PNAS
www.pnas.org

Supplementary Information for

Evidence of within-species specialization by soil microbes and the implications for plant community diversity

Jenalle L. Eck, Simon M. Stump, Camille S. Delavaux, Scott A. Mangan, and Liza S. Comita

Corresponding author: Jenalle L. Eck
Email: jenalle.eck@ieu.uzh.ch

This PDF file includes:

Supplementary text
Figs. S1 to S10
Tables S1 to S8
Captions for Datasets S1 to S20
References for SI citations

Other supplementary materials for this manuscript include the following:

Datasets S1 to S20

www.pnas.org/cgi/doi/10.1073/pnas.1810767116

Supplementary Text

Supplementary Methods 1: Extended Methods and Results for the Shadehouse Experiment

Seed and soil collection

Spatial distances between pairs of adult trees in the experiment range from ~25 m to ~1.6 km (*SI Appendix*, Fig. S6). Maternal trees were large, ~50-100 cm DBH (diameter at breast height). Most seeds were collected with the arils still on (meaning they were not dispersed by frugivores). We avoided collecting any seeds that showed signs of having been defecated (and thus dispersed). All soil was collected in September 2014.

Experimental process and design

Full replication of all eleven maternal seed sources in all sixteen soil inocula was not possible due to limited seedling availability. The number of replicate seedlings in each treatment varies among maternal seed sources, as seedling availability was limited by seed production and relatively low germination of field-collected seeds in the shade-house. Within each maternal seed source, seedlings were assigned at random to the soil inoculum of several non-parent conspecific trees and to the soil inoculum of several heterospecific trees. Initial dry biomass was estimated for each experimental seedling based on height at the time of transplant using an allometric linear regression model ($F_{(1,42)} = 338.1$, $p < 0.001$, $R^2 = 0.887$). This allometric model was built based on measurements of height and dry biomass of a randomly harvested sample of the potential

experimental seedlings at the beginning of the experiment. During harvest, roots were washed gently under running water to remove soil.

Quantification and analysis of colonization by AMF in shadehouse seedlings

To quantify colonization by arbuscular mycorrhizal fungi (AMF) in the seedlings that survived to the end of the experiment, we used a magnified root intersect method (51). To prepare roots for quantification, we used a modified root clearing and staining protocol (51-54). For each seedling, we cleared a random 0.2 mg subsample of the seedling's fine root mass in 10% bleach, 10% KOH, and 1% HCL. We then stained the subsamples using direct blue stain and mounted them on glass slides using lactic acid. We quantified the colonization of AMF in the subsamples at 200x magnification under a compound light microscope by recording the presence or absence of visible AMF structures (hyphae, vesicles, arbuscules, or arbuscular coils) in a minimum of 35 root intersects per seedling (see *SI Appendix,* Fig. S8 for a micrograph of the structures quantified). We obtained high-quality AMF colonization data for 155 of the 192 surviving experimental seedlings. We calculated AMF colonization for each of these seedlings as the proportion of root intersects quantified in the subsample that contained any visible AMF structure. As a more conservative estimate of AMF colonization, we also calculated the proportion of root intersects that contained visible arbuscules only.

We analyzed whether AMF colonization varied among the three experimental soil microbial treatments (maternal soil, non-parent conspecific soil, and heterospecific soil) using a generalized linear mixed-effects model. In this model, initial seedling dry biomass was included as a covariate, and maternal seed source, soil microbial inoculum

source, and shadehouse bench were included as random effects. Overall, colonization by AMF was similar between heterospecific soil microbial inocula and conspecific soil microbial inocula (*SI Appendix*, Fig. S3 and Table S3).

We then analyzed the effect of AMF colonization on seedling total dry biomass at harvest using a separate linear mixed-effects model. In this model, initial seedling dry biomass was included as a covariate, and maternal seed source, soil microbial inoculum source, and shadehouse bench were included as random effects. We found that AMF colonization was not a significant predictor of seedling biomass (*SI Appendix*, Fig. S2 and Table S2). Our conservative AMF colonization model (i.e. the model including arbuscules only as the measure of AMF colonization) yielded similar results (*SI Appendix*, Table S7), therefore we focus here on the results of the model utilizing data on all visible structures.

Quantification and analysis of nutrients in the soil inocula in the shadehouse experiment

Our shadehouse experiment was designed to minimize soil nutrient differences among soil microbial inoculum sources (i.e. by using only 20% by volume of field soil inoculum and 80% by volume of a common growing medium in each pot). However, because soil variables have previously been shown to be a key driver of tree species distributions and seedling growth in central Panama (55, 56), we also assessed the effect of soil nutrients on seedling biomass. At the time of harvest (April 2015), a sample of the soil surrounding the roots of each surviving seedling was collected, refrigerated, and stored. In June 2016, a small amount of soil from each of the samples in an inocula was combined and homogenized to create one 50 g soil sample per inocula. These samples

were air dried and taken to the Soils Lab at Smithsonian Tropical Research Institute (Panama) for nutrient quantification. Soil pH (in water and $CaCl_2$), soil phosphorous (total P mg/soil kg and plant-available P, Bray 1-P mg/soil kg), soil organic matter (% loss on ignition), soil carbon (% total C), soil nitrogen (% total N), and soil C:N ratio were quantified for each of the inocula.

To reduce the dimensionality of our soil nutrient variables before using them in experimental analyses, we first standardized all variables, and then conducted a principal component analysis (PCA). Seven of the eight soil variables loaded on the first principal component (PC1), which explained 50.2% of the variance; the second principal component (PC2) also loaded seven variables and explained an additional 28.1% of variance (*SI Appendix*, Fig. S9 and Table S8). We used PC1 as the response variable in a linear regression model to test whether soil nutrients varied among seedlings in the three experimental soil microbial treatments (maternal soil, non-parent-conspecific soil, and heterospecific soil) (*SI Appendix*, Table S4). To test whether soil nutrients affected seedling biomass, we then performed principal components regression (PCR) by building a linear mixed-effects model with seedling dry biomass at harvest as the response variable, with PC1 as a fixed effect, initial seedling biomass as a covariate, and maternal seed source and shadehouse bench as random effects. Soil inoculum source was not included as a random effect in this model because our soil variables were measured at the inocula level (i.e., all seedlings planted in a given inocula received the same nutrient values). Our model revealed that soil variables are not a significant predictor of seedling biomass (*SI Appendix*, Table S5). In addition, because nitrogen measurements might not be meaningful after soil storage, we also performed an identical PCA and PCR as

described above, but after excluding our nitrogen-related variables (% total N and C:N ratio). PCA and PCR results were similar when nitrogen-related variables were excluded vs. included, so we only present here the results of the PCR including nitrogen, utilizing all available soil information.

Analysis of seedling growth in heterospecific soil microbial communities

We analyzed whether seedling total dry biomass at harvest differed among seedlings planted in the soil microbial communities of the five heterospecific species in the study (simultaneously with seedlings planted in maternal soil and in non-parent conspecific soil microbial communities) using a linear mixed-effects model (*SI Appendix*, Table S6). In this model, initial seedling dry biomass was included as a covariate, and maternal seed source and shadehouse bench were included as random effects.

Supplementary Methods 2: Extended Methods & Results for the Plant Community Simulation

Our model is based on Stump and Chesson (26) and examines a community of n plant species. Each time step, each adult produces Y seeds. Each adult keeps $Y(1–d)$ seeds at its site, where d is the fraction of seeds dispersed. The remaining Yd seeds disperse uniformly around the environment, landing under plant species in proportion to that species' density, $N_j(t)$ (for species j at time t). Seeds exposed to species-specific pathogens are attacked/infected with probability α_S, and seeds exposed to species- and genotype-specific pathogens are attacked with probability $\alpha_S+\alpha_G$. Conspecific seedlings and adults will share a genotype with probability p. Seeds will die unless they can recruit to be an adult in that time step. If an adult dies (which it does with probability δ), then a seed's probability of recruitment is inversely proportional to $C_j(x,t)$, the number of seeds at that site post-consumption (at site x, time t, in a site previously held by species j). Thus, an individual adult's expected contribution to the next time step through survival and reproduction, $\lambda_j(x,t)$ (species j, site x, time t), is

$$\lambda_j(x,t) = (1-\delta) + \delta Y(1-d)\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} + \delta Y\mathbf{E}\left[dN_j(t)p\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right] + \delta Y\mathbf{E}\left[dN_j(t)(1-p)\frac{1-\alpha_S}{C_j(x,t)}\right] + \sum_{k\neq j}\delta Y\mathbf{E}\left[dN_k(t)\frac{1}{C_k(x,t)}\right],$$

(1)

where $\mathbf{E}[z]$ is the spatial average of z across sites, the summation is over all heterospecific species k, and the number of seeds at a site is

$$C_j(x,t) = Y(1-d)(1-\alpha_S-\alpha_G) + YN_j(t)d(1-\alpha_S-p\alpha_G) + \sum_{k\neq j} YN_k(t)d\,.$$

(2)

This equation breaks fitness into survival, (1–δ), plus four terms for recruitment (sites held by the parent, non-parent conspecifics with the same genotype, non-parent conspecifics of a different genotype, and heterospecifics, respectively). The population-level growth rate, $\tilde{\lambda}_j$, is $\lambda_j(x,t)$ averaged across all individuals (48). As described in detail below, we then used this model to examine species coexistence and the evolution of seed dispersal in communities with and without genotype-specific pathogens.

Analysis of species coexistence

We tested how genotype-specific pathogens affect species coexistence using an invasion analysis (47, 57). We select one species to be the invader (indicated with subscript *i*) and set its density to 0. We then allow the *n*–1 other species (the "residents", subscript *r*) to come to equilibrium, and calculate $\tilde{\lambda}_{\mathrm{i}}$. The assumptions in our model make all species identical; therefore, every species will have the same invader growth rate, and all species will have the same dynamics and densities as residents. Previous results have suggested that similar results will hold if species are non-identical (26).

We determine which factors boost invader growth rates using the method of Chesson (48, 58). Each resident is at equilibrium (i.e., $\tilde{\lambda}_{\mathrm{r}} = 1$); therefore,

$$\tilde{\lambda}_{\mathrm{i}} = \tilde{\lambda}_{\mathrm{i}} - \frac{1}{n-1}\sum_{r \neq i}\left(\tilde{\lambda}_{\mathrm{r}} - 1\right) \tag{A. 1}$$

where the summation is over the *n*–1 residents. This can be taken a step further by partitioning $\tilde{\lambda}_{\mathrm{j}}$ as a series of additive terms, $\tilde{\lambda}_{\mathrm{j}} = 1 + A_j + B_j + \cdots$, each with some biological meaning. Using this partitioning, the above equation becomes (47, 57)

$$\tilde{\lambda}_{\mathrm{i}} = 1 + \left(A_i - \frac{1}{n-1} \sum_{r \neq i} A_r \right) + \left(B_i - \frac{1}{n-1} \sum_{r \neq i} B_r \right) + \cdots \tag{A. 2}$$

For example, we show below that one term in the partitioning is the fraction of seeds consumed by genotype-specific pathogens; that term is slightly lower for invaders, and therefore slightly boosts invader growth rates. Similarly, we show that some terms are the same for all species, and thus have no effect on stability.

To simplify our analysis, we used the perturbation analysis method described by Chesson (48, 58). If α_S and α_G both equal 0, then $\tilde{\lambda}_{\mathrm{j}} = 1$, for all species, regardless of density. If α_S and α_G are small, then $\tilde{\lambda}_{\mathrm{j}}$ will be approximately a linear function of α_S and α_G. In a technical sense, we say that these terms are $O(\alpha)$, and we ignore any terms that are smaller than this (e.g., $\alpha_S\alpha_G$ and α_G^2 are both $O(\alpha^2)$), and thus can be ignored. Computer simulations suggest that ignoring those terms causes minor errors when α_S and α_G are as big as 0.4 or 0.5, and that this method slightly underestimates the stabilizing effect of pathogens (*SI Appendix*, Fig. S10). However, the qualitative results remain the same. In the work below, we will write X≈Y if $X=Y+O(\alpha^2)$.

We first calculate the probability that an individual's seeds will recruit in their parent's site,

$$\mathbf{P}(\text{parent site}) = \delta(1-d)\frac{Y(1-\alpha_S-\alpha_G)}{C_j(x,t)}. \tag{A. 3}$$

The fraction will make calculations hard to manage, and interpretation challenging. However, we can simplify it by using a Taylor Series approximation around the point $\alpha_S = 0$, $\alpha_G = 0$,

$$\mathbf{P}(\text{parent site}) = \mathbf{P}(\text{parent site}) + \alpha_S \frac{\partial \mathbf{P}(\text{parent site})}{\partial \alpha_S} + \alpha_G \frac{\partial \mathbf{P}(\text{parent site})}{\partial \alpha_G} + O(\alpha^2)$$

$$= \delta(1-d)\left[1 - \alpha_S\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) - \alpha_G\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right] + O(\alpha^2).$$

(A. 4)

We will leave the effect of α_S and α_G on competition as $\partial C_j(x,t)/\partial \alpha_S$ and $\partial C_j(x,t)/\partial \alpha_G$ for now, as it will simplify calculations, and to separate the effect of predation on seed survival from its effect on competition.

A similar procedure to the one above can be used to approximate each other birth term. Substituting these values into equation (A.1), and ignoring any $O(\alpha^2)$ terms, we have that

$$\begin{aligned}\tilde{\lambda}_j \approx\ & (1-\delta) + \delta(1-d)\left[1 - \alpha_S\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) - \alpha_G\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right] \\ & + \delta d N_j(t) p\left[1 - \alpha_S\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) - \alpha_G\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right] \\ & + \delta d N_j(t)(1-p)\left[1 - \alpha_S\left(1 + \frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) - \alpha_G\left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right] \\ & + \sum_{k \neq j} \delta d N_k(t)\left[1 - \alpha_S\left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) - \alpha_G\left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right].\end{aligned}$$

(A. 5)

This simplifies to

$$\tilde{\lambda}_j \approx 1 + \delta \Bigg[-\alpha_S \Big(1 - d + dN_j(t)\Big) - \alpha_G \big(1 - d + dN_j(t)p\big) - (1-d)\left[\alpha_S \left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) + \alpha_G \left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right] - \sum_{\text{all } k} dN_k(t)\left[\alpha_S \left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right) + \alpha_G \left(\frac{1}{Y}\frac{\partial C_j(x,t)}{\partial \alpha_S}\right)\right]\Bigg], \quad \text{(A. 6)}$$

where the summation is over all species (including *j*). While complex, this splits the effect of pathogens into four terms. The first two terms represent the fraction of seeds consumed by species-specific and genotype-specific pathogens, respectively. A seed is exposed to species-specific pathogens if it does not disperse (with probability 1–*d*), or if it disperses to a site held by a non-parent conspecific (with probability $dN_j(t)$); therefore, a seed will be consumed by species-specific pathogens with probability $\alpha_S(1–d+dN_j(t))$. The result is similar for genotype-specific pathogens, except that dispersed seeds are only in danger if they disperse to a site held by a non-parent conspecific of the same genotype (which happens with probability $dN_j(t)p$). The third and fourth term represent how pathogens change the strength of competition experienced by a seed that does not disperse, and a seed that disperses, respectively.

Next, we use equation (A.3) to calculate the invader growth rates as resident-invader differences. The fourth term in (A.6) is the same for all species, and therefore it cancels out. Using the fact that residents are identical, and that their densities ($N_r(t)$) must sum to 1, this simplifies to

$$\tilde{\lambda}_i \approx 1 + \delta\left(\frac{d\alpha_S}{n-1} + \frac{dp\alpha_G}{n-1} - (1-d)\left[\frac{\alpha_S}{Y}\left(\frac{\partial C_i(x,t)}{\partial \alpha_S} - \frac{\partial C_r(x,t)}{\partial \alpha_S}\right) + \frac{\alpha_G}{Y}\left(\frac{\partial C_i(x,t)}{\partial \alpha_G} - \frac{\partial C_r(x,t)}{\partial \alpha_G}\right)\right]\right).$$

(A. 7)

The first two terms are intuitive: if a rare seed disperses then it will likely never encounter a conspecific, if a common seed disperses then it might encounter a conspecific, therefore the predation risk of an invader's seed is $d(\alpha_S+p\alpha_G)/(n–1)$ less than the predation risk for a common seed. The last term represents how pathogens alter the competition experienced by non-dispersed seeds.

The partial derivatives of $C_j(x,t)$ are

$$\frac{\partial C_j(x,t)}{\partial \alpha_S} = -Y\big(1 - d + dN_j(t)\big)$$

(A. 8)

and

$$\frac{\partial C_j(x,t)}{\partial \alpha_G} = -Y\big(1 - d + dN_j(t)p\big).$$

(A. 9)

Thus,

$$-\frac{\alpha_S}{Y}\left(\frac{\partial C_i(x,t)}{\partial \alpha_S} - \frac{\partial C_r(x,t)}{\partial \alpha_S}\right) = -\frac{\alpha_S d}{n-1}$$

(A. 10)

and

$$-\frac{\alpha_G}{Y}\left(\frac{\partial C_i(x,t)}{\partial \alpha_G} - \frac{\partial C_r(x,t)}{\partial \alpha_G}\right) = -\frac{\alpha_S dp}{n-1}.$$

(A. 11)

Therefore, invaders are harmed slightly due to increased competition in their sites, an effect previously noted in Stump and Chesson (26).

Combining everything, we find that an invader's growth rate will be

$$\tilde{\lambda}_i \approx 1 + \delta\left(\frac{d^2\alpha_S}{n-1} + \frac{d^2 p\alpha_G}{n-1}\right).$$

(A. 12)

Fig. 2 was generated by simulating this model, using code that was adapted from previous work (13, 26). A community with 10,000 sites was initiated. The invader was removed, and the remainder of the community was given 500 time-steps to come to equilibrium. The invader was introduced, occupying 0.2% of community biomass. Its density was held at between 0.1% and 0.5% of community biomass. The invader was given 500 time-steps to come to equilibrium, after which we recorded its growth. Growth, $\tilde{\lambda}_i(t)$, was calculated as the number of sites occupied at the end of the time step (before populations were changed) divided by the number of sites occupied at the start of the time step. We then averaged this value across 2000 time-steps to give the $\tilde{\lambda}_i$ estimate for the simulation. The invader growth rate in Fig. 2 was the average across 200 simulations.

We can use our simulation results to check our approximation. *SI Appendix*, Fig. S10 shows the simulation results compared against the expectation from approximation (A.12). Thus, our approximation is good, though slightly underestimates invader growth rates.

We then used our model to explore the implications of genotype-specific microbes for species coexistence under three additional scenarios: 1) if seedlings perform better in non-parent conspecific sites than in heterospecific sites, 2) if seedlings perform better in heterospecific sites than in non-parent conspecific sites, and 3) if seedlings

perform similarly in non-parent conspecific sites and heterospecific sites (as per our experimental finding). In all three scenarios, seedlings in parent soils perform worse than seedlings in either heterospecific or non-parent conspecific soils. We found that if non-parent conspecific sites have a positive effect relative to heterospecific sites, it creates a priority effect that causes one random species to become dominant (*SI Appendix*, Fig. S4*C*). If heterospecific sites have a positive effect on survival relative to non-parent conspecific sites, then we get stable coexistence where all species stay near an equilibrium (*SI Appendix*, Fig. S4*A*). If non-parent conspecific sites have a similar effect as heterospecific sites, then we get neutrality, which over time will look like a neutral model random walk, with species randomly gaining or losing abundance (*SI Appendix*, Fig. S4*B*).

Additionally, we examined a model where 100% of the pathogens are genotype-specific (i.e. where α_S=0). The results depend on the relatedness parameter (p), which represents the chance that two conspecific individuals share genotype-specific pathogens (*SI Appendix*, Fig. S5). In this case, the total stabilizing effect is proportional to relatedness: when relatedness is zero, the community behaves neutrally; when relatedness is high, the stabilizing mechanism increases.

Evolution of seed dispersal

We next show how pathogens affect the evolution of seed dispersal. We assume that every species is equivalent in an n-species community, and that all species are residents at equilibrium. We then calculate the selection coefficient on dispersal, $\partial\tilde{\lambda}_i/\partial d$. This can allow us to determine the direction of selection on d.

As we will show, pathogens will cause monotonic selection for increased dispersal. Thus, in the absence of any cost, *d* should evolve to 1. To make our analysis more interesting and realistic, we will assume that an individual's seed yield, *Y*, will be some (decreasing) function of *d*.

The growth rate of a given individual can be simplified to

$$\tilde{\lambda}_j(t) = (1-\delta) + \delta Y(1-d)\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} + \delta Y d\frac{1-\alpha_S N_j(t)-\alpha_G N_j(t)p}{C_r(x,t)},$$

(A. 13)

i.e., survival plus recruitment of non-dispersed seeds plus recruitment of dispersed seeds. We write $C_j(x,t)$ to indicate the number of seeds in the parent's site, and $C_r(x,t)$ as the number of seeds in a non-parent site. We assume that enough seeds are dispersed around the environment that if individuals change their own level of dispersal or seed-production, then it will only affect the number of seeds at their own site. With this assumption,

$$\frac{\partial\tilde{\lambda}_j(t)}{\partial d} = \delta\frac{\partial Y}{\partial d}(1-d)\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} - \delta Y\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} + \delta Y(1-d)\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d}$$
$$+ \delta\frac{\partial Y}{\partial d}d\frac{1-\alpha_S N_j(t)-\alpha_G N_j(t)p}{C_r(x,t)} + \delta Y\frac{1-\alpha_S N_j(t)-\alpha_G N_j(t)p}{C_r(x,t)}.$$

(A. 14)

We can simplify this slightly by noting that if species *j* is at equilibrium, then the first and third terms must sum to $1/Y$. Thus,

$$\frac{\partial \tilde{\lambda}_j(t)}{\partial d} = \frac{1}{Y}\delta\frac{\partial Y}{\partial d} - \delta Y\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} + \delta Y(1-d)\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d}$$

$$+ \delta Y\frac{1-\alpha_S N_j(t)-\alpha_G N_j(t)p}{C_r(x,t)}.$$

(A. 15)

Second, as we showed previously, we can approximate the second and fourth terms using the perturbation analysis from before,

$$\frac{1-\alpha_S-\alpha_G}{C_j(x,t)} \approx \frac{1}{Y}\left[-\alpha_S d\left(1-N_j(t)\right) - \alpha_G d\left(1-N_j(t)p\right)\right]$$

$$\frac{1-\alpha_S N_j(t)-\alpha_G N_j(t)p}{C_r(x,t)} \approx \frac{1}{Y}\left[\alpha_S(1-d)\left(1-N_j(t)\right) + \alpha_G(1-d)\left(1-N_j(t)p\right)\right]$$

(A. 16)

Thus, the growth rate simplifies further to

$$\frac{\partial \tilde{\lambda}_j(t)}{\partial d} = \frac{1}{Y}\delta\frac{\partial Y}{\partial d} + \delta\left(\alpha_S\left(1-N_j(t)\right) + \alpha_G\left(1-N_j(t)p\right)\right)$$

$$+ \delta Y(1-d)\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d}.$$

(A. 17)

We simplify the last term for the sake of mathematical rigor below; however, first we analyze what each of these terms mean. The first term represents how *d* affects yield. Simply, if dispersing seeds is costly, then there may be some selection to reduce seed dispersal. This is the only term that can be negative. The second term represents selection to avoid pathogens. A seed that does not disperse from its parent always encounters distance-responsive pathogens, but a seed that disperses may escape them; this term captures that effect. It will often be stronger for rare species, and for species with high

genetic diversity, because the chance of encountering a conspecific or individual with the same genotype will be low. Finally, the last term is selection to decrease competition for the parent's site. Simply, if a seed disperses away from its parent, then that is one less seed competing for that site, which slightly increases the probability that another non-dispersed seed can capture that site. As we show below, this term will be positive (selecting for increased dispersal) unless $\partial Y/\partial d$ is extremely large.

Here, we calculate the last term. First,

$$\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d} = -\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^2}\frac{\partial C_j(x,t)}{\partial d}\ .$$

(A. 18)

We assume that

$$C_j(x,t) = (1-d)Y(1-\alpha_s-\ \alpha_G) + d^*N_j(t)Y^*(1-\alpha_S-\ \alpha_G p) + d^*\left(1-N_j(t)\right)Y^*,$$

(A. 19)

where d^* is the d value of the rest of the community (and thus not under selection), and $Y^*=Y(d^*)$. Thus,

$$\frac{\partial C_j(x,t)}{\partial d} = (1-d)(1-\alpha_s-\ \alpha_G)\frac{\partial Y}{\partial d} - Y(1-\alpha_s-\ \alpha_G).$$

(A. 20)

This is a complicated quantity, so we again simplify it using the perturbation analysis method described above. The overall calculations are tedious but are mentioned here briefly. By taking the α_G derivatives of (A.18) and using the chain rule,

$$\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d}$$

$$\approx\left(-\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^2}\frac{\partial C_j(x,t)}{\partial d}\bigg|_{\alpha_G=\alpha_S=0}\right)$$

$$+\alpha_G\left(\frac{1}{\left(C_j(x,t)\right)^2}\frac{\partial C_j(x,t)}{\partial d}\bigg|_{\alpha_G=\alpha_S=0}\right)$$

$$+\alpha_G\left(-\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^2}\frac{\partial^2 C_j(x,t)}{\partial d\partial\alpha_G}\bigg|_{\alpha_G=\alpha_S=0}\right)$$

$$+\alpha_G\left(2\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^3}\frac{\partial C_j(x,t)}{\partial d}\frac{\partial C_j(x,t)}{\partial\alpha_G}\bigg|_{\alpha_G=\alpha_S=0}\right)$$

(A. 21)

Each term then simplifies as follows when α_G=α_S=0:

$$\left(-\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^2}\frac{\partial C_j(x,t)}{\partial d}\bigg|_{\alpha_G=\alpha_S=0}\right)=-\frac{1}{Y}\left((1-d)\frac{\partial Y}{\partial d}\frac{1}{Y}-1\right)$$

(A. 22)

$$\alpha_G\left(\frac{1}{\left(C_j(x,t)\right)^2}\frac{\partial C_j(x,t)}{\partial d}\bigg|_{\alpha_G=\alpha_S=0}\right)=\alpha_G\frac{1}{Y}\left((1-d)\frac{\partial Y}{\partial d}\frac{1}{Y}-1\right)$$

(A. 23)

$$\alpha_G\left(-\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^2}\frac{\partial^2 C_j(x,t)}{\partial d\partial\alpha_G}\bigg|_{\alpha_G=\alpha_S=0}\right)=\alpha_G\frac{1}{Y}\left((1-d)\frac{\partial Y}{\partial d}\frac{1}{Y}-1\right)$$

(A. 24)

$$\alpha_G\left(-2\frac{1-\alpha_S-\alpha_G}{\left(C_j(x,t)\right)^3}\frac{\partial C_j(x,t)}{\partial d}\frac{\partial C_j(x,t)}{\partial \alpha_G}\bigg|_{\alpha_G=\alpha_S=0}\right)$$

$$=-2\alpha_G\frac{1}{Y}\Big(1-d+dN_j(t)\Big)\left((1-d)\frac{\partial Y}{\partial d}\frac{1}{Y}-1\right)$$

(A. 25)

The α_S derivatives are essentially identical. Thus, combining these terms,

$$\frac{\partial\left(\frac{1-\alpha_S-\alpha_G}{C_j(x,t)}\right)}{\partial d}$$

$$\approx\left(1+2\alpha_G d\left(1-pN_j(t)\right)+2\alpha_s d\left(1-N_j(t)\right)\right)\frac{1}{Y}\left(1-(1-d)\frac{\partial Y}{\partial d}\frac{1}{Y}\right).$$

(A. 26)

We generated Fig. 5 through simulations, following the general framework of adaptive dynamics and evolutionary game theory (49, 50). We started by assuming that d = 0.4 for every species in the community and allowed the community to come to equilibrium. We calculated $\frac{\partial\tilde{\lambda}_j(t)}{\partial d}$ as

$$\frac{\partial\tilde{\lambda}_j(t)}{\partial d}\approx\frac{\left(\tilde{\lambda}_j(t)\big|_{d=d+0.01}\right)-\left(\tilde{\lambda}_j(t)\big|_{d=d-0.01}\right)}{0.02}.$$

(A. 27)

We then increased the value of d by an amount $0.1\frac{\partial\tilde{\lambda}_j(t)}{\partial d}$, and repeated the process. The community tended to converge after repeating this process a few hundred times, so we repeated it 1000 times when generating Fig. 5.

Although simple, our dispersal kernel captures two important processes motivating our model: 1) some seeds stay near their parent and 2) seeds of rare species are less likely to encounter conspecifics than seeds of common species. Simple, spatially-implicit models that use the dispersal kernel presented here have previously been shown to re-produce the qualitative results of more complex, spatially-explicit models. Specifically, Stump & Chesson (26) produced qualitatively the same results as Adler & Muller-Landau (59), Muller-Landau & Adler (27), and Murrell (28), and Stump & Comita (34) produced qualitatively the same results as Mack & Bever (59).

Supplementary Figures

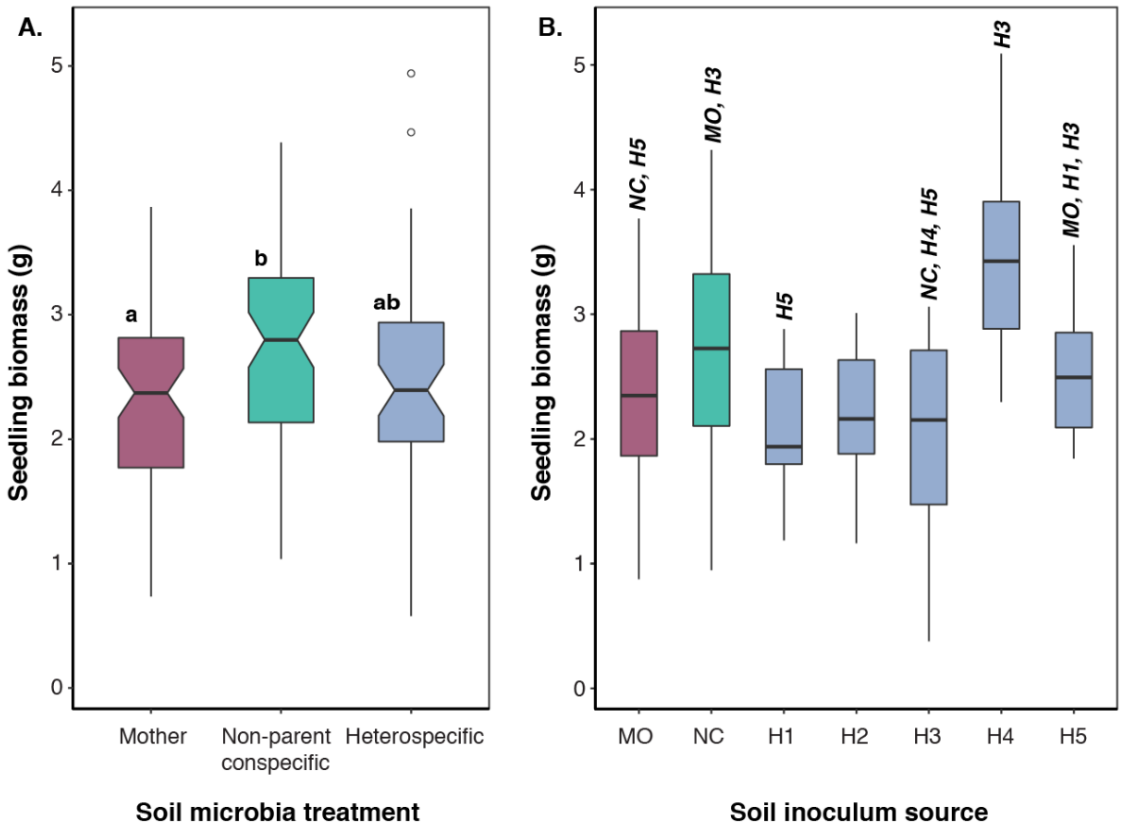

Fig. S1: Seedling biomass in the experimental soil microbial treatments. (*A*) Seedling biomass was similar between heterospecific and conspecific soil microbial treatments overall (*SI Appendix*, Table S1). Differing letters indicate a difference in predicted biomass between soil microbial treatment groups ($p < 0.05$; N = 192). (*B*) Seedling biomass varied in the five heterospecific soil microbial inocula in the study (*SI Appendix*, Table S6). MO = Mother soil treatment, NC = non-parent conspecific soil treatment, H1 = *Spondias mombin*, H2 = *Ormosia macrocalyx*, H3 = *Anacardium excelsum*, H4 = *Platypodium elegans*, H5 = *Protium tenuifolium*. Acronyms above the boxes indicate the soil inoculum sources for which seedling biomass varied significantly from the indicated group (e.g., the mother soil treatment varied significantly from the non-parent conspecific treatment and from H5; $p < 0.05$). In both panels, box belts represent the predicted median biomass for seedlings in each group. In (*A*), box notches represent a roughly 95% confidence interval for comparing predicted medians. In both panels, box hinges correspond to the first and third quartiles of predicted biomass, and box whiskers extend to the largest and smallest value no further than 1.5 × the interquartile range from hinges.

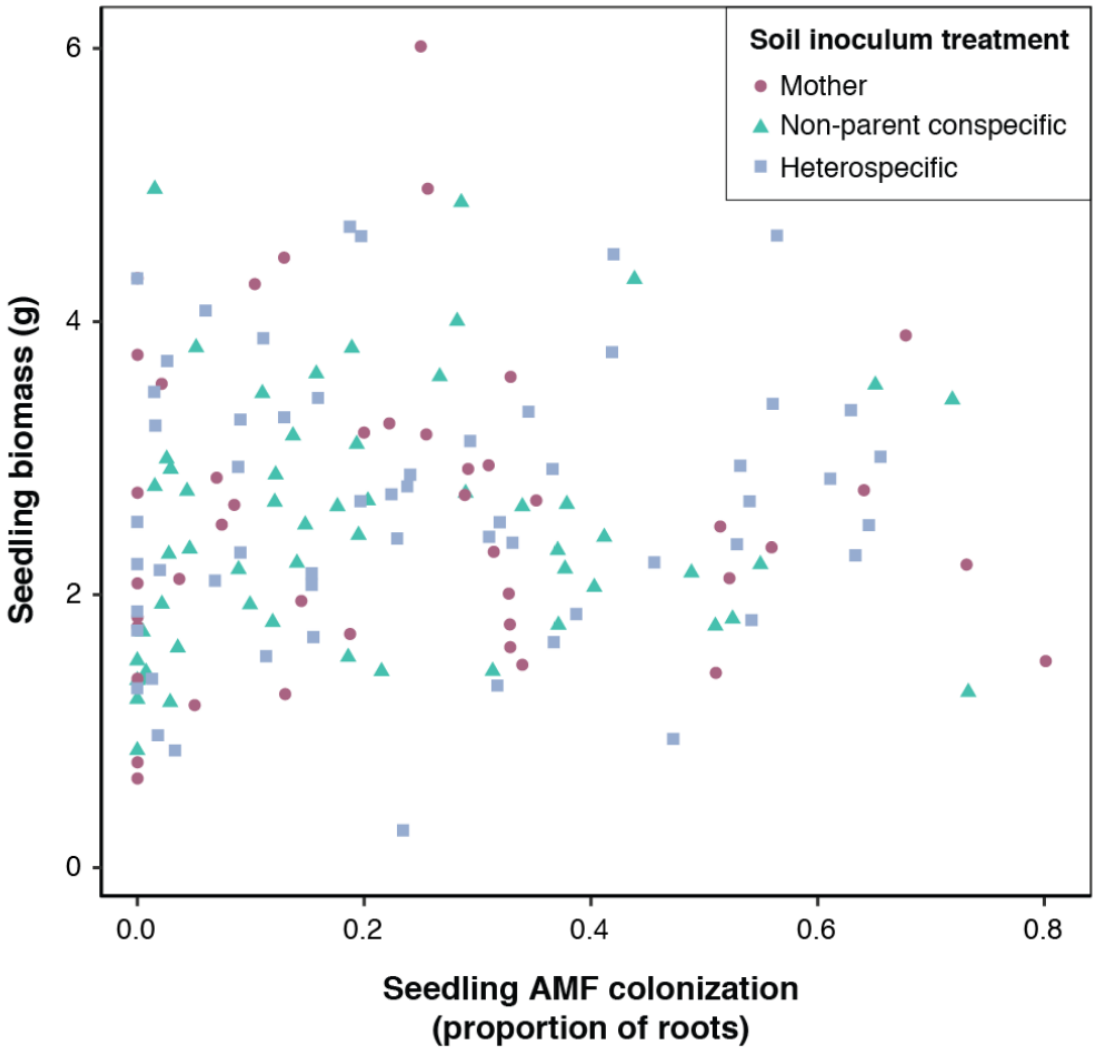

Fig. S2: Seedling biomass in the shadehouse as a function of colonization by arbuscular mycorrhizal fungi. The total above- and below-ground biomass of *V. surinamensis* seedlings in the shadehouse experiment is plotted as a function of their proportion colonization by arbuscular mycorrhizal fungi (AMF) (N = 155). A linear mixed-effects model showed that seedling biomass did not vary as a function of colonization by AMF (*SI Appendix*, Table S2).

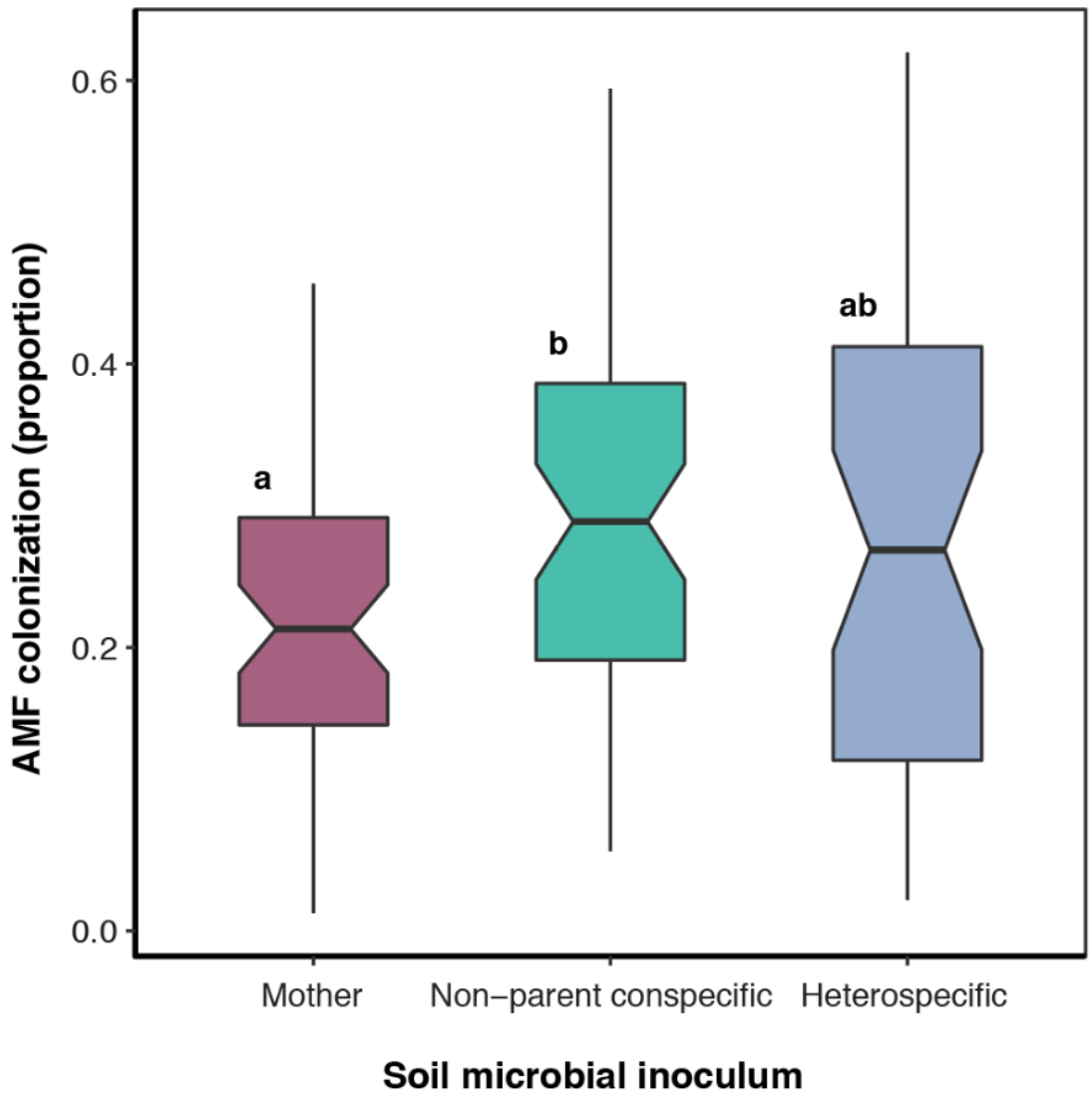

Fig. S3: Seedling colonization by arbuscular mycorrhizal fungi in the experimental soil treatments. Seedling colonization by arbuscular mycorrhizal fungi differed between the mother and non-parent conspecific soil microbial treatments but was similar between heterospecific and conspecific soil microbial treatments. Differing letters indicate a significant difference in predicted colonization between groups of seedlings ($p < 0.05$; N = 155; *SI Appendix*, Table S3). Box belts represent the predicted median colonization rate for seedlings in each group, while box notches represent a roughly 95% confidence interval for comparing predicted medians. Box hinges correspond to the first and third quartiles, and box whiskers extend to the largest and smallest value no further than 1.5 × the interquartile range from hinges.

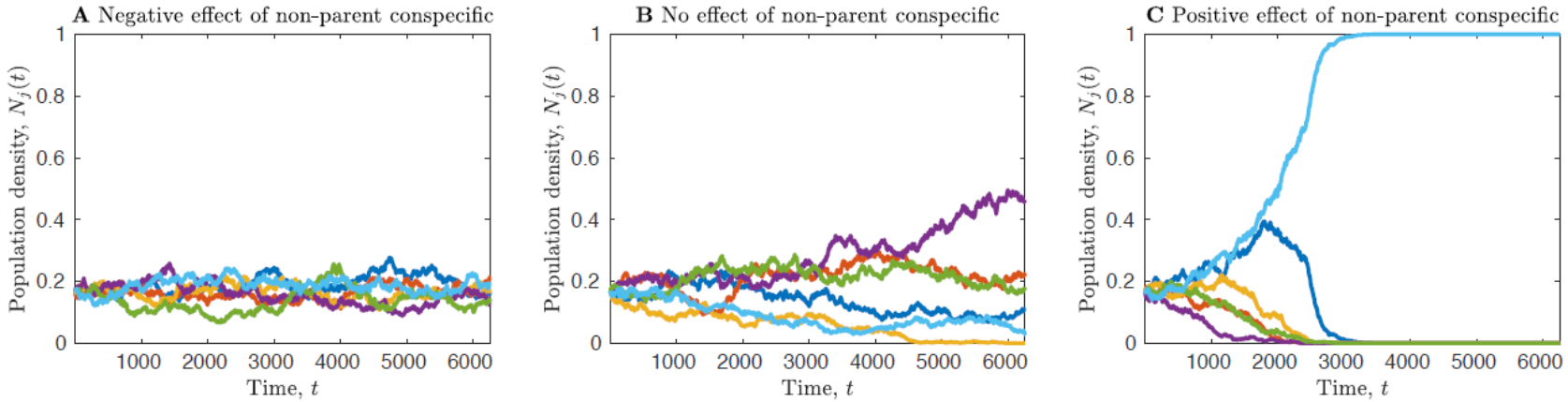

Fig. S4: Three possible effects of a non-parent conspecific on seedling survival. (*A*) Non-parent conspecifics may reduce seedling growth (i.e. non-offspring conspecific seedlings grow worse than heterospecific seedlings). In this case, genotype-specific soil microbes promote plant species coexistence. (*B*) Non-parent conspecifics may have no impact on seedling growth and survival (i.e. non-offspring conspecific seedlings grow the same as heterospecific seedlings). In this case, genotype-specific soil microbes have no impact on coexistence, and all else being equal, species will neutrally co-occur. (*C*) Non-parent conspecifics may increase seedling growth (i.e. non-offspring conspecific seedlings grow better than heterospecific seedlings). In this case, genotype-specific soil microbes create a destabilizing mechanism, such that the first species to become abundant will rapidly exclude all others. Parameters: 6 species, d=0.655, δ=0.4, α_G=0.3, and p=0. In (*A*) α_S=0.1, in (*B*) α_S=0, in (*C*) α_S=–0.1.

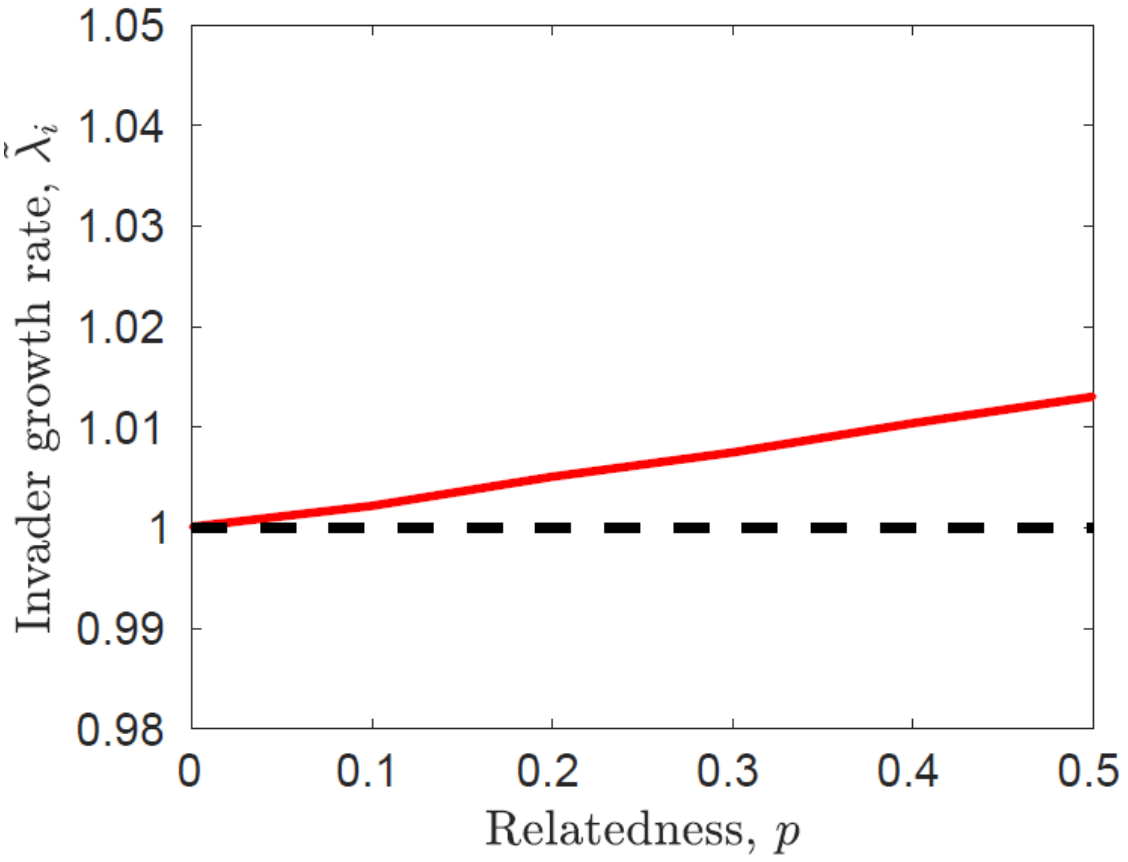

Fig. S5: Impact of relatedness in a community without species-specific pathogens. Here we show how the invader growth rate, $\tilde{\lambda}_i$ (a measure of community stability), changes as conspecifics share a greater degree of genotype-specific pathogens, p. When p=0, seedlings only share pathogens with their parents, and the community is neutral. When p is large, then many conspecific individuals share genotype-specific pathogens, and the community is stabilized. Parameters: 4 species, d=0.655, δ=0.4, α_G=0.35, and α_S=0.

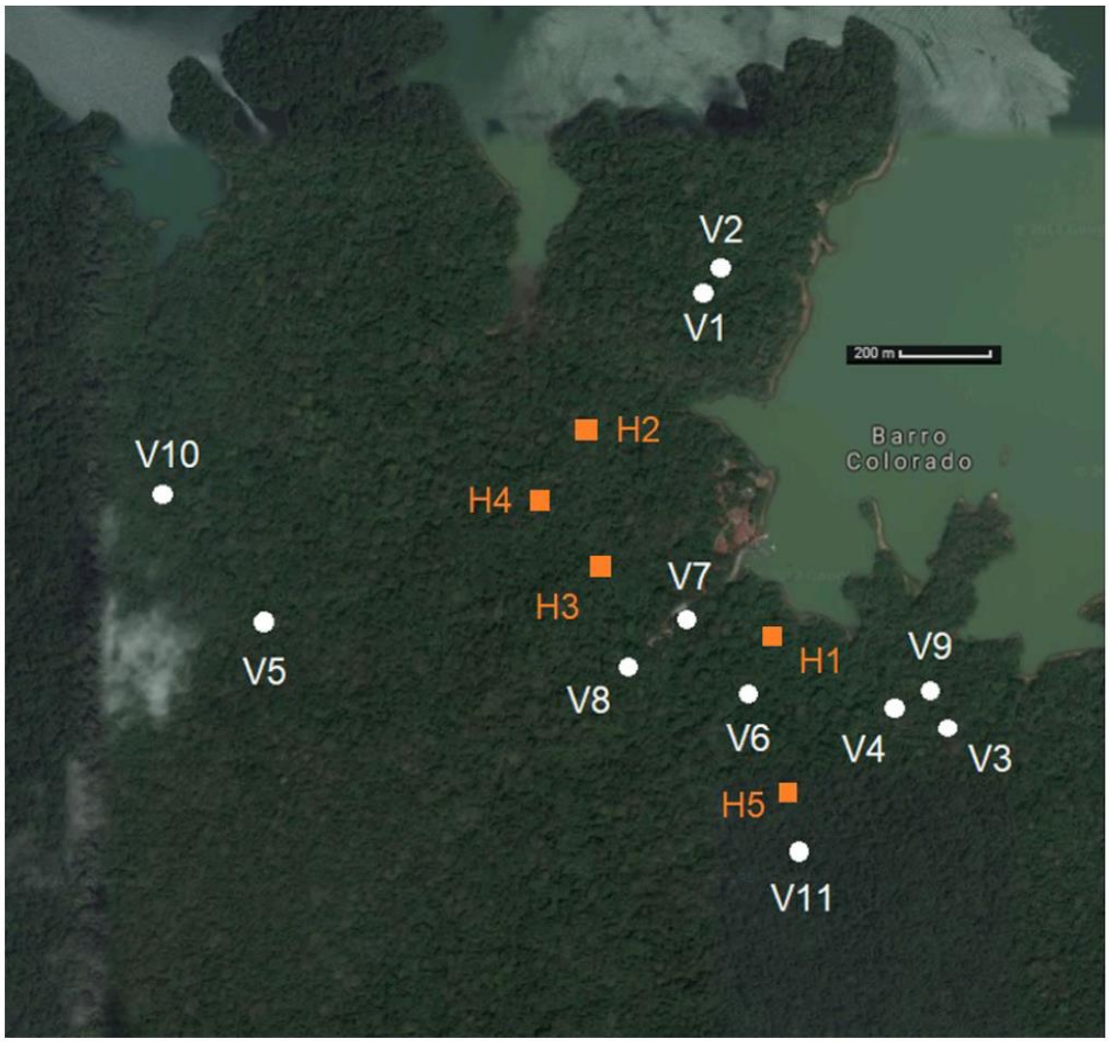

Fig. S6: A map of the location on Barro Colorado Island, Panama of the adult trees used as sources of seed and/or soil microbial inocula in the shadehouse experiment. The approximate locations of the eleven maternal *V. surinamensis* trees are shown as white circles (V1 through V11). The approximate locations of the five heterospecific trees that provided soil microbial inocula are shown as orange squares (H1 through H5). H1 = *Spondias mombin*, H2 = *Ormosia macrocalyx*, H3 = *Anacardium excelsum*, H4 = *Platypodium elegans*, H5 = *Protium tenuifolium*. Satellite image obtained from Google Maps ® 2017.

Fig. S7: *Virola surinamensis* seedlings growing in their experimental soil microbial treatments in a shadehouse on Barro Colorado Island, Panama. Seedlings were planted in pots containing soil microbial inoculum from one of the following sources: the seedling's mother tree, an adult in the *V. surinamensis* population that was not the seedling's parent, or a heterospecific tree within the study area.

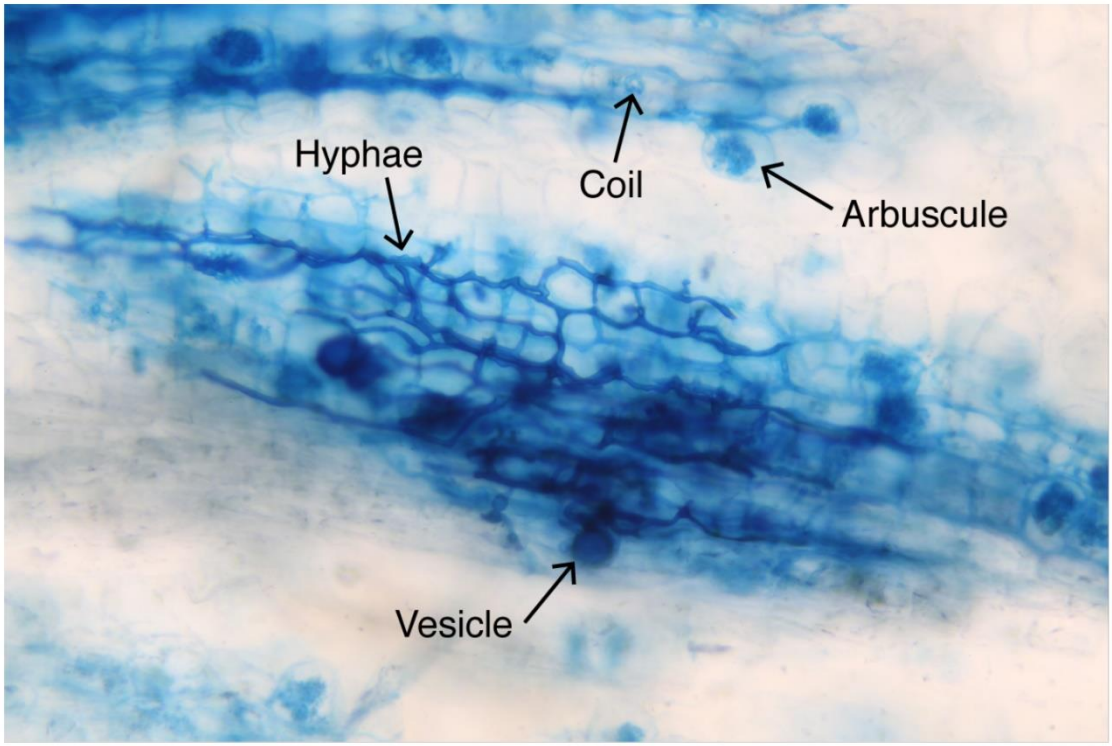

Fig. S8: A micrograph of fungal structures in the root of a *V. surinamensis* seedling. A photograph taken at 200x magnification shows examples of the four structures quantified in the arbuscular mycorrhizal fungi study (hyphae, arbuscules, vesicles, and coils), stained blue inside the root of an experimental seedling.

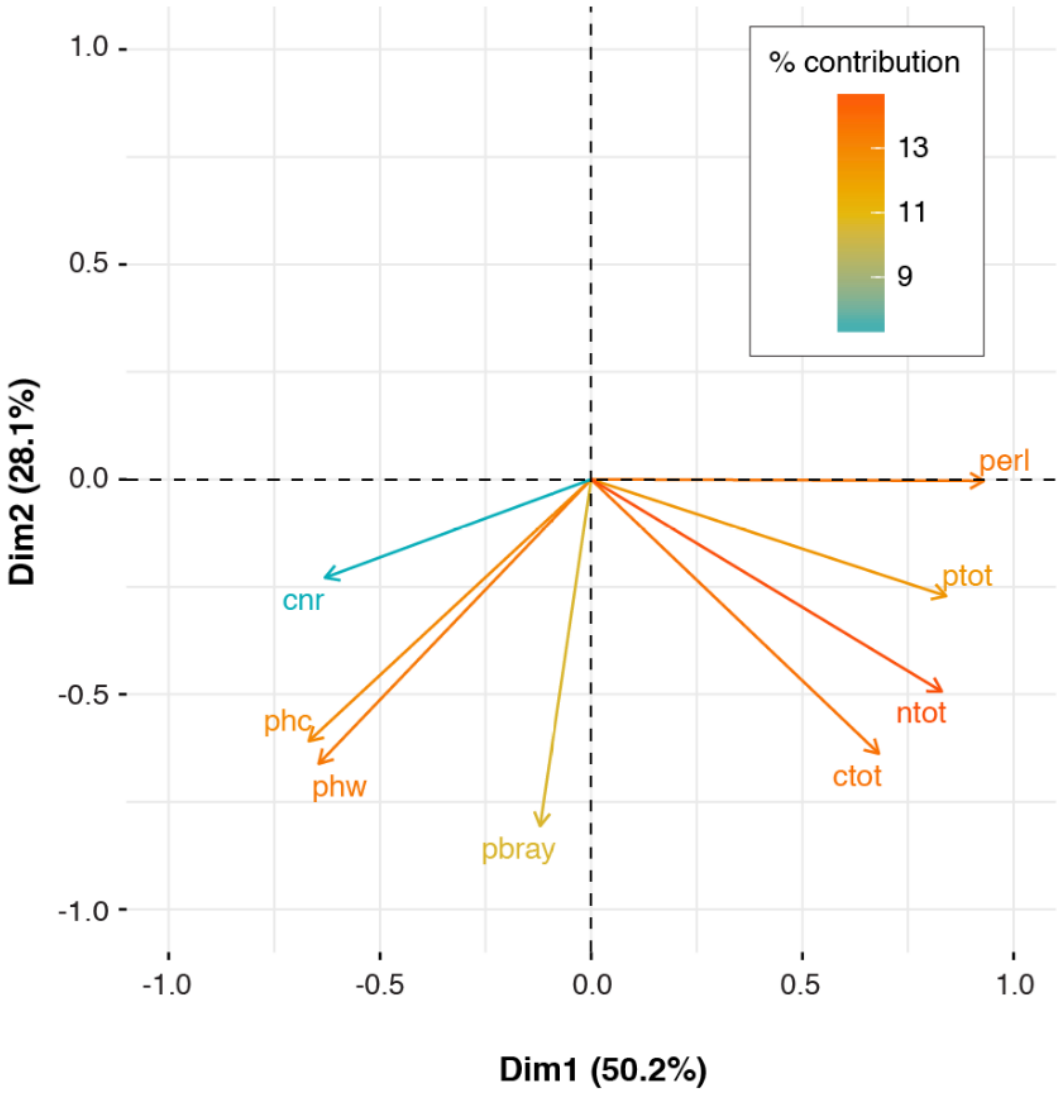

Fig. S9: PCA of soil nutrient variables. Soil nutrient variable abbreviations are as follows: cnr = C:N ratio, phc = pH in $CaCl_2$, phw = pH in water, pbray = plant-available P (Bray 1-P mg/soil kg), ctot = % total C, ntot = % total N, ptot = total P (mg/soil kg), and perl = organic matter (% loss on ignition).

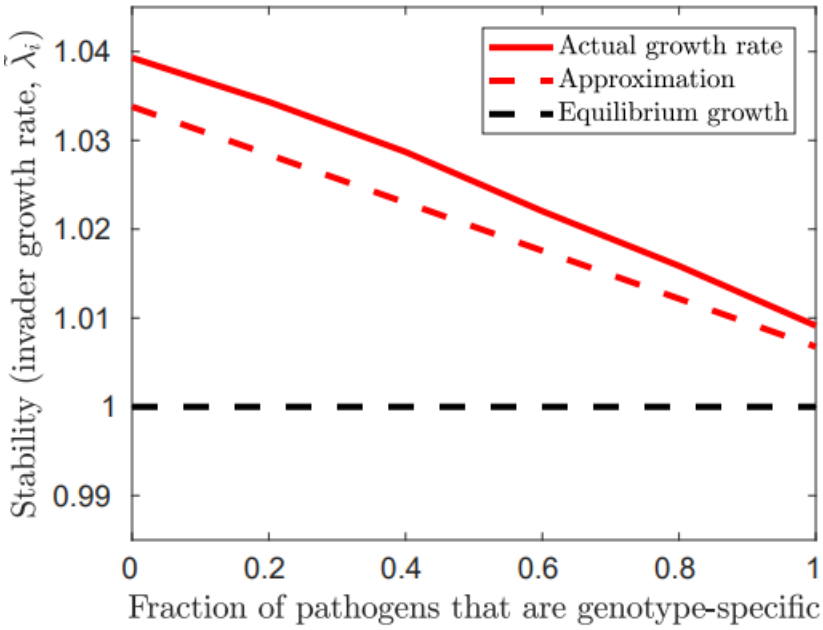

Fig. S10. Accuracy of approximations. We display the results of simulations (Fig. 2) next to the predictions based on equation (A.12). The fact that the lines are so close suggests that the approximation is nearly accurate, but slightly underestimates the stabilizing effect of pathogens.

Supplementary Tables

Table S1: Mixed-model and ANCOVA summary of seedling biomass in each experimental soil microbial treatment. We analyzed the biomass of *V. surinamensis* seedlings at the end of the shadehouse experiment as a function of their soil microbial inoculum treatment (i.e., maternal soil, non-parent conspecific soil, or heterospecific soil) and initial biomass using a linear mixed-effects model. Seedlings in maternal soil microbial inoculum had significantly lower biomass than seedlings in non-parent conspecific soil microbial inoculum. Seedlings in heterospecific soil microbial inoculum had similar biomass as seedlings in the two conspecific treatments.

Fixed effects	**β**	***SE***	***p***
Intercept (maternal soil)	0.946	0.392	0.062
Non-parent conspecific soil	0.267	0.121	**0.027**
Heterospecific soil	0.106	0.160	0.516
Initial seedling biomass	3.712	0.365	**< 0.001**

Random effects	***var(X)***
Seed source	0.084
Soil source	0.033
Shadehouse bench	0.468
Residual	0.466

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
Soil treatment	2.33	1.16	2	25.23	2.50	0.10
Initial seedling biomass	48.24	48.24	1	182.28	103.57	**< 0.001**

Table S2: Mixed-model and ANCOVA summary of seedling biomass in the shadehouse experiment as a function of colonization by AMF. The biomass of *V. surinamensis* seedlings did not vary as a function of their colonization by AMF.

Fixed effects	**β**	***SE***	***p***
Intercept	1.38	0.44	0.03
AMF colonization	-0.22	0.27	0.43
Initial seedling biomass	3.14	0.39	**< 0.001**

Random effects	***var(X)***
Maternal seed source	0.11
Soil inoculum source	0.03
Shadehouse bench	0.59
Residual	0.39

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
AMF colonization	0.25	0.25	1	139.92	0.62	0.43
Initial seedling biomass	25.3	25.3	1	144.19	62.26	**< 0.001**

Table S3: Mixed-model and ANCOVA summary of seedling colonization by arbuscular mycorrhizal fungi in each experimental soil microbial treatment. We analyzed the colonization by arbuscular mycorrhizal fungi (AMF) of *V. surinamensis* seedlings at the end of the shadehouse experiment as a function of their soil inoculum treatment (i.e., maternal soil, non-parent conspecific soil, or heterospecific soil) and initial biomass using a generalized linear mixed-effects model (*SI Appendix*). Seedlings in maternal soil inoculum had lower colonization by AMF than seedlings in non-parent conspecific soil inoculum. Seedlings in heterospecific soil inoculum had similar colonization by AMF as seedlings in the two conspecific treatments.

Fixed effects	**β**	***SE***	***p***
Intercept (maternal soil)	-2.010	0.435	< 0.001
Non-parent conspecific soil	0.604	0.059	**< 0.001**
Heterospecific soil	0.079	0.557	0.887
Initial seedling biomass	1.581	0.174	**< 0.001**

Random effects	***var(X)***
Seed source	0.230
Soil source	1.043
Shadehouse bench	0.261

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
Soil treatment	98.98	49.49	2	29.64	49.49	**< 0.001**
Initial seedling biomass	82.64	82.64	1	141.76	82.64	**< 0.001**

Table S4: Soil nutrients varied between heterospecific and conspecific soil treatments in the shadehouse experiment but were similar between conspecific treatments.

Fixed effects	**β**	***SE***	***p***
Intercept (maternal soil)	0.22	0.23	0.33
Non-parent conspecific soil	0.38	0.32	0.25
Heterospecific soil	-1.26	0.34	**< 0.001**

ANCOVA	***SS***	***MS***	**Num *df***	***F***	***p***
Soil treatment	87.18	43.58	2	12.10	**< 0.001**
Initial seedling biomass	680.92	3.60	189		

Table S5: Mixed-model ANCOVA of seedling biomass as a function of soil nutrients. A linear mixed effects model revealed that the biomass of *V. surinamensis* seedlings in the shadehouse experiment did not vary as a function of the first principal component of soil nutrients.

Fixed effects	**β**	***SE***	***p***
Intercept	1.01	0.39	0.05
Soil nutrients	-0.05	0.03	0.09
Initial seedling biomass	3.84	0.37	**< 0.001**

Random effects	***var(X)***
Maternal seed source	0.09
Shadehouse bench	0.48
Residual	0.50

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
Soil nutrients	1.46	1.46	1	183.49	2.94	0.09
Initial seedling biomass	53.68	53.68	1	183.66	108.2	**< 0.001**

Table S6: Mixed-model and ANCOVA summary of seedling biomass in five heterospecific soil microbial inocula relative to conspecific soil treatments. We analyzed the biomass of *V. surinamensis* seedlings at the end of the shadehouse experiment as a function of their soil inoculum treatment (i.e., maternal soil, non-parent conspecific soil, or heterospecific species H1, H2, H3… etc.) and initial biomass using a linear mixed-effects model. H1 = *Spondias mombin*, H2 = *Ormosia macrocalyx*, H3 = *Anacardium excelsum*, H4 = *Platypodium elegans*, H5 = *Protium tenuifolium*. Seedling biomass varied among individual heterospecific soil inocula and between individual heterospecific soil inocula and conspecific soil treatments.

Fixed effects	**β**	***SE***	***p***
Intercept (maternal soil)	1.000	0.387	0.050
Non-parent conspecific soil	0.274	0.119	**0.023**
H1 soil	-0.150	0.229	0.513
H2 soil	0.073	0.211	0.732
H3 soil	-0.241	0.238	0.311
H4 soil	0.419	0.236	0.077
H5 soil	0.502	0.250	**0.046**
Initial seedling biomass	3.554	0.379	**< 0.001**

Random effects	***var(X)***
Seed source	0.082
Shadehouse bench	0.458
Residual	0.482

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
Soil treatment	6.73	1.12	6	173.68	2.330	**0.035**
Initial seedling biomass	42.32	42.32	1	178.53	87.82	**< 0.001**

Table S7: A more conservative model of seedling biomass as a function of colonization by AMF. In this model, we considered arbuscules only as our metric of AMF colonization. Like the model where all AMF structures were considered, final biomass of *V. surinamensis* seedlings in the shadehouse did not vary as a function of their colonization by AMF when only arbuscules were considered.

Fixed effects	**β**	***SE***	***p***
Intercept	1.36	0.43	0.03
AMF colonization (arbuscules only)	-0.29	0.49	0.55
Initial seedling biomass	3.10	0.39	**< 0.001**

Random effects	***var(X)***
Maternal seed source	0.11
Soil inoculum source	0.03
Shadehouse bench	0.58
Residual	0.39

ANCOVA	***SS***	***MS***	**Num *df***	**Den *df***	***F***	***p***
AMF colonization (arb.)	0.13	0.13	1	146.3	0.35	0.55
Initial seedling biomass	24.54	24.54	1	144.7	62.56	**< 0.001**

Table S8: **A summary of PCA of soil nutrients**. *Top:* a summary of the importance of components. *Bottom:* the loadings on each component. Soil variable abbreviations are as follows: cnr = C:N ratio, phc = pH in $CaCl_2$, phw = pH in water, pbray = plant-available P (Bray 1-P mg/soil kg), ctot = % total C, ntot = % total N, ptot = total P (mg/soil kg), and perl = organic matter (% loss on ignition).

Summary	**Comp. 1**	**Comp. 2**	**Comp. 3**	**Comp. 4**	**Comp. 5**	**Comp. 6**	**Comp. 7**	**Comp. 8**
Standard deviation	1.99	1.50	0.88	0.73	0.59	0.25	0.12	0.04
Proportion of variance	0.50	0.28	0.10	0.07	0.04	< 0.01	< 0.01	< 0.01
Cumulative variance	0.50	0.78	0.88	0.95	0.99	1.00	1.00	1.00

Soil variables	**Comp. 1**	**Comp. 2**	**Comp. 3**	**Comp. 4**	**Comp. 5**	**Comp. 6**	**Comp. 7**	**Comp. 8**
phw	-0.32	-0.44	-0.40	-0.15		0.10	0.71	
phc	-0.33	-0.41	-0.41	-0.23	0.17		-0.68	
pbray		-0.54	0.30	0.69		0.36	-0.10	
ptot	0.42	-0.18		0.14	0.74	-0.45	0.12	
perl	0.47		-0.14	-0.32	0.26	0.77		
ctot	0.34	-0.43	0.20	-0.31	-0.33	-0.17		-0.65
ntot	0.42	-0.33			-0.40	-0.17		0.72
cnr	-0.32	-0.15	0.72	-0.47	0.28			0.23

Captions for Datasets S1 to S20

Dataset S1. This dataset provides several types of data on the seedlings in the shadehouse experiment. First, data is provided related to the experimental design: seedling maternal seed source, soil inoculum source, soil treatment type (mother, non-parent conspecific, or heterospecific), shadehouse table, etc. Second, we provide growth data on the total dry biomass of the seedlings at the end of the experiment, as well as the estimate of initial dry biomass at the beginning of the experiment. Third, data on colonization by arbuscular mycorrhizal fungi in seedling roots is provided. Fourth, data on soil nutrient variables is provided.

Dataset S2. This dataset provides the R code used to analyze the shadehouse experiment and generate the figures related to the shadehouse experiment.

Dataset S3. This dataset is a README file pertaining to all datasets related to the simulation model (*SI Appendix,* Datasets S4 through S20). Please read for help using these datasets.

Dataset S4. This dataset is our simulation of the model.

Dataset S5. This dataset calculates the optimal dispersal fraction.

Dataset S6. This dataset is the dispersal-fecundity trade-off function.

Dataset S7. This dataset generates data for Figure 2.

Dataset S8. This dataset generates data for Figure S4 in the *SI Appendix*.

Dataset S9. This dataset generates data for Figure S5 in the *SI Appendix*.

Dataset S10. This dataset generates Figure 2.

Dataset S11. This dataset generates Figure 4.

Dataset S12. This dataset generates Figure 5.

Dataset S13. This dataset generates Figure S10 in the *SI Appendix*.

Dataset S14. This dataset generates Figure S4 in the *SI Appendix*.

Dataset S15. This dataset generates Figure S5 in the *SI Appendix*.

Dataset S16. This dataset contains the data used to generate Figure 2.

Dataset S17. This dataset contains the data used to generate Figure S4a in the *SI Appendix*.

Dataset S18: This dataset contains the data used to generate Figure S4b in the *SI Appendix*.

Dataset S19. This dataset contains the data used to generate Figure S4c in the *SI Appendix*.

Dataset S20. This dataset contains the data used to generate Figure S5 in the *SI Appendix*.

***SI Appendix*: References**

51. McGonigle TP, Miller MH, Evans DG, Fairchild GL, Swan JA (1990) A new method which gives an objective measure of colonization of roots by vesicular-arbuscular mycorrhizal fungi. *New Phytologist* 115:495-495.

52. Giovannetti M, Mosse B (1980) An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. *New Phytologist* 84:489-500.

53. Vierheilig H, Schweiger P, Brundrett M (2005) An overview of methods for the detection and observation of arbuscular mycorrhizal fungi in roots. *Physiol. Plant.* 125:393-404.

54. International Culture Collection of (Vesicular) Arbuscular Mycorrhizal Fungi (INVAM) (2015). Methods for measuring plant and mycorrhiza root lengths. http://invam.wvu.edu/methods/mycorrhizae/mycorrhiza-root-length

55. Condit R, Engelbrecht BMJ, Pino D, Perez R, Turner BL (2013) Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *PNAS* 110(13):5064-5068.

56. Zalamea P-C, et al. (2016) Seedling growth responses to phosphorus reflect adult distribution patterns of tropical trees. *New Phytologist* 212(2):400-408.

57. Chesson P (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology & Systematics* 31:343-363.

58. Chesson P (1994) Multispecies competition in variable environments. *Theoretical Population Biology* 45:227-276.

59. Adler FR, Muller-Landau HC (2005) When do localized natural enemies increase species richness? *Ecology Letters* 8:438-447.

59. Mack KML, Bever JD (2014) Coexistence and relative abundance in plant communities are determined by feedbacks when the scale of feedback and dispersal is local. *Journal of Ecology* 102:1195-1201.