Host-plant genetic variation dominates phenotypic plasticity in structuring above and belowground communities

Matthew A. Barbour[1,2,*]

Sonya Erlandson[3]

Kabir Peay[3]

Brendan Locke[4]

Erik S. Jules[5]

Gregory M. Crutsinger[5]

1. Department of Zoology, University of British Columbia, Vancouver, Canada.

2. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.

3. Department of Biology, Stanford University, Stanford, CA, USA.

4. Department of Biological Sciences, Humboldt State University, Arcata, CA, USA.

5. Pix4D Inc, San Francisco, CA, USA.

$*$ Corresponding author; e-mail: matthew.barbour@ieu.uzh.ch.

Manuscript elements: Figure 1, figure 2, figure 3, figure 4, table 1, online appendix A (including table A1, table A2, table A3, table A4, table A5, table A6, table A7, and table A8). Figure 3 and figure 4 are to print in color.

Keywords: genotype-by-environnment, community genetics, plant traits, phenotypic plasticity,

rhizosphere, intraspecific variation.

Manuscript type: Article.

Prepared using the suggested LaTeX template for *Am. Nat.*

Abstract

Host-plant genetic variation affects the diversity and composition of associated above and belowground communities. Most evidence supporting this view is derived from studies within a single common garden, thereby constraining the range of biotic and abiotic environmental conditions that might directly or indirectly (via phenotypic plasticity) affect communities. If natural variability in the environment renders host-plant genetic effects on associated communities unimportant, then studying the community-level consequences of genetic variation may not be warranted. We addressed this knowledge gap by planting a series of common gardens consisting of 10 different clones (genotypes) of the willow *Salix hookeriana* in a coastal dune ecosystem and manipulated natural variation in ant-aphid interactions (biotic) and wind exposure (abiotic) in two separate experiments. We then quantified the responses of associated species assemblages both above (foliar arthropods) and belowground (rhizosphere fungi and bacteria). In addition, we quantified plant phenotypic responses (plant growth, leaf quality, and root quality) to tease apart the effects of genetic variation, phenotypic plasticity, and direct environmental effects on associated communities. In the ant-aphid experiment, we found that willow genotype explained more variation in foliar arthropod communities than aphid additions and proximity to aphid-tending ant mounds. However, aphid additions modified willow genotype effects on arthropod community composition by attracting other aphid species to certain willow genotypes. In the wind experiment, wind exposure explained more variation than willow genotype in structuring communities of foliar arthropods and rhizosphere bacteria. Still, willow genotype had strong effect sizes on several community properties of arthropods and fungi, indicating that host-plant genetic variation remains important. Across both experiments, genetic variation in plant traits was more important than phenotypic plasticity in structuring associated communities. The relative importance of genetic variation vs. direct environmental effects though depended on the type of environmental gradient ($G > E_{aphid}$ but $E_{wind} > G$). Taken together, our results suggest that host-plant genetic variation is an important driver of above and belowground biodiversity,

despite natural variation in the biotic and abiotic environment.

Introduction

Intraspecific genetic variation is a key driver of trait variation within host plants, which in turn can have cascading effects on associated species and entire communities of organisms (Antonovics 1992; Fritz and Price 1988; Lamit et al. 2016; Maddox and Root 1990). For example, genetic variation in the leaf chemistry of cottonwoods (Whitham et al. 2006) and in the plant architecture of coyote bush (Crutsinger et al. 2014) structures diverse assemblages of species, from foliar arthropods aboveground to soil microbes below. While community-level consequences of genetic variation (commonly referred to as 'community genetics', sensu Antonovics 1992) have been documented in a variety of host-plant taxa (Whitham et al. 2012), evidence comes primarily from common garden experiments that minimize environmental variation. These controlled environments likely limit effects of the biotic and abiotic environment on the expression of host plant traits (Gratani 2014) as well as the diversity and composition of species assemblages ((Gaston, 2000; MacArthur, 1972)). Therefore, the importance of host-plant genetic variation in structuring ecological communities in naturally varying environments remains unclear (Crutsinger 2015; Hersch-Green et al. 2011; Tack et al. 2011).

A key challenge for advancing studies of community genetics beyond the common garden is to identify important biotic and abiotic environmental factors that structure communities associated with host plants. For example, a series of experiments in common milkweed (*Asclepias syriaca*) have shown that a diversity of biotic factors, such as light competition from neighboring plants (Agrawal and Zandt 2003), caterpillar herbivory (Abdala-Roberts et al. 2012), and aphid-tending ants (Abdala-Roberts et al. 2012; Mooney and Agrawal 2008) can act independently or interact with milkweed genotype to shape its associated community of foliar arthropods. Similarly, abiotic factors, such as soil nutrient availability, can act independently or modify the effects of host-plant genotype on arthropod assemblages (Barrios-Garcia et al. 2016; Orians and Fritz 1996) and tri-trophic interactions (Abdala-Roberts and Mooney 2012; Rossi and Stiling 1998). Still, we are lacking explicit comparisons of host-plant genotypic effects across natural gradi-

ents in both biotic and abiotic factors, so their relative importance in structuring communities is unclear.

Although host plants provide essential resources for a diverse array of taxa both above- and belowground, the majority of community genetics studies have focused on aboveground assemblages (Whitham et al. 2012). Studies that have simultaneously examined above- and belowground communities have found variable results, with host-plant genetic effects on aboveground communities being stronger (Bailey et al. 2009; Crutsinger et al. 2008) or comparable and coupled (Crutsinger et al. 2014) with those belowground. Above- and belowground linkages can have important consequences for both plant fitness (Whitham et al. 2006) and terrestrial ecosystem processes (Wardle 2004). In addition, feedbacks between above- and belowground assemblages may depend strongly on the biotic and/or abiotic environment (Wardle 2004). Consequently, a rising challenge for community genetics is to understand the linkages between above- and belowground communities (Crutsinger et al. 2014; Lamit et al. 2015) and whether these linkages are modified by environmental variation.

Host-plant traits determine the quantity and quality of resources for the diverse organisms that colonize them; therefore, measuring functional trait responses of host-plant genotypes to different environments can give insight to mechanisms of community assembly in genotype-by-environment studies. Phenotypic traits can vary in their plasticity (change in trait expression of a genotype in response to the environment: Scheiner 1993) and may even be plastic in response to one environmental gradient but not another (Garbutt and Bazzaz 1987; Scheiner 1993; Scheiner and Goodnight 1984). In addition, multiple plant traits can be important in structuring associated communities on host plants (Agrawal 2004, 2005; Agrawal and Fishbein 2006; Barbour et al. 2016, 2015). Simultaneous measurements of multiple functional traits and community-level patterns in genotype-by-environment studies can distinguish the effects of genetic variation (proportion of variance in a trait explained by genotype: Lynch et al. 1998), phenotypic plasticity, and direct environmental effects on species assemblages.

Here, we use common garden experiments to examine how host-plant genotypic variation

as well as the biotic and abiotic environment structure communities associated with the willow *Salix hookeriana* in a coastal dune ecosystem. Prior work in this system has shown that willow genotypes host distinct arthropod communities and that multiple plant traits are important in determining community assembly (Barbour et al. 2016, 2015). Importantly, these traits varied substantially in their degree of heritability (plant growth, mean H^2= 0.26; leaf quality, mean H^2 = 0.72), suggesting that the environment may influence them in different ways. We sought to answer the following questions: (1) What is the relative importance of willow genotype vs. the biotic and abiotic environment in structuring aboveground communities? (2) Do willow genetic and environmental variation have different effects on above and belowground communities? (3) What is the relative importance of genetic variation, phenotypic plasticity, and direct environmental effects in structuring communities?

Methods

Study site

We conducted this research at Lanphere Dunes (40°53'29.85"N, 124°8'49.06"W), a pristine coastal dune ecosystem managed by US Fish and Wildlife service in Humboldt County, California, USA. Coastal willow (*Salix hookeriana* ex Barratt ex Hooker) naturally occurs in nearshore dune swales — seasonal freshwater wetlands that form in depressions between dune ridges (Pickart and Barbour 2007). Aside from coastal willow (hereafter willow), the dominant vegetation in these swales consists of beach pine (*Pinus contorta* ssp. *contorta*) and slough sedge (*Carex obnupta*).

During preliminary surveys, we qualitatively identified two important sources of environmental variation for willows in the dunes – one biotic (the presence of ant-aphid mutualisms) and one abiotic (wind exposure). We observed that the aphid *Aphis farinosa* was an abundant herbivore at Lanphere Dunes. *Aphis farinosa* is usually found at the tips of new shoot growth where it feeds on willow phloem. As with many other aphid species, *A. farinosa* excretes carbohydrate-rich honeydew while feeding, which attracts ants that tend the aphids and feed on the honeydew.

This ant-aphid interaction is often mutualistic, because the ants will defend aphids from predatory arthropods and also eat other herbivores that may be competing with the aphids (Floate and Whitham 1994; Mooney and Agrawal 2008). The ant species we observed most frequently tending *A. farinosa* was the western thatching ant, *Formica obscuripes*. Western thatching ants create distinct dome-shaped mounds from nearby plant-material and are known to reduce herbivory from leaf chewing arthropods on *S. hookeriana* at our study site (CRUTSINGER and SANDERS 2005), presumably by deterring ovipositing females or predating young larva. This work suggests that the presence of aphids and the proximity to ant mounds influences associated communities through three non-mutually exclusive mechanisms: (i) increased abundance of aphid-tending ants, which could deter other arthropods; (ii) attraction of predators or deterrence of other herbivores, by aphids; (iii) alteration of plant-growth or leaf quality traits by aphids. We also observed that willows growing in wind-exposed habitats often exhibit reduced growth, especially at their leading edge, appearing to be "swept back" by the wind. We hypothesized that wind exposure may influence associated communities through three non-mutually exclusive mechanisms: (i) reduced plant size due to wind pruning; (ii) altered soil characteristics due to increased evaporation; and (iii) direct inhibition of ovipositing female arthropods.

Experimental design

Prior to bud break in February 2012, we took shoot cuttings (40 cm length & $\sim$0.5 cm diameter) from one to two replicates of 10 different willow genotypes from a pool of 26 locally collected willow genotypes planted in a large common garden experiment. Details about the establishment of this common garden are given in Barbour et al. (2015). These 10 genotypes displayed substantial variation in both plant-growth and leaf-quality traits (Barbour et al. 2015). Shoot cuttings were soaked in water overnight and then planted in a mixture of 80% perlite, 20% peat moss (dolomite lime added to balance pH) inside "Cone-tainers" (Stuewe & Sons, Inc.). We grew cuttings under ambient weather conditions outside the greenhouse at Humboldt State University until we transplanted willows into multiple common gardens at Lanphere Dunes.

Ant-aphid experiment

To examine how the presence of aphids, proximity to ant mounds, and willow genotype affected associated communities, we established common gardens around 5 different ant mounds (treated as blocks) in late May 2012. Within each block, we randomly planted 20 cuttings (2 replicates of each of 10 genotypes) with 0.5 m spacing in plots that were at a distance of 1, 6, and 12 meters from the edge of the ant mound, for a total of 60 cuttings per ant mound (300 cuttings for entire experiment). Within each plot, we randomly assigned the aphid treatment (aphid presence vs. absence) to one of the two replicates for each genotype. On May 22, we collected aphids (*Aphis farinosa*) from a single willow patch at Lanphere Dunes and placed 5 adult apterate aphids on the tips of willow cuttings in the aphid treatment using a moist paintbrush. We bagged aphids onto the apical shoots of cuttings using organza bags to promote aphid establishment on plants. Similarly, we placed organza bags on all control plants. On May 27, we checked aphid treatments to ensure there were 5 adult aphids and removed bags from all cuttings. If necessary, we added aphids to these treatments until there were 5 adults and we removed any aphid nymphs that were produced since initial establishment. We checked plants for aphids on June 6, June 13, June 24, July 4, July 14, and July 20, 2012. If plants in the aphid treatment had less than 5 apterate aphids, we noted their abundance and added aphids until there were at least 5 individuals. The ant-aphid experiment was restricted to the summer of 2012, because in the summer of 2013 there was high willow mortality induced by drought and *A. farinosa* was too low in abundance on naturally occurring willows to allow us to repeat the experiment.

Wind experiment

To examine how wind exposure and willow genotype affected associated communities in the coastal dunes, we planted 200 willow cuttings in a split-plot experimental design in late May of 2012. At each of 10 different willow patches (treated as blocks), we established an 'exposed' and a 'unexposed' common garden with exposed gardens facing prevailing winds during the growing season. Each garden consisted of one replicate cutting of each of 10 genotypes randomly planted

in 2 m by 0.5 m grid with 0.5 m spacing between plants. The center of exposed and unexposed gardens within each block were the same distance (2 m) from the edge of the willow patch to control for insect accessibility. To estimate the difference in wind conditions experienced by exposed vs. unexposed plants, we went out on a representative windy afternoon in September 2012. A nearby weather station estimated wind speeds of 22 km/h during this period (Arcata, CA). We used a hand-held anemometer (Kestrel 1000) to measure wind speed at a height of 37 cm aboveground (approximate height of tallest plants in the garden in 2012) in each plot of our experiment. For each block, we randomly selected the order in which exposed and unexposed plots were measured and took maximum wind speed measurements over a 30 s period. We found that willows growing in wind-exposed plots experienced up to 3.7-fold higher wind speeds compared to unexposed plots ($F_{1,9} = 187.32, P < 0.001$), suggesting that the location of our plots were effective manipulations of wind exposure.

What is the relative importance of willow genotype vs. the biotic and abiotic environment in structuring aboveground communities?

To address this question, we visually surveyed plants for arthropods to determine the abundances of different (morpho)species. For the ant-aphid experiment, we surveyed arthropods on 5 different occasions between early June and late July 2012. For the wind experiment, we surveyed arthropods once at the end of July 2012 and then once a month in May, June, and July of 2013. So that individuals were not counted twice between sampling dates, we took the maximum abundance for each arthropod (morpho)species from each plant across all sampling dates within each year. This approach provides a conservative estimate of the total number of individuals of each (morpho)species that occurred on individual plants through the summer. Given the relatively low abundances of individual (morpho)species, we grouped arthropods at the Family-level for insects and at the Order-level for all other arthropods prior to analyzing community composition (details in Statistical Analyses section below).

Do willow genetic and environmental variation have different effects on above and belowground communities?

To address this question, we dug up the willows from the wind experiment to sample fungal and bacterial communities associated with willow roots in late July of 2013. We did not sample belowground communities of plants in the ant-aphid experiment due to the high mortality of plants in 2013. To sample these belowground communities, we removed willows with the surrounding soil intact to preserve root systems, separated shoots and roots, then brushed soil off root systems and stored roots in separate plastic bags. Within 6 hours of excavation, root systems were stored at 4°C. To process roots, we gently rinsed them in tap water until free of visible soil. In order to randomly select roots for molecular analysis, second order roots were cut up into 2 cm lengths, spread out on a grid, and then, using a random number generator, a total of 30 cm of root length was picked from numbered grid cells. These random root subsamples were flash frozen in liquid N, and kept at -80°C until DNA extraction. To increase efficiency of DNA extraction, roots were physically disrupted with 2 beads per 2 mL tube (3.0 mm Yttria stabilized Zirconia Grinding Media) for 30 seconds at 1500 strokes per minute (SPEX SamplePrep 200 geno/grinder). Total DNA was extracted from the ground, frozen root samples using MoBio PowerSoil 96 sample DNA extraction kits following the manufacture's instructions.

To identify fungal and bacterial OTUs, we used custom Illumina-compatible barcode primer sets ITS1f/ITS4 and 515f/806r (Caporaso et al. 2012) to amplify via PCR the fungal ITS1, 5.8S, and ITS2 region of ribosomal DNA and the V4 region of bacterial 16S ribosomal DNA from total root DNA extractions. Product quality was assessed by gel electrophoresis. PCR products were cleaned with house-made magnetic bead solution, quantified with a Qubit fluorometric kit, then sample libraries were pooled at a bacteria:fungi concentration ratio of 2:1. Pooled amplicon libraries were sequenced as single-index (the reverse barcode was uniquely indexed) 300 base pair reads at Stanford Functional Genomics Facility on one lane of an Illumina MiSeq. Quality control of reads consisted of these steps: trimming bases with quality score less than 20 phred;

trimming sequenced adaptors; and removing reads with average error rates greater than 0.25 using UPARSE (Edgar 2013). Only high quality, paired forward and reverse reads were used for OTU clustering at 97% identity, then OTUs were checked for chimeras against the GOLD 16s rRNA database (Reddy et al. 2014) and UNITE fungal ITS database ver6_97_13.05.2014 (Kõljalg et al. 2005) with UPARSE. Taxonomy was assigned using the RDP Classifier (Wang et al. 2007) and UNITE (ver6_97_13.05.2014) in QIIME (Caporaso et al. 2010). We normalized datasets to account for differences in each sample's library size (number of reads obtained for each sample). Finally, we discarded some OTUs and samples based on the following conditions: OTUs with no known taxonomy (any OTU that did not blast to at least Kingdom Fungi, Bacteria or Archaea); root samples with fewer than 6000 fungal reads and 9000 bacterial reads; and mitochondrial and chloroplast OTUs.

What is the relative importance of genetic variation, phenotypic plasticity, and direct environmental effects in structuring communities?

Teasing apart the relative importance of these factors requires simultaneous measurements of plant functional traits, correlated environmental gradients, and community-level patterns within a genotype-by-environment experiment.

Plant traits

Prior work in this study system demonstrated that variation in both plant growth and leaf quality traits affects the likelihood of willows being colonized by foliar arthropods (Barbour et al. 2015). To quantify plant-growth traits, we measured plant height, the number of shoots produced, and average shoot length in late July of each year (end of growing season) for both experiments. We quantified plant height as the distance (mm) from the ground to the tip of the tallest shoot. We quantified average shoot length by measuring every shoot on each plant to the nearest millimeter and calculating the average shoot length for each plant. We also measured several traits that could shape leaf quality for herbivores, including water content, trichome den-

sity, specific leaf area (SLA), percentage carbon (C) and nitrogen (N), and C:N. To measure these traits, we excised fully expanded and undamaged leaves from plants in late July of each year, stored leaf samples with a moist paper towel in separate plastic bags within a cooler and immediately brought them back to the laboratory. We then weighed leaves to obtain fresh mass (g), digitally scanned them to measure leaf area (mm^2) using ImageJ (Abràmoff et al. 2004), and oven-dried them at 60 °C for 72 h to obtain dry weight (g)(Cornelissen et al. 2003). We calculated SLA as $(leaf\ area)/(dry\ mass)$ (Cornelissen et al. 2003). We calculated leaf water content as the $(fresh\ mass - dry\ mass)/(dry\ mass)$ (Cornelissen et al. 2003). To measure trichome density, we counted the number of trichomes along an 11 mm by 1 mm transect in the center of the leaf, halfway between the leaf edge and the mid-vein, under a dissecting scope. To measure percentage C and N, we ground oven-dried leaves to a fine powder using a ball mill (Mixer/Mill 8000D, SPEX SamplePrep; Metuchen, NJ, USA). Subsamples of each material were then analyzed for percentage C and N on an elemental analyzer (ECS 4010; Costech Analytical Technologies, Valencia, California, USA) using atropine (4.84% N and 70.56% C) as a reference standard. For root-associated communities, we hypothesized that variation in root C:N may affect community assembly. To measure root C and N, a subsample of oven-dried roots were crushed with a razor blade, then analyzed for percent C and N on an elemental analyzer (Carlo-Erba NA 1500) using atropine (4.84% N and 70.56% C) as a reference standard.

Soil characteristics

Soil nutrients, total organic matter, and moisture may all influence plant traits and the assembly of fungal and bacterial communities on plant roots (Erlandson et al. 2015). Moreover, we expected that wind exposure would affect these soil characteristics (Lortie and Cushman 2007); therefore, we measured soil nutrients, percent organic matter, and moisture within each plot of the wind experiment (one exposed and one unexposed plot per block). To estimate soil nutrient uptake by willows, we installed Plant Root Simulator (PRS) Probes (Western Ag Innovations, Saskatchewan, Canada) at three randomly selected locations within each plot for 11 days in September 2012.

PRS Probes estimate nutrient supply rates to roots by continuously adsorbing charged ionic elements over the burial period. For our study, we estimated potential root uptake of NO_3^+, NH_4^-, Ca^{2+}, Mg^{2+}, K^+, $H_2PO_4^-$, Fe^{3+}, Mn^{2+}, Cu^{2+}, Zn^{2+}, $B(OH)_4^-$, SO_4^{2-}, Pb^{2+}, Al^{3+}, Cd^{2+}. From this nutrient data, we calculated total N as NO_3^+ + NH_4^-, and then used principal components analysis to condense these nutrients into a single axis (nutrients PC1) that explained 34% of the variation. Nutrients PC1 described the negative correlation between nitrogen compounds (NO_3^+, NH_4^-) and the rest of the ionic elements, with positive values indicating high supply rates of all ionic elements except for the nitrogen compounds. To measure percent organic matter content (%OM), we used a trowel to collect soil (depth = 0 – 15 cm) adjacent to the randomly positioned PRS probes in September 2012. Soils were transported back to the lab in plastic bags, sieved into fragments less than 2 mm, randomly subsampled using a soil splitter, and dried at 105 °C for 72 hours. We then weighed a subsample of the oven dried soil into an oven dried crucible and placed the crucible and soil into a furnace to be combusted at 375 °C for 16 hours. We then weighed the mass (g) of combusted samples, placed them in a desiccator for 20 minutes, and weighed them again. We calculated percent organic matter as $\%OM = (\textit{oven dry mass} - \textit{combusted dry mass})/(\textit{oven dry mass}) \times 100$. To measure soil moisture (volumetric water content, m^3/m^3), we used a 5TE soil sensor coupled to an EM50 Digital/Analog Data Logger (Decagon Devices, Pullman, Washington, USA). In September 2012, while PRS probes were in the ground, we measured soil moisture at a depth of 5 cm in three random locations within each plot on three different days between 1100 – 1500 hours. We repeated this same sampling scheme in early July 2013. Plot levels measurements of soil moisture were highly correlated between years (Pearson's $r = 0.93,\ t_{18} = 10.91,\ P < 0.001$), so we averaged these soil moisture estimates to determine a single soil moisture value per plot.

Statistical analyses

Community responses

To determine how willow genotype, the environment, and their interaction, influenced richness, abundance, and rarefied richness of aboveground arthropods as well as rarefied richness of root fungi and bacteria, we used separate generalized linear mixed-effect models (GLMMs) (Bolker et al. 2009). For the ant-aphid experiment, we omitted *A. farinosa* and *F. obscuripes* from our calculations of arthropod community properties because we expected our treatments to manipulate their abundances. We specified block (ant mound) and plots nested within block (the 3 different distances from ant mound) as random effects. We specified willow genotype, aphid treatment, distance from ant mound, and their 3-way interaction as fixed effects in the model. For the wind experiment, we specified block (willow patch) and plots nested within block (the 2 wind exposure treatments) as random effects. We specified willow genotype, wind treatment, sampling year, and their 3-way interaction as fixed effects in the model. Plant mortality in each experiment resulted in unbalanced designs, so we used Type II sum-of-squares to test the significance of fixed effects. For continuous responses (rarefied richness) we specified Gaussian error distributions in our models and tested the significance of fixed effects using *F*-tests with Kenward-Roger approximated degrees of freedom. For count responses (richness and arthropod abundances), we specified Poisson error distributions in our models and tested the significance of fixed effects using likelihood-ratio tests. If necessary, we accounted for overdispersion in these Poisson models by specifying an individual-level random effect.

To examine whether community composition depended on willow genotype, the environment, or their interaction, we applied a Hellinger transformation to our community data (square root of proportional abundance of species found on each willow; Legendre and Gallagher 2001) and conducted separate redundancy analyses (RDA, 1000 permutations on Euclidean distances) for the arthropod, fungal, and bacterial communities. A Hellinger transformation was appropriate because calculating Euclidean distances on raw abundance data can sometimes result in two

communities that do not share the same species being more similar than two communities that do share the same species Legendre and Gallagher (2001). We incorporated the same fixed-effects model structure as we used to analyze the univariate community responses for each experiment. To test the significance of each effect, we used Type II sum-of-squares and compared the observed community dissimilarities to the dissimilarities we would expect by random chance with a permutation test that controls for the blocked design of our experiment. To test the significance of treatments that varied at the plot-level (wind exposure and distance from ant mound), we first calculated the community's centroid in multivariate space for each plot. We then included block as a covariate and ran the same permutation test as previously described. This ensured that our significance tests of treatments that varied at the plot-level were based on the appropriate residual degrees of freedom (wind exposure residual $df = 9$; distance from ant mound residual $df = 4$).

To compare the relative importance of willow genotype vs. different environmental factors, we fit reduced GLMMs with significant factors as random effects. We then used the variance components estimated for these random effects to calculate the percentage of variance (hereafter σ^2) explained by willow genotype and specific environmental factors. This method has been proposed as appropriate for comparing the relative importance of plant genotype and the environment in genotype-by-environment studies Johnson and Agrawal (2005)Hersch-Green et al. (2011). For models of community composition, we fit reduced RDAs with only significant factors and calculated the adjusted redundancy statistic for each significant factor, which has been shown to be an unbiased estimator of explained variance Peres-Neto et al. (2006). This adjusted redundancy statistic is an analogue of adjusted R^2 but for multivariate responses and we refer to it hereafter as R^2_{adj}.

Plant traits

To analyze how willow genotype, the environment, and their interaction, influenced willow phenotypes, we used separate GLMMs with the same structure described in the "Community

responses" section. For the wind experiment, we lacked multiple years of data on leaf trichome density (2012 only), SLA (2013 only), leaf C:N (2013 only), and root C:N (2013 only); therefore, we removed sampling year, and its interactions, from the fixed effects structures of these GLMMs.

Soil characteristics

To examine the effect of wind exposure on soil characteristics (total N, nutrients PC1, %OM, and soil moisture), we used separate mixed effect models with wind treatment as a fixed effect and block (willow patch) as a random effect. Since all soil characteristics were continuous responses, we specified Gaussian error distributions in our models and tested the significance of fixed effects using *F*-tests with Kenward-Roger approximated degrees of freedom.

Genetic variation vs. phenotypic plasticity vs. direct environmental effects

To identify potential mechanisms by which willow genetic and environmental variation affect community responses, we used piecewise structural equation models (SEMs)Lefcheck (2015). We only modeled potential mechanisms and community responses that were statistically significant in our prior analyses. For example, if we did not observe a G$\times$E effect on a community response variable, then we did not model this interactive effect in the piecewise SEM. An advantage of piecewise SEMs is that they are flexible, allowing users to account for correlated structure (i.e. random effects) in their experimental design. However, as with any technique that relies on multiple regression, structural equation models can give misleading results if there is collinearity among predictor variables. To mitigate the effects of collinearity, we used principal components analysis (PCA) to condense aboveground willow phenotypes as well as soil properties into a small number of uncorrelated variables. For aboveground willow traits in the wind experiment, we analyzed separate PCAs for 2012 and 2013 since we did not always have data on the same traits in each year. At times, we lacked data for all traits on each plant or all soil properties measured in each plot. Therefore, we used a regularized iterative PCA algorithm to impute missing values Josse et al. (2012). For each PCA, we retained principal components with eigenvalues

360 greater than 1.

To calculate standardized coefficients (β) in our piecewise SEM, we scaled all predictor and response variables to mean = 0 and SD = 1 prior to analyzing them with GLMMs (error distri-
363 bution = Gaussian). For willow genotype, we specified the average effect for the 10 genotypes as the reference level (i.e. deviation contrasts) and calculated the standard deviation of the coefficients to determine its standardized coefficient. We then multiplied standardized coefficients
366 across a given pathway to calculate the strength of each mechanism we modelled. To evaluate the explanatory power of our separate GLMMs, we report marginal R^2Nakagawa and Schielzeth (2012). Marginal R^2 values do not adjust for the variance explained by our random effects; there-
369 fore, they give us a truer sense of the explanatory power of our models. To evaluate the fit of the full structural equation model, we used a test of directed separation Shipley (2000). This test identifies missing paths in the model, calculates the P-value for each missing pathway, and
372 then calculates a test statistic, Fisher's C, using the following equation: $C = -2 \times \ln(P_i)$, where P_i is the P-value of the i-th missing pathway. Fisher's C can then be compared to a chi-square distribution with $2k$ degrees of freedom, where k is the total number of missing pathways. If
375 there are many missing pathways with low P-values, this will result in a lower P-value for the structural equation model. Therefore, a P-value < 0.05 indicates a poor fit for the structural equation model, whereas a P-value > 0.05 indicates a good fit. Note that if we have included
378 the key plant traits as well as biotic and abiotic factors, then there should be no missing paths between willow genotype and our environmental treatments. All analyses were conducted in R version 3.2.4 R Core Team (2016).

Results

What is the relative importance of willow genotype vs. the biotic and abiotic environment in structuring aboveground communities?

Ant-aphid experiment

Willow genotype tended to be more important than the biotic environment in structuring the arthropod community (table A1 in Supporting Information). We found that average arthropod richness varied from 1.2 to 3.2 species among genotypes ($\sigma^2 = 11\%,\ \chi_9 = 41.35,\ P < 0.001$), while arthropod abundance varied 4-fold among the different clones (figure 1A; $\sigma^2 = 7\%,\ \chi_9 = 34.86,\ P < 0.001$). The effect of willow genotype on arthropod richness was explained by correlated responses in arthropod abundance, as there was no difference in rarefied richness among genotypes ($F_{9,138.8} = 0.83,\ P = 0.586$). Aphid treatment was the only factor that affected rarefied richness ($\sigma^2 = 4\%,\ F_{1,139.2} = 5.34,\ P = 0.022$), resulting in a 16% decrease in rarefied richness when aphids were added to willows (figure 1C); however, this effect of aphid treatment did not translate into an effect on total richness ($\chi_1 = 0.45,\ P = 0.504$). Willows in the aphid treatment also had 2-fold more arthropods, but only at the furthest distance from ant mounds (figure 1B, $\sigma^2 = 9\%,\ \chi_1 = 8.12,\ P = 0.004$). Proximity to ant mounds did not influence any other aspect of the arthropod community (table A1). In contrast to richness and abundance responses, arthropod community composition was influenced by an interaction between willow genotype and the aphid treatment (figure 2A; $R^2_{adj} = 6\%,\ F_{9,157} = 1.45,\ P = 0.022$). This G×E effect was primarily due to the differential response of other aphids to a single willow genotype (figure 2A solid line). If we remove this genotype from the analysis, we find independent effects on community composition, with the effect of willow genotype ($R^2_{adj} = 3\%,\ F_{8,156} = 1.66,\ P = 0.007$) being relatively more important than the addition of aphids ($R^2_{adj} = 1\%,\ F_{1,156} = 2.93,\ P = 0.017$).

Wind experiment

In contrast to the ant-aphid experiment, we found that the abiotic environment tended to be more important than willow genotype in structuring the arthropod community (table A3). In particular, willows growing in wind-exposed plots hosted 51% fewer species ($\sigma^2 = 11\%,\ \chi_1 = 13.55,\ P < 0.001$), 47% fewer individuals (figure 1B; $\sigma^2 = 2\%,\ \chi_1 = 5.48,\ P = 0.019$), and 60% fewer rarefied species (figure 1D; $\sigma^2 = 23\%,\ F_{1,7.8} = 22.82,\ P = 0.001$) compared to unexposed willows. In spite of the effects of wind exposure, willow genotype had clear effects on both the richness ($\sim$3-fold differences, $\sigma^2 = 6\%,\ \chi_9 = 28.01,\ P < 0.001$) and abundance (figure 1B; $\sim$5-fold differences, $\sigma^2 = 5\%,\ \chi_9 = 25.25,\ P = 0.003$) of arthropods, but only a marginal effect on rarefied richness (figure 1D; $\sigma^2 = 5\%,\ F_{9,71.1} = 1.96,\ P = 0.058$). Arthropod communities on willows had both more species ($\sigma^2 = 4\%,\ \chi_1 = 10.33,\ P = 0.001$) and more individuals ($\sigma^2 = 2\%,\ \chi_1 = 6.72,\ P = 0.010$) in the second year of the experiment compared to the first; however, we also conducted more arthropod surveys for the wind experiment in 2013 vs. 2012. In terms of community composition, we observed strong effects of wind exposure by the end of experiment (figure 2B; $R^2_{adj} = 40\%,\ F_{1,68} = 12.80,\ P = 0.001$). These compositional differences were driven primarily by gall midges (Family: Cecidomyiidae) being relatively less abundant on wind-exposed willows ($\chi_1 = 16.28,\ P < 0.001$), whereas leaf-tiering moths (Family: Tortricidae) were insensitive to wind exposure (and therefore relatively more abundant; $\chi_1 = 1.34,\ P > 0.05$). Although several arthropod taxa varied in total abundance among willow genotypes (table A54), we did not detect an effect of genotype on community composition in either year of the experiment (2012: $F_{9,51} = 0.96,\ P = 0.502$; 2013: $F_{9,68} = 1.17,\ P = 0.271$).

Do willow genetic and environmental variation have different effects on above and belowground communities?

Willow genotype and wind exposure had distinct effects on root-associated fungal and bacterial communities compared to foliar arthropods (table A3). Neither wind exposure ($F_{1,8.8} = 0.90,\ P =$

0.369) nor willow genotype ($F_{9,95.0} = 1.21,\ P = 0.295$) influenced the rarefied richness of fungal OTUs (figure 1E). However, the composition of the fungal community did vary among willow genotypes (figure 1C; $R^2_{adj} = 1\%,\ F_{9,117} = 1.00,\ P = 0.005$) with no detectable effect of wind-exposure ($F_{1,9} = 1.19,\ P = 0.162$). Genotypic differences in fungal community composition were due to variation in the relative abundance of at least eight different OTUs, most of which were saprotrophs (table A6) and not arbuscular or ectomycorrhiza.

In contrast to the fungal community, wind exposure influenced the bacterial community (table A3), but in the opposite direction of foliar arthropods. For example, the rarefied richness of bacterial OTUs was 6% higher on the roots of wind-exposed vs. unexposed plants (figure 1F; $\sigma^2 = 7\%,\ F_{1,7.9} = 6.52,\ P = 0.034$). While wind exposure had a marginal effect on the composition of the bacteria community (figure 1D; $R^2_{adj} = 3\%,\ F_{1,9} = 1.38,\ P = 0.092$), there was no detectable effect of willow genotype on any aspect of the bacterial community (rarefied richness: $F_{9,100.3} = 1.63,\ P = 0.117$; community composition: $F_{9,120} = 0.93,\ P = 0.536$).

What is the relative importance of genetic variation, phenotypic plasticity, and direct environmental effects in structuring communities?

Ant-aphid experiment

We hypothesized that the effect of willow genetic variation and the biotic environment on arthropod communities would be mediated, in part, by variation in the abundance of *A. farinosa* and *F. obscuripes*. While distance from ant mounds had little effect on *A. farinosa* ($\chi_1 = 0.55,\ P = 0.460$), willow genotype had a strong effect, with the average number of aphids ranging from 0.05 to 7 among the most disparate willow genotypes in the aphid treatment ($\chi_9 = 20.83,\ P = 0.013$). This strong effect of willow genotype on *A. farinosa* in the aphid treatment resulted in a $\mathrm{G}\times\mathrm{E}_{\mathrm{aphid}}$ effect on the abundance of *F. obscuripes* ($\chi_2 = 6.26,\ P = 0.044$), with ant abundance varying from 0 to $\sim$0.5 individuals (on average) among clones in the aphid treatment, whereas they were virtually absent in the absence of aphids. Proximity to ant mounds had no effect on the abundance

of *F. obscuripes* ($\chi_1 = 1.68,\ P = 0.195$).

In addition to ant-aphid interactions, we hypothesized that the effect of willow genetic variation and the biotic environment on arthropod communities would be mediated by plant traits. We observed both direct and interactive effects of willow genotype and the biotic environment on plant traits (table A1). All of the plant-growth traits we measured varied approximately 2-fold among the most disparate willow genotypes (plant height $F_{9,204.2} = 15.83,\ P < 0.001$; shoot count $\chi_9 = 65.84,\ P < 0.001$; shoot length $F_{9,204.2} = 7.27,\ P < 0.001$). Willows did appear to produce 28% more shoots in the absence of aphids, but only at the furthest distance from ant mounds ($E_{aphid} \times E_{ant}$ $\chi_1 = 4.20,\ P = 0.040$). While there was little apparent effect of willow genotype and the biotic environment on leaf water content (table A1), we found that the addition of aphids modified the effect of certain willow genotypes on leaf trichome density ($G \times E_{aphid}$ $\chi_8 = 23.17,\ P = 0.003$). Specifically, two clones (S and T) produced $\sim$4-fold more trichomes when aphids were absent, whereas genotype L produced 3-fold more trichomes when aphids were present (table A1).

Using structural equation models (for richness, abundance, and rarefied richness) and redundancy analysis (for community composition), we found that genetic variation in plant traits and direct effects of aphid additions were the primary determinants of the arthropod community rather than phenotypic plasticity (table 1). The effect of genetic variation on arthropod richness and abundance was mediated primarily by plant trait PC1 (Fig.3A,B). Plant height, shoot count, and shoot length all had strong, positive loadings on trait PC1 (table A7), indicating that larger willows hosted more arthropod species and individuals. Arthropod abundance was also positively influenced by the addition of aphids, primarily because *A. farinosa* attracted individuals of other ant species (Pearson's $r = 0.42,\ t_{282} = 7.74,\ P < 0.001$) and these other ants were the second most abundant taxonomic group in the community. In contrast to total abundance, the addition of aphids negatively affected rarefied richness (figure 3C). This negative effect was due in part to aphid additions attracting more *F. obscuripes*, an active generalist predator that likely consumed or inhibited the colonization of other arthropods. In terms of composition, we found

that the abundance of *A. farinosa* was the only factor (of the mechanisms we modeled) influencing the arthropod community ($F_{1,183} = 2.86,\ P = 0.025$). Specifically, higher abundance of *A. farinosa* resulted in an increase in the proportional abundance of other ant species in the community (figure 3D).

Despite our detailed analysis of potential mechanisms, our structural equation models revealed multiple missing paths (dotted lines, figure 3A,B,C), resulting in rather poor fits for most models (richness: $C_2 = 24.84,\ P < 0.001$; abundance: $C_{32} = 48.88,\ P = 0.029$; rarefied richness: $C_4 = 10.52,\ P = 0.032$). For example, after accounting for the traits we measured, willow genotype still had a strong effect on arthropod richness (figure 3A) and *A. farinosa* abundance (figure 3B), indicating that we failed to identify key pathways (likely unmeasured traits) by which genetic variation influenced these responses. Similarly, we failed to fully identify the $\mathrm{E_{aphid}} \times \mathrm{E_{ant}}$ effect on arthropod abundance (figure 3B) as well as how the addition of aphids negatively affected rarefied richness (figure 3C). For our redundancy analysis of community composition, we found that *A. farinosa* abundance explained the effect of the aphid treatment ($F_{1,173} = 090,\ P = 0.447$), but we still failed to detect the effect of both willow genotype ($F_{9,173} = 1.53,\ P = 0.014$) and the $\mathrm{G} \times \mathrm{E_{aphid}}$ effect ($F_{9,164} = 1.71,\ P = 0.004$), suggesting that we failed to measure important constitutive and inducible plant traits.

Wind experiment

One of the mechanisms by which wind exposure could influence willow-associated communities is through accumulated effects on soil properties; however, we observed only modest effects of wind exposure on soil properties (table A42). Specifically, soil in wind-exposed plots was marginally drier ($F_{1,9.0} = 3.52,\ P = 0.093$) with higher amounts of total Nitrogen ($F_{1,9.0} = 5.08,\ P = 0.051$) than in unexposed plots, but there was no clear difference in either percent organic matter ($F_{1,8.4} = 0.68,\ P = 0.434$) or nutrient composition ($F_{1,9.0} = 1.31,\ P = 0.282$).

As with the ant-aphid experiment, we hypothesized that the effects of wind exposure and willow genotype on associated communities would be mediated by plant traits. Interestingly,

we found that plant-growth and leaf quality traits responded differently to wind exposure and willow genetic variation (table A4). For example, wind exposure negatively affected all plant-growth traits (plant height $F_{1,9.0} = 29.10,\ P < 0.001$; shoot count $\chi_1 = 9.91,\ P = 0.002$; shoot length $F_{1,9.0} = 10.44,\ P = 0.010$). Moreover, the negative effects of wind exposure were magnified by the end of the experiment for both plant height ($F_{1,158.4} = 16.69,\ P < 0.001$) and the number of shoots produced ($\chi_1 = 12.53,\ P < 0.001$). Still, willow genotype had a pronounced effect on all plant-growth traits, resulting in willows that varied over 2-fold in height ($F_{9,145.3} = 9.13,\ P < 0.001$), number of shoots ($\chi_9 = 47.42,\ P < 0.001$), and shoot length ($F_{9,144.2} = 4.97,\ P < 0.001$) among the most disparate genotypes. While the effect of willow genotype on the number of shoots changed by the end of the experiment ($\chi_9 = 18.26, P = 0.032$), this G×$E_{year}$ effect was relatively small ($R^2 = 0.05$) compared to the effect of genotype alone ($R^2 = 0.13$). In contrast to plant-growth traits, willow genotype was the primary factor in determining leaf traits across both years of the experiment (table A4). The leaves of willow genotypes varied 46-fold in trichome density ($\chi_9 = 67.31,\ P < 0.001$), 1.5-fold in SLA ($F_{9,122.5} = 4.21,\ P < 0.001$), and 1.6-fold in C:N ($F_{9,70.5} = 4.88,\ P < 0.001$). We had data available on leaf water content for 2012 and 2013, and we found that the amount of variation explained by willow genotype depended on the sampling year ($F_{9,141.6} = 2.80,\ P = 0.005$; 2012, $R^2 = 0.11$; 2013, $R^2 = 0.16$). Unlike aboveground plant traits, root C:N did not appear to be influenced by either wind exposure ($F_{1,8.7} = 0.31,\ P = 0.590$) or willow genotype ($F_{9,107.0} = 0.85,\ P = 0.569$).

Similar to the ant-aphid experiment, we found that genetic variation and direct effects of wind exposure were the primary factors influencing the arthropod community rather than phenotypic plasticity (table 1). Both trait PC1 and PC2 mediated the indirect effects of willow genetic variation and wind exposure (negative) on the arthropod community (figure 4A), but the effects of genetic variation were relatively stronger (table 11). Trait PC1 had a strong, positive effect on arthropod richness ($\beta = 0.37$), abundance ($\beta = 0.28$), and rarefied richness ($\beta = 0.37$, figure 4A). Similar to the ant-aphid experiment, trait PC1 had strong, positive associations with plant height, shoot count, and shoot length (table A7), indicating that larger willows hosted

more arthropod individuals and species. Trait PC2 had a smaller, but negative effect on arthropod richness ($\beta = -0.13$), abundance ($\beta = -0.15$), and rarefied richness ($\beta = -0.12$, figure 4A). Trait PC2 had a strong positive correlation with leaf C:N, but strong negative correlations with leaf water content and SLA (table A7), indicating that willows with poorer quality leaf tissue hosted fewer arthropod individuals and species. Although the community-level effects of wind exposure due to phenotypic plasticity were relatively weak, wind exposure had strong and direct, negative effects on arthropod richness ($\beta = -0.22$), abundance ($\beta = -0.08$), and rarefied richness ($\beta = -0.26$, figure 4A). The qualitative effects of genetic variation, phenotypic plasticity, and direct environmental effects held for the richness, abundance, and rarefied richness of foliar arthropods in the first year of the experiment as well ($C_{22} = 28.70,\ P = 0.154$), except that trait PC2 was determined by different traits (table A7) and did not appear to affect any aspect of the arthropod community (richness, $P = 0.657$; abundance, $P = 0.104$; rarefied richness, $P = 0.850$). For community composition, we only analyzed the data from the second year of the experiment because this was the only year for which we detected a significant effect of wind exposure. We found that the effects of wind exposure on community composition were primarily mediated by plant trait PC1 ($F_{1,76} = 12.05,\ P = 0.001$). Positive values of trait PC1 (i.e. larger plants) had greater proportional abundance of gall midges, leaf-mining moths, and spiders, whereas leaf-tiering moths were insensitive to plant size (figure 4B).

As with the aboveground community, we found that genetic variation and direct environmental effects were the primary determinants of rhizosphere communities, although we did a poorer job of identifying the specific mechanisms. For example, we previously showed that fungal community composition varied among willow genotypes (table A3), but we failed to identify the root traits mediating the effect of willow genetic variation ($F_{9,106} = 1.03,\ P = 0.002$). Our failure to identify this process is not surprising though, given that we measured only one belowground plant trait (root C:N) and it was not strongly influenced by willow genotype. While wind exposure only had a modest effect on soil properties, soil characteristics had strong direct effects on the rarefied richness of rhizosphere communities. For example, soil PC1 ($\beta = -0.28$), and to

a lesser extent root C:N ($\beta = -0.22$), negatively affected fungal rarefied richness (figure 4A). Soil PC1 had strong positive correlations with soil moisture and organic matter, but negative correlations with NO_3^- and NH_4^+ (table A8), indicating that fungal communities were more diverse in drier environments with more available nitrogen. In contrast, soil PC2 was the primary factor in determining bacterial rarefied richness ($\beta = 0.25$). Micronutrients such as Ca^{2+}, Mg^{2+}, and Cd^{2+} had strong positive loadings on soil PC2 (A8), indicating that bacterial richness was greater in environments with more of these micronutrients. Although we detected clear effects of soil properties and root C:N on rarefied richness of root-associated communities, none of these soil and plant characteristics were strong predictors of rhizosphere community composition (table A9).

In contrast to the ant-aphid experiment, our structural equation models provided good fits to our data (i.e., $P > 0.05$), indicating that we identified the key processes affecting the richness ($C_{10} = 2.94,\ P = 0.983$), abundance ($C_{10} = 6.73,\ P = 0.751$), and rarefied richness (figure 4A; $C_{32} = 23.32,\ P = 0.899$) of willow-associated communities.

Discussion

What is the relative importance of willow genotype vs. the biotic and abiotic environment in structuring aboveground communities?

The relative importance of host-plant genetic effects on its associated arthropod community depended on the type of environmental gradient. In the ant-aphid experiment, willow genetic variation tended to have a stronger effect on arthropod community structure compared to aphid additions and proximity to ant mounds. This result supports an emerging generalization that host-plant genetic variation is often more important than the presence of arthropod herbivores (Cronin and Abrahamson 2001; Fritz et al. 1986; Johnson 2008; Mcguire and Johnson 2006) or predatory ants (Abdala-Roberts et al. 2012; Johnson 2008; Mooney and Agrawal 2008) in structuring arthropod communities. This is likely due to host plants providing all of the habitat for the associated communities, compared to the smaller spatial scale at which ant-aphid interactions

occur. Still, we did find that aphid additions reduced arthropod diversity (rarefied richness) and modified the effect of willow genotype on the composition of the arthropod community, suggesting that accurately predicting community assembly requires an understanding of both factors.

In the wind experiment, we found that wind exposure dominated willow genotype in the strength of its effect on foliar arthropods, supporting the notion that wind is a key environmental factor in structuring communities associated with host plants in coastal dune ecosystems (Crutsinger et al. 2014, 2010; Miller and Weis 1999). For example, Crutsinger et al. (2014) found that there were more arthropod species and individuals on prostrate vs. erect morphs of coyote bush (*Baccharis piluaris*), due to their low-lying growth form which enabled them to be more productive than erect morphs at their windy coastal dunes site. Unfortunately, few studies have examined community-level responses to natural variability in the abiotic environment and host-plant genetic variation, making it difficult to draw other useful comparisons. The majority of genotype-by-abiotic environment studies to date have used fertilizers to manipulate soil nutrient availability (Abdala-Roberts et al. 2012; Barrios-Garcia et al. 2016; Orians and Fritz 1996; Rossi and Stiling 1998), but it is unclear whether these manipulations reflect natural variation in soil nutrients. This may explain why the effects of variation in soil nutrients range from being independent and weak (Abdala-Roberts et al. 2012; Barrios-Garcia et al. 2016) to being strong modifiers (Orians and Fritz 1996) of host-plant genotype on arthropod communities. The only other genotype-by-abiotic environment study that we are aware of manipulated sun exposure to sea daisies (*Borrichia frutescens*, Rossi and Stiling 1998) and, similar to our study, observed strong, independent effects of sun exposure on densities of the gall midge, *Asphondylia borrichiae*. If we are to make progress on understanding the relative importance of willow genotype vs. environment for associated communities, future experimental work should focus on manipulating natural variation in specific abiotic factors, or across natural gradients in the abiotic environment.

Do willow genetic and environmental variation have different effects on above and belowground communities?

Although diverse assemblages of above and belowground taxa colonize host plants, there are no genotype-by-environment studies, to our knowledge, that have simultaneously measured the responses of above and belowground assemblages. We found that foliar arthropods responded differently than root-associated fungi and bacteria to willow genetic and environmental variation, suggesting that these communities are responding to different plant traits and environmental correlates of wind exposure. Similarly, Lamit et al. (2015) found that communities of foliar arthropods and ectomycorrhizal fungi did not covary across genotypes of narrowleaf cottonwood (*Populus angustifolia*). The lack of covariation between these communities is likely because these assemblages are responding to different plant traits, although we did not fully identify the root traits mediating fungal and bacteria communities. While we have made substantial progress in the past decade understanding plant-arthropod interactions, it is time that community genetics research turns its attention belowground to understand the plant traits influencing these diverse assemblages. This will have the added benefit of understanding associations between above and belowground traits which will be important for predicting when we would expect linkages between above and belowground communities.

What is the relative importance of genetic variation, phenotypic plasticity, and direct environmental effects in structuring communities?

Across both experiments, we found that the effects of host plant traits on foliar arthropod communities were primarily mediated by genetic variation rather than phenotypic plasticity. Although the indirect effect of environmental variation was of minor importance, we found that the biotic and abiotic environment could have strong direct effects on community structure.

While willow genotype had a strong effect on leaf quality traits, the effects of genetic variation on arthropod community structure were primarily mediated by plant size in both experiments.

We found similar results in our prior work with *Salix hookeriana* (Barbour et al. 2015), although the effects of plant size appeared much more important in this study. This is likely because the absolute size of the plants in the current study were small (<60 cm tall), therefore size could be a limiting resource in determining arthropod community structure. Genotype-by-environment experiments with small-statured plants (forbs, herbs, etc.) have similarly found that plant size is a key factor in structuring arthropod communities (Crutsinger et al. 2014; Johnson and Agrawal 2005). Therefore, the effects of host-plant size on arthropod communities may be especially strong for smaller plants, with leaf quality traits becoming more important as plants grow. While the importance of plant size corresponds with other work in coastal dune systems (Crutsinger et al. 2014), studies in ant-aphid systems often find that the effects of genetic variation on arthropod communities are mediated by indirect genetic effects on ant abundance. One possible reason for this discrepancy is that the absolute variation in *F. obscuripes* abundance was quite low in our experiment (max. genotype average = $\sim$0.5 individuals) compared to other studies (max. genotype average = $\sim$3 individuals), which was likely due to the rather low abundance of aphids we observed (max. genotype average = $\sim$7 individuals). While we failed to fully identify the mechanisms (likely unmeasured plant traits) explaining arthropod community responses in this experiment, our results do suggest that host-plant evolution could have strong effects on arthropod community structure, despite variability in this biotic factor.

Conclusion

Overall, our study reinforces the importance of host-plant genetic variation in shaping associated communities and extends this finding to natural biotic and abiotic gradients in coastal dune ecosystems. Our findings also suggest that predicting community responses to genetic and environmental variation is a complex task that may depend on historical processes that have shaped the genetic architecture for the populations of interest (i.e. sensitivity to specific environmental factors). Still, the effects of willow genetic variation were clear at both the level of plant traits

and the community structure of foliar arthropods and fungi. Importantly, this suggests that host-plant evolution could have a strong influence on the biodiversity of above and belowground communities. Future studies should work toward understanding how these diverse communities feedback to impose selection pressures on host plants as well as how host-plant traits mediate interactions between above and belowground communities. In doing so, we will be able to work toward a more synthetic understanding of the evolutionary ecology of host plants and their diverse associated communities.

Acknowledgments

We thank A. Pickart and the staff of Humboldt Bay National Wildlife Refuge (U.S. Fish and Wildlife Service) for permission to work at Lanphere Dunes and for facilitating experimental logistics. L. Mackas-Burns assisted with the fieldwork. M.A.B. was supported by multiple fellowships from the University of British Columbia (BRITE Fellowship, James Robert Thompson Fellowship, and a Four-Year Fellowship). G.M.C. was supported by the Miller Institute for Basic Research in Science and a NSERC Discovery grant.

References

Abdala-Roberts, L., A. A. Agrawal, and K. A. Mooney. 2012. Ant-aphid interactions on Asclepias syriaca are mediated by plant genotype and caterpillar damage. Oikos 121:1905–1913.

Abdala-Roberts, L., and K. A. Mooney. 2012. Environmental and plant genetic effects on tritrophic interactions. Oikos 122:1157–1166.

Abràmoff, M. D., P. J. Magalhães, and S. J. Ram. 2004. Image processing with ImageJ. Biophotonics international 11:36–42.

Agrawal, A. A. 2004. Resistance and Susceptibility of Milkweed: Competition, Root Herbivory, and Plant Genetic Variation. Ecology 85:2118–2133.

———. 2005. Natural selection on common milkweed (Asclepias syriaca) by a community of specialized insect herbivores. Evolutionary Ecology Research 7:651–667.

Agrawal, A. A., and M. Fishbein. 2006. Plant defense syndromes. Ecology 87:S132–S149.

Agrawal, A. A., and P. A. V. Zandt. 2003. Ecological play in the coevolutionary theatre: genetic and environmental determinants of attack by a specialist weevil on milkweed. Journal of Ecology 91:1049–1059.

Antonovics, J. 1992. Toward community genetics. Plant resistance to herbivores and pathogens: ecology, evolution, and genetics pages 426–449.

Bailey, J. K., J. A. Schweitzer, F. Ubeda, J. Koricheva, C. J. LeRoy, M. D. Madritch, B. J. Rehill, R. K. Bangert, D. G. Fischer, G. J. Allan, and T. G. Whitham. 2009. From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. Philosophical Transactions of the Royal Society B: Biological Sciences 364:1607–1616.

Barbour, M. A., M. A. Fortuna, J. Bascompte, J. R. Nicholson, R. Julkunen-Tiitto, E. S. Jules, and G. M. Crutsinger. 2016. Genetic specificity of a plant–insect food web: Implications for linking

genetic variation to network complexity. Proceedings of the National Academy of Sciences 113:2128–2133.

Barbour, M. A., M. A. Rodriguez-Cabal, E. T. Wu, R. Julkunen-Tiitto, C. E. Ritland, A. E. Miscampbell, E. S. Jules, and G. M. Crutsinger. 2015. Multiple plant traits shape the genetic basis of herbivore community assembly. Functional Ecology 29:995–1006.

Barrios-Garcia, M. N., M. A. Rodriguez-Cabal, J. A. Rudgers, and G. M. Crutsinger. 2016. Soil fertilization does not alter plant architectural effects on arthropod communities. Journal of Plant Ecology page rtw087.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution 24:127–135.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J. A. Gilbert, G. Smith, and R. Knight. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal 6:1621–1624.

Cornelissen, J., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D. Gurvich, P. Reich, H. Ter Steege, H. Morgan, M. Van Der Heijden, et al. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian journal of Botany 51:335–380.

Cronin, J. T., and W. G. Abrahamson. 2001. Goldenrod stem galler preference and performance: effects of multiple herbivores and plant genotypes. Oecologia 127.

Crutsinger, G. M. 2015. A community genetics perspective: opportunities for the coming decade. New Phytologist 210:65–70.

Crutsinger, G. M., W. N. Reynolds, A. T. Classen, and N. J. Sanders. 2008. Disparate effects of plant genotypic diversity on foliage and litter arthropod communities. Oecologia 158:65–75.

Crutsinger, G. M., M. A. Rodriguez-Cabal, A. B. Roddy, K. G. Peay, J. L. Bastow, A. G. Kidder, T. E. Dawson, P. V. A. Fine, and J. A. Rudgers. 2014. Genetic variation within a dominant shrub structures green and brown community assemblages. Ecology 95:387–398.

CRUTSINGER, G. M., and N. J. SANDERS. 2005. Aphid-tending Ants Affect Secondary Users in Leaf Shelters and Rates of Herbivory on Salix hookeriana in a Coastal Dune Habitat. The American Midland Naturalist 154:296–304.

Crutsinger, G. M., S. Y. Strauss, and J. A. Rudgers. 2010. Genetic variation within a dominant shrub species determines plant species colonization in a coastal dune ecosystem. Ecology 91:1237–1243.

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996–998.

Erlandson, S. R., J. A. Savage, J. M. Cavender-Bares, and K. G. Peay. 2015. Soil moisture and chemistry influence diversity of ectomycorrhizal fungal communities associating with willow along an hydrologic gradient. FEMS Microbiology Ecology 92:fiv148.

Floate, K. D., and T. G. Whitham. 1994. Aphid-ant interaction reduces chrysomelid herbivory in a cottonwood hybrid zone. Oecologia 97:215–221.

Fritz, R. S., and P. W. Price. 1988. Genetic Variation Among Plants and Insect Community Structure: Willows and Sawflies. Ecology 69:845–856.

Fritz, R. S., C. F. Sacchi, and P. W. Price. 1986. Competition versus host plant phenotype in species composition: Willow sawflies. Ecology 67:1608–1618.

Garbutt, K., and F. A. Bazzaz. 1987. Population niche structure. Oecologia 72:291–296.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405:220–227.

Gratani, L. 2014. Plant Phenotypic Plasticity in Response to Environmental Factors. Advances in Botany 2014:1–17.

Hersch-Green, E. I., N. E. Turley, and M. T. J. Johnson. 2011. Community genetics: what have we accomplished and where should we be going? Philosophical Transactions of the Royal Society B: Biological Sciences 366:1453–1460.

Johnson, M. T. J. 2008. Bottom-up effects of plant genotype on aphids, ands, and predators. Ecology 89:145–154.

Johnson, M. T. J., and A. A. Agrawal. 2005. Plant Genotype and Environment Interact to Shape a Diverse Arthropod Community on Evening Primrose (Oenothera biennis). Ecology 86:874–885.

Josse, J., M. Chavent, B. Liquet, and F. Husson. 2012. Handling missing values with regularized iterative multiple correspondence analysis. Journal of classification 29:91–116.

Kõljalg, U., K.-H. Larsson, K. Abarenkov, R. H. Nilsson, I. J. Alexander, U. Eberhardt, S. Erland, K. Høiland, R. Kjøller, E. Larsson, et al. 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytologist 166:1063–1068.

Lamit, L., L. Holeski, L. Flores-Rentería, T. Whitham, and C. Gehring. 2016. Tree genotype influences ectomycorrhizal fungal community structure: Ecological and evolutionary implications. Fungal Ecology 24:124–134.

Lamit, L. J., P. E. Busby, M. K. Lau, Z. G. Compson, T. Wojtowicz, A. R. Keith, M. S. Zinkgraf, J. A. Schweitzer, S. M. Shuster, C. A. Gehring, and T. G. Whitham. 2015. Tree genotype mediates covariance among communities from microbes to lichens and arthropods. Journal of Ecology 103:840–850.

Lefcheck, J. S. 2015. piecewiseSEM: Piecewise structural equation modelling inr for ecology evolution, and systematics. Methods in Ecology and Evolution 7:573–579.

Legendre, P., and E. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

Lortie, C. J., and J. H. Cushman. 2007. Effects of a directional abiotic gradient on plant community dynamics and invasion in a coastal dune system. Journal of Ecology 95:468–481.

Lynch, M., B. Walsh, et al. 1998. Genetics and analysis of quantitative traits, vol. 1. Sinauer Sunderland, MA.

MacArthur, R. H. 1972. Geographical ecology: patterns in the distribution of species. Princeton University Press.

Maddox, G. D., and R. B. Root. 1990. Structure of the Encounter between Goldenrod (Solidago Altissima) and Its Diverse Insect Fauna. Ecology 71:2115–2124.

Mcguire, R. J., and M. T. J. Johnson. 2006. Plant genotype and induced responses affect resistance to herbivores on evening primrose (oenothera biennis). Ecological Entomology 31:20–31.

Miller, W. B., and A. E. Weis. 1999. Adaptation of Coyote Brush to the Abiotic Environment and Its Effects on Susceptibility to a Gall-Making Midge. Oikos 84:199.

Mooney, K. A., and A. A. Agrawal. 2008. Plant Genotype Shapes Ant-Aphid Interactions: Implications for Community Structure and Indirect Plant Defense. The American Naturalist 171:E195–E205.

Nakagawa, S., and H. Schielzeth. 2012. A general and simple method for obtaining R 2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4:133–142.

Orians, C. M., and R. S. Fritz. 1996. Genetic and soil-nutrient effects on the abundance of herbivores on willow. Oecologia 105:388–396.

792 Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation Partitioning of Species Data Matrices: Estimation and Comparison of Fractions. Ecology 87:2614–2625.

Pickart, A. J., and M. G. Barbour. 2007. Beach and dune. Barbour, MG & al.(eds.), Terrestrial
795 Vegetation of California, pages 155–173.

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

798 Reddy, T., A. D. Thomas, D. Stamatis, J. Bertsch, M. Isbandi, J. Jansson, J. Mallajosyula, I. Pagani, E. A. Lobos, and N. C. Kyrpides. 2014. The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta) genome project classification. Nucleic acids
801 research page gku950.

Rossi, A. M., and P. Stiling. 1998. The interactions of plant clone and abiotic factors on a gall-making midge. Oecologia 116:170–176.

804 Scheiner, S. M. 1993. Genetics and Evolution of Phenotypic Plasticity. Annual Review of Ecology and Systematics 24:35–68.

Scheiner, S. M., and C. J. Goodnight. 1984. The Comparison of Phenotypic Plasticity and Genetic
807 Variation in Populations of the Grass Danthonia spicata. Evolution 38:845.

Shipley, B. 2000. A new inferential test for path models based on directed acyclic graphs. Structural Equation Modeling 7:206–218.

810 Tack, A. J. M., M. T. J. Johnson, and T. Roslin. 2011. Sizing up community genetics: it's a matter of scale. Oikos 121:481–488.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naive Bayesian classifier for rapid
813 assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology 73:5261–5267.

Wardle, D. A. 2004. Ecological Linkages Between Aboveground and Belowground Biota. Science 304:1629–1633.

Whitham, T. G., J. K. Bailey, J. A. Schweitzer, S. M. Shuster, R. K. Bangert, C. J. LeRoy, E. V. Lonsdorf, G. J. Allan, S. P. DiFazio, B. M. Potts, D. G. Fischer, C. A. Gehring, R. L. Lindroth, J. C. Marks, S. C. Hart, G. M. Wimp, and S. C. Wooley. 2006. A framework for community and ecosystem genetics: from genes to ecosystems. Nature Reviews Genetics 7:510–523.

Whitham, T. G., C. A. Gehring, L. J. Lamit, T. Wojtowicz, L. M. Evans, A. R. Keith, and D. S. Smith. 2012. Community specificity: life and afterlife effects of genes. Trends in Plant Science 17:271–281.

Tables

Table 1: Net effect sizes of genetic variation, phenotypic plasticity, and direct environmental effects on community structure.

Response	Genetic Variation	Phenotypic plasticity	Direct Environmental effects
Ant-Aphid 2012			
Arthropod Richness	0.57	—	—
Arthropod Abundance	0.11	—	0.25
Arthropod Rarefied Richness	—	—	-0.2
Wind 2013			
Arthropod Richness	0.19	-0.13	-0.22
Arthropod Abundance	0.16	-0.11	-0.08
Arthropod Rarefied Richness	0.19	0.12	-0.26
Fungal Rarefied Richness	0.05	—	0.07
Bacterial Rarefied Richness	—	—	0.05

Note: Values represent the sum of effect sizes calculated across different pathways of the structural equation models (figure 3 and 4). Pathways that were not modelled because of lack of *a priori* evidence are denoted by '—'. Note that effects sizes include significant missing pathways (e.g. dotted lines in figure 3).

Figures

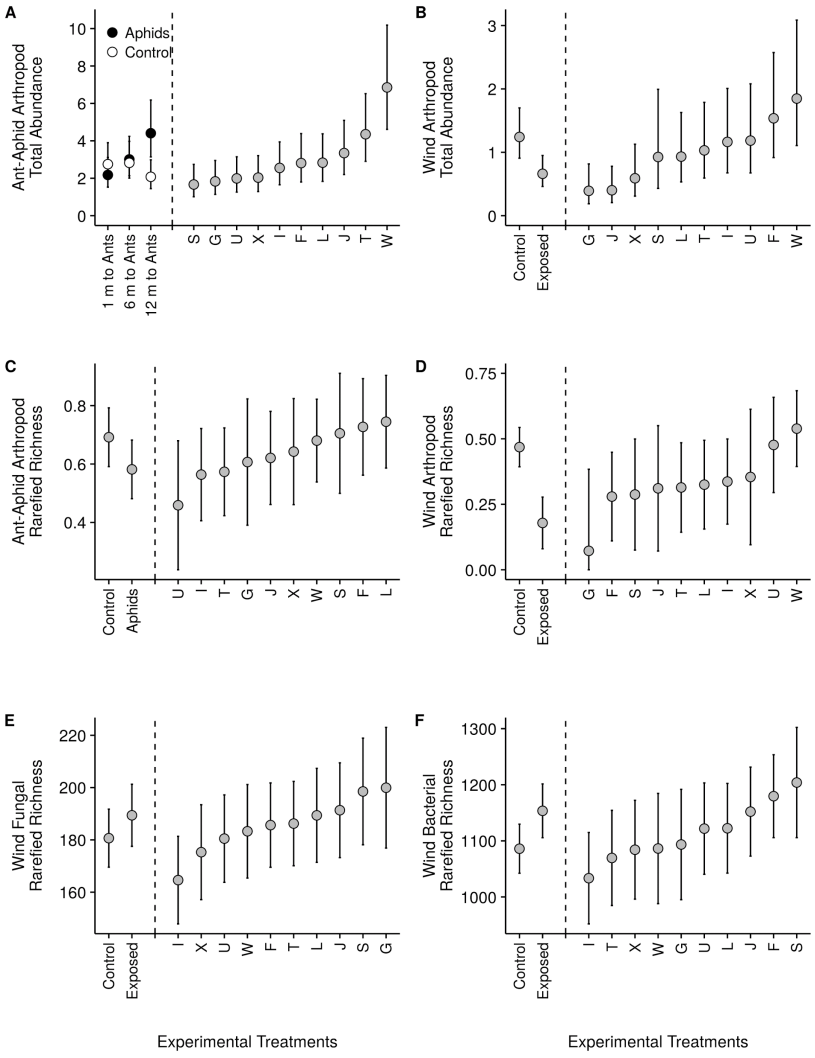

Figure 1: Effects of genetic variation within the willow *Salix hookeriana* as well as biotic (aphid additions and proximity to ant mounds) and abiotic (wind exposure) factors on the structure of above and belowground communities. Willow genotype had strong effects on arthropod abundance in both the ant-aphid (A) and wind exposure (B) experiments. In the ant-aphid experiment (A), arthropod abundance was influenced by the addition of the aphid *Aphis farinosa*, but only at the furthest distance from mounds of the ant *Formica obscuripes*. In the wind experiment (B), wind exposure reduced arthropod abundance. In contrast to arthropod abundance, willow genotype had weak effects on arthropod rarefied richness in both the ant-aphid (C) and wind exposure (D) experiments. In the ant-aphid experiment (C), the addition of aphids reduced the probability of encountering a different arthropod species (rarefied richness). In the wind experiment (D), wind exposure dramatically reduced arthropod rarefied richness. Similar to arthropod rarefied richness, willow genotype had weak effects on the rarefied richness of rhizosphere fungi (E) and bacteria (F). While wind exposure had a weak effect on fungal rarefied richness (E), we found that wind exposed willows hosted a more diverse community of rhizosphere bacteria (F). Points and error bars correspond to the response variable's mean $\pm$ 95% confidence interval. We calculated mean and confidence intervals based on the full models (Tables A1,A3) using the *effects* package in R.

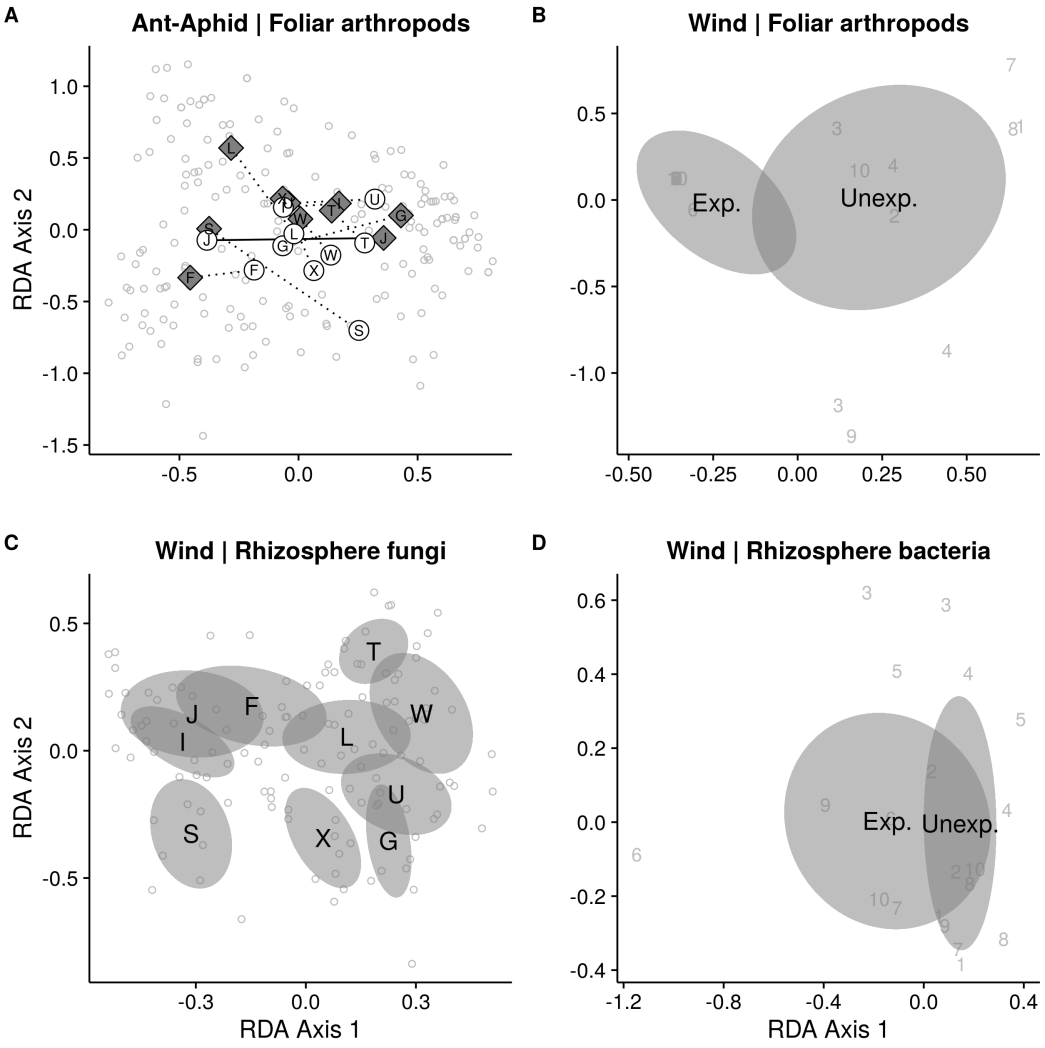

Figure 2: Effects of genetic variation within the willow *Salix hookeriana* as well as biotic (aphid additions) and abiotic (wind exposure) factors on the composition of above and belowground communities. In the ant-aphid experiment (A), the addition of the aphid *Aphis farinosa* (grey diamonds) modified the effect of willow genotype on the composition of the arthropod community. This interaction between willow genotype and aphid treatment was solely due to the differential effect of genotype J (solid line) on the abundance of non-*A. farinosa* aphids in the aphid treatment (table A2). For arthropods in the wind experiment (B), we found that wind exposure had a strong effect on community composition in 2013 (but not 2012, table A3), with no clear effect of willow genotype. Belowground, we found that willow genotype influenced the composition of rhizosphere fungi (C), while the composition of rhizosphere bacteria was only influenced by wind exposure (D). Black text and grey ellipses correspond to the community centroid ± 95% confidence interval. Grey numbers denote blocks and each unique number is the community centroid for the plot within each block. Grey circles mark the location of individual willow communities in multivariate space. We calculated the locations of centroids ± 95% confidence interval and individual samples using redundancy analysis on Hellinger-transformed community data.

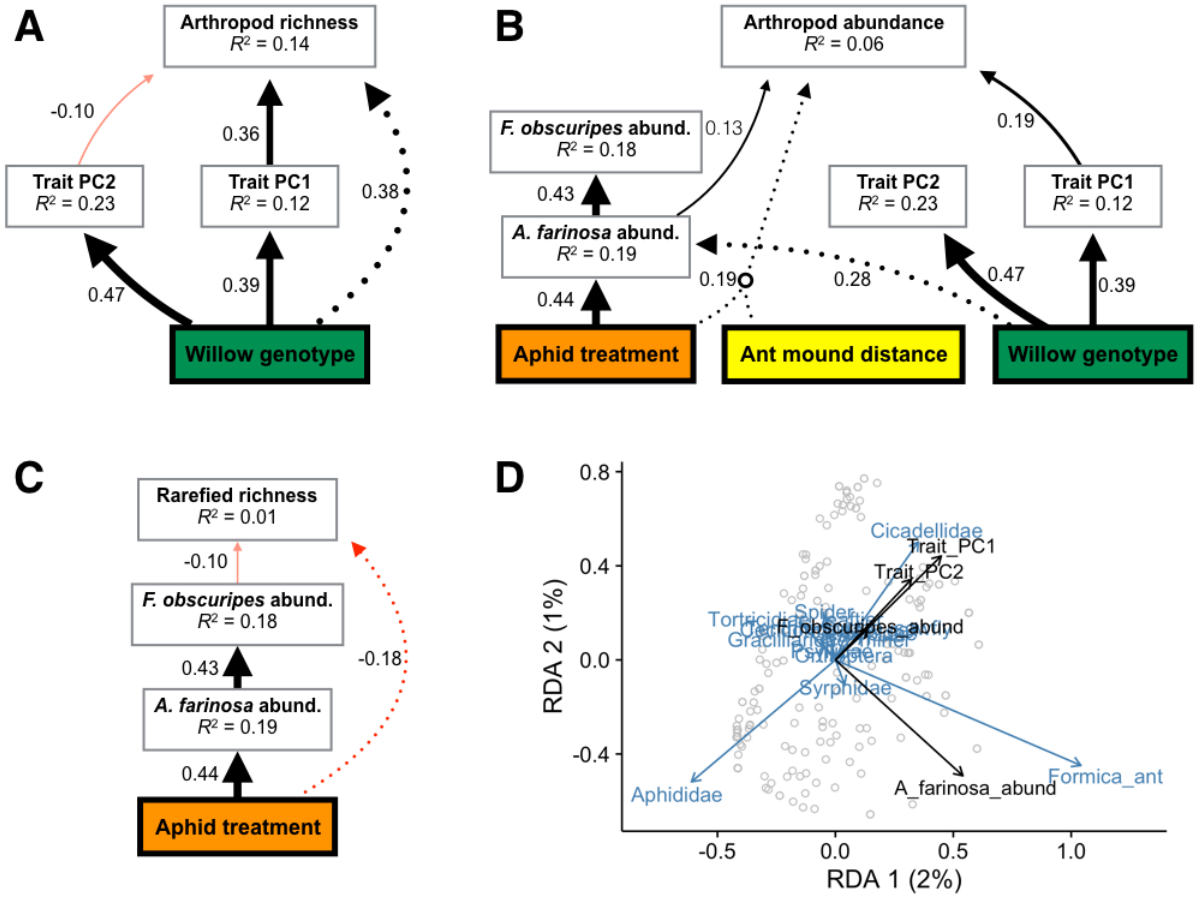

Figure 3: Statistical models of the processes mediating arthropod community assembly in the ant-aphid experiment. Piecewise structural equation models of arthropod richness (A), abundance (B), and rarefied richness (C). Colored and white boxes represent exogenous and endogenous variables, respectively. Solid, single-headed arrows correspond to modeled pathways between predictor and response variables, and may be either positive (black) or negative (red). For clarity, we only plotted paths with standardized coefficients > 0.10, but marginally significant effects ($0.05 < P < 0.10$) are transparent. Numbers next to all arrows represent the standardized path coefficient, which also corresponds to the thickness of arrows. (D) Redundancy analysis illustrating the effect of plant traits (Trait PC1, PC2) and *A. farinosa* abundance on the arthropod community (Hellinger-transformed = square root of proportional abundances of species found on each willow). Black and blue arrows correspond to plant traits and species, respectively, while grey dots represent the position of individual willow communities.

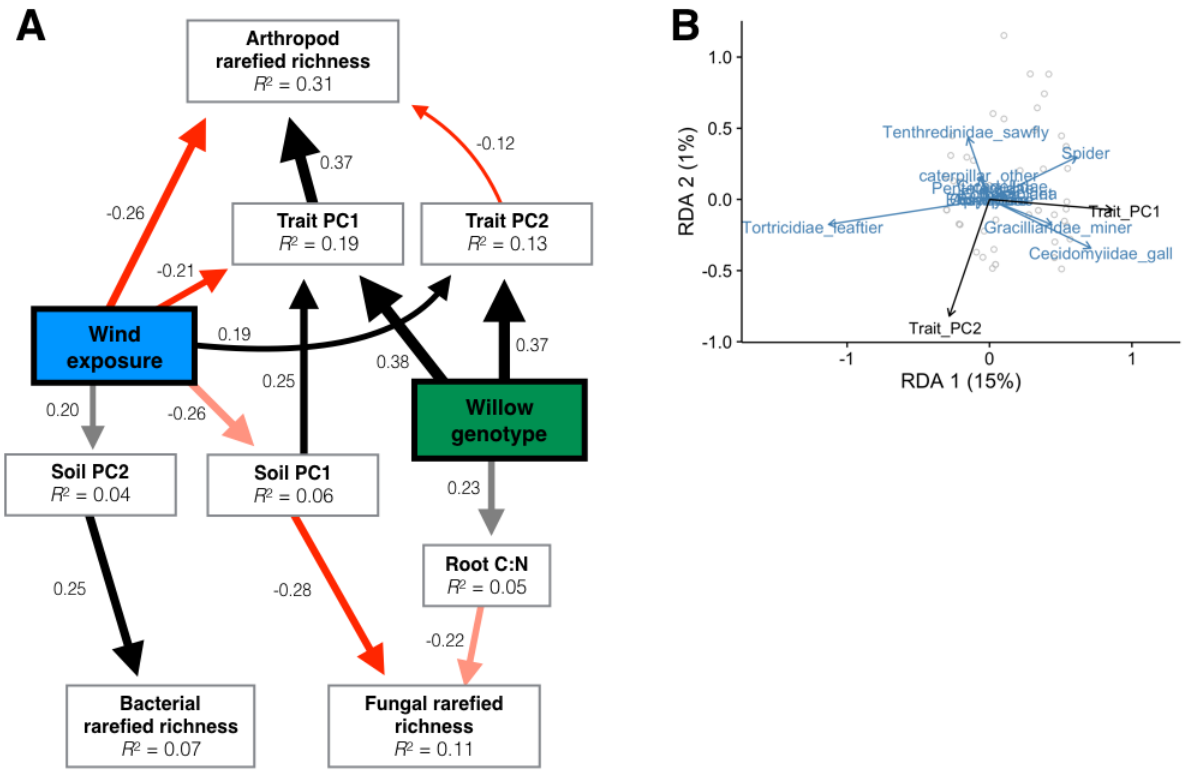

Figure 4: Statistical models of the processes mediating community assembly in the wind experiment for 2013. (A) Piecewise structural equation model of the rarefied richness of foliar arthropods as well as rhizosphere fungi and bacteria. Colored and white boxes represent exogenous and endogenous variables, respectively. Solid, single-headed arrows correspond to modeled pathways between predictor and response variables, and may be either positive (black) or negative (red). For clarity, we only plotted paths with standardized coefficients > 0.10, but marginally significant effects ($0.05 < P < 0.10$) are transparent. Numbers next to all arrows represent the standardized path coefficient, which also corresponds to the thickness of arrows. (B) Redundancy analysis illustrating the effect of plant traits (Trait PC1, PC2) on the arthropod community (Hellinger-transformed = square root of proportional abundances of species found on each willow). Black and blue arrows correspond to plant traits and species, respectively, while grey dots represent the position of individual willow communities.

Online Appendix A: Supplementary Tables

Table A1: Summary of statistical models that analyze the effects of willow genotype, aphid treatment, and distance from ant mounds on the arthropod community, ant-aphid interactions, and plant traits.

Responses	Genotype (G)	E_{aphid}	E_{ant}	G × E_{aphid}	G×E_{ant}	E_{aphid} × E_{ant}	G× E_{aphid} × E_{ant}
Foliar arthropods							
Total richness[a]	**41.35**$_{(9)}$	0.45$_{(1)}$	1.12$_{(1)}$	7.66$_{(9)}$	7.84$_{(9)}$	1.15$_{(1)}$	6.17$_{(9)}$
Total abundance[a]	**34.86**$_{(9)}$	1.79$_{(1)}$	1.42$_{(1)}$	**14.61**$_{(9)}$	9.00$_{(9)}$	8.12$_{(1)}$	9.18$_{(9)}$
Rarefied richness[b]	0.83$_{(9,138.8)}$	**5.34**$_{(1,139.2)}$	0.46$_{(1,8.2)}$	1.81$_{(9,139.7)}$	0.94$_{(9,139.7)}$	0.41$_{(1,140.6)}$	0.70$_{(9,139.4)}$
Community composition[c]	1.52$_{(9,176)}$	2.04$_{(1,176)}$	1.01$_{(1,9)}$	**1.45**$_{(9,157)}$	0.97$_{(9,157)}$	0.69$_{(1,157)}$	0.93$_{(9,148)}$
Ant-aphid interactions							
A. farinosa abund.[a]	**20.83**$_{(9)}$	—	0.55$_{(1)}$	—	10.25$_{(9)}$	—	—
F. obscuripes abund.[a]	2.42$_{(1*)}$	9.77$_{(1)}$	1.68$_{(1)}$	**6.26**$_{(2*)}$	—	—	—
Plant traits							
Height[b]	**15.83**$_{(9,204.2)}$	0.63$_{(1,204.3)}$	0.31$_{(1,9.1)}$	0.93$_{(9,204.5)}$	0.98$_{(9,204.4)}$	0.07$_{(1,204.3)}$	1.62$_{(9,204.7)}$
Shoot count[b]	**65.84**$_{(9)}$	2.76$_{(1)}$	0.21$_{(1)}$	12.11$_{(9)}$	8.80$_{(9)}$	**4.20**$_{(1)}$	9.21$_{(9)}$
Shoot length[b]	**7.27**$_{(9,204.2)}$	2.39$_{(1,204.2)}$	0.10$_{(1,9.1)}$	1.05$_{(9,204.5)}$	0.70$_{(9,204.3)}$	1.24$_{(1,204.3)}$	0.56$_{(9,204.6)}$
Leaf trichome density[a]	38.17$_{(9)}$	0.44$_{(1)}$	0.81$_{(1)}$	**23.17**$_{(8)}$	8.41$_{(9)}$	0.84$_{(1)}$	—
log(Leaf water content)[b]	1.33$_{(9,69.6)}$	0.01$_{(1,69.4)}$	1.02$_{(1,7.1)}$	0.48$_{(8,70.4)}$	0.79$_{(9,69.5)}$	0.36$_{(1,70.6)}$	1.02$_{(7,72.0)}$

Note: We report the test statistic and include the degrees of freedom for each test in parentheses. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$). [a]Likelihood-ratio test and degrees of freedom calculated using a generalized linear mixed-effect model (error distribution = Poisson, link function = log); [b]F-test and Kenward-Roger approximated degrees of freedom calculated using a linear mixed-effect model; [c]F-test calculated using redundancy analysis on Hellinger-transformed community data; *indicates that predictor was modeled as a random effect and its significance was determined using a likelihood ratio test.

Table A2: Summary of abundance responses of key arthropod guilds in the ant-aphid experiment.

Arthropod abundances	Genotype $(G)_{(9)}$	$E_{aphid(1)}$	$E_{ant(1)}$	$G \times E_{aphid(9)}$	$G \times E_{ant(9)}$	$E_{aphid} \times E_{ant(1)}$	$G \times E_{aphid} \times E_{ant(9)}$
Leaf-mining moth (Gracilliaridae)	**26.78**	0.32	2.35	13.56	**20.31**	**4.32**	—
non-*A. farinosa* aphids (Aphididae)	**24.43**	0.01	0.04	**23.16**	6.99	*3.63*	8.16
Leafhopper (Cicadellidae)	**21.92**	0.84	0.01	7.29	11.54	1.67	—
Spiders (Araneae)	*16.24*	0.01	0.10	*15.39*	11.34	0.01	—
non-*F. obscuripes* ants (Formicidae)	**22.43**	**17.70**	1.52	5.21	7.07	0.73	—
Leaf-tiering moth(Tortricidae)	**23.79**	0.81	**9.79**	—	—	*3.77*	—

Note: We analyzed all of these responses using generalized linear mixed-effect models (error distribution = Poisson, link function = log). We report the likelihood-ratio test statistic and include the degrees of freedom for each test as a subscript next to each predictor. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$).

Table A3: Summary of statistical models that analyze the effects of willow genotype and wind exposure on associated communities.

Responses	Genotype (G)	E_{wind}	E_{year}	$G \times E_{wind}$	$G \times E_{year}$	$E_{wind} \times E_{year}$	$G \times E_{wind} \times E_{year}$
Foliar arthropods							
Richness[a]	**28.01**$_{(9)}$	**10.33**$_{(1)}$	**13.55**$_{(1)}$	3.74$_{(9)}$	9.85$_{(9)}$	0.92$_{(1)}$	7.04$_{(9)}$
Abundance[a]	**25.25**$_{(9)}$	**5.48**$_{(1)}$	**6.72**$_{(1)}$	7.33$_{(9)}$	8.22$_{(9)}$	1.65$_{(1)}$	11.85$_{(9)}$
Rarefied richness[b]	1.96$_{(9,71.1)}$	**22.82**$_{(1,7.8)}$	1.13$_{(1,82.7)}$	0.66$_{(9,80.9)}$	—	0.67$_{(1,81.9)}$	—
Community composition$_{2012}$	0.96$_{(9,51)}$	1.26$_{(1,7)}$	—	0.91$_{(9,42)}$	—	—	—
Community composition$_{2013}$	1.17$_{(9,68)}$	**5.70**$_{(1,9)}$	—	0.69$_{(6,62)}$	—	—	
Root-associated Mycorrhiza							
Rarefied richness$_{2013}$[b]	0.87$_{(9,95.1)}$	0.88$_{(1,8.8)}$	—	0.93$_{(9,95.9)}$	—	—	—
Community composition$_{2013}$	**1.01**$_{(9,117)}$	1.18$_{(1,9)}$	—	0.87$_{(9,108)}$	—	—	—
Root-associated Bacteria					—	—	—
Rarefied richness$_{2013}$[b]	1.48$_{(9,99.9)}$	**6.03**$_{(1,7.8)}$	—	1.35$_{(9,100.5)}$	—	—	—
Community composition$_{2013}$	0.93$_{(9,120)}$	*1.38*$_{(1,9)}$	—	0.87$_{(9,111)}$	—	—	—

Note: We report the test statistic and include the degrees of freedom for each test in parentheses. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$). [a]Likelihood-ratio test and degrees of freedom calculated using a generalized linear mixed-effect model (error distribution = Poisson, link function = log); [b]F-test and Kenward-Roger approximated degrees of freedom calculated using a linear mixed-effect model; [c]F-test calculated using redundancy analysis on Hellinger-transformed community data.

Table A4: Summary of statistical models that analyze the effects of willow genotype and wind exposure on soil characteristics and plant traits.

Responses	Genotype (G)	E_{wind}	E_{year}	$G \times E_{wind}$	$G \times E_{year}$	$E_{wind} \times E_{year}$	$G \times E_{wind} \times E_{year}$
Soil characteristics							
Total N[b]	—	$5.08_{(1,9)}$	—	—	—	—	—
Soil moisture[b]	—	$3.52_{(1,9)}$	—	—	—	—	—
Percent organic matter[b]	—	$0.68_{(1,8.4)}$	—	—	—	—	—
Nutrient PC1[b]	—	$1.31_{(1,9)}$	—	—	—	—	—
Plant traits							
Height[b]	$\mathbf{9.13}_{(9,145.3)}$	$\mathbf{29.10}_{(1,9.0)}$	$210.09_{(1,156.3)}$	$0.71_{(9,147.9)}$	$0.80_{(9,157.8)}$	$\mathbf{16.69}_{(1,158.4)}$	$1.84_{(9,160.9)}$
Shoot count[a]	$\mathbf{47.42}_{(9)}$	$\mathbf{9.91}_{(1)}$	$5.68_{(1)}$	$10.70_{(9)}$	$\mathbf{18.26}_{(9)}$	$\mathbf{12.53}_{(1)}$	$5.76_{(9)}$
Shoot length[b]	$\mathbf{4.97}_{(9,144.2)}$	$\mathbf{10.44}_{(1,9.0)}$	$75.36_{(1,158.5)}$	$0.84_{(9,146.9)}$	$1.61_{(9,160.1)}$	$0.05_{(1,160.7)}$	$0.70_{(9,163.2)}$
Leaf water content[b]	$4.90_{(9,129.0)}$	$0.97_{(1,8.7)}$	$2.93_{(1,139.7)}$	$0.47_{(9,132.0)}$	$\mathbf{2.80}_{(9,141.6)}$	$2.03_{(1,141.5)}$	$1.56_{(9,144.1)}$
Leaf trichome density$_{2012}$[b]	$\mathbf{67.31}_{(9)}$	$0.02_{(1)}$	—	$10.45_{(9)}$	—	—	—
SLA_{2013}[b]	$\mathbf{4.21}_{(9,122.5)}$	$0.34_{(1,8.9)}$	—	$1.19_{(9,123.4)}$	—	—	—
Leaf $C:N_{2013}$[b]	$\mathbf{4.88}_{(9,70.48)}$	$1.54_{(1,7.8)}$	—	$1.31_{(9,71.6)}$	—	—	—
Root $C:N_{2013}$[b]	$0.85_{(9,107.0)}$	$0.31_{(1,8.7)}$	—	$0.33_{(9,107.5)}$	—	—	—

Note: We report the test statistic and include the degrees of freedom for each test in parentheses. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$). [a]Likelihood-ratio test and degrees of freedom calculated using a generalized linear mixed-effect model (error distribution = Poisson, link function = log); [b]*F*-test and Kenward-Roger approximated degrees of freedom calculated using a linear mixed-effect model.

Table A5: Summary of abundance responses of key arthropod guilds in the wind experiment.

Arthropod abundances	Genotype $(G)_{(9)}$	$E_{wind(1)}$	$G \times E_{wind(9)}$	$E_{year(1)}$	$G \times E_{year(9)}$	$E_{wind} \times E_{year(1)}$	$G \times E_{wind} \times E_{year(9)}$
Leaf-mining moths (Gracilliaridae)	**17.15**	**9.42**	—	**4.26**	—	0.09	—
Gall midges (Cecidomyiidae)	**19.26**	**17.59**	—	**38.07**	—	—	—
Leaf-tiering moths (Tortricidae)	**24.78**	1.34	11.50	**117.19**	—	2.65	—
$Aphids_{2012}$ (Aphididae)	4.31	0.32	—	—	—	—	—
Spiders (Araneae)	8.27	**6.04**	—	**6.54**	—	0.20	—

Note: We analyzed all of these responses using generalized linear mixed-effect models (error distribution = Poisson, link function = log). We report the likelihood-ratio test statistic and include the degrees of freedom for each test as a subscript next to each predictor. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$).

Table A6: Summary of abundance responses of key fungal operational taxonomic units (OTUs) in the wind experiment.

Fungal abundances	$E_{wind(1)}$	Genotype $(G)_{(9)}$	$G \times E_{wind(9)}$
Humicola nigrescens (OTU 54)	0.28	**36.41**	—
Ascomycota sp. (OTU 68)	0.18	**24.38**	—
Tremellomycetes sp. (OTU 134)	**11.12**	**33.41**	—
Cryptococcus terricola (OTU 66)	0.86	**46.81**	**20.38**
Cryptococcus albidus (OTU 119)	2.16	**35.54**	7.43
Cadophora finlandica (OTU 262)	0.29	**43.37**	—
Pseudogymnoascus roseus (OTU 1606)	0.01	**22.70**	13.58
Mortierella sp. (OTU 113)	**4.14**	**21.08**	9.21

Note: We analyzed all of these responses using generalized linear mixed-effect models (error distribution = Poisson, link function = log). We also specified an offset in the model (log(total fungal abundance)) to account for differences in sampling effort. We report the likelihood-ratio test statistic and include the degrees of freedom for each test as a subscript next to each predictor. Pathways that were not modelled because of insufficient data are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, *italic* $P < 0.10$, normal $P > 0.10$).

Table A7: Summary of loadings and variance explained by first two components from separate principal components analysis (PCA) of aboveground plant traits.

Individual traits	Trait PC1	Trait PC2
Ant-aphid experiment		
Plant height	0.51	-0.49
Shoot count	0.43	0.25
Shoot length	0.64	-0.14
Leaf trichome density	-0.12	-0.72
Leaf water content	-0.36	-0.39
Variance explained	39%	29%
Wind experiment, 2012		
Plant height	0.55	-0.18
Shoot count	0.47	0.29
Shoot length	0.68	-0.07
Leaf trichome density	-0.08	0.64
Leaf water content	0.09	0.69
Variance explained	36%	24%
Wind experiment, 2013		
Plant height	0.46	-0.29
Shoot count	0.49	-0.27
Shoot length	0.45	-0.20
Leaf water content	-0.36	-0.52
Specific Leaf Area (SLA)	-0.46	-0.42
Leaf C:N	-0.04	0.59
Variance explained	45%	26%

Note: All traits were scaled to mean = 0 and SD = 1 prior to PCA to give them equal weight in the analysis.

Table A8: Summary of loadings and variance explained by first two components from principal components analysis (PCA) of soil properties.

Soil properties	Soil PC1	Soil PC2
NO_3^-	-0.25	-0.09
NH_4^+	-0.15	-0.02
Ca^{2+}	0.23	0.45
Mg^{2+}	0.29	0.34
K^+	0.23	-0.20
$H_2PO_4^-$	0.23	0.09
Fe^{3+}	0.33	-0.20
Mn^{2+}	0.19	0.16
Cu^{2+}	0.24	-0.30
Zn^{2+}	0.17	-0.40
$B(OH)_4^-$	0.20	-0.26
SO_4^{2-}	0.26	0.10
Pb^{2+}	0.21	-0.10
Al^{3+}	0.27	-0.23
Cd^{2+}	0.05	0.40
Organic matter (%)	0.35	0.14
Soil moisture	0.30	0.04
Variance explained	38%	14%

Note: All soil characteristics were scaled to mean = 0 and SD = 1 prior to PCA to give them equal weight in the analysis.

Table A9: Redundancy analyses of foliar arthropods and root-associated fungi and bacteria in the last year of the wind experiment.

Community composition	E_{wind}	Trait PC1	Trait PC2	Root C:N	Soil PC1	Soil PC2
Arthropods	$0.83_{(1,9)}$	$\mathbf{12.05}_{(1,76)}$	$0.65_{(1,76)}$	—	—	—
Mycorrhiza	—	—	—	$1.17_{(1,115)}$	$1.89_{(1,8)}$	$0.85_{(1,8)}$
Bacteria	—	—	—	$1.31_{(1,116)}$	$1.90_{(1,8)}$	$0.81_{(1,8)}$

Note: We report F-statistics and degrees of freedom in parenthesis. Effects that were not modelled because we did not hypothesize their direct effects are denoted by '—'. Font type denotes statistical significance (**bold** $P < 0.05$, italic $P < 0.10$, normal $P > 0.10$).