion of the variation in pathogen diversity and composition was explained by plant species composition. With more time for microbial communities to establish in response to plant treatments, we expect this trend to strengthen and explain more of the variation in pathogen diversity and composition (*see* Figure 1). Moreover, we did not observe significant effects of phylogenetic dispersion on plant pathogen richness, which may manifest over longer periods of time.

Our results are consistent with the dynamics of soil-borne fungal pathogens contributing to plant community diversity maintenance, supporting the literature on negative soil feedbacks

due to pathogen accumulation inhibiting single-species dominance (Bever *et al.* 1997, Crawford *et al.* 2019). We also have support for plant pathogen host-specificity having a phylogenetic pattern (Gilbert & Parker 2016). The next step in incorporating this study's findings to the bigger picture of microbial mediation of plant community diversity and productivity is a feedback test, growing the plants in their own soil (presumably with accumulated pathogens) as well as soil from other plant species mixtures. In this way, we can gather evidence to support the dilution effect and the impact of host-specific pathogen accumulation in the field plots. In addition, re-sampling the soil two years after this initial sampling with allow us to assess the development of these relationships over time. While fungal pathogen diversity increased with plant species richness, we expect the impact of pathogens to be lower due to the dilution effect. Similarly, differences in fungal pathogen composition expose likely host-specificity and spill-over, further supporting the need of diverse plant communities to sustain productivity. This supports the need of diverse plant communities to maintain healthy ecosystems, which is of ever-increasing importance in restoration and conservation of prairies nationwide.

References

Abarenkov, Kessy; Zirk, Allan; Piirmann, Timo; Pöhönen, Raivo; Ivanov, Filipp; Nilsson, R.
Henrik; Kõljalg, Urmas (2020): UNITE QIIME release for Fungi. Version 04.02.2020.
UNITE Community. https://doi.org/10.15156/BIO/786385

Bagchi, R. *et al.* Pathogens and insect herbivores drive rainforest plant diversity and
composition. *Nature* **506**, 85–88 (2014).

Bever, J. D., Mangan, S. A. & Alexander, H. M. Maintenance of Plant Species Diversity by
Pathogens. *Annu. Rev. Ecol. Evol. Syst.* **46**, 305–325 (2015).

Bever, J. D., Platt, T. G. & Morton, E. R. Microbial Population and Community Dynamics on
Plant Roots and Their Feedbacks on Plant Communities. *Annu. Rev. Microbiol.* **66**, 265–283
(2012).

Bever, J. D., Westover, K. M. & Antonovics, J. Incorporating the Soil Community into Plant
Population Dynamics: The Utility of the Feedback Approach. *The Journal of Ecology* **85**,
561 (1997).

Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science
using QIIME 2. *Nat Biotechnol* **37**, 852–857 (2019).

Collins, C. D., Bever, J. D. & Hersh, M. H. Community context for mechanisms of disease
dilution: insights from linking epidemiology and plant–soil feedback theory. *Ann. N.Y.
Acad. Sci.* nyas. 14325 (2020) doi:10.1111/nyas.14325.

Crawford, K. M. *et al.* When and where plant-soil feedback may promote plant coexistence: a
meta-analysis. *Ecology Letters* **22**, 1274–1284 (2019).

Delavaux, CS, Bever, JD, Karppinen, EM, Bainard, LD. Soil sample cold pack storage and DNA shipment up to 1 month does not impact metabarcoding results. *Ecol Evol*. 00: 1– 13 (2020). https://doi.org/10.1002/ece3.6219

Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26**(19) 2460-2461 (2010).

Eppinga, M. B. *et al.* Frequency-dependent feedback constrains plant community coexistence. *Nat Ecol Evol* **2**, 1403–1407 (2018).

Gilbert, G. S. Evolutionary Ecology of Plant Diseases in Natural Ecosystems. *Annual Review of Phytopathology* **40**, 13 (2002).

Gilbert, G. S. & Webb, C. O. Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Sciences* **104**, 4979–4983 (2007).

Gilbert, G. S. & Parker, I. M. The Evolutionary Ecology of Plant Disease: A Phylogenetic Perspective. *Annu. Rev. Phytopathol.* **54**, 549–578 (2016).

Hantsch, L., Braun, U., Scherer-Lorenzen, M. & Bruelheide, H. Species richness and species identity effects on occurrence of foliar fungal pathogens in a tree diversity experiment. *Ecosphere* **4**, art81 (2013).

Ingold C.T., Hudson H.J. (1993) Fungi as plant pathogens. In: The Biology of Fungi. Springer, Dordrecht

Ihrmark, K., Bödeker, I., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandström-Durling, M., Clemmensen, K.E., Lindahl, B. D. New primers to amplify the fungal ITS2 region–evaluation by 454-sequencing of artificial and natural communities. *FEMS microbiology ecology*, *82*(3), 666-677 (2012).

Keesing, F. & Ostfeld, R. S. Is biodiversity good for your health? *Science* **349**, 235–236 (2015).

Kettle, W.D., Rich, P.M., Kindscher, K., Pittman, G.L., Fu, P. Land-Use History in Ecosystem Restoration: A 40-Year Study in the Prairie-Forest Ecotone. Restoration Ecology 8, 307–317 (2000). https://doi.org/10.1046/j.1526-100x.2000.80043.x

Mangan, S. A. *et al.* Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**, 752–755 (2010).

Maron, J. L., Marler, M., Klironomos, J. N. & Cleveland, C. C. Soil fungal pathogens and the relationship between plant diversity and productivity. *Ecology Letters* **14**, 36–41 (2011).

Martin, Marcel. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**(1) 10-12, (2011). doi:https://doi.org/10.14806/ej.17.1.200

Mitchell, C. E., Tilman, D. & Groth, J. V. Effects of Grassland Plant Species Diversity, Abundance, and Composition on Foliar Fungal Disease. *Ecology* **83**, 1713–1726 (2002).

Nilsson, R.H., Larsson, K., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K.,Glöckner, F.O., Tedersoo, L., Saar, I., Kõljalg, U., Abarenkov, K. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research*. **47**(D1) D259–D264 (2019). https://doi-org.www2.lib.ku.edu/10.1093/nar/gky1022

Nguyen, N. H. *et al.* FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* **20**, 241–248 (2016).

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., Wanger, H. vegan: Community Ecology Package. *R package version 2.5-6* (2019). https://CRAN.R-project.org/package=vegan

Parker, I. M. *et al.* Phylogenetic structure and host abundance drive disease pressure in communities. *Nature* **520**, 542–544 (2015).

Rizzo, D. M. & Garbelotto, M. Sudden Oak Death: Endangering California and Oregon Forest Ecosystems. *Frontiers in Ecology and the Environment* **1**, 197–204 (2003).

Rottstock, T., Joshi, J., Kummer, V. & Fischer, M. Higher plant diversity promotes higher diversity of fungal pathogens, while it decreases pathogen infection per plant. *Ecology* **95**, 1907–1917 (2014).

Schnitzer, S. A. *et al.* Soil microbes drive the classic plant diversity–productivity pattern. *Ecology* **92**, 296–303 (2011).

Stephenson, S. L. Changes in a Former Chestnut-Dominated Forest after a Half Century of Succession. *The American Midland Naturalist* **116**, 173–179 (1986).

Tilman, D., Isbell, F. & Cowles, J. M. Biodiversity and Ecosystem Functioning. *Annu. Rev. Ecol. Evol. Syst.* **45**, 471–493 (2014).

Vogelsang, K. M., Reynolds, H. L. & Bever, J. D. Mycorrhizal fungal identity and richness determine the diversity and productivity of a tallgrass prairie system. *New Phytol* **172**, 554–562 (2006).

Wang, G., Schultz, P., Tipton A., Zhang, J. Zhang, F., Bever, J. D. Soil microbiome mediates positive plant diversity-productivity relationships in late successional grassland species. *Ecol Lett* ele.13273 (2019). doi:10.1111/ele.13273.

Appendix A: Plant species details

Poaceae	Details
Schizachrium scoparium	Late successional, 0.6-1.2 m tall, C4
Andropogon gerardii	Late successional, 1.2-1.8 m tall, C4
Koeleria macrantha	Early successional, 0.3-0.6 m tall, C3
Elymus canadense	Late successional, 0.6-1.5 m tall, C3
Bouteloua gracilis	Early successional, 0.2-0.7 m tall, C4
Panicum virgatum	Late successional, 0.9-1.8 m tall, C4

Fabaceae	Details
Amorpha canescens	Flowers July-September, 0.6-0.9 m tall
Dalea candida	Flowers May-July, 0.3-0.6 m tall
Dalea purpureum	Flowers June-August, 0.3-0.9 m tall
Desmanthus illinoensis	Flowers June, 0.6-0.9 m tall
Desmodium canadense	Flowers July-September, 0.6-1.8 m tall
Chamaecrista fasticulata	Flowers June-September, 0.3-0.9 m tall

Asteraceae	Details
Liatris pycnostachya	Flowers July-August, 0.6-1.5 m tall
Coreopsis tinctoria	Flowers June-September, 0.6-1.2 m tall
Echinacea pallida	Flowers June-July, 0.6-0.9 m tall
Eupatorium altissimum	Flowers August-October, 0.6-1.8 m tall
Silphium integrifolium	Flowers July-September, 0.6-0.9 m tall
Helianthus mollis	Flowers July-September, 0.6-1.2 m tall

Missouribotanicalgarden.org/PlantFinder/

Appendix B: Plant cover species

species	code
Abutilon theophrasti	abuthe
Acalypha virginiana	acavir
Allium canadense	allcan
Allium cernuum	allcer
Amaranthus rudis	amarud
Ambrosia artemisiifolia	ambart
Ambrosia trifida	ambtri
Amorpha canescens	amocan
Anagalis arvensis	anaarv
Andropogon gerardii	andger
Andropogon virginicus	andvir
Antenneria neglecta	antneg
Apocynum cannabinum	apocan
Aristida oligantha	arioli
Artemesia ludoviciana	artlud
Asclepias syriaca	ascsyr
Asclepias verticillata	ascver
Astragalus canadense	astcan
Baptisia alba	bapalb
Baptisia bracteata	bapbra
Bidens aristosa	bidari
Bouteloua curtipendula	boucur
Bouteloua gracilis	bougra
Bromus inermis	broine
Buchloe dactyloides	bucdac
Cardiospermum halicacabum	carhal
Carex spp.	carex
Cerastium fontanum	cerfon
Chaemacrista fasticulata	chafas
Chenopodium album	chealb
Chloris verticillata	chlver
Cichorium intybus	cicint
Cirsium altissima	ciralt
Commelia erecta	comere
Convolvulus arvensis	conarv
Coreopsis lanceolata	corlan
Coreopsis tinctoria	cortin
Coreposis tripteris	cortri

Cornus drummondii	cordru
Croton capitatus	crocap
Cynanchum laeve	cynlae
Cyperus esculentus	cypesc
Dalea candida	dalcan
Dalea purpureum	dalpur
Daucus carota	daucar
Desmanthus illinoensis	desill
Desmodium canadense	descan
Desmodium paniculatum	despan
Dichanthelium acuminatum	dicacu
Dichanthelium oligosanthes	dicobt
Digitaria spp.	digitaria spp.
Diospyros virginiana	diovir
Echinacea angustifolia	echang
Echinacea pallida	echpal
Echinochloa crus-galli	echcru
Eleusine indica	eleind
Elymus canadense	elycan
Eragrostis spectabalis	eraspe
Erectites hieracifolia	erehie
Erigeron annuus	eriann
Eryngium yuccifolium	eryyuc
Eupatorium altissimum	eupalt
Eupatorium seratina	eupser
Eupatorim spp	Eupatorim spp
Euphorbia corralata	eupcor
Euphorbia dentata	eupden
Euphorbia maculata	eupmac
Euphorbia marginata	eupmar
Euphorbia nutans	eupnut
Euthamia gymnospermoides	eutgym
Festuca arundicacea	fesaru
Fragaria virginiana	fravir
Fraxinus spp	fraxinus spp.
Gaura parviflora	gaupar
Geranium spp.	geranium spp.
Geum canadense	geucan
Helianthus annuss	helann

Helianthus grosseserratus	helgro
Helianthus hirsutus	helhir
Helianthus mollis	helmol
Helianthus tuberosa	heltub
Heliopsis helianthoides	helhel
Heuchera richardii	heuric
Hibiscus trionum	hibtri
Hypericum perforatum	hypper
Hypericum punctata	hyppun
Ipomea lacinatum	ipolac
Iva annua	ivaann
Juncus spp.	juncus
Koeleria macrantha	koemac
Lactuca canadensis	laccan
Lactuca serriola	lacser
Leptoloma cognatum	lepcog
Lespedeza capitata	lescap
Lespedeza cuneata	lescun
Lespedeza striata	lessti
Leucanthemum vulgare	leuvul
Liatris punctata	liapun
Liatris pycnostachya	liapyc
Liatris spicata	liaspi
Linum sulcatum	linsul
Lobelia spicata	lobspi
Lysimachina nummularia	lysnum
Melilotus officialis	meloff
Monarda fistulosa	monfis
Morus spp.	morus
Muhlenbergia frondosa	muhfro
Mullugo verticillata	mulver
Oenothera biennis	oenbie
Oenothera speciosa	oenspe
Oxalis stricta	oxastr
Panicum capillare	pancap
Panicum dichotomiflorum	pandic
Panicum virgatum	panvir
Parthenium integrifolium	parint
Parthenocissus quinquefolia	parqui
Paspalum spp.	paspalum spp.

Penstemon digitalis	pendig
Phlox pillosa	phlpil
Physalis spp.	physalis spp.
Phytolacca americana	phyame
Plantago major	plamaj
Plantago patagonica	plapat
Poa pratense	poapra
Polyganum pennsylvanica	polpen
Polygonaceae	Polyganaceae
Populus deltoides	popdel
Potentilla arguta	potarg
Potentilla simplex	potsim
Pycnanthemum tenufolium	pycten
Ratibida columnifera	ratcol
Ratibida pinnata	ratpin
Rhus glabra	rhugla
Rosa carolina	roscar
Rubus spp	rubus spp
Rudbeckia hirta	rudhir
Rudbeckia subtomentosa	rudsub
Rumex crispus	rumcri
Salvia azurea	salazu
Sambucus nigra	samnig
Schizachyrium scoparium	schsco
Securigera varia	secvar
Senna hebecarpa	senheb
Senna merilandica	senmer
Setaria faberii	setfab
Setaria glauca	setgla
Sida spinosa	sidspi
Silphium integrifolium	silint
Silphium laciniatum	sillac
Silphium terebinthinaceum	silter
Sisyrinchium angustifolium	sisang
Sisyrinchium campestre	siscam
Solanum carolinense	solcar
Solanum ptycanthum	solpty
Solidago canadense	solcan
Solidago nemoralis	solnem
Solidago rigida	solrig
Sorgastrum nutans	sornut

Sorghum halpense	sorhal
Sporabolis cryptandrus	spocry
Sporabolis heterolepis	spohet
Symphoricarpos orbiculatus	symorb
Symphyotrichum ericoides	symeri
Symphyotrichum novae-angliae	symnov
Sympyotrichum pilose	sympil
Taraxacum officialis	taroff
Teucreum canadense	teucan
Thlapsi arvense	thlarv
Toxicodendron radicans	toxrad
Tradescantia ohiensis	traohi
Tridens flavus	trifla
Trifolium pretense	tripre
Trifolium repens	trirep
Tripsicum dactyloides	tridac
Verbascum thapsus	vertha
Verbena hastata	verhas
Verbena stricta	verstr
Vernonia baldwinii	verbal
Vernonia fasticulata	verfas
Veronicastrum virginicum	vervir
Vitis vulpina	vitvul
Zizia aurea	zizaur

Appendix C: Bioinformatics Pipeline

```
# use 1 mismatch fastq files
# separate R1 into new folder
cd ~/work/BioinformaticsClass/4Afungproj/1mismatch
mkdir R1
mv *R1_001.fastq R1/

# zip files
gzip -9 R1/*R1_001.fastq

# import into qiime2 artifact
module load qiime2
qiime tools import \
--type 'SampleData[SequencesWithQuality]' \
--input-path R1 \
--input-format CasavaOneEightSingleLanePerSampleDirFmt \
--output-path 4AfungR1.qza

##### remove primers #####
qiime cutadapt trim-single \
--i-demultiplexed-sequences 4AfungR1.qza \
--p-front GTATGYYTGTATCAGTG \
--o-trimmed-sequences 4AfungR1_trimmed.qza

##### visualize feature table #####
qiime demux summarize \
--i-data 4AfungR1_trimmed.qza \
--o-visualization 4AfungR1_trimmed.qzv
# go to https://view.qiime2.org/ interactive quality plot

# decide parameters
##### dada2 script #####
#!/bin/bash
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=122gb
#SBATCH --time=00-15:00:00
#SBATCH --partition=kbs
#SBATCH --mail-type=ALL
#SBATCH --mail-user=burrill.haley@ku.edu
#SBATCH --
output=/home/h128b273/work/BioinformaticsClass/4Afungproj/dada2fungi.error.log
#SBATCH --job-name=dada2denoise_fungi
```

```
module load qiime2
cd ~/work/BioinformaticsClass/4Afungproj/1mismatch

qiime dada2 denoise-single \
--i-demultiplexed-seqs 4AfungR1_trimmed.qza \
--p-trim-left 0 \
--p-trunc-len 260 \
--o-representative-sequences rep-seqs-4AfungR1.qza \
--o-table table-dada2_4AfungR1.qza \
--o-denoising-stats stats-dada2_4AfungR1.qza

##### filter features #####
# filter OTUs that occur less than 5 times
qiime feature-table filter-features \
--i-table table-dada2_4AfungR1.qza \
--p-min-frequency 5 \
--o-filtered-table feature-frequency-filtered-table_4AfungR1.qza

# assign taxonomy using qiime
# make UNITE database
cd sh_qiime_release_04.02.2020!
qiime tools import \
--input-path sh_refs_qiime_ver8_dynamic_02.02.2019.fasta \
--output-path UNITECLUSTERDATABASE \
--type 'FeatureData[Sequence]'

qiime tools import \
--type FeatureData[Taxonomy] \
--input-path sh_taxonomy_qiime_ver8_dynamic_02.02.2019.txt \
--output-path UNITEtaxonomy2.19.qza \
--input-format HeaderlessTSVTaxonomyFormat

#!/bin/bash
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=122gb
#SBATCH --time=15-00:00:00
#SBATCH --partition=kbs
#SBATCH --mail-type=ALL
#SBATCH --mail-user=burrill.haley@ku.edu
#SBATCH --output=/home/h128b273/work/BioinformaticsClass/4Afungproj/tax.error.log
#SBATCH --job-name=tax_fung

module load qiime2
cd ~/work/BioinformaticsClass/4Afungproj/1mismatch/sh_qiime_release_04.02.2020!
```

```
qiime feature-classifier classify-consensus-vsearch \
--i-query rep-seqs-4AfungR1.qza \
--i-reference-reads UNITECLUSTERDATABASE.qza \
--i-reference-taxonomy UNITEtaxonomy2.19.qza \
--p-perc-identity 0.97 \
--o-classification taxonomy-vsearch-UNITE_fung.qza

# export taxonomy table
qiime tools export \
--input-path taxonomy-vsearch-UNITE_fung.qza \
--output-path exported-tax.table

cd exported-tax.table
use gui to export taxonomy.tsv

# prep for Funguild
##### merge in R, pairing using OTU_ID #####
# first remove line 1 in notepad++
# add _ between OTU and ID
# import table
otu <- read.table("otu_table4Afung_trans.txt", header=TRUE)
# Make row 1 colnames
rownames(otu) <- otu$OTU_ID
otu$OTU_ID <- NULL
# import taxonomy
tax <- read.table("4Afung_TAX.tsv", skip=1, col.names = c("OTU_ID", "taxonomy",
"confidence"))
head(tax)
otutax <- merge(otu, tax, by="OTU_ID", all.x=TRUE)
head(otutax)
# export merged file for funguild
write.table(otutax, file= "4Afung_otutax.txt")

# in excel remove 'confidence' column
# upload to funguild "matched and unmatched"
```

Appendix D: Statistics and Figures in R

```
setwd("C:/Users/hburrill/Desktop/KU/Bever Lab/Dimensions/4A/data")
allfung <- read.table("4Afung_allfung_matched.txt", sep="\t", header=T)
head(allfung)
rownames(allfung) <- allfung$OTU_ID
allfung$OTU_ID <- NULL
library(vegan)
#rarefy
allfung <- t(allfung)
allfung.rare <- rrarefy(x=allfung, sample=54)
#vegdist
allfungdist <- vegdist(allfung.rare, method="bray")
# creates matrix of OTUs per sample
pathfung <- read.table("otu_table4Afung-qiimefiltpath.txt", sep="\t", header=T)
rownames(pathfung) <- pathfung$OTU_ID
pathfung$OTU_ID <- NULL
pathfung <- t(pathfung)
pathfung.rare <- rrarefy(x=pathfung, sample=54)
#vegdist
pathdist <- vegdist(pathfung.rare, method="bray")
#matrix of OTUs per sample
# merge with meta
meta<-read.table("DPMeta.csv", head=T, sep=",", row.names = 1)
head(meta)
Fungdiv <- diversity(allfung)
meta$Fungdiv <- Fungdiv
Fungpathdiv <- diversity(pathfung)
meta$Fungpathdiv <- Fungpathdiv
#######################################################
# PLNT RICH LM
plntdiff <- meta$SpRichnessAvg - meta$PlntDiv
meta$weedsrich <- plntdiff
meta$plntsprich <- meta$PlntDiv + meta$weeds
#######################################################
# Fig2 - all fung
plot(x = meta$PlntDiv, y = meta$Fungdiv, xlab="Plant species richness", ylab="Fungal
diversity", col="orange", pch=16, cex.lab=1.2, ylim=c(0,6)) # main="Figure 1"
fitLM2 <- lm(meta$Fungdiv ~ (meta$PlntDiv + meta$weeds) + meta$Phylo)
abline(fitLM2, col="orange", lwd=3)
summary(fitLM2) # plnt div 0.58 # weeds 0.762 # phylo 0.775
par(new=T)
# Fig 2 - fung path
plot(x = meta$PlntDiv, y = meta$Fungpathdiv, xlab="Plant species richness", ylab="",
col="purple", pch=16, cex.lab=1.2, ylim=c(0,6))
fitLM1 <- lm(meta$Fungpathdiv ~ (meta$PlntDiv + meta$weeds) + meta$Phylo)
```

```
summary(fitLM1) # plnt div 0.0247 # weeds 0.42 # phylo 0.4118
abline(fitLM1, col="purple", lwd=3)
fitLM11 <- lm(meta$Fungpathdiv ~ meta$PlntDiv)
summary(fitLM11) # only plnt div 0.0444
legend("topright",c("All fungi","Pathogens"), col=c("orange","purple"), lwd=3, cex=0.75)
############################################################
# PHYLO DISPERSION
# permanova for all fung and plnt fam
adonis2(allfungdist ~ Block + Phylo + PlntDiv*PropPoa*PropLeg*PropComp*PropWeeds,
data=meta)
adonis2(allfungdist ~ Block + Phylo + PlntDiv*RpropPoa*RpropLeg*RpropComp, data=meta)
adonis2(pathdist ~ Block + Phylo + PlntDiv*PropPoa*PropLeg*PropComp*PropWeeds,
data=meta)
adonis2(pathdist ~ Block + Phylo + PlntDiv*RpropPoa*RpropLeg*RpropComp, data=meta)
# PERMANOVA all sp species
adonis2(allfungdist ~ Block + Phylo + PlntDiv*PropPoa*PropLeg*PropComp*PropWeeds +
SCHSCO + ANDGER + KOEMAC + ELYCAN + BOUGRA + PANVIR + AMOCAN +
DALCAN + DALPUR + DESILL + DESCAN + CHAFAS + LIAPYC + CORTIN + ECHPAL
+ EUPALT + SILINT + HELMOL, data=meta)
adonis2(allfungdist ~ Block + Phylo + PlntDiv*RpropPoa*RpropLeg*RpropComp + Rschsco +
Randger + Rkoemac + Relycan + Rbougra + Rpanvir + Ramocan + Rdalcan + Rdalpur + Rdesill
+ Rdescan + Rchafas + Rliapyc + Rcortin + Rechpal + Reupalt + Rsilint + Rhelmol, data=meta)
# PERMANOVA pathogens
adonis2(pathdist ~ Block + Phylo + PlntDiv*PropPoa*PropLeg*PropComp*PropWeeds +
SCHSCO + ANDGER + KOEMAC + ELYCAN + BOUGRA + PANVIR + AMOCAN +
DALCAN + DALPUR + DESILL + DESCAN + CHAFAS + LIAPYC + CORTIN + ECHPAL
+ EUPALT + SILINT + HELMOL, data=meta)
adonis2(pathdist ~ Block + Phylo + PlntDiv*RpropPoa*RpropLeg*RpropComp + Rschsco +
Randger + Rkoemac + Relycan + Rbougra + Rpanvir + Ramocan + Rdalcan + Rdalpur + Rdesill
+ Rdescan + Rchafas + Rliapyc + Rcortin + Rechpal + Reupalt + Rsilint + Rhelmol, data=meta)
###############################################################
# fungal family taxonomy
setwd("C:/Users/hburrill/Desktop/KU/Bever Lab/Dimensions/4A/data/Taxoutput")
pathfam <- read.table("Pathfam.csv", header=T, sep=",", row.names=1)
head(pathfam)
# subset genus taxa
setwd("C:/Users/hburrill/Desktop/KU/Bever Lab/Dimensions/4A/data/Taxoutput")
genus.taxa <- read.table("sum_gwmeta.csv", header=T, sep=",", row.names=1)
# Fig 4
# Hyponectriaceae
plot(pathfam$Hyponectriaceae ~ meta$PropLeg, ylab="Hyponectriaceae RA", xlab="Proportion
plant family", col="blue", ylim=c(0,140))
HyL <- lm(pathfam$Hyponectriaceae ~ meta$PropLeg)
summary(HyL) # 0.000183 # 14.93 # increase
abline(HyL, col="blue", lwd=2)
par(new=T)
```

```
plot(pathfam$Hyponectriaceae ~ meta$PropPoa, ylab="", xlab="", col="aquamarine",
ylim=c(0,140))
HyP <- lm(pathfam$Hyponectriaceae ~ meta$PropPoa)
summary(HyP) # 0.04675 # 4.039 # decrease
abline(HyP, col="aquamarine", lwd=2)
par(new=T)
plot(pathfam$Hyponectriaceae ~ meta$PropComp, ylab="", xlab="", col="red", ylim=c(0,140))
HyC <- lm(pathfam$Hyponectriaceae ~ meta$PropComp)
summary(HyC) # 0.0933 # 2.862 # decrease
abline(HyC, col="red", lwd=2)
par(new=T)
legend("topright",c("Fabaceae","Poaceae","Asteraceae"), col=c("blue","aquamarine","red"),
lwd=3, cex=0.75)
# Figure 5 Mycosphaerellaceae
plot(pathfam$Mycosphaerellaceae ~ meta$PropLeg, ylab="Mycosphaerellaceae RA",
xlab="Proportion plant family", col="blue", ylim=c(0,90))
MyL <- lm(pathfam$Mycosphaerellaceae ~ meta$PropLeg)
summary(MyL) # 0.007878 # 7.308 # increase
abline(MyL, col="blue", lwd=2)
par(new=T)
plot(pathfam$Mycosphaerellaceae ~ meta$PropPoa, ylab="", xlab="", col="aquamarine",
ylim=c(0,90))
MyP <- lm(pathfam$Mycosphaerellaceae ~ meta$PropPoa)
summary(MyP) # 0.0616 # 3.561 # decrease
abline(MyP, col="aquamarine", lwd=2)
legend("topright",c("Fabaceae","Poaceae"), col=c("blue","aquamarine"), lwd=3, cex=0.75)
MyAst <- lm(pathfam$Mycosphaerellaceae ~ meta$PropComp)
summary(MyAst)
# Fig 6 Pleosporaceae
plot(pathfam$Pleosporaceae ~ meta$PropLeg, ylab="Pleosporaceae RA", xlab="Proportion plant
family", col="blue", ylim=c(0,50))
PleL <- lm(pathfam$Pleosporaceae ~ meta$PropLeg)
summary(PleL) # 0.00877 # 7.105 # decrease
abline(PleL, col="blue", lwd=2)
par(new=T)
plot(pathfam$Pleosporaceae ~ meta$PropPoa, ylab="", xlab="", col="aquamarine",
ylim=c(0,50))
PleoP <- lm(pathfam$Pleosporaceae ~ meta$PropPoa)
summary(PleoP) # 0.0578 # 3.671 # increase
abline(PleoP, col="aquamarine", lwd=2)
legend("topright",c("Fabaceae","Poaceae"), col=c("blue","aquamarine"), lwd=3, cex=0.75)
PleoAst<- lm(pathfam$Pleosporaceae ~ meta$PropComp)
summary(PleoAst)
# Figure 7 Monographella
plot(genus.taxa$Monographella ~ meta$PropLeg, ylab="Monographella RA", xlab="Proportion
plant family", col="blue", ylim=c(0,100))
```

```
MonL <- lm(genus.taxa$Monographella ~ meta$PropLeg)
summary(MonL) # 0.000183 # 14.93 # increase
abline(MonL, col="blue", lwd=2)
par(new=T)
plot(genus.taxa$Monographella ~ meta$PropPoa, ylab="", xlab="", col="aquamarine",
ylim=c(0,100))
MonP <- lm(genus.taxa$Monographella ~ meta$PropPoa)
summary(MonP) #0.04675 # 4.039 # decrease
abline(MonP, col="aquamarine", lwd=2)
par(new=T)
plot(genus.taxa$Monographella ~ meta$PropComp, ylab="", xlab="", col="red", ylim=c(0,100))
MonC <- lm(genus.taxa$Monographella ~ meta$PropComp)
summary(MonC) # 0.0933 # 2.862 # decrease
abline(MonC,col="red",lwd=2)
legend("topright",c("Fabaceae","Poaceae","Asteraceae"), col=c("blue","aquamarine","red"),
lwd=3, cex=0.75)
# Fig 8 Erysiphe
plot(genus.taxa$Erysiphe ~ meta$PropLeg, ylab="Erysiphe RA", xlab="Proportion plant
family", col="blue", ylim=c(0, 10))
EryL <- lm(genus.taxa$Erysiphe ~ meta$PropLeg)
summary(EryL) # 0.00281 # 9.315 # increase
abline(EryL, col="blue", lwd=2)
par(new=T)
plot(genus.taxa$Erysiphe ~ meta$PropComp, ylab="", xlab="", col="red", ylim=c(0, 10))
EryC <- lm(genus.taxa$Erysiphe ~ meta$PropComp)
summary(EryC) # 0.09137 # 2.796 # decrease
abline(EryC, col="red", lwd=2)
legend("topright",c("Fabaceae","Asteraceae"), col=c("blue","red"), lwd=3, cex=0.75)
plot(genus.taxa$Erysiphe ~ meta$PropPoa, ylab="Erysiphe", xlab="Proportion Poaceae")
EryPo <- lm(genus.taxa$Erysiphe ~ meta$PropPoa)
summary(EryPo)
# Fig 9 Cercospora
plot(genus.taxa$Cercospora ~ meta$PropLeg, ylab="Cercospora RA", xlab="Proportion plant
family", col="blue")
CerL <- lm(genus.taxa$Cercospora ~ meta$PropLeg)
summary(CerL) # 0.00507 # 8.158 # increase
abline(CerL, col="blue", lwd=2)
par(new=T)
plot(genus.taxa$Cercospora ~ meta$PropPoa, ylab="", xlab="", col="aquamarine")
CerP <- lm(genus.taxa$Cercospora ~ meta$PropPoa)
summary(CerP) # 0.0456 # 4.081 # decrease
abline(CerP, col="aquamarine", lwd=2)
legend("topright",c("Fabaceae","Poaceae"), col=c("blue","aquamarine"), lwd=3, cex=0.75)
CerAst <- lm(genus.taxa$Cercospora ~ meta$PropComp)
summary(CerAst)
```

Appendix E: Usearch10 code

```
# download usearch10.exe to find relative abundances of fungal OTUs at each taxonomic rank
for i in p c o f g;do
>usearch10 -sintax_summary sintax-example.txt \
>-otutabin first-pathogen-example.txt \
>-rank ${i} \
>-output output/sum_${i}.txt
done
```