A suite of rare microbes interacts with a dominant, heritable, fungal endophyte to influence plant trait expression

Joshua G. Harrison[1], Lyra P. Beltran[2], C. Alex Buerkle[1], Daniel Cook[3], Dale R. Gardner[3], Thomas L. Parchman[2], Simon R. Poulson[4], Matthew L. Forister[2]

[1]Department of Botany, University of Wyoming, Laramie, WY 82071, USA

[2]Ecology, Evolution, and Conservation Biology Program, Biology Department, University of Nevada, Reno, NV 89557, USA

[3]Poisonous Plant Research Laboratory, Agricultural Research Service, United States Department of Agriculture, Logan, UT 84341, USA

[4]Department of Geological Sciences & Engineering, University of Nevada, Reno, NV 89557, USA

Corresponding author: Joshua G. Harrison
1000 E. University Ave.
Department of Botany, 3165
University of Wyoming
Laramie, WY 82071, USA
joshua.harrison@uwyo.edu

Keywords: endophytes, plant traits, *Astragalus*, locoweed, swainsonine, microbial ecology

Running title: Endophytes affect plant traits

Author contributions: JGH, LPB, and MLF conducted the field experiment. LPB and JGH performed culturing. JGH executed analyses. Analytical chemistry was conducted by DC and DRG. All authors contributed to experimental design and manuscript preparation.

Abstract

Endophytes are microbes that live, for at least a portion of their life history, within plant tissues. Endophyte assemblages are often composed of a few abundant taxa and many infrequently-observed, low-biomass taxa that are, in a word, rare. The ways in which most endophytes affect host phenotype are unknown; however, certain dominant endophytes can influence plants in ecologically meaningful ways–including by affecting growth and immune system functioning. In contrast, the effects of rare endophytes on their hosts have been unexplored, including how rare endophytes might interact with abundant endophytes to shape plant phenotype. Here, we manipulate both the suite of rare foliar endophytes (including both fungi and bacteria) and *Alternaria fulva*–a vertically-transmitted and usually abundant fungus–within the fabaceous forb *Astragalus lentiginosus*. We report that rare, low-biomass endophytes affected host size and foliar %N, but only when the heritable fungal endophyte (*A. fulva*) was not present. *A. fulva* also reduced plant size and %N, but these deleterious effects on the host could be offset by a negative association we observed between this heritable fungus and a foliar pathogen. These results demonstrate how interactions among endophytic taxa determine the net effects on host plants and suggest that the myriad rare endophytes within plant leaves may be more than a collection of uninfluential, commensal organisms, but instead have meaningful ecological roles.

Introduction

Plants are intimately associated with numerous fungi and bacteria that live within leaves, roots, stems and other tissues [1, 2]. These microbes, termed endophytes [3] are ubiquitous and occur in hosts representing all major lineages of plants [4]. Over the last twenty years, it has become clear that dominant endophytic taxa can have dramatic ecological consequences–a finding demonstrated particularly well in studies manipulating the abundance of vertically-transmitted fungi occurring within cool-season, perennial grasses [5, 6]. For example, these fungi can influence successional trajectories of vegetation [7, 8], reshape host-associated arthropod assemblages [9], and mediate host reproductive output [10]. In contrast, the ecological roles of rare endophytes–which we define as those taxa that are infrequently encountered and of low biomass–remain largely unexamined, despite that fact that these rare taxa constitute the bulk of biodiversity present within endophyte assemblages. Here, we manipulate both rare and dominant endophytes living within a perennial forb to characterize how these taxa interact and affect host phenotype.

Most endophytes are horizontally-transmitted among mature hosts via rainfall, air currents, or arthropods [11, 12] and colonize only a few cubic millimeters of host tissue [13]. Given the low biomass of these rare taxa, it is tempting to downplay their importance. However, examples from macroorganism community ecology demonstrate that certain "keystone" species, despite relatively low abundance, can exert community-wide influence [14]. For instance, beavers are uncommon mammals, yet, by reshaping fluvial geomorphology, they have profound influence on co-occurring aquatic animals, waterfowl, and riparian plants [15]. Similarly, rare endophytes could function as keystone species via several mechanisms,

including by influencing the host phenotype, catabolism of low-concentration compounds into products required by other microbial taxa, or synthesis of potent bioactive compounds [16–18].

However, the ecological influence of rare endophytes need not be the purview of just a few species. Instead, minor effects of individual taxa could accrue to the point of assemblage-wide relevance–just as numerous genetic variants, each of minimal influence, commonly underlie phenotypes [19]. For example, an individual endophytic bacterium may trigger a highly-localized immune response of negligible importance for the host and co-occurring endophytes. But the combined effects of many bacteria might initiate systemic acquired resistance within plants, with important implications for pathogen resistance and endophyte community assembly [20, 21].

Ascribing ecological influence to endophytic taxa, rare or otherwise, is complicated by a lack of understanding regarding how endophytes mediate plant trait expression [22, 23]. While the effects of certain endophytes on host growth promotion [24] and pathogen resistance [25–27] have attracted attention, few studies have examined endophyte mediation of other traits–including, for example, functional traits such as specific leaf area (e.g. [28]), phenology [29], and foliar elemental concentration [30] (for more, see reviews by [22, 23, 31]). Nevertheless, the handful of studies demonstrating plant trait mediation by endophytes are impressive. For instance, Mejía et al. 2014 [32] reported that inoculation of *Theobroma cacao* trees with the widespread, horizontally-transmitted, fungal endophyte *Colletotrichum tropicale* affected expression of hundreds of host genes, including up-regulation of some involved in the ethylene-driven immune response. These authors also found that inoculation decreased photosynthetic rate, increased leaf cellulose and lignin content, and shifted foliar isotopic ratios of nitrogen (N) and carbon (C). Similarly impressive results were reported by Dupont et al.[33] who found colonization of the grass *Lolium perenne* by the *Epichloë festucae* endophyte affected transcription of one third of host genes (for slightly more tempered, but still dramatic, results see [34]). These studies demonstrate the importance of systemic, or otherwise abundant, endophytes on their hosts, but we are unaware of any studies that manipulate the presence of low biomass, non-systemic endophytes to determine the extent to which they have similar effects on host phenotype.

Here, we perturb the microbial consortium within the fabaceaous forb *Astragalus lentiginosus* (spotted locoweed) to understand how endophytes belonging to different abundance categories affect plant trait expression. *A. lentiginosus* is a widespread, perennial forb that grows throughout the arid regions of the western United States of America [35]. *A. lentiginosus* exhibits extreme phenotypic variation and has over forty varietal designations [35], making it the most taxonomically rich plant species in North America [36]. A dominant fungal endophyte present within *A. lentiginosus* is *Alternaria fulva* (Ascomycota: Dothideomycetes: Pleosporaceae: *Alternaria* section Undifilum [37–39]). *A. fulva* is a seed-borne endophyte that grows systemically through its host and synthesizes the bioactive alkaloid swainsonine [40]. Consumption of swainsonine-laced tissues by mammalian herbivores can lead to extreme toxicosis and even death [41]. *A. fulva* is prevalent throughout the range of its host, though not all populations of *A. lentiginosus* are colonized by the fungus, and intrapopulation variation in fungal colonization has also been reported [42].

Alternaria section Undifilum fungi have been observed in numerous swainsonine-containing

taxa within *Astragalus* and *Oxytropis* that are colloquially called "locoweeds" [39, 43–45]. The nature of the relationship between locoweeds and their seed-borne fungi is somewhat unclear. Swainsonine does not seem to influence certain specialist arthropod herbivores [46, 47], which is suggestive of commensalism [48]. However, recent work supports a more mutualistic relationship between plant and fungus. For instance, Harrison et al.[42] demonstrated, via a DNA sequence-based survey, that swainsonine concentrations and *A. fulva* relative abundance were inversely related to fungal endophyte richness, potentially reducing exposure of hosts to pathogens. In a culture-based survey, Lu et al. [49] reported similar results in two other locoweed species (also see [50] for an analogous phenomenon in a grass-*Neotyphodium* endophyte system). The results from these surveys suggest that vertically-transmitted *Alternaria* endophytes can shape fungal endophyte assemblages, though effects on bacterial endophytes are unknown. Additionally, Cook et al. 2017 [51] demonstrated that *Alternaria* section Undifilum endophytes can affect the biomass and protein content of several locoweed taxa, including *A. lentiginosus*. These results suggest *A. fulva* may mediate other host traits as well.

By removing embryos from the seed coat, the abundance of *A. fulva* in plant tissues can be greatly reduced or eliminated [52]. We used this approach to manipulate the abundance of *A. fulva* in *A. lentiginosus* plants to experimentally test the aforementioned antagonistic relationship between *A. fulva* and co-occurring endophytes and explore how *A. fulva* affects various host traits, including size, leaflet area, specific leaf area, foliar C and N, phenology, and nitrogen fixation in the rhizosphere. For a subset of focal plants, we applied an inoculum slurry to leaf surfaces to boost exposure to rare, horizontally-transmitted endophytes. We applied these manipulations in a full factorial design to compare how endophytes of differing abundance categories shape host plant phenotype and explore how a dominant endophyte might influence co-occurring fungi and bacteria.

Methods

Field experiment

During the early spring of 2017, seeds of *A. lentiginosus* var. *wahweapensis* from the Henry Mts. in Utah, U.S.A. (collected in 2005 from a population known to possess *A. fulva*) were lightly scarified, left to imbibe deionized water overnight and then germinated indoors in a mix of humus, compost, and topsoil sourced from the Reno, NV region. To reduce the relative abundance of the vertically-transmitted fungal endophyte, *A. fulva* embryos were excised from a subset of seeds prior to planting, as per [52]. Notably, this treatment can eliminate the presence of *A. fulva* from plants altogether, but occasionally only reduces *A. fulva* abundance. Seedlings were grown at ambient temperature under a 16:8 (light:dark) daily lighting regime and watered with dechlorinated tap water. Individuals from different treatments were interspersed haphazardly and not allowed to touch one another. Seedlings were periodically reorganized to avoid any influence of subtly differing conditions across the growth area. To speed growth, Miracle-Gro (Scotts Miracle-Gro Company, Marysville, OH) was applied several times to all replicates during the first month of growth. To control for

possible confounding effects of seed coat removal, embryos were excised from a subset of seeds and planted along with potato dextrose agar (PDA) that was sterile, or that was colonized by *A. fulva*, which had been cultured from intact seeds. These control seedlings were planted several weeks later than other seedlings, due to slow growth of *A. fulva* cultures.

In early June, seedlings were installed in five gallon pots filled with equal parts locally sourced humus and topsoil and placed in an abandoned, largely denuded field near the University of Nevada, Reno. No *Astragalus* taxa were observed growing within this field and the plants that were present in the field were mostly non-native forbs. *A. lentiginosus* grows in a wide variety of settings, including roadsides and disturbed areas not unlike our site. A total of 300 plants were installed (between 54–68 per treatment group, see Table 1). Pots were organized randomly with respect to treatment and were placed one meter apart so plants never touched one another. Dechlorinated water was applied as needed to all plants at the same time (typically every other day, except during the heat of summer when watering was conducted daily). Every two weeks a slurry of microbial inocula (described below) was sprayed on leaves of half of the plants. A solution with identical surfactant, but no microbial inoculum, was applied to untreated plants. Plants were left in the field from early June through mid-September, at which point leaves were removed for sequencing and culturing.

Inoculum preparation

Twenty morphologically unique, reproductive fungal isolates were obtained from the following woody shrubs growing near Reno, NV: *Artemisia tridentata*, *Ericameria nauseosa*, *Prunus andersonii*, and *Tetradymia canescens*. These shrubs are abundant throughout the Great Basin Desert and, consequently, we reasoned they contained horizontally-transmitted, foliar microbes likely to be regularly encountered by *A. lentiginosus*. Indeed, sequencing revealed that most of the fungal taxa within the inoculum were observed in the wild-collected *A. lentiginosus* individuals examined in Harrison *et al.* [42] (for details see the Supplemental Material). Individual shrubs to be sampled were selected haphazardly. We did not use leaves from *Astragalus* species to avoid inoculating plants with *A. fulva* and thus obviating our treatment to reduce this fungus. Leaves were cut into sections (of several mm^2) and placed on PDA and the resulting microbial growth isolated and subcultured over two weeks. Spores from isolates were removed and suspended in deionized water and 0.0001% TWEEN 20 (Sigma-Aldrich), a detergent that functioned as a surfactant. A haemocytometer was used to dilute the suspension to ~100,000 spores ml^{-1}. This concentration was chosen because it produced no obvious negative symptoms in *A. lentiginosus* seedlings during preliminary experiments. Aseptic technique was used throughout culturing and inoculum preparation. Two aliqouts of inoculum were sequenced to identify the constituent microbial taxa. We specifically targeted fungi for culturing, but sequencing revealed that bacteria were also present within the inoculum.

Plant trait measurement

All plant traits were measured concomitant with sample collection for foliar microbiome characterization. Plant size was measured as the minimum size box that would enclose the

plant. This was calculated as the product of the width of the plant at its widest point, the width of the plant perpendicular to that point, and plant height. Phenological state and number of leaves were characterized for each plant. Area and specific leaf area (SLA; leaflet area divided by mass of leaflet) were measured for three dried leaflets per plant and averaged. Two healthy leaflets were removed from 8–12 leaves per plant, rinsed with tap water, dried in a laminar flow hood (<12 h total) and frozen until further processing. These leaflets were then parsed for microbiome characterization, swainsonine quantification, and carbon (C) and nitrogen (N) analysis. Swainsonine concentration in ~50 mg of dried, ground foliar tissue was measured using an LC-MS/MS approach described in Gardner *et al.* [53]. Briefly, an 18 h extraction in 2% acetic acid with agitation was followed by centrifugation. Supernatant was added to 20 mM ammonium acetate and subjected to LC-MS/MS analysis. Percent C and N and ^{15}N:^{14}N isotopic ratios in 3–4 mg dried foliar tissue, were measured by the Nevada Stable Isotope Laboratory using a Micromass Isoprime stable isotope ratio mass spectrometer (Elementar, Stockport, UK) and a Eurovector elemental analyzer (Eurovector, Pavia, Italy). The percentage of nitrogen in tissues due to fixation alone (NDFA) was calculated as per Högberg [54] through comparison with samples from co-occurring *Chenopodium album*, which is not known to harbor nitrogen fixing rhizosphere bacteria.

Sequence and culture-based characterization of the foliar microbiome

We characterized endophytic assemblages through both culturing and DNA sequencing, thus affording us insight into the effects of treatment on microbial assemblages via two complimentary measurement techniques. For our culture-based assay, we choose three leaflets per plant. Leaflets were surface sterilized, cut into 3–4 pieces, and plated onto PDA using aseptic technique. Surface sterilization involved rinsing in 95% ethanol for 30 s, followed by 2 min in 10% sodium hypochlorite solution, 2 min in 70% in ethanol, and a final rinse with deionized water. Preliminary experiments confirmed the success of this surface sterilization technique. Cultures were grown in the dark at ambient temperatures for 1.5 months. Microbial growth (either fungal or bacterial) was isolated, subcultured, and the number of morphologically unique cultures and the percentage of leaf pieces colonized recorded. Cultures corresponding to *A. fulva* were identified visually through comparison to *A. fulva* cultures grown from seeds used for this experiment and through sequencing (we did not sequence the other cultures for logistical reasons).

DNA was extracted from three surface-sterilized, dried, and ground leaflets per plant using DNeasy plant mini kits (Qiagen, Hilden, Germany). Extraction blanks for each kit were used as negative controls. Dual-indexed libraries were made at the University of Wyoming and were sequenced on the Illumina NovaSeq platform (paired-end 2x250; San Diego, CA, U.S.A.) by psomagen, Inc. (Rockville, Maryland, USA). To characterize bacterial assemblages, the 16S (V4) locus was amplified using the 515-806 primer pair [55], while for fungal assemblages the ITS1 locus was amplified using the ITS1f-ITS2 primer pair [56]. A synthetic DNA internal standard (ISD) was added to template DNA prior to library creation [57]. Additionally, unique synthetic DNAs of our own design were added to each sample to allow cross-contamination to be detected (*sensu* Tourlousse *et al.* [58]; we refer to these oligos as "coligos", which is short for cross-contamination checking oligos). A mock community consisting of eight bacteria and two fungi was also sequenced (Zymo Research, Irvine, CA. U.S.A.) as a positive control. PCR

was performed in duplicate and unique index sequences were ligated onto each PCR replicate, thus allowing us to determine technical variation due to PCR. For full library preparation details see the Supplemental Material.

Sequence data were demultiplexed using a custom perl script that used Levenshtein distances to correct errors in index sequences and assign reads to samples. Primers and Illumina adaptors were removed using cutadapt v1.13-py27 (Martin 2011) and poly-G tails removed using fastp v0.21.0 (poly-G tails occurred due to a lack of signal for very short-template molecules, such as our coligos [see above]; Chen et al. 2018). Paired-end reads were merged using vsearch v2.9.0 [59, 60] with staggering allowed, a minimum of ten overlapping bases, and a maximum of twelve mismatches in the overlapping region. The probability of base-calling errors within merged reads was estimated and those reads expected to include more than one error were discarded.

Using vsearch, unique reads were identified and clustered into OTUs (operational taxonomic units) using the 'cluster unoise' algorithm [61] with a minimum of twelve sequences required, each with a minimum length of 56 nt. OTUs clustered via this algorithm may differ by as little as a single nucleotide and are sometimes referred to as exact sequence variants, or ESVs [62]. There is an ongoing dialogue regarding the use of ESVs for fungi; for a justification of our approach see the Supplemental Material.

Chimeras were detected using the 'uchime3 denovo' algorithm and removed. For ITS sequences, those reads that did not merge were concatenated and processed separately, but identically, as those that did merge. OTUs made from shorter, merged reads were aligned to those made using the longer, concatenated reads and any short OTUs that aligned were not considered. We chose this approach for ITS reads because in preliminary work we discovered that, for some taxa, the ITS1 region was too lengthy to allow paired-end sequences to merge. After removing chimeras, OTUs, both long and short, were combined and an OTU table was made via aligning both merged and concatenated, unmerged reads to OTUs using the 'usearch global' algorithm.

Taxonomic hypotheses for OTUs were generated using the SINTAX algorithm [63] and the UNITE (v7.2; [64]) and Ribosomal Database Project database (RDP; v16; [65]) for fungi and bacteria, respectively. Read counts for OTUs corresponding to the ISD were summed for each replicate to allow normalization by the ISD. For 16S data, host plastid DNA was identified through comparison to all fabaceous chloroplasts available from the NCBI nucleotide database (accessed Feb 9, 2016; Harrison *et al.* [42]). For ITS data, host DNA was identified through matching known *A. lentiginosus* and *Oxytropis* sp. (sister taxon to *Astragalus*) ITS sequences to OTUs. For both 16S and ITS datasets, reads for all plant OTUs were summed. Additionally, for the 16S data, OTUs for mitochondrial DNA were identified using the MIDORI database [66] and removed from the OTU table. For both 16S and ITS data, those few OTUs that were not classified to any taxon were removed from the data. If >5% of the total reads for an OTU were in negative controls, then these OTUs were deemed possible contaminants and discarded (30 fungal OTUs and 2 bacterial OTUs were discarded). OTUs corresponding to *A. fulva* were identified through comparison to GenBank accession JX827264.1 and those corresponding to *Leveillula taurica* with accession MT472005.1.

Statistical analysis

We analyzed sequence count data via a hierarchical Bayesian modeling (HBM) framework that provides estimates of proportional relative abundance for each microbial taxon [67, 68]. The model estimates parameters of replicate-specific, multinomial distributions that describe taxon proportions (p parameters) and Dirichlet parameters that describe proportion estimates for the entire sampling group. This method shares information among replicates for more accurate parameter estimation and allows propagation of uncertainty in parameter estimates to downstream analyses (for a full description, see [68]). Rarefaction is not needed when using this modeling approach, because proportion estimates are used for downstream analyses and because estimates for each replicate are informed by data from all other replicates due to the hierarchical nature of the model. Modeling was conducted in the R computing environment (R Core Team 2019) using the CNVRG v0.2 R package[70]. CNVRG is a user-friendly wrapper for implementing Dirichlet-multinomial modeling with rStan [71], which itself is an interface to the Stan model specification software [72]. We used the Hamiltonian Monte Carlo sampling algorithm to characterize posterior probability distributions. We took 1,000 samples from posteriors, with a thinning rate of two, after a burn-in of 500 samples. Model convergence was confirmed via the Gelman-Rubin statistic (all parameters had a statistic very near one; [73]). To account for compositionality, we divided the proportion estimate for each taxon in a replicate by the proportion of reads assigned to the ISD for that replicate. By placing the relative abundances of all taxa on the scale of the ISD, we were able to better compare taxon abundances among treatment groups (for more details of the problem of compositionality and how an ISD can help see: [57, 74]).

To measure the extent to which OTUs differed in relative abundance among treatment groups, posterior probability distributions (PPDs) for Dirichlet parameters for each OTU and treatment group were subtracted. This generated a PPD of the difference in that parameter between any two treatment groups. If 5% or less of the density of that PPD was on either side of zero, then we deemed a treatment-associated shift in microbial relative abundance was credible. Means of PPDs for parameters of interest were used as point estimates for those parameters.

Species equivalents of Shannon's entropy and Simpson's diversity [75] were calculated using CNVRG for the treatment group as a whole and for each replicate. To estimate diversity equivalencies for a treatment group, the equivalency was calculated for *each* sample of the Dirichlet distribution characterizing microbial relative abundances within that treatment group. This generated a PPD of diversity, thus propagating uncertainty in relative abundance estimates into estimates of treatment group diversity (for a similar approach see [76]). To determine how diversity equivalents differed between treatment groups, the overlap of PPDs for each group was examined (as per above). Diversity equivalents were also estimated for each replicate so that these estimates could function as the response in a linear model testing for associations between plant trait variation and shifts in microbial diversity (see below). To estimate diversity for each replicate, the means of PPDs of multinomial parameters for that replicate were calculated (recall that these parameters estimated proportional microbial relative abundance) and diversity equivalencies were calculated for the resulting vector.

Hierarchical Bayesian modeling (HBM) was also used to estimate differences among

treatment groups in the mean values of plant traits, sequence-based estimates of microbial diversity, and culture richness. Each response variable was modeled as a draw from a normal distribution characteristic of the sampling group as per [77]. The mean (μ) and variance (σ^2) of this distribution was estimated through sharing of information among replicates. The prior distribution for μ was a normal distribution centered at zero with a precision of 0.0001 (variance = 10000). The prior distribution for σ^2 was a uniform distribution from 0–100 (for full model specification see provided R code). MCMC sampling and tests for credible effects of treatment via PPD overlap were conducted as described above. For these analyses, we used the JAGS model specification language [78] as implemented via rjags v4-6 [79]

To evaluate associations between plant traits and microbial diversity, linear models were created in a HBM framework. Beta coefficients for plant traits were estimated for each treatment group, with a prior sampled from a normal distribution centered at the estimated across-treatment effect of each trait and a precision estimated across all treatments. Hyperpriors for beta coefficients were normal distributions centered at zero with a precision of 0.0001 (for full model specification see R code provided; also see [42]). Means of PPDs for each beta coefficient were used as point estimates of the effect of that covariate. The proportion of the PPD for each beta coefficient that did not overlap zero was used to determine certainty of a non-zero effect. Prior to modeling, missing values in covariates were imputed using the random forest algorithm [80] as implemented by the randomForest R package [81]. When models were run without imputing data, results were similar to those reported here. To determine effects of treatment on microbial assemblages, as a whole, principal coordinates analysis (PCoA) and PERMANOVA were conducted on Bray-Curtis transformed tables of point estimates of proportions (derived as described above) for microbial taxa.

We chose not to report effects of treatment on endophyte richness using sequence data. To explain, when a dominant taxon is present (such as *A. fulva* or *L. taurica*) within a replicate it captures much of the sequencer's bandwidth for that replicate. Therefore, that replicate would have fewer reads available to allocate to the other taxa present, which would result in spuriously lower richness. Consequently, to assay effects of treatment on richness, we relied on culturing data.

We omitted from all analyses ten plants for which seed coat removal did not reduce *A. fulva* as ascertained via swainsonine concentration (this compound is not known to be produced by the host plant), culturing, or molecular data. We also omitted plants with no evidence of *A. fulva* occurrence from the *A. fulva* positive treatment group, because *A. fulva* is known to be incompletely transmitted between generations [82]. Analyses were repeated while retaining all of these plants in their original treatment groups and the results obtained were qualitatively similar to those presented here.

Results

Sequencing summary and microbial diversity description

After removing host, ISD, coligo (oligos for accounting of cross-contamination), and contaminant reads and applying our stringent quality control approach, we retained for analysis

7,417,832 reads from 2,292 fungal OTUs and 76,900 reads from 657 bacterial OTUs (from over 20 million 16S reads, most of which were from host organelles; for full details see the Supplemental Material).

The majority of fungal OTUs (which were specified at single nucleotide variation resolution) belonged to the Ascomycota (86%). Many OTUs were assigned to *Leveillula taurica* (991 OTUs) and *A. fulva* (101 OTUs). A variety of bacterial taxa were observed, including many members of Acidobacteria, Actinobacteria, Proteobacteria, Firmicutes, and Chloroflexi. The most abundant bacterial OTUs were assigned to the Lactobacillaceae, Bacillaceae, Listeriaceae, and Staphylococcaceae.

Effects of the vertically-transmitted fungus on the host and co-occurring microbes

Treatment to reduce the relative abundance of the dominant, heritable fungus *A. fulva* from *A. lentiginosus* plants was successful as evidenced by read counts (Fig. S1, S1), swainsonine concentrations (Fig. 1), and culturing (Fig. 2). *A. fulva* presence influenced plant phenotype–colonized plants were much smaller and had fewer leaves than uncolonized plants (Figs. 1; S3). The negative effect of *A. fulva* on plant size was observed in the second year of monitoring as well (Fig. S4). Foliar N was affected by both *A. fulva* and rare microbes–plants without *A. fulva* and that were untreated with inoculum had elevated %N in their leaves. Moreover, *A. fulva* generally increased the $^{15}\delta$ (ratio of N isotopes) and reduced NDFA, a proxy for rhizosphere nitrogen fixation activity, though these effects were less pronounced than some other effects on host phenotype (Table 2). The effects of *A. fulva* colonization were generally similar for plants grown from embryos planted alongside *A. fulva* infected agar (see Methods, for details of this control treatment), thus the observed results are not due to the confounding influence of seed coat removal (Fig. S3). We did not observe an effect of *A. fulva* colonization on %C, SLA, or phenology (using Fisher's exact test to examine flowering status at time of harvest).

A. fulva presence modestly affected diversity of co-occurring bacterial and fungal endophytes (Fig. 3, Table S3). Species equivalents of Shannon's entropy for both bacteria and fungi increased in plants colonized by *A. fulva* (Figs. 3), but the opposite was true for equivalents of Simpson's diversity. Simpson's diversity index places more weight on abundant taxa than does the Shannon index [83].

We also observed a negative association between *A. fulva* and *Leveillula taurica*. *L. taurica* is a powdery mildew (Erysiphaceae) known to colonize numerous plant species, including *A. lentiginosus* [42]. *L. taurica* was the most abundant fungus sequenced and dropped in relative abundance when *A. fulva* was present (Fig. 2, Table S1). This negative association was also observed visually, as we noted a powdery mildew infection on the leaves of a subset of the plants used for this experiment, and infections were less severe in plants colonized by *A. fulva* (Fig. S5).

For less abundant fungi (those less than 0.04% of total reads), which accounted for all fungal OTUs aside from *A. fulva* and *L. taurica*, we observed a modest increase in absolute abundance when *A. fulva* was present (Fig. 4), which was likely responsible for the increase in Shannon's diversity with *A. fulva* infection.

The inoculum slurry applied to plants was created from twenty morphologically distinct cultures derived from shrubs common in the Great Basin Desert. Sequencing revealed the slurry was composed of 20 fungal and 260 bacterial OTUs, including many taxa that have been observed in wild-collected plants (for a full description, see the Supplemental Material). Inoculation was successful, as shown by the effects of treatment discussed elsewhere. Additional evidence for inoculation success was that 55 of the bacterial OTUs and 3 of the fungal OTUs were sequenced from plants used in this experiment and in almost all cases, inoculum application led to an increase in the read counts obtained for those taxa (in 87% of bacterial taxa and two out of the three fungal taxa). Fungi present in inoculum that were also observed in treated plants included *Preussia* sp. (the most abundant fungus in the inoculum), *L. taurica*, and *Penicillium* sp., while successful bacterial colonizers were predominantly members of Proteobacteria, Actinobacteria, and Firmicutes. Inoculum application had modest effects on overall Shannon's and Simpson's diversity (Fig. 3). Inoculum application did not lead to significantly different centroids in PCoA ordinations (Figs. S1, S2).

Inoculum application had no visibly pathogenic effects on plants–they appeared healthy and leaves had no evidence of necrosis. However, inoculum application did influence plant phenotype, but only when the dominant fungus *A. fulva* was not present. For instance, inoculum application reduced leaf count (by approximately 50%) and foliar %N (Table 2), but this was only apparent for plants without *A. fulva* (Fig. 1). Inoculum treatment had minor, idiosyncratic effects on trait variation, sometimes reducing and other times increasing variation (Table S2). For %N and plant size, the directionality of the effect of inoculum on trait variation depended on *A. fulva* treatment, with inoculation increasing trait variation when *A. fulva* was reduced, but decreasing it otherwise. In general, associations between the diversity entropies of horizontally-transmitted endophytes, either bacterial or fungal, with plant trait variation were weak and often limited to a specific treatment group (Tables S4, S5).

Discussion

Foliar endophyte assemblages are typically composed of a few dominant taxa and numerous taxa of low relative abundances that occupy a small proportion of their host's tissues (we refer to these as rare taxa; [13, 84]). Ecological relevance is often considered the domain of abundant microbial taxa, because of their greater biomass and prevalence. However, our results suggest that characterizing the overall effects of endophyte assemblages may require study of rare taxa. Indeed, we report that a suite of rare endophytes affected host size and foliar N content, among other traits. We also observed that the influence of these taxa was attenuated by the presence of a dominant fungal endophyte, which itself mediated host plant phenotype.

It is important to note that, regardless of treatment, host plants appeared healthy to the eye–their tissues were green, and, in many cases, they fruited successfully during the first and second years of growth. Thus, both *A. fulva* and co-occurring microbes meet the criterion of

living asymptomatically within plant tissues necessary for designation as endophytic taxa rather than obligate pathogens [3].

Aside from *A. fulva* and *L. taurica*, all other microbial taxa observed, be they fungal or bacterial, were rare, as shown through read counts (Fig. 4) and culture-determined infection rate (Fig. 2b). Microbial taxa within the inoculum mixture were similarly very low-abundance in plants (Fig. 2). Indeed, the abundance of all these taxa combined was generally about that of the internal standard, which was spiked into samples with the minuscule concentration of $0.03\,\mathrm{pg}\,\mu\mathrm{L}^{-1}$. Thus, it was surprising that we saw fairly dramatic effects of treatment with the inoculum.

We were not able to attribute the effects of the inoculum to specific taxa, but this was by design. We were curious what the overall effects of a complex endophyte assemblage might be for plants grown in nature—thus, we were attempting to provide a different perspective than that offered by gnotobiotic studies in laboratory settings. Many of the taxa in our inoculum mixture have been observed in wild plants, with low relative abundances [42], including possible pathogens such as *Preussia*. Thus, it seems very plausible that the effects of the inoculum mixture that we observed could occur in wild populations of *A. lentiginosus*.

It seems unlikely that unintended consequences of treatment induced the results we observed since we applied a mock treatment to account for the effects of leaf wetting and surfactant application and we controlled for the effect of embryo excision from seed coats. However, we considered two caveats to our results. First, it is possible that our inoculum mixture was biased towards endophytes that grow rapidly in culture. If that was the case, then perhaps an inoculum mixture containing slower-growing taxa, which plausibly could include more biotrophic taxa, would have different effects on the host than those we observed. Second, our plants were reared with regular watering. This was by necessity because the plants were grown in pots as opposed to planted directly in the soil. Pots provided many logistical advantages, but their use led to a heating of the soil and a rapid draining of water that would have killed many plants if they did not receive water that was supplemental to precipitation. It is unclear what effects watering might have had on our results, though we note that we were careful to avoid wetting leaves during watering as leaf wetness can influence microbial colonization.

Ecological roles of rare endophytes

The effects of rare endophytes that we observed likely have implications for host fitness. Certainly the approximately 50% reduction in plant size and leaf number that we report could lead to reduced seed output (Table 2) and it seems likely that the effects on foliar N and leaf morphology we observed could also influence host fitness. Beyond the direct influence of trait variation to plant persistence and reproductive output, the marked shifts in phenotype that we observed could have indirect effects on host-associated organisms, such as arthropods, which also affect host fitness. For instance, size and foliar %N are often strong predictors of variation in insect assemblages and herbivory across plant species [85, 86], thus shifts in these traits induced by low-biomass microbial taxa could have cascading effects on arthropod communities.

We considered two possibilities for the distribution of ecological influence among rare endophytic taxa. Specifically, influence could be limited to several keystone taxa or could be cumulative, such that a quorom must be reached before the combined effect of rare endophytes induces a response by the host. We were unable to satisfactorily resolve these two non-mutually exclusive hypotheses. However, the quorom hypothesis might lead to a negative association between microbial diversity and plant size, because higher microbial diversity would occur when taxa were more evenly distributed, each with a role to play. We did not unequivocally observe such a negative association (Tables S4, S5). Indeed, we found that most rare endophytes occurred infrequently in samples, suggesting that if a quorom was present and responsible for the shifts in host phenotype observed, then that quorom must be easily met and be composed of very little total biomass that does not manifest in notable shifts in diversity.

Alternatively, infrequently-observed, keystone taxa could have caused the treatment effects we report, as these taxa, by definition, exert greater influence than would be predicted from their low biomass. For instance, a localized infection by a keystone taxon could have effects that spread throughout the host (e.g., through hormone stimulation), yet that taxon would not be present in the majority of leaves sequenced from that host. This concept suggests limitations of the common practice of *in silico* identification of keystone taxa as those taxa that are prevalent among samples, such that their removal from co-occurrence networks causes a shift in network topology [87–89]. We reiterate the non-exclusivity of the keystone and quorom hypotheses and suggest disentangling the two represents a profitable line of inquiry for future work.

Effect of Alternaria fulva, the heritable fungus

In addition to the influence of rare endophytes, we observed that the dominant, vertically-transmitted fungal endophyte *A. fulva* also reduced host size and foliar %N (Fig. 1; consistent with Cook et al. 2017[51]). Inhibition of host growth by *A. fulva* is perplexing, because the fungus is vertically-transmitted in seeds and, therefore, its fitness is tied to that of its host. Larger *A. lentiginosus* plants generally produce more seeds (J. Harrison, personal observation) and, thus, selection should operate against mechanisms by which *A. fulva* reduces host growth. On the other hand, *A. fulva* grows very slowly in culture [90, 91] and fast-growing plants could possibly outpace hyphal growth. If the fungus cannot grow fast enough to reach seeds before they mature, then its direct fitness is zero. Consequently, constraining host growth may improve fungal fitness, because it would allow time for hyphae to reach reproductive structures. This hypothesis awaits further testing.

Another intriguing possibility is that a fungal-induced reduction in plant size could actually improve plant longevity in extreme conditions and thereby lead to a positive, time-averaged, effect on fungal fitness. For several native plants in the Great Basin, small stature paradoxically facilities the ability to withstand drought and competition from invasive annual grasses [92, 93]. Thus, it is possible that plants colonized by *Alternaria* endophytes could better survive the harsh desert climate, providing both the plant and the fungus more opportunity to reproduce. Interestingly, previous work has shown that *Alternaria* endophytes do not reduce plant size in locoweeds that are drought-stressed [94] and that swainsonine

concentration can increase during drought stress. Thus, perhaps the negative affect of *A. fulva* on host size we observed here, in a well-watered, controlled setting, would not play out in drought-stressed wild populations.

The potential fitness costs imposed by *A. fulva* on its host may be ameliorated by the negative association we observed between *A. fulva* and the most abundant pathogen present, *Leveillula taurica* (Figs. 2, S5). These results support the hypothesis posed by Lu et al.[49] that the *Alternaria* spp. occurring within *Astragalus* and *Oxytropis* act as mutualists to their hosts by restricting pathogen exposure. *A. lentiginosus* is a plant of frequently disturbed, climatically variable, arid landscapes and it is likely that pathogen pressure in such locales is particularly damaging, because the lack of resources could impede recovery from tissue loss. The same rationale has inspired the growth-rate hypothesis in the literature characterizing interactions between plants and insect herbivores [95]. This hypothesis predicts plants growing in resource poor conditions will recuperate from herbivory slowly, and thus benefit from investment in phytochemical defenses that would otherwise be too costly. Similarly, tolls imposed by *A. fulva* on *A. lentiginosus* may be acceptable to the host given the harshness of the arid American West.

Our results compliment recent research presented in Christian et al. [30] showing that endophytes, and interactions between endophytes and pathogens, can alter N distribution and uptake in plants. The study did not demonstrate endophyte-induced shifts in %N content at the whole plant level (such as those we observed here), but it did show that endophytes increased ^{15}N in plants and affected ^{15}N distribution among leaves (also see [32]). We also observed credible, treatment-induced shifts in δ^{15}N and nitrogen fixation in the rhizosphere (NDFA) (Fig. 1; Table 2), but we also found that rare endophytes reduced total foliar %N when *A. fulva* was not present to attenuate their effects. When taken together with previous work [30, 32], our results suggest that the effects of horizontally-transmitted endophytes on foliar N depend on host taxon and individual, abiotic context (e.g. N availability), and interactions with other microbiota, and, when these factors align, the effect of endophytes on N allocation within hosts can be noteworthy. Notably, we saw some subtle trends that bear further examination; specifically, reducing *A. fulva* colonization was associated with higher NDFA. Thus, the work we present here, when coupled with the results in Christian et al.[30] suggest that foliar endophytes may affect N fixation, which typically happens below-ground. This suggests that interactions among microbes in the phyllosphere can influence what happens in the rhizosphere.

Other considerations: a milieu of interactions and experimental design

Our study demonstrates the ecological consequences of interactions among microbes [18], as evidenced by a negligible effect of inoculum application for plants colonized by *A. fulva* and the negative association between *A. fulva* and *L. taurica*. We suggest that these results are not likely due to direct competition for resources between *A. fulva* and other microbes–the disparity in leaf size and microbe size is too great and there seems to be enough healthy leaf tissue for all parties (based on the lack of visual infection symptoms in our plants). Even for *A. fulva*, which grows systemically through its host, physical encounters with co-occurring microbes are probably rare–with the possible exception of encounters with *L. taurica*. We

note that it is plausible that swainsonine could inhibit growth of competing microbes, as the molecule inhibits the action of alpha-mannosidase [40], which is used by fungi to process oligosaccharides [96]. However, it is unknown if swainsonine is exuded into plant tissues or is instead retained within fungal cells, if the latter, then it is not clear how the compound would affect co-occurring microbes. We suggest that indirect mediation of microbe-microbe interactions by the host is more likely. Indeed, gene expression studies in several perennial grasses [33, 34, 97] and in *Theobroma cacao* [32] have demonstrated an upregulation in the host immune response after colonization by endophytes (see [98]). To speculate, it is possible that *A. fulva* similarly primes the host immune response, which could negatively affect co-occurring microbes.

To account for any adverse effects of seed coat removal, we planted control seeds alongside sterile agar or agar inoculated with *A. fulva*. This technique was successful as shown by culturing results (Fig. 2), swainsonine concentrations (Fig. S3), and sequencing output (Fig. 2). The results we observed from control plants were very similar to those from treated plants, except for foliar C and N concentrations, which were more variable among controls. Most manipulative studies of vertically-transmitted fungal endophytes attempt to kill fungi within seeds through either heat treatment (e.g. [99]), long-term storage (e.g. [7]), or fungicide application (e.g. [100]). While studies manipulating endophytes via these techniques have been of critical importance, it is possible that these treatments could have undesirable consequences that are hard to control for and that could obscure effects of endophyte reduction. Consequently, we suggest others consider the approach we used here when seeds from endophyte-free plants are not available.

Conclusion

Our results suggest that rare, low-biomass endophytic taxa can have marked influence on their hosts and that these effects may be mediated by co-occurring dominant microbial taxa. It remains to be seen how often rare endophytic taxa affect host phenotypes in other systems. However, given that every study of endophytic biodiversity with which we are familiar shows a steep rank-abundance curve, with many taxa of low relative abundances, it seems plausible that the cumulative role of rare microbes across hosts could be substantial.

We also hope that a more careful consideration of different taxa (common and rare) within endophyte assemblages will illuminate parallels with diversity-ecosystem function studies of macroscopic organisms where community-wide effects of rare taxa have been demonstrated [101]. As biodiversity declines, such connections across scales of observation could provide impetus for conserving rare taxa, large and small, as important contributors to ecosystem processes.

Acknowledgments and conflict of interest statement

We received three very insightful reviews that improved this manuscript substantially. Our thanks go to these reviewers. We declare no conflicts of interest. Funding was provided by a National Science Foundation (NSF) Doctoral Dissertation Grant awarded to JGH. JGH and CAB were supported by the Microbial Ecology Collaborative with funding from

NSF award #EPS-1655726. Computing was performed in the Teton Computing Environment at the Advanced Research Computing Center, University of Wyoming, Laramie (https://doi.org/10.15786/M2FY47).

Data availability

All scripts, plant trait data, and processed sequence data are available at: `https://github.com/JHarrisonEcoEvo/RareMicrobes`

Raw data are hosted by the University of Wyoming. MiSeq data used during preliminary work for this experiment can be found at: `https://doi.org/10.15786/r9xy-6x03` while NovaSeq data analyzed herein can be found at: `https://hdl.handle.net/20.500.11919/7166`

References

1. Ryan, R. P., Germaine, K., Franks, A., Ryan, D. J. & Dowling, D. N. Bacterial endophytes: recent developments and applications. *FEMS Microbiology Letters* **278,** 1–9 (2008).
2. Rodriguez, R., White Jr, J., Arnold, A. & Redman, R. Fungal endophytes: diversity and functional roles. *New Phytologist* **182,** 314–330 (2009).
3. Wilson, D. Endophyte: the evolution of a term, and clarification of its use and definition. *Oikos,* 274–276 (1995).
4. Harrison, J. G. & Griffin, E. A. The diversity and distribution of endophytes across biomes, plant phylogeny and host tissues: how far have we come and where do we go from here? *Environmental Microbiology* **22,** 2107–2123 (2020).
5. Clay, K. & Schardl, C. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *The American Naturalist* **160,** S99–S127 (2002).
6. Rudgers, J. A. *et al.* A fungus among us: broad patterns of endophyte distribution in the grasses. *Ecology* **90,** 1531–1539 (2009).
7. Clay, K. & Holah, J. Fungal endophyte symbiosis and plant diversity in successional fields. *Science* **285,** 1742–1744 (1999).
8. Afkhami, M. E. & Strauss, S. Y. Native fungal endophytes suppress an exotic dominant and increase plant diversity over small and large spatial scales. *Ecology* **97,** 1159–1169 (2016).
9. Rudgers, J. A. & Clay, K. An invasive plant–fungal mutualism reduces arthropod diversity. *Ecology Letters* **11,** 831–840 (2008).
10. Gorischek, A. M., Afkhami, M. E., Seifert, E. K. & Rudgers, J. A. Fungal symbionts as manipulators of plant reproductive biology. *The American Naturalist* **181,** 562–570 (2013).

11. Malloch, D. & Blackwell, M. *The fungal community: its organization and role in the ecosystem* Second Edition, 147–171 (Marcel Dekker, Inc, New York, NY, 1992).

12. Devarajan, P. & Suryanarayanan, T. Evidence for the role of phytophagous insects in dispersal of non-grass fungal endophytes. *Fungal Diversity* **23,** 111–119 (2006).

13. Lodge, D. J., Fisher, P. & Sutton, B. Endophytic fungi of *Manilkara bidentata* leaves in Puerto Rico. *Mycologia,* 733–738 (1996).

14. Paine, R. T. A note on trophic complexity and community stability. *The American Naturalist* **103,** 91–93 (1969).

15. Jenkins, S. H. & Busher, P. E. *Castor canadensis. Mammalian Species,* 1–8 (1979).

16. Hajishengallis, G., Darveau, R. P. & Curtis, M. A. The keystone-pathogen hypothesis. *Nature Reviews Microbiology* **10,** 717–725 (2012).

17. Jousset, A. *et al.* Where less may be more: how the rare biosphere pulls ecosystems strings. *The ISME Journal* **11,** 853–862 (2017).

18. Hassani, M. A., Durán, P. & Hacquard, S. Microbial interactions within the plant holobiont. *Microbiome* **6,** 58 (2018).

19. Rockman, M. V. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66,** 1–17 (2012).

20. Beckers, G. J. & Conrath, U. Priming for stress resistance: from the lab to the field. *Current Opinion in Plant Biology. Special Issue on Biotic Interactions* **10,** 425–431 (2007).

21. Hartmann, A., Rothballer, M., Hense, B. A. & Schröder, P. Bacterial quorum sensing compounds are important modulators of microbe-plant interactions. *Frontiers in Plant Science* **5** (2014).

22. Friesen, M. L. *et al.* Microbially mediated plant functional traits. *Annual Review of Ecology, Evolution, and Systematics* **42,** 23–46 (2011).

23. Hardoim, P. R. *et al.* The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiology and Molecular Biology Reviews* **79,** 293–320 (2015).

24. Doty, S. L. *Endophytes of Forest Trees: Biology and Applications* (eds Pirttilä, A. M. & Frank, A. C.) 151–156 (Springer Netherlands, Dordrecht, 2011).

25. Arnold, A. E. & Herre, E. A. Canopy cover and leaf age affect colonization by tropical fungal endophytes: ecological pattern and process in *Theobroma cacao* (Malvaceae). *Mycologia* **95,** 388–398 (2003).

26. Busby, P. E., Peay, K. G. & Newcombe, G. Common foliar fungi of *Populus trichocarpa* modify *Melampsora* rust disease severity. *New Phytologist* **209,** 1681–1692 (2016).

27. Christian, N., Herre, E. A., Mejia, L. C. & Clay, K. Exposure to the leaf litter microbiome of healthy adults protects seedlings from pathogen damage. *Proc. R. Soc. B* **284,** 20170641 (2017).

28. Cheplick, G. P. & Cho, R. Interactive effects of fungal endophyte infection and host genotype on growth and storage in *Lolium perenne*. *New Phytologist* **158,** 183–191 (2003).

29. Zahn, G. & Amend, A. S. Foliar fungi alter reproductive timing and allocation in *Arabidopsis* under normal and water-stressed conditions. *bioRxiv,* 519678 (2019).

30. Christian, N., Herre, E. A. & Clay, K. Foliar endophytic fungi alter patterns of nitrogen uptake and distribution in *Theobroma cacao*. *New Phytologist* **222,** 1573–1583 (2019).

31. Rosado, B. H. P., Almeida, L. C., Alves, L. F., Lambais, M. R. & Oliveira, R. S. The importance of phyllosphere on plant functional ecology: a phyllo trait manifesto. *New Phytologist* **219,** 1145–1149 (2018).

32. Mejía, L. C. *et al.* Pervasive effects of a dominant foliar endophytic fungus on host genetic and phenotypic expression in a tropical tree. *Frontiers in microbiology* **5** (2014).

33. Dupont, P.-Y. *et al.* Fungal endophyte infection of ryegrass reprograms host metabolism and alters development. *New Phytologist* **208,** 1227–1240 (2015).

34. Dinkins, R. D., Nagabhyru, P., Graham, M. A., Boykin, D. & Schardl, C. L. Transcriptome response of *Lolium arundinaceum* to its fungal endophyte *Epichloë coenophiala*. *New Phytologist* **213,** 324–337 (2017).

35. Welsh, S. *North American species of Astragalus Linnaeus (Leguminosae): a taxonomic revision.* (Brigham Young University, 2007).

36. Knaus, B. J. Morphometric architecture of the most taxon-rich species in the U.S. flora: *Astragalus lentiginosus* (Fabaceae). *American Journal of Botany* **97,** 1816–1826 (2010).

37. Baucom, D. L., Romero, M., Belfon, R. & Creamer, R. Two new species of *Undifilum*, fungal endophytes of *Astragalus* (locoweeds) in the United States. *Botany* **90,** 866–875 (2012).

38. Woudenberg, J. H. C., Groenewald, J. Z., Binder, M. & Crous, P. W. *Alternaria* redefined. *Studies in Mycology* **75,** 171–212 (2013).

39. Cook, D., Gardner, D. R., Martinez, A., Robles, C. A. & Pfister, J. A. Screening for swainsonine among South American *Astragalus* species. *Toxicon* **139,** 54–57 (2017).

40. Molyneux, R. J. & James, L. F. Loco intoxication: indolizidine alkaloids of spotted locoweed (*Astragalus lentiginosus*). *Science* **216,** 190–191 (1982).

41. Cook, D. *et al.* Swainsoninine concentrations and endophyte amounts of *Undifilum oxytropis* in different plant parts of *Oxytropis sericea*. *Journal of Chemical Ecology* **35,** 1272–1278 (2009).

42. Harrison, J. G., Parchman, T. L., Cook, D., Gardner, D. R. & Forister, M. L. A heritable symbiont and host-associated factors shape fungal endophyte communities across spatial scales. *Journal of Ecology* **106,** 2274–2286 (2018).

43. Grum, D. S. *et al.* Production of the alkaloid swainsonine by a fungal endophyte in the host *Swainsona canescens*. *Journal of Natural Products* **76,** 1984–1988 (2013).

44. Cook, D., Gardner, D. R. & Pfister, J. A. Swainsonine-containing plants and their relationship to endophytic fungi. *Journal of Agricultural and Food Chemistry* **62,** 7326–7334 (2014).

45. Panaccione, D. G., Beaulieu, W. T. & Cook, D. Bioactive alkaloids in vertically transmitted fungal endophytes. *Functional Ecology* **28,** 299–314 (2014).

46. Thompson, D. C., Knight, J. L., Sterling, T. M. & Murray, L. W. Preference for specific varieties of woolly locoweed by a specialist weevil, *Cleonidius trivittatus* (Say). *Southwestern Entomologist* **20,** 325–325 (1995).

47. Parker, J. E. *Effects of insect herbivory by the four-lined locoweed weevil, Cleonidius trivittatus (say) (Coleoptera: Curculionidae), on the alkaloid swainsonine in locoweeds Astragalus mollissimus and Oxytropis sericea* PhD thesis (New Mexico State University, Las Cruces, New Mexico, 2008).

48. Creamer, R. & Baucom, D. Fungal endophytes of locoweeds: a commensal relationship? *Journal of Plant Physiology and Pathology* **1** (2013).

49. Lu, H. *et al.* Endogenous fungi isolated from three locoweed species from rangeland in western China. *African Journal of Microbiology Research* **11,** 155–170 (2017).

50. Schulthess, F. M. & Faeth, S. H. Distribution, abundances, and associations of the endophytic fungal community of Arizona fescue (*Festuca arizonica*). *Mycologia* **90,** 569–578 (1998).

51. Cook, D. *et al.* Effects of elevated CO2 on the swainsonine chemotypes of *Astragalus lentiginosus* and *Astragalus mollissimus*. *Journal of Chemical Ecology* **43,** 307–316 (2017).

52. Oldrup, E. *et al.* Localization of endophytic *Undifilum* fungi in locoweed seed and influence of environmental parameters on a locoweed in vitro culture system. *Botany* **88,** 512–521 (2010).

53. Gardner, D. R., Molyneux, R. J. & Ralphs, M. H. Analysis of swainsonine: extraction methods, detection, and measurement in populations of locoweeds (*Oxytropis* spp.) *Journal of Agricultural and Food Chemistry* **49,** 4573–4580 (2001).

54. Högberg, P. 15N natural abundance in soil–plant systems. *The New Phytologist* **137,** 179–203 (1997).

55. Wang, Y. & Qian, P.-Y. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLOS ONE* **4,** e7401 (2009).

56. White, T. J., Bruns, T., Lee, S. & Taylor, J. *Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. A. Innis, D. H. Glefand, J. J. Sninsky, and T. J. White [eds.], PCR protocols: A guide to methods and applications.* (Academic Press, London, UK, 1990).

57. Harrison, J. G., Calder, W. J., Shuman, B. & Buerkle, C. A. The quest for absolute abundance: the use of internal standards for DNA-based community ecology. *Molecular Ecology Resources* **Online early** (2020).

58. Tourlousse, D. M. *et al.* Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. *Nucleic Acids Research* **45,** e23–e23 (2017).

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

60. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4** (2016).

61. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv,* 081257 (2016).

62. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* **11,** 2639 (2017).

63. Edgar, R. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv,* 074161 (2016).

64. Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research* (2018).

65. Cole, J. R. *et al.* Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research* **42,** D633 (2014).

66. Machida, R. J., Leray, M., Ho, S.-L. & Knowlton, N. Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data* **4,** 170027 (2017).

67. Fordyce, J. A., Gompert, Z., Forister, M. L. & Nice, C. C. A hierarchical Bayesian approach to ecological count data: a flexible tool for ecologists. *PLOS ONE* **6,** e26785 (2011).

68. Harrison, J. G., Calder, W. J., Shastry, V. & Buerkle, C. A. Dirichlet-multinomial modelling outperforms alternatives for analysis of microbiome and other ecological count data. *Molecular Ecology Resources* **20,** 481–497 (2020).

69. R Core Team. *R: A language and environment for statistical computing* (Vienna, Austria, 2019).

70. Harrison, J., Shastry, V., Calder, W. J. & Buerkle, C. A. *CNVRG: Dirichlet-multinomial modelling of relative abundance data* 2020.

71. Team, S. D. *Stan modeling language users guide and reference manual* 2020.

72. Team, S. D. *RStan: the R interface to Stan. R package* 2020.

73. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science* **7,** 457–472 (1992).

74. Gloor, G. B. & Reid, G. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Canadian Journal of Microbiology* **62,** 692–703 (2016).

75. Jost, L. Entropy and diversity. *Oikos* **113,** 363–375 (2006).

76. Marion, Z. H., Fordyce, J. A. & Fitzpatrick, B. M. A hierarchical Bayesian model to incorporate uncertainty into methods for diversity partitioning. *Ecology* **99,** 947–956 (2018).

77. Harrison, J. G. *et al.* The many dimensions of diet breadth: phytochemical, genetic, behavioral, and physiological perspectives on the interaction between a native herbivore and an exotic host. *PLOS ONE* **11,** e0147971 (2016).

78. Plummer, M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd international workshop on distributed statistical computing* **124,** 1–8 (2003).

79. Plummer, M. *rjags: bayesian graphical models using MCMC. R package version 3-15. https://CRAN.R-project.org/package=rjags* 2015.

80. Breiman, L. Random forests. *Machine Learning* **45,** 5–32 (2001).

81. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R news* **2,** 18–22 (2002).

82. Grum, D. S. *et al.* Influence of seed endophyte amounts on swainsonine concentrations in *Astragalus* and *Oxytropis* locoweeds. *Journal of Agricultural and Food Chemistry* **60,** 8083–8089 (2012).

83. Marion, Z. H., Fordyce, J. A. & Fitzpatrick, B. M. Extending the concept of diversity partitioning to characterize phenotypic complexity. *The American Naturalist* **186,** 348–361 (2015).

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

85. Strong, D. R., Lawton, J. H. & Southwood, S. R. *Insects on plants. Community patterns and mechanisms.* (Blackwell Scientific Publicatons, Oxford, UK, 1984).

86. Carmona, D., Lajeunesse, M. J. & Johnson, M. T. Plant traits that predict resistance to herbivores. *Functional Ecology* **25,** 358–367 (2011).

87. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology* **5** (2014).

88. Trosvik, P. & de Muinck, E. J. Ecology of bacteria in the human gastrointestinal tract—identification of keystone and foundation taxa. *Microbiome* **3,** 44 (2015).

89. Banerjee, S., Schlaeppi, K. & Heijden, M. G. A. v. d. Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology* **16,** 567 (2018).

90. Braun, K., Romero, J., Liddell, C. & Creamer, R. Production of swainsonine by fungal endophytes of locoweed. *Mycological Research* **107,** 980–988 (2003).

91. Noor, A. I., Nava, A., Cooke, P., Cook, D. & Creamer, R. Evidence for nonpathogenic relationships of *Alternaria* section *Undifilum* endophytes within three host locoweed plant species. *Botany* **96,** 187–200 (2018).

92. Kulpa, S. M. & Leger, E. A. Strong natural selection during plant restoration favors an unexpected suite of plant traits. *Evolutionary Applications* **6,** 510–523 (2013).

93. Leger, E. A. & Baughman, O. W. What seeds to plant in the Great Basin? comparing traits prioritized in native plant cultivars and releases with those that promote survival in the field. *Natural Areas Journal* **35,** 54–68 (2015).

94. Klypina, N., Pinch, M., Schutte, B. J., Maruthavanan, J. & Sterling, T. M. Water-deficit stress tolerance differs between two locoweed genera (*Astragalus* and *Oxytropis*) with fungal endophytes. *Weed Science,* 1–13 (2017).

95. Stamp, N. Out of the quagmire of plant defense hypotheses. *The Quarterly Review of Biology* **78,** 23–55 (2003).

96. Eades, C. J. & Hintz, W. E. Characterization of the α-mannosidase gene family in filamentous fungi: N-glycan remodelling for the development of eukaryotic expression systems. *Biotechnology and Bioprocess Engineering* **5,** 227 (2000).

97. Schmid, J. *et al.* Host tissue environment directs activities of an *Epichloë* endophyte, while it induces systemic hormone and defense responses in its native perennial ryegrass host. *Molecular Plant-Microbe Interactions* **30,** 138–149 (2016).

98. Zamioudis, C. & Pieterse, C. M. J. Modulation of host immunity by beneficial microbes. *Molecular Plant-Microbe Interactions* **25,** 139–150 (2011).

99. Kannadan, S. & Rudgers, J. A. Endophyte symbiosis benefits a rare grass under low water availability. *Functional Ecology* **22,** 706–713 (2008).

100. Barillas, J. R. V., Paschke, M. W., Ralphs, M. H. & Child, R. D. White locoweed toxicity is facilitated by a fungal endophyte and nitrogen-fixing bacteria. *Ecology* **88,** 1850–1856 (2007).

101. Isbell, F. *et al.* High plant diversity is needed to maintain ecosystem services. *Nature* **477,** 199–202 (2011).

102. Hawkins, J. A., Jones, S. K., Finkelstein, I. J. & Press, W. H. Indel-correcting DNA barcodes for high-throughput sequencing. *Proceedings of the National Academy of Sciences* **115,** E6217–E6226 (2018).

103. Tourlousse, D. M., Ohashi, A. & Sekiguchi, Y. Sample tracking in microbiome community profiling assays using synthetic 16S rRNA gene spike-in controls. *Scientific Reports* **8,** 1–9 (2018).

104. Nilsson, R. H. *et al.* Mycobiome diversity: high-throughput sequencing and identification of fungi. *Nature Reviews Microbiology,* 1 (2018).

105. Lofgren, L. A. *et al.* Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology* **28,** 721–730 (2019).

106. Kembel, S. W., Wu, M., Eisen, J. A. & Green, J. L. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLOS Computational Biology* **8,** e1002743 (2012).

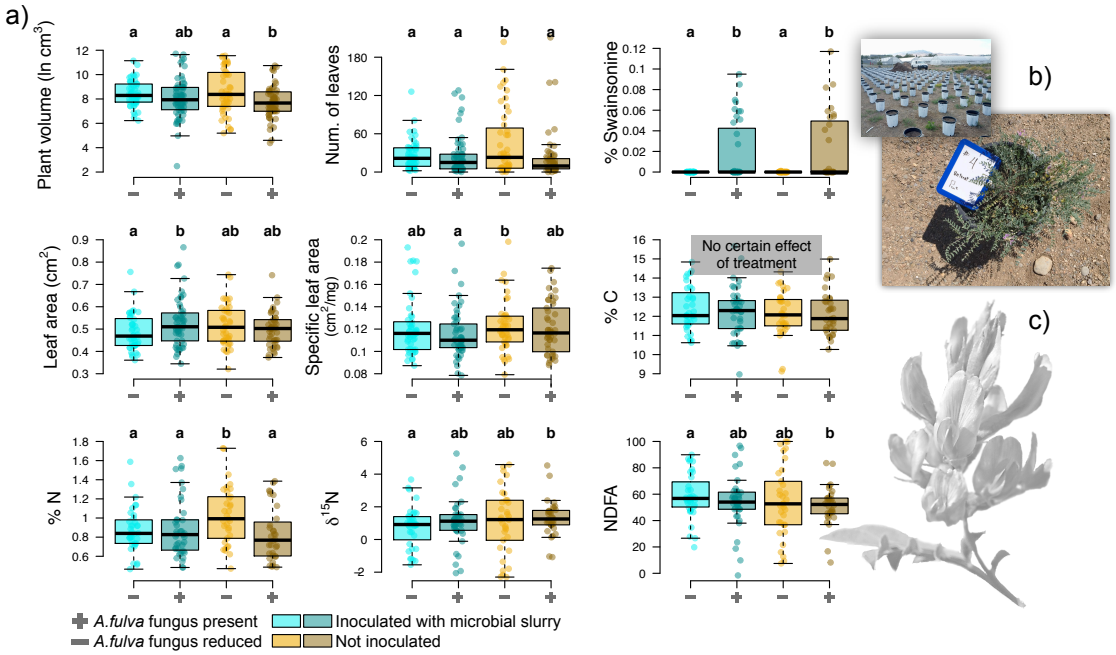

Figure 1: Variation in *Astragalus lentiginosus* traits among treatment groups (a). + and - symbols on the x-axis represent treatment to reduce the relative abundance of the vertically-transmitted fungus, *Alternaria fulva*. Boxes shaded blue denote treatment with endophyte inoculum slurry and boxes shaded yellow denote plants that did not receive the slurry. Percentage of N, C, and swainsonine refer to foliar dry mass composition. Differences in mean trait values among treatment groups were determined through a hierarchical Bayesian analysis. Credible differences ($\geq$ 95% difference in posterior probability distributions of parameter estimates) among treatment groups are designated through the letters above each boxplot. For estimates of mean trait values for each treatment group see Table 2. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles. Panel b shows an overview of the experimental installation and panel c depicts the inflorescence of *A. lentiginosus*. All photos by J. Harrison.

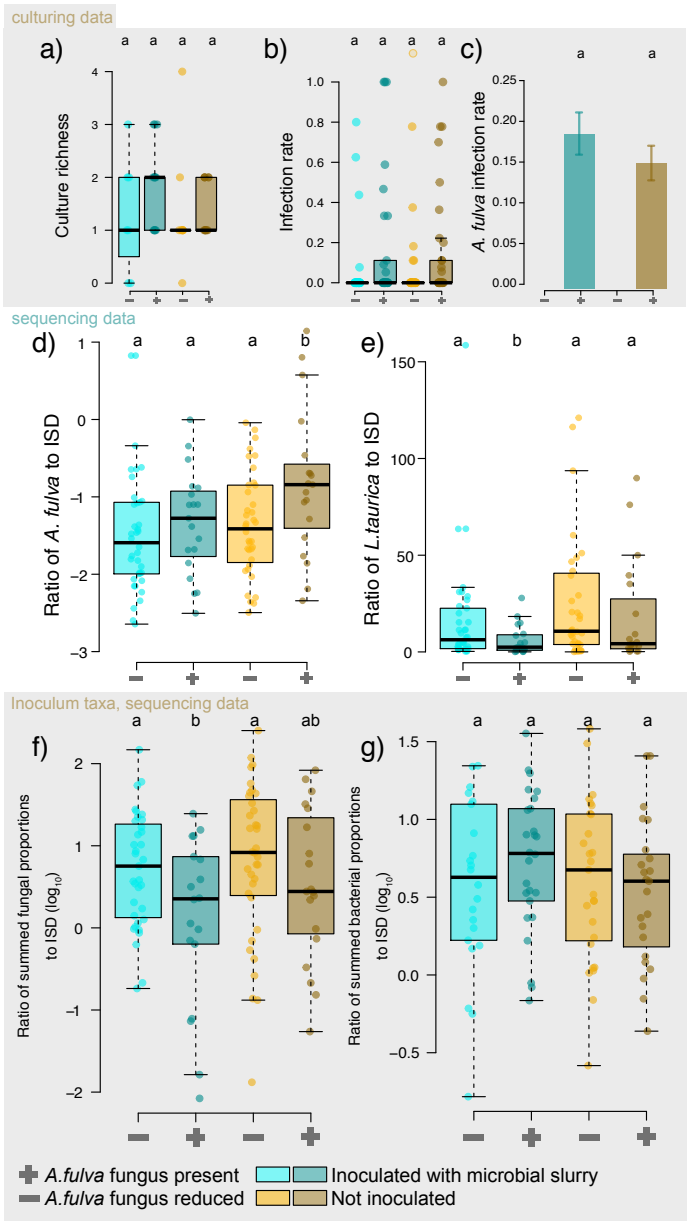

Figure 2: Influence of treatment on microbial assemblages within *Astragalus lentiginosus*. Boxes shaded blue denote treatment with endophyte inoculum slurry; boxes shaded yellow denote plants that did not receive the slurry. The top row contains data from culturing microbes from leaves of host plants. Data shown include culture richness (a), the proportion of leaves infected with microbes other than *A. fulva* (b), and the rate of *A. fulva* infection (c). Cultured taxa were not identified to species, but were nearly all fungi. The middle row depicts the effect of treatment on *Alternaria fulva* (d) and *Leveillula taurica* (e) abundance (estimated as the ratio of the proportion of reads for either taxon to the internal standard [ISD]); these were the two most abundant fungal taxa present. The bottom row shows abundances of fungal (f) and bacterial (g) taxa that were in the inoculum mixture used to treat plants. Credible differences among treatment groups are denoted through the letters above each boxplot. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles. Treatment to reduce *Alternaria fulva* is shown as *A. fulva* –, with a + denoting lack of treatment to reduce this taxon.

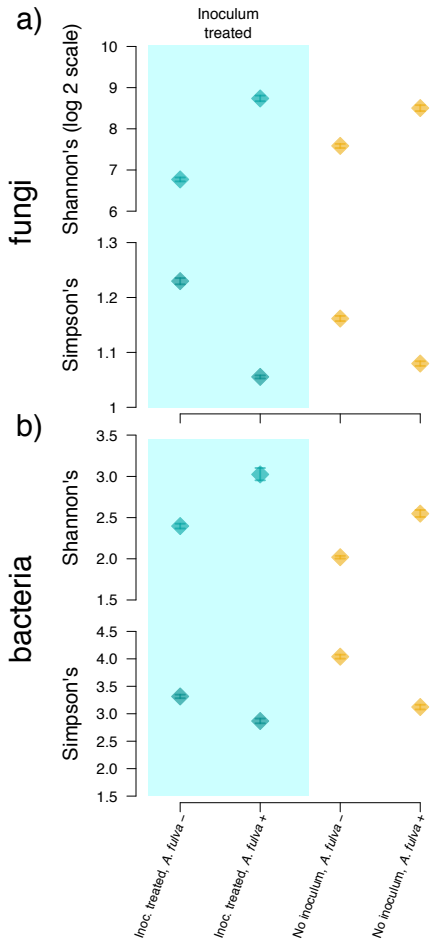

Figure 3: Effect of treatment on fungal (a) and bacterial diversity (b) as estimated from sequencing data. Treatments are noted on the x axis. Points denote the means of posterior probability distributions (PPD) of diversity entropy equivalencies. 95% high density intervals for these PPDs were very narrow, but are shown superimposed on each estimate. Points shaded blue denote treatment groups receiving the endophyte inoculum slurry; points shaded yellow denote groups that did not receive the slurry. Treatment to reduce *Alternaria fulva* is shown as *A. fulva* −, with a + denoting lack of treatment to reduce this taxon.

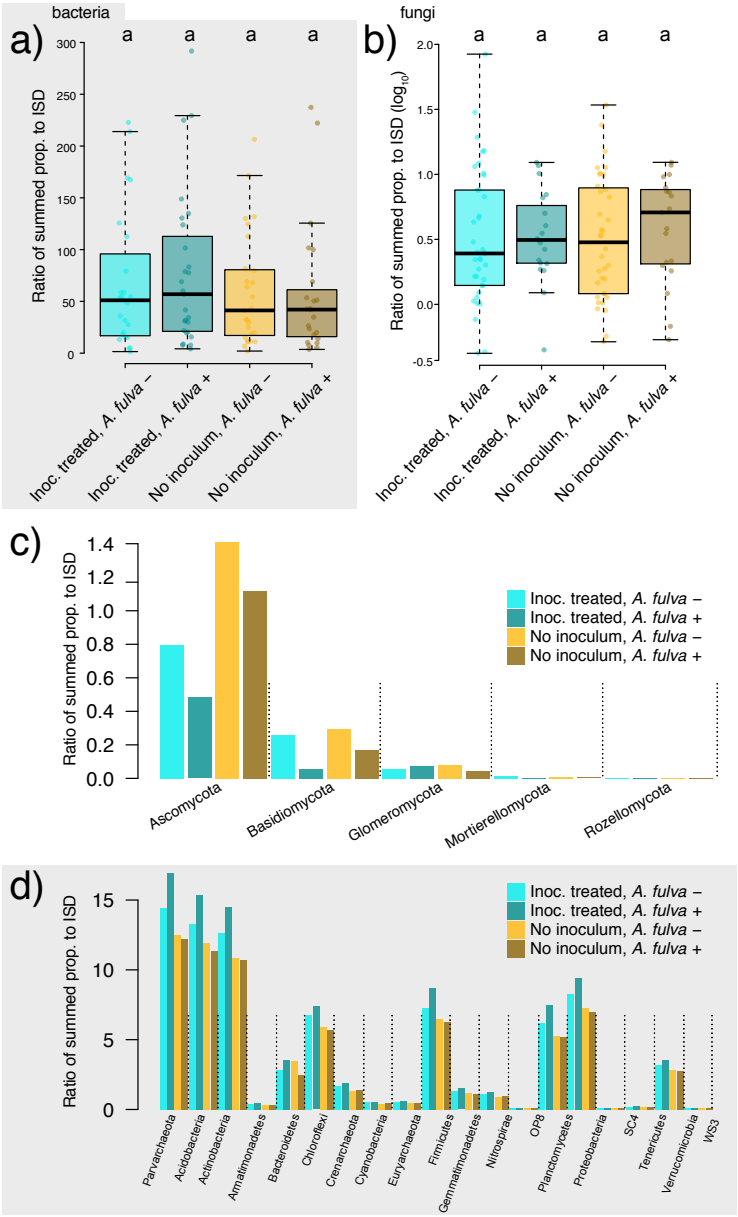

Figure 4: Most of the microbial biodiversity present in this experiment was low abundance. The top row shows the ratio of bacterial (a) and fungal (b) proportions to the internal standard (ISD; not including *A. fulva* and *L. taurica*, which were more abundant), with proportions for all taxa summed by treatment group. The proportion of reads for all bacterial taxa was less than 3% for each treatment group (the other reads were from host organelle DNA). As for fungi, the most abundant taxon, besides *A. fulva* and *L. taurica*, was represented by less than 5% of the fungal reads in any given treatment group and, when averaged across groups, the relative abundance of this taxon was less than a percent. Most fungal taxa in most treatment groups composed less than 0.02% of reads. The abundances of fungal phyla by treatment group are shown on the middle row (c) and abundances of bacterial phyla are shown on the bottom row (d). Blue shading denotes treatment groups receiving the endophyte inoculum slurry whereas groups that did not receive the slurry are shaded yellow. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles. Treatment to reduce *Alternaria fulva* is shown as *A. fulva* –, with a + denoting lack of treatment to reduce this taxon.

Table 1: Number of *Astragalus lentiginosus* plants installed in field experiment for each treatment group. Treatments were applied to plants in a full factorial design and included reduction of the relative abundance of the vertically-transmitted fungal endophyte, *Alternaria fulva* (through embryo excision) and foliar application of regionally-sourced, endophyte inocula. To control for effects of seed coat removal, embryos were excised from seed coats and planted alongside agar that was either sterile or contained *A. fulva*.

Experimental plants		
A. fulva reduced	Inoculum applied	Sample size
Yes	Yes	62
Yes	No	68
No	Yes	54
No	No	54
Control plants installed with agar		
Yes	Yes	13
Yes	No	13
No	Yes	18
No	No	18

Table 2: Trait values for *Astragalus lentiginosus* individuals in each treatment group. Values shown are means of posterior probability distributions of the mean for each trait with 95% credible intervals in parentheses. A plus after *Alternaria fulva* means we did not attempt to reduce *A. fulva* abundance through embryo excision. Minus symbols mean *A. fulva* was reduced.

	Inoc. treated, *A. fulva* −	Inoc. treated, *A. fulva* +	No inoculum, *A. fulva* −	No inoculum, *A. fulva* +
Size (cm^3)	9456 (9430,9485)	12417 (12393,12443)	20932 (20904,20962)	6502 (6478,6528)
Leaves	26.84 (20.17,33.32)	24.78 (16.74,32.54)	45.21 (30.29,60.24)	21.86 (11.55,31.43)
Leaflet area (cm^2)	0.49 (0.46,0.51)	0.53 (0.5,0.56)	0.52 (0.49,0.55)	0.5 (0.48,0.52)
SLA (cm^2mg^{-1})	0.12 (0.11,0.13)	0.11 (0.11,0.12)	0.12 (0.11,0.13)	0.12 (0.11,0.13)
δ N15	0.78 (0.33,1.18)	1.09 (0.59,1.59)	1.1 (0.42,1.79)	1.32 (0.89,1.74)
% swainsonine	0 (0,0)	0.02 (0.01,0.03)	0 (0,0)	0.02 (0.01,0.03)
% N	0.87 (0.8,0.96)	0.89 (0.79,1)	1.03 (0.92,1.13)	0.83 (0.72,0.93)
% C	12.4 (12.01,12.78)	12.23 (11.82,12.65)	12.12 (11.71,12.53)	12.15 (11.74,12.55)
NDFA	58.56 (52.6,64.45)	54.24 (47.4,60.68)	54 (44.7,63.79)	51.24 (45.79,56.7)

Supplementary Material

Library preparation details

A two-step PCR was used to prepare sequencing libraries. In the first step, the locus of interest was amplified and in a subsequent step Illumina adapters were added to amplicons. To characterize bacterial assemblages, the 16S (V4) locus was amplified using the 515-806 primer pair [55], while for fungal assemblages the ITS1 locus was amplified using the ITS1f-ITS2 primer pair [56]. A synthetically designed internal standard (ISD; Tourlousse *et al.* [58]) was added to each sample to allow estimation of absolute abundances. The ITS ISD was "CTTGGTCATTTAGAGGAAGTAATGCCACAGATACGTACCGCTCATAACGCGAACCGAAGCGCAGTAGAAGTACTCCGTATCCTACCTCGGTCGTGGTTTAGGCTATCGACATCTTGCATGGGCTTCCCTAGTGAACTCTTGGGATGTGCATCGATGAAGAACGCAGC" and the 16S ISD was "GTGCCAGCAGCCGCGGTAAGCCACAGATACGTACCGCTCATAACGCGAACCGAAGCGCAGTAGAAGTACTCCGTATCCTACCTCGGTCGTGGTTTAGGCTATCGACATCTTGCATGGGCTTCCCTAGTGAACTCTTGGGATGTATTAGATACCCTAGTAGTCC". These sequences are shortened variants of those presented in Tourlousse *et al.* [58]. Additionally, unique synthetic DNA molecules of our own design were added to each well of the 96-well plates that we used. By tracking occurrence of these DNAs among samples, we were able to account for cross-contamination. We will be describing these sequences in a forthcoming paper, briefly, they are 13 nucleotide long sequences taken from [102] and we refer to them as 'coligos', which is short for cross-contamination checking oligo (for a similar approach see [103]).

Both 16S and ITS libraries were created using duplicate 15 µl PCR reactions. We added 6 µl of 0.01 pg μL^{-1} coligos and 0.03 pg μL^{-1} of the ISD to 30 ul of template. Template was then normalized to a concentration of 10 ng μL^{-1} across all samples. Constituents of the first round of PCR were: 6 µl each of 0.25 µM forward and reverse primers, 0.3 µl Kapa HiFi HotStart DNA polymerase (New England BioLabs, Ipswich, MA, U.S.A.), 0.45 µl 10 M dNTPs, 3 µl 5x KAPA HiFi HotStart PCR buffer (New England BioLabs, Ipswich, MA, U.S.A.), 3.25 µl of water, and 2 µl of template. Conditions for the first round of amplification were as follows: an initial denaturation at 95°C for 3 min, followed by 15 cycles of 98°C for 30 s, 62°C for 30 s, and 72°C for 30 s, and a final 5 min extension at 72°C. Amplicons were cleaned using Axygen AxyPrep MagBead (24 µl; Corning; Glendale, Arizona, U.S.A) prior to the second round of PCR. Reaction volume for the second round of PCR was 15 µl. Constituents of the second PCR were: 0.5 µl each of 10 µM flowcell primers, 0.3 µl Kapa HiFi HotStart DNA polymerase (New England BioLabs, Ipswich, MA), 0.45 µl 10 M dNTPs, 3 µl 5x KAPA HiFi HotStart PCR buffer (New England BioLabs; Ipswich, MA, U.S.A.), 0.75 µl of water, and 10 µl of cleaned product from the first round of PCR. Conditions for the second round of amplification were as follows: an initial denaturation at 95°C for 3 min, followed by 19 cycles of 98°C for 30 s, 55°C for 30 s, and 72°C for 30 s, and a final 5 min extension at 72°C. Products were cleaned again using magnetic beads, as above. Amplicons were pooled in equimolar fashion and library success determined via a Bio-Analyzer (Agilent, Santa Clara, CA, U.S.A) and qPCR prior to sequencing on the Illumina NovaSeq system (San Diego, CA, U.S.A.) by psomagen, Inc. (Rockville, MA, U.S.A.). Aside from sequencing, all laboratory

work was conducted at the Genome Technologies Laboratory at the University of Wyoming (Laramie, WY, U.S.A.).

Use of high-resolution OTUs, also referred to as exact sequence variants (ESVs)

We relied on OTUs that were clustered so as to identify single-nucleotide variants within the data. These OTUs are often referred to as exact sequence variants (ESVs; we have also seen the terms amplicon sequence variants, and zero-radius OTUs in the literature; [62]). Using finely-resolved OTUs has become accepted practice for 16S data [62], however there is ongoing discussion regarding how appropriate they are for fungal data [104]. The criticism brought to bear is that fungi can be multicellular and multinucleate and that the ITS region can have multiple variants within taxa and even within individual fungal organisms [105]. Thus, the concern regarding the use of highly-resolved OTUs for fungi is that multiple OTUs could be obtained from the same organism (this could also plausibly happen for bacteria, as they can posses copy number variation for 16S [106], but it seems less likely to be as extreme as what might be expected for fungi). It is important to remember that read counts do not translate well into counts of individuals, even in best case scenarios, because of copy number variation among microbial taxa (the consequences of this are discussed in [57]). That said, if multiple OTUs were obtained for the same organism then it could influence richness estimates in undesirable ways. However, there is no guarantee that the sequence variation observed would not cause the same biases to richness if OTUs were clustered differently (e.g., at 97%), in other words the effects on richness of sequence clustering threshold would need to behave non-linearly to have untoward influence. We do not present analyses of richness in this manuscript, because we consider such analyses fraught for microbial sequence data that are dominated by several taxa (because of the compositional nature of the data) and because, in preliminary work, richness description did not influence the inferences we describe herein.

Additional justification for our use of high-resolution OTUs is that one of our primary goals was to characterize and analyze the genetic variation present via diversity entropy analysis and ordinations. For such analyses, using all genetic variation present seemed appropriate, rather than arbitrarily lumping variants by sequence similarity. For taxon-specific analyses, such as our characterization of how *A. fulva* and *L. taurica* were influenced by treatment, we summed reads across OTUs for each taxon prior to analysis; thus, the results we present would be no different if we had used other OTU clustering thresholds, because reads from those OTUs would have been summed in the same way.

There are mathematical reasons to prefer high-resolution OTUs as well. For instance, a problem with OTUs clustered via percent sequence similarity is that the sequences lumped into an OTU depend upon the length of the OTU, because the length determines the denominator in percent calculations. Thus, OTU specification changes over the phylogeny, assuming length variation of the marker locus used (length polymorphism is very common for fungi [104, 105]). Also, the choice of which data to include when calling OTUs (i.e., which replicates) and the order of the sequences encountered by the clustering algorithm can affect read counts assigned to an OTU and OTU richness, whereas high-resolution OTUs are more consistently called. This improves sharing information among research groups that have used different methods (e.g., to see where a genotype appears on the landscape) and makes analyses that

use high-resolution OTUs more reproducible than analyses that use OTUs created through clustering reads by similarity.

A final advantage of high-resolution OTUs is that they have a clear definition: a genotype at a locus. Whereas OTUs of lower resolution could correspond to a species or subspecies (or some other taxonomic level. For all these reasons, we suggest that using high-resolution OTUs (ESVs) is appropriate and desirable for many studies of fungal assemblages

Inoculum sequence details

Inoculum was created through culturing microbes from foliar tissue of various Great Basin shrubs. Twenty morphologically distinct cultures were obtained. For details of inoculum preparation see main text. We sequenced two aliquots of the inoculum to identify its constituent microbes. We performed sequencing of the aliquots via the NovaSeq methods described here and via the MiSeq platform, using the same DNA extracts and very similar library preparation procedures. Libraries for the MiSeq sequencing were created by the Genome Sequencing and Analysis Facility at the University of Texas (see Harrison *et al.* [42] for description of these methods, though they are very similar to those described here). We combined the data from both sequencing runs and recovered 20 fungal OTUs and 260 bacterial OTUs. Many of the OTUs were assigned to taxa with known plant pathogens and endophytes. The most abundant fungal taxa in the inoculum were *Preussia* and *Cladosporium*, (both of which occur in wild plants, see below). The most abundant 16S sequences appeared to be mitochondrial and were mapped to several fungi that appeared in the ITS data, including *Cladosporium*, *Penicillium* and *Aureobasidium pullulans*. The most abundant bacteria in the inoculum mixture mapped to *Bacillus* and *Propionibacterium acnes*.

To determine if the taxa within the inoculum occurred in wild *Astragalus lentiginosus* we aligned the OTUs from our inoculum with the OTUs in Harrison *et al.* [42]. In this work, the authors describe the fungal microbiome of *A. lentiginosus* plants collected from throughout the Great Basin desert. We found that OTUs corresponding to *Preussia*, *Cladosporium*, *Alternaria*, and *Leveillula* were present in both the inoculum and wild-collected plants (99% match). When we relaxed our match criterion to 80% we found that 15 of the 20 fungal OTUs in our inoculum were present in the data collected by Harrison *et al.* [42], including members of *Aspergillus*, *Penicillium* and *Endoconidioma*, among others, that are known to grow endophytically. This suggests that many of the taxa, or their closely-related congeners, within our inoculum occur within wild-collected plants. Harrison *et al.* [42] did not sequence bacteria so we were unable to verify that the bacteria in our inoculum can often be found in wild *A. lentiginosus*.

Sequencing of mock community

Sequencing the ITS from the mock community resulted in 34 OTUs, all of which were assigned to either *Cryptococcus neoformans* or *Saccharomyces cerevisiae*, the two fungi expected to occur in the ZymoBIOMICS mock community (Zymo Research, Irvine, CA. U.S.A.). Notably, we only could recover *S. cerevisiae* when we used concatenated reads, because this taxon apparently has a long enough ITS1 to prevent merging, at least using our software parameters. This observation was the justification for the concatenation approach we used herein.

We observed 28 OTUs in the 16S data from the mock community. Bacterial taxa in the mock community included: *Listeria monocytogenes, Pseudomonas aeruginosa, Bacillus subtilis, Escherichia coli, Salmonella enterica, Lactobacillus fermentum, Enterococcus faecalis* and *Staphylococcus aureus*. Our taxonomic hypothesis generation method accurately captured the genus of each of these taxa, and in many cases species-level taxonomic hypotheses matched expectations. These results, confirm that our method can recover the taxonomic breadth within the ZymoBIOMICS mock community.

Sequencing results

A total of 17,616,644 ITS1 reads and 26,398,005 16S reads were generated. Of these 20,260,185 for 16S and 10,611,127 for ITS were merged, passed filters, and were mapped to OTUs used to generate OTU tables. Many of the reads were assigned to the *A. lentiginosus* host, however we did recover on average 20,491 fungal and 422 bacterial reads per sample. This low number of bacterial reads per sample reflects a preponderance of reads being assigned to the host plant organelles. This is noteworthy as it confirms that host organelles were vastly more abundant within samples than bacterial endophytes. Cross-contamination was extremely rare as evidenced by our coligos (see above), only 0.2% of coligo sequences were recovered from unexpected wells. Therefore, all samples were retained for analysis. Read counts from PCR duplicates were highly correlated (R = 0.87 for ITS and R = 0.99 for 16S) and so data for duplicate pairs were combined.

During preliminary sequencing, we extracted three aliquots of DNA from five plants, passed these through library preparation, and sequenced them on an Illumina MiSeq (using methods very similar to those mentioned here) to determine the technical variation inherent to our method. We observed very little technical variation, which concurs with our observations of sequence data from PCR duplicates.

Table S1: Effect of treatment on *A. fulva* and *L. taurica* abundance. Plus symbols denote *A. fulva* presence, while minus symbols refer to plants that were treated to reduce *A. fulva*. Estimates of multinomial p parameters and of Dirichlet π parameters are provided. Both of these estimates can be thought of as proportions of the total read count. The multinomial parameters were averaged over all plants in the treatment group. All parameters are also shown as ratios to the ISD, to place estimates on a scale proportional to absolute abundance.

	A. fulva - Inoc. treated	*A. fulva* + Inoc. treated	*A. fulva* – No inoculum	*A. fulva* + No inoculum
A. fulva multinomial p	0.008 (-0.000–0.016)	0.015 (0.002–0.027)	0.005 (0.003–0.007)	0.047 (0.001–0.093)
A.fulva Dirichlet π	0.012 (0.011–0.012)	0.018 (0.017–0.019)	0.014 (0.013–0.014)	0.018 (0.017–0.019)
A.fulva multinomial p / ISD	0.25 (-0.12–0.63)	0.14 (0.03–0.25)	0.13 (0.06–0.20)	1.16 (0.01–2.3)
A. fulva Dirichlet π / ISD	0.15 (0.14–0.15)	0.13 (0.11–0.13)	0.25 (0.24–0.26)	0.24 (0.22–0.26)
L. taurica multinomial p	0.58 (0.50–0.66)	0.42 (0.29–0.56)	0.58 (0.49–0.68)	0.50 (0.36–0.63)
L. taurica Dirichlet π	0.08 (0.08–0.08)	0.14 (0.14–0.14)	0.06 (0.06–0.06)	0.08 (0.08–0.08)
L. taurica multinomial p / ISD	17.2 (7.4–26.9)	6.3 (2.7–9.9)	30.9 (14.3–47.5)	18.4 (6.2–30.7)
L. taurica Dirichlet π / ISD	6.3 (6.1–6.4)	2.1 (2.1–2.2)	8.5 (8.2–8.8)	5.0 (4.7–5.2)

Table S2: Estimated coefficient of variation in *A. lentiginosus* traits for each treatment group. Coefficients were estimated using a hierarchical Bayesian approach that modeled trait values as normally distributed. The coefficient of variation was the standard deviation of the posterior probability distribution (PPD) for the mean, divided by the mean of that distribution, and thus is a measure of trait variation. Coefficients were calculated for every sample from the PPDs for the means, thus providing a PPD of the coefficient of variation. 95% high density intervals of this distribution are given parenthetically. By subtracting these distributions we determined if trait variation differed between groups. If 95% or more of these distributions did not overlap then we considered this evidence for a credible difference in trait variation, which we denoted by letters prefixed to each estimate, similarly to the approach we used in Fig. 1. Note that the NA value for swainsonine reflects an estimated zero concentration of swainsonine in this treatment group.

Trait	*A. fulva* − Inoc. treated	*A. fulva* + Inoc. treated	*A. fulva* − No inoculum	*A. fulva* + No inoculum
Size (ln)	a 0.1005 (0.0813,0.1218)	b 0.0723 (0.0586,0.086)	c 0.0611 (0.048,0.075)	d 0.0875 (0.0714,0.1044)
Num. of leaves	a 0.0016 (0.0011,0.0022)	a 0.0013 (9e-04,0.002)	b 4e-04 (3e-04,6e-04)	a 0.0013 (8e-04,0.0023)
Leaf area	a 24.817 (19.5497,30.5745)	b 17.1858 (13.8143,20.7423)	ab 20.2617 (15.9141,24.8689)	a 25.8478 (20.3218,31.8292)
Specific leaf area	a 318.9053 (252.2565,392.4747)	a 377.1165 (306.4815,451.7356)	a 340.0024 (265.1219,418.5106)	a 340.3893 (265.5111,415.8287)
delta N15	a 1.0928 (0.5828,2.2532)	a 0.6507 (0.377,1.102)	a 0.5474 (0.2458,1.1806)	a 0.6797 (0.4284,0.989)
% swainsonine	NA	a 2016.1664 (1079.6788,3819.3327)	NA	a 1523.6243 (787.5588,2975.1052)
% N	a 4.8262 (3.7128,6.0467)	a 3.5636 (2.6828,4.5083)	b 3.1097 (2.3358,3.9722)	a 4.2724 (3.1363,5.5376)
% C	a 0.0691 (0.0537,0.0857)	a 0.0608 (0.0476,0.0746)	a 0.0695 (0.0535,0.0877)	a 0.0705 (0.053,0.0891)
NDFA	ac 0.001 (7e-04,0.0012)	ab 9e-04 (7e-04,0.0011)	b 7e-04 (5e-04,9e-04)	c 0.0013 (9e-04,0.0017)

Table S3: Effect of treatment on microbial relative abundance. Values shown are the number of microbial taxa differing in relative abundance for each pairwise comparison. The upper half of the table (shaded) describes results for fungi (2292 taxa considered), and the bottom half of the table describes results for bacteria (657 taxa considered). OTUs corresponding to *A. fulva* genotypes were omitted during analysis, however these OTUs were also influenced by treatment (see Fig. 2). Differences in relative abundance were determined through a hierarchical Bayesian modeling approach (see main text). Plus symbols denote *A. fulva* colonization, while minus symbols refer to plants that were treated to reduce *A. fulva*.

	A. fulva + Inoc. treated	*A. fulva* + No inoculum	*A. fulva* − Inoc. treated	*A. fulva* − No inoculum
A. fulva + Inoc. treated	–	4	24	10
A. fulva + No inoculum	3	–	18	12
A. fulva − Inoc. treated	20	3	–	8
A. fulva − No inoculum	642	16	6	–

Table S4: Associations between plant traits and species equivalencies of Shannon's entropy for fungi (shaded, top rows) and bacteria (unshaded, bottom rows). Values in cells are the mean of posterior probability distributions (PPDs) of beta coefficient estimates from a hierarchical linear model calculated using a Bayesian approach. Values in parentheses denote the proportion of each PPD greater than zero, which is the certainty of a non-zero effect of the covariate on the response. Plus symbols denote *A. fulva* colonization, while minus symbols refer to plants that were treated to reduce *A. fulva*.

Treatment	Intercept	Plant volume	SLA	%N	%C	NDFA	Swainsonine
A. fulva – Inoc. treated	**19.49 (97.8%)**	1.25 (57.2%)	-3.15 (28.4%)	-0.67 (43.5%)	1.56 (62.1%)	1.89 (65.0%)	–
A. fulva + Inoc. treated	**20.24 (99.9%)**	-1.08 (48.7%)	-3.10 (26.9%)	-1.05 (40.6%)	1.46 (61.6%)	2.08 (66.5%)	-3.4 (37.2%)
A. fulva – No inoculum	**20.8 (100%)**	-0.22 (49.7%)	-0.67 (41.6%)	-1.07 (41.0%)	1.95 (66.4%)	2.31 (68.3%)	–
A. fulva + No inoculum	**21.47 (99.5%)**	3.18 (62.7%)	-2.35 (30.9%)	-0.68 (43.6%)	1.56 (63.0%)	1.79 (65.6%)	-0.48 (47.8%)
A. fulva – Inoc. treated	**1.90 (100%)**	0.05 (59.8%)	0.11 (81.0%)	**-0.31 (4.7%)**	0.18 (87.6%)	-0.09 (28.7%)	–
A. fulva + Inoc. treated	**2.08 (100%)**	-0.04 (45.0%)	-0.09 (42.1%)	-0.28 (17.5%)	0.27 (80.2%)	-0.13 (32.3%)	-0.11 (47.0%)
A. fulva – No inoculum	**1.81 (100%)**	-0.07 (23.4%)	0.06 (77.0%)	-0.07 (28.6%)	0.11 (86.0%)	0.03 (70.6%)	–
A. fulva + No inoculum	1.85 (100%)	0.17 (73.8%)	**0.24 (96.2%)**	-0.18 (16.8%)	-0.05 (40.0%)	-0.04 (38.8%)	-0.03 (44.0%)

Table S5: Associations between plant traits and species equivalencies of Simpson's diversity for fungi (shaded, top rows) and bacteria (unshaded, bottom rows). Values in cells are the mean of posterior probability distributions (PPDs) of beta coefficient estimates from a hierarchical linear model calculated in a Bayesian framework. Values in parentheses denote the proportion of each PPD greater than zero, which is the certainty of a non-zero effect of the covariate on the response. Plus symbols denote *A. fulva* colonization, while minus symbols refer to plants that were treated to reduce *A. fulva*.

Treatment	Intercept	Plant volume	SLA	%N	%C	NDFA	Swainsonine
A. fulva – Inoc. treated	**2.06 (100%)**	-0.07 (42.1%)	-0.05 (42.2%)	0.0 (51.5%)	-0.08 (25.2%)	-0.14 (32.8%)	–
A. fulva + Inoc. treated	**1.90 (100%)**	**-0.14 (9.9%)**	0.0 (48.9%)	**-0.12 (10.0%)**	-0.07 (17.1%)	0.06 (79.0%)	**-0.16 (2.6%)**
A. fulva – No inoculum	**2.08 (100%)**	0.04 (72.0%)	-0.03 (34.7%)	**-0.15 (6.2%)**	**0.11 (92.7%)**	0.08 (85.2%)	–
A. fulva + No inoculum	**1.87 (100%)**	0.04 (66.2%)	0.03 (65.6%)	0.0 (51.0%)	0.01 (55.0%)	-0.07 (13.1%)	0.04 (65.5%)
A. fulva – Inoc. treated	**1.31 (100%)**	0.03 (68.8%)	0.02 (74.3%)	**-0.08 (5.5%)**	**0.05 (91.3%)**	-0.04 (16.0%)	–
A. fulva + Inoc. treated	**1.40 (100%)**	-0.03 (27.1%)	0.01 (56.6%)	-0.06 (18.5%)	**0.08 (94.2%)**	0.01 (59.5%)	0.01 (52.0%)
A. fulva – No inoculum	**1.28 (100%)**	**-0.04 (8.4%)**	0.02 (81.7%)	-0.01 (44.4%)	0.02 (73.5%)	0.01 (69.5%)	–
A. fulva + No inoculum	**1.31 (100%)**	0.05 (66.6%)	**0.10 (97.8%)**	-0.04 (33.0%)	-0.06 (20.1%)	0.00 (48.8%)	-0.01 (47.0%)

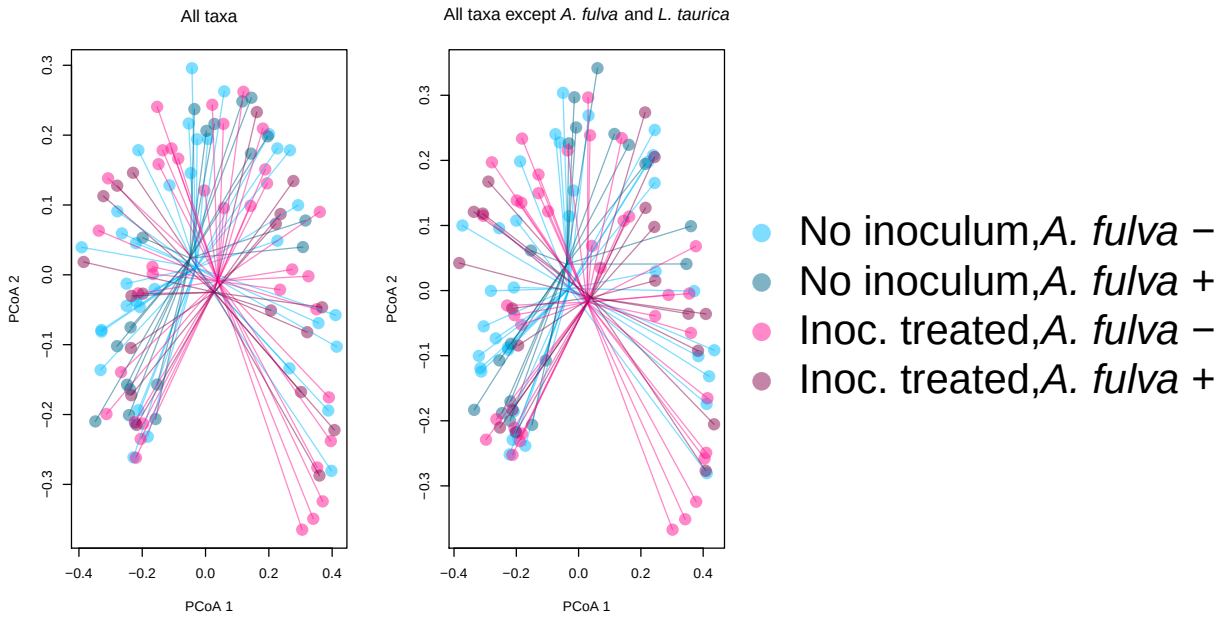

Figure S1: Principal coordinates analysis (PCoA) of Bray-Curtis transformed maximum likelihood estimates of fungal relative abundances for each sample. Ordinations are color-coded to show the effect of treatment on fungal assemblage composition for a) all taxa, b) all taxa except *A. fulva* and *L. taurica*, which were two of the most abundant taxa in most plants. Axes refer to values from PCoA. Points are samples and are connected by lines to make it easier to identify the centroid of treatment groups. PERMANOVA analysis revealed no significant effects of treatment for either dataset.

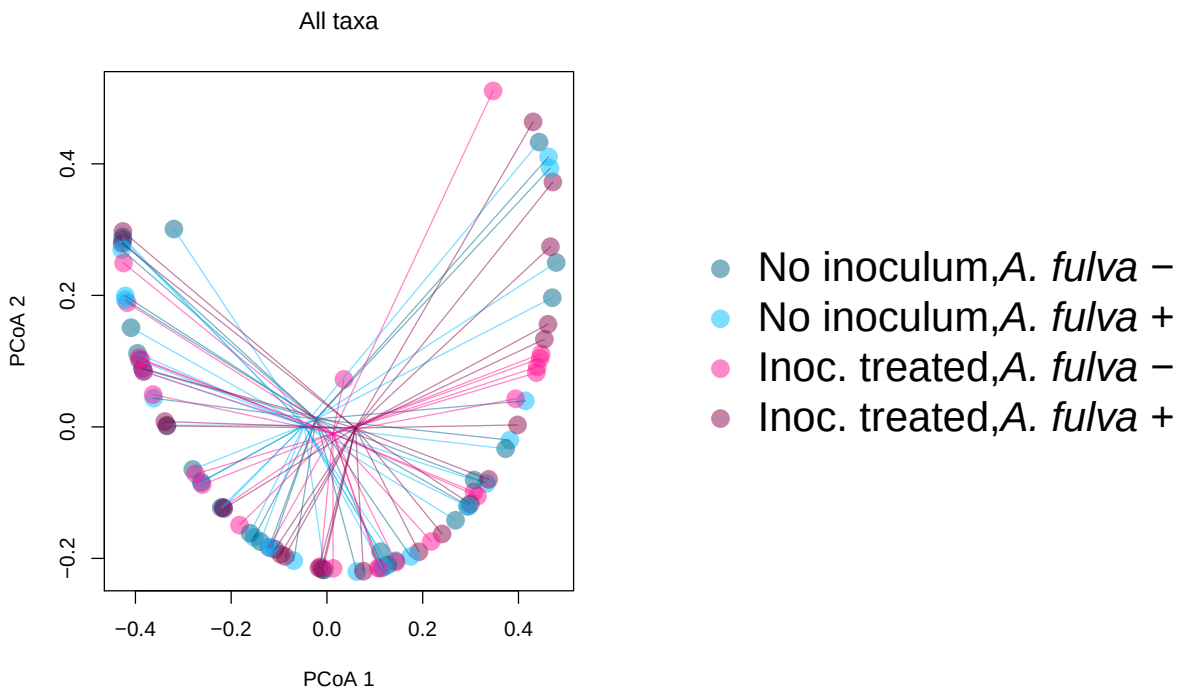

Figure S2: Principal coordinates analysis (PCoA) of Bray-Curtis transformed maximum likelihood estimates of bacterial relative abundances for each sample. Ordinations are color-coded to show the effect of treatment on bacterial assemblage composition for all taxa. Axes refer to values from PCoA. Points are samples and are connected by lines to make it easier to identify the centroid of treatment groups. PERMANOVA analysis revealed no significant effect of treatment.

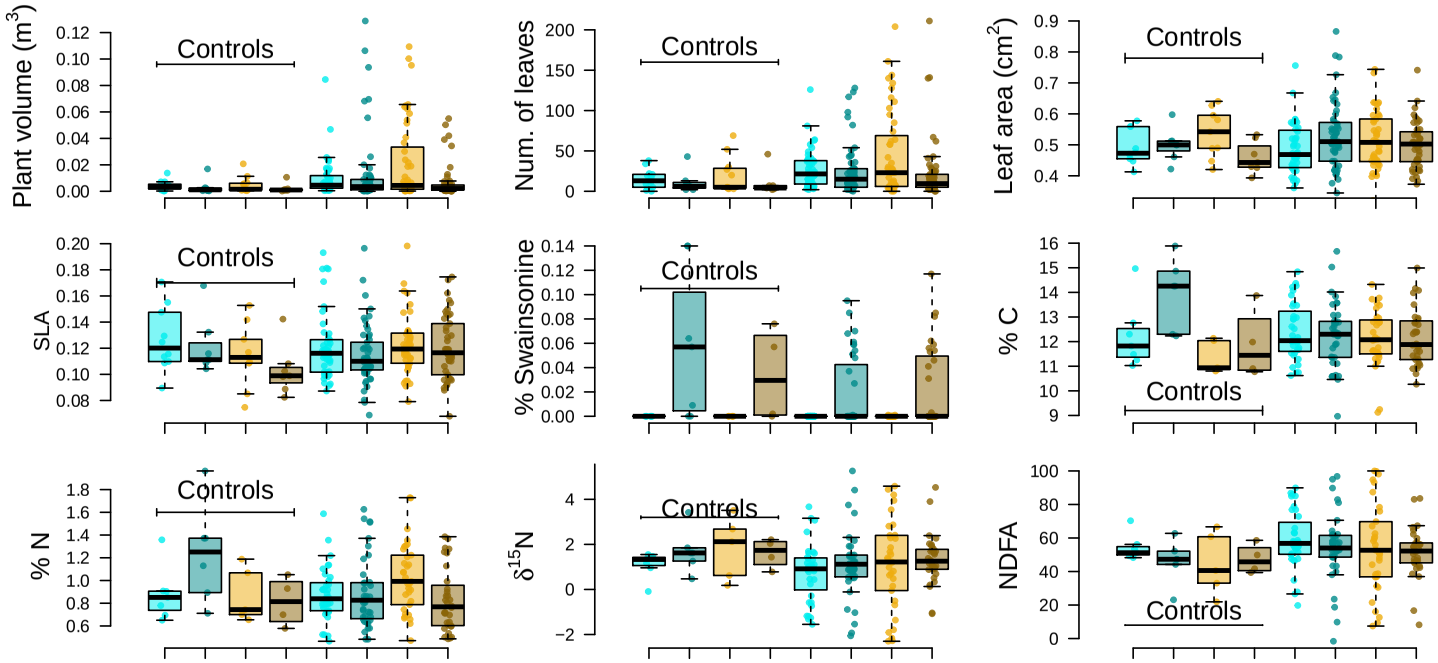

Figure S3: Influence of treatment on plant traits. This figures mirrors Fig. 1, except that control treatment groups are also shown. Control treatment was to plant embryos alongside agar chunks that differed in *A. fulva* colonization. % C and N is the % of dry mass of foliar tissue composed of either element. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles.

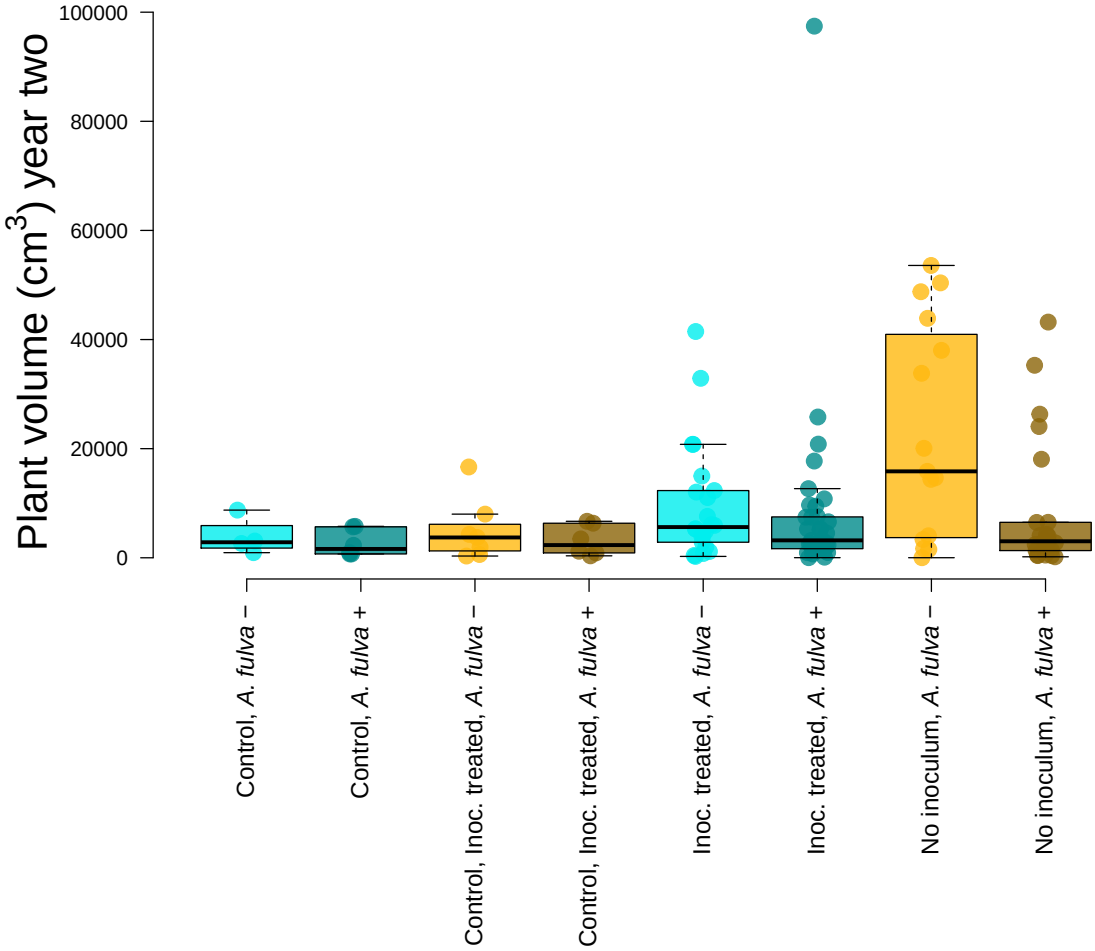

Figure S4: Size of *A. lentiginosus* plants in the year following initial experiment set up and sample collection. The smaller size of *A. fulva* colonized plants is apparent and mirrors what was observed during the year the experiment was conducted. Control treatment was to plant embryos alongside agar chunks that differed in *A. fulva* colonization. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles. Plus symbols refer to plants colonized by *A. fulva* and minus symbols refer to plants that were treated to reduce *A. fulva*. Inoculation treatment is denoted by blue shading and lack of inoculation by yellow.

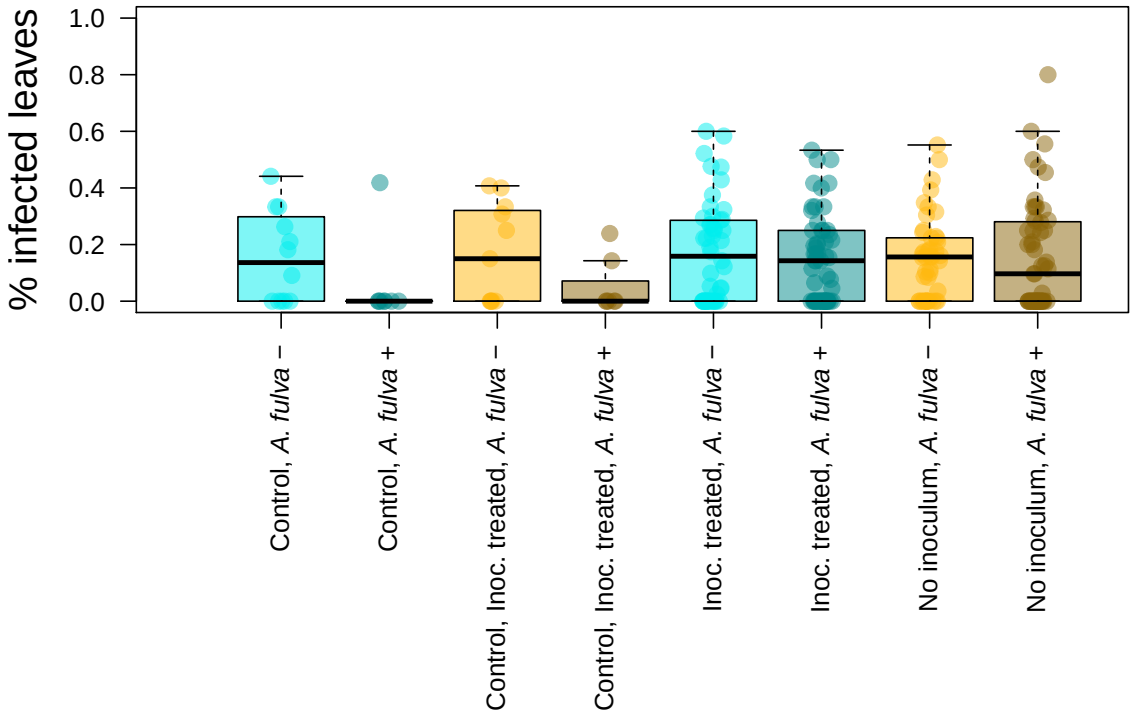

Figure S5: Differences in powdery mildew infection among treatment groups. The proportion of infected leaves observed is shown on the y-axis. Infection was determined visually (a whitish growth on leaf surfaces). Control treatment was to plant embryos alongside agar chunks that differed in *A. fulva* colonization. Boxplots summarize the data and describe interquartile range with a horizontal line denoting the median. Whiskers extend to the 10th and 90th percentiles. Posterior probability distributions (PPDs) of the mean difference between groups differed for control seedlings with high certainty (>80% of the PPDs did not overlap). Much more overlap was observed for non-control plants, though the directionality of effect was consistent for all treatment groups. Specifically, *A. fulva* colonization was associated with reduced mildew presence. Plus symbols refer to plants colonized by *A. fulva* and minus symbols refer to plants that were treated to reduce *A. fulva*. Inoculation treatment is denoted by blue shading and lack of inoculation by yellow.