Insect herbivory reshapes a native leaf microbiome

Parris T Humphrey[1,3,✉] **and Noah K Whiteman**[2,3]

[1]Department of Organismic & Evolutionary Biology, Harvard University, Cambridge, MA 02138 USA.
[2]Department of Integrative Biology, University of California, Berkeley, CA 91604 USA.
[3]Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA.

Abstract

Insect herbivory is pervasive in natural plant communities, but its impact on microbial plant colonizers has not been well-studied outside of laboratory or agricultural systems. By calibrating sequencing-based bacterial detection with an absolute standard of bacterial load, we find that insect herbivory on a native herbaceous plant amplifies within-host abundance of most leaf microbiome (phyllosphere) taxa, most dramatically so for the phytopathogen *Pseudomonas syringae*. This insect-associated amplification of absolute bacterial abundance was reflected in community-wide compositional shifts towards lower ecological diversity. Such community-level analyses using relative abundances mask the extent and direction of underlying abundance changes, as these are non-identifiable in the presence of variation in bacterial load. In parallel, using 12 genetically diverse field isolates of the dominant bacterial genus *Pseudomonas*, we found that anti-herbivore plant defenses amplify the growth of only a subset of *P. syringae* strains, and such strain-level variation in within-host fitness are masked at coarser taxonomic levels. In the field, herbivore damage was inversely correlated with plant reproductive success and was highly clustered across plant patches, which predicts extensive co-clustering with the *P. syringae* population across plants. Together, these results establish that herbivores can alter the phylogenetic and spatial structure of plant microbiomes by generating variation in within-host bacterial fitness. Plant disease surveillance and control efforts may also benefit from considering how herbivorous insects shape the population dynamics and epidemiology of phytopathogens.

co-infection, *Pseudomonas*, Brassicaceae, phyllosphere, microbiome

Correspondence: `phumphrey@fas.harvard.edu`

AUTHOR CONTRIBUTIONS
PTH and NKW designed the study. PTH carried out the study and analyzed the data. PTH and NKW wrote the manuscript. The authors declare that they have no conflicts of interest.

PRE-PRINT

Introduction

For many organisms, attack by multiple enemies is inevitable and often occurs sequentially during the lifetime of individual hosts. While hosts can vary in genetically-based resistance or tolerance to each enemy, prior attack can alter the course of future attacks and generate knock-on effects at larger spatial and temporal scales (1–4). Accounting for co-infection, or its after-effects, is crucial for understanding the biology and epidemiology of focal host–parasite interactions (4–6) given the large effects co-infection can have on host health, parasite fitness, and population-level disease prevalence. But rarely has this perspective been extended to studies of diverse host-associated microbial communities ('microbiomes'). Instead, microbiome studies tend to focus on effects of genotype or abiotic variation on microbiome diversity patterns (7–13), in relative isolation from additional biotic interactions that may otherwise impact the phenotype or fitness of the host. It therefore remains an important question how biotic interactions with other colonizers shape the assembly, abundance, and diversity of host-associated microbiomes.

For plants, there is tremendous interest in understanding the structure and function of the microbiome both for applied purposes, such as engineering growth promotion and disease resistance (14, 15), and as a model system for host–microbial symbioses. Insect herbivory represents a pervasive threat to both native and agricultural plants (16). Herbivory alters plant phenotypes through tissue damage and induction of plant defenses, which can alter susceptibility of plants to attack by insects (17) as well as microbes (18–20). Thus, factors that influence the impact of herbivores on hosts will likely affect the colonization and growth of plant-associated microbes. While insect herbivores (16) and plant-associated microbes have clear effects on plant phenotypes and fitness (21), these two factors are generally considered independently. Our study addresses this gap by explicitly considering how patterns of abundance and diversity of leaf-colonizing bacterial taxa are altered in the presence of insect herbivory and to explore the associations among herbivory, bacterial infection, and plant fitness in natural plant populations.

We tested how insect herbivory alters the distribution and abundance of leaf-colonizing bacteria in native bittercress (*Cardamine cordifolia*, Brassicaceae). We first used marker gene sequencing (16S rRNA) coupled to paired leaf culturing to quantitatively assess changes in the distribution and abundance of bacterial taxa. We expressed microbial abundance relative to internal 16S sequences from host-derived organelles in order to link sequence counts of bacteria to their predicted absolute abundance within host tissue. This approach allowed our 16S sequence data to provide a quantitative measure of absolute abundance across the population of sampled hosts for entire suites of bacterial taxa. We then assessed the extent of co-clustering between microbial abundance and intensity of insect herbivory at the plant patch scale across our study population, and we related microbe–herbivore co-aggregation to fruit set, a component of plant fitness. In parallel, we directly examined variation in sensitivity to inducible anti-herbivore defenses among 12 genetically diverse, bittercress-derived isolates of *Pseudomonas* spp. bac-

teria. We did so by experimentally infecting accessions of native bittercress in which prior herbivory was simulated by exogenously pre-treating plants the hormone jasmonic acid (JA).

Our experiments reveal that herbivory—and the defense phenotypes typically induced by herbivory—can strongly impact bacterial diversity and abundance within leaf microbiomes and amplify specific strains of potentially phytopathogenic bacterial taxa. Crucially, this degree of abundance variation cannot be revealed by traditional compositional analysis, which masks the extent and direction of within-host variation in bacterial load detected by our analyses. By linking marker gene counts to an absolute standard our study revealed how insect herbivory generates variation in bacterial loads at leaf and patch scales within a natural plant population. Together, our results highlight the importance of (*a*) accounting for prevalent but "non-focal" biotic interactions hosts have with other colonizers in their natural contexts, and (*b*) using detection and analytical approaches to quantify these effects on components of microbial fitness.

Methods

Field studies of herbivore–bacteria co-infection. The main goal of our field studies was to measure how the distribution and abundance of leaf-colonizing bacteria is altered in plant tissues affected by insect herbivory. To achieve this, we surveyed herbivore damage arising from the specialist leaf-mining insect herbivore *Scaptomyza nigrita* (Drosophilidae; figure S1f–g) across transects of plants within two natural populations of native bittercress (*Cardamine cordifolia*, Brassicaceae). We then sampled leaf tissues to detect and enumerate bacteria to examine how their abundance varies across damaged and undamaged leaves.

Although not our primary focus, these field studies were also designed to test how mid-season pre-treatment with the exogenous plant defense hormones jasmonic acid (JA) or salicylic acid (SA) impact plant attack rates by *S. nigrita*. JA induces canonical anti-herbivore defenses in plants (20), including bittercress, and JA-induced bittercress can locally deter adult *S. nigrita* and reduce larval feeding rates (22). SA treatment canonically suppresses anti-herbivore defenses (20) and can increase local susceptibility to *S. nigrita* (23). Thus, treatment with either hormone can modify the foraging behavior of *S. nigrita* in this system. We therefore hypothesized that eliciting early-season defenses in the plants would perturb patterns of herbivory across the growing season.

To conduct these experiments, we marked plots of native bittercress along sub-alpine and alpine streams near the Rocky Mountain Biological Laboratory (RMBL) in each of two years (2012, site Emerald Lake, EL; 2013, site North Pole Basin, NP; figure S1a–b). We marked plant plots ($n = 24$, site EL; $n = 60$, site NP) separated by at least 1 m and no more than 100 m along 0.5–1 km transects at each site (figure S1c–d). The design of this hormone treatment field experiment was described and conducted previously for site EL (23) and was then replicated, with slight modification, at site NP in the present study (see Supplemental Methods). At site NP, near the end of the bittercress growing season (13–18 August 2013), we monitored the accumulation of *S. nigrita* leaf-miner damage by counting leaf-mines on all leaves from $n = 16$ stems per plant patch, for both plant patches per plot (figure S1e). This gave a total stem-level sample size of 768 for EL and 1920 for NP. We then examined whether hormone pre-treatment of field-grown bittercress perturbed rates of herbivore attack across plant patches in our study population (NP); results for site EL have been previously published (23). A description of our statistical approach for site NP can be found in the Supplemental Methods.

In both studies, we collected leaf tissue to quantify the abundance and diversity of bacteria that had colonized the leaf interior. While our selection of leaf tissue was randomized in all cases, our sampling strategy differed between study sites. At EL, we collected a single leaf from a randomly determined leaf position along each of four stems chosen randomly from the 16 stems surveyed per patch. We excluded the top two leaves from sampling, which were generally $< 1\text{cm}^2$—too small to offer adequate amounts of leaf material for DNA extraction. Leaves along each stem were included irrespective of herbivore damage status. Thus, leaves with herbivore damage in our leaf sample set from EL reflected the leaf-level prevalence of herbivore damage in the EL bittercress population.

At site NP, we randomly sampled a damaged and undamaged leaf from each of four stems to ensure we achieved even representation from both damage classes from all plant patches. Thus, our selection of damaged and undamaged leaf tissue was constrained by the availability of leaves of both types along a single stem. For stems that had either only one type of leaf (i.e., damaged or undamaged), we continued to sample adjacent stems until one harbored both leaf types. An additional distinction from site EL was that, at NP, tissue from individual leaves from four separate stems were pooled to comprise a single sample, resulting in one pair of matched damaged and undamaged samples from each plant patch. In parallel, we counted the number of non-aborted fruits produced per stem at the end of the growing season at site NP, allowing us to relate plot-level variation in herbivore and/or bacterial loads to this metric of plant fitness.

Amplicon sequencing of bacteria in leaf tissues. We quantified bacterial abundance in leaf tissues using next-generation amplicon sequencing of the bacterial 16S rRNA locus using the Illumina MiSeq platform. Following surface sterilization and DNA extraction of 192 leaf discs ($\sim 0.8\text{cm}^2$) from site EL and 192 tissue pools from site NP (4 four discs per pool), we amplified bacterial 16S following protocols established for the Earth Microbiome Project (24). We amended this protocol by including peptic nucleic acid (PNA) PCR clamps into reaction mixtures to reduce amplification of chloroplast- and mitochondria-derived 16S amplicons, following (25), which was highly effective at reducing the proportion of host-derived 16S reads per library in our sample sets (figure S2). Library preparation and sequencing was conducted at Argonne National Laboratory, and a link to detailed sample preparation and sequencing protocols can be

PRE-PRINT

found in the Supplemental Materials.

We used DADA2 (26) to error-correct, trim, quality-filter, and merge the paired-end sequencing reads that passed error thresholds off the sequencer. Of the approximately 4 million raw reads, $\sim 90\%$ were retained following quality control via DADA2 (figure S3), and these reads were then delineated into exact amplicon sequencing variants (ASVs). We assigned provisional taxonomy to all ASVs via the Ribosomal Database Project Bayesian classifier (27) to closely inspect putatively plant- or fungus-derived sequences. All such ASVs matching non-Brassicaceae plant- or fungi-derived 16S were discarded, while those annotated as *Cardamine* (for chloroplast) or *Arabidopsis* (for mitochondria) were retained and pooled as host-derived counts. We then manually verified monophyly of putative bittercress-derived ASVs after constructing maximum likelihood trees of all non-bacterial bASVs (see Supplemental Methods). We discarded ASVs seen in only one library or with fewer than 10 counts across the dataset. At no point did we perform down-sampling (i.e., rarefaction) on our 16S counts.

Quantifying bacterial abundance. Using our cleaned 16S datasets, we devised and validated an estimator of the abundance of each bacterial ASV within host tissues. To establish the estimator, we hypothesized that the likelihood that a 16S sequencing read derives from the host, as opposed to a bacterium, depends on the concentration of bacteria within the starting material. Thus, the composition of the sequencing data, in terms of host versus bacteria-derived 16S reads, may provide information about the underlying density of bacteria. This occurs, we reasoned, because DNA templates of the two sources compete as targets during the amplification reaction, and biases towards one or the other will accrue exponentially. By this logic, the logarithm of the relative abundance of bacteria–to–host 16S counts should capture information about the density of bacterial cells in the starting material. Accordingly, for each sample, we calculated $\gamma = \ln(r_B/r_H)$, where r_B and r_H are the read counts of bacteria- and curated host-derived 16S counts for a given sample, respectively. r_B can be calculated at any taxonomic level, ranging from the single bASV to all of the bacteria present in the sample, by summing the sequence counts at the desired taxonomic scale.

We then validated this estimator by empirically testing the relationship between γ and an independent measure of bacterial abundance (i.e., density) in leaf tissue. To do this, we randomly selected a sub-set of $n = 101$ leaf samples from the sample set from site EL to conduct bacterial culturing of paired leaf discs alongside 16S marker gene amplification (figure 1a). These samples were surface-sterilized, homogenized, and plated on non-selective King's B media to enumerate bacterial colony forming units (CFU) per g starting leaf material, following (23). This approach is valid given that a majority of bacterial taxa typically found to colonize leaf tissues can be cultivated in the laboratory on rich media (23, 28).

We then estimated the slope and intercept in a linear regression between the directly observed $\log_{10}$ CFU g^{-1} leaf tissue (hereafter log CFU) and the predictor variable γ for the set of $n = 101$ calibration samples. We fit a Bayesian regression model in order to generate sampling distributions of the parameter values of slope, intercept, and residual variation, so that uncertainty in the fit could be incorporated into downstream analyses. We fit this model with Stan in its interface to **R** (29) via Rstan (30) using package brms (31, 32). Posterior distributions of model parameters were generated from 8000 iterations (following 8000 burn-in iterations) spread across four chains.

Overview of statistical modeling approach. Having established that γ represents an empirically valid estimator of absolute bacterial density in leaf tissues, we devised a two-stage Bayesian regression modeling approach that allowed us to link bacterial 16S sequencing counts to absolute bacterial abundance via the relationship between γ and log CFU leaf tissue described above. The first stage involved fitting and comparing a series of models which estimate how γ varies as a function of herbivore damage in leaves, which was our focal predictor variable. We then took the candidate best stage-1 model and used it to generate sets of simulated response values ($\tilde{\gamma}$) predicted by the model parameters that were fit to the original γ input data. This is technically known as the posterior predictive distribution of the input data, and it can be thought of as representing the range of values that the response variable is likely to take under the specified model. The posterior predictive distribution, $p(\tilde{\gamma}|\gamma, \theta)$, reflects both uncertainty in the parameter estimates ($p(\theta|\gamma)$) in the model as well as the extent of group-level and residual variation present in the data itself.

The next stage involved combining the information in $\tilde{\gamma}$ with information on the relationship between γ and observed log CFU. We did this in order to report our bacterial abundance estimates, based initially on 16S count data, on the scale of predicted log CFU per unit leaf mass, which is a more directly interpretable measure that relates to within-host bacterial fitness. To achieve this, we used the distribution of $\tilde{\gamma}$ generated from the stage-1 model as input to the model we fit between γ and observed log CFU from the set of 101 calibration samples described above. Instead of using point estimates of slope and intercept parameters from the calibration CFU model, we incorporated uncertainty in the parameter values by drawing intercept, slope, and residual error parameters from their joint posterior distribution for each simulated leaf sample separately. We then used these posterior parameter values to calculate a (posterior) predicted log CFU value of for each value of the input predictor variable, $\tilde{\gamma}$.

This two-stage modeling approach maximally incorporates uncertainty in the model fit for γ as well as in the relationship between γ and observed log CFU. The product of this two-stage modeling procedure, for each focal model, was a distribution of posterior predicted abundance values on the log CFU scale, and we used this as the proximal response variable when calculating differences in predicted abundance of bASVs across damaged and undamaged samples.

PRE-PRINT

Details of statistical modeling approach. *Stage 1: γ models.* We ran hierarchical Bayesian (log) normal regression models for data collected at sites EL and NP separately. To calculate γ, we took r_B at the bASV-level for all bASVs within each of the 14 most abundant bacterial Families, which together comprised $> 95\%$ of total bASV counts across the datasets. We ran a series of increasingly flexible models for each bacterial family (at each site, separately), which included group-level model terms to capture variation in the intercept and slope of the relationship between bASV abundance and herbivore damage. We also included terms to model heterogeneity of variances between sample classes, as well as overall skew. We compared models using approximate leave-one-out Bayesian information criterion (LOO-IC) (33), and we heuristically took the model displaying the lowest LOO-IC as the candidate best-fitting model for downstream purposes. Further details of model specification can be found in the Supplemental Methods.

Stage 2: absolute abundance models. From each candidate best stage-1 model, we generated $\tilde{\gamma}$ values by conducting 200 replicate simulations, each the size of the original input data, from the posterior distributions of model parameters. This procedure generated 200 replicate estimates per input sample of the predicted logCFU for each individual bASV contained within the specified bacterial family. We then used these posterior predicted estimates of log CFU to calculate the difference in bacterial abundance of each bacterial family between damaged and undamaged leaf sets. To calculate abundance values at the higher taxonomic scale of family, we summed the predicted counts (on the linear scale) of the individual bASVs at the level of the leaf sample, and we did this for each of the 200 posterior simulations separately. The decision to include a bASV in the family-level sum with an observed count of zero for a particular leaf sample was determined probabilistically after modeling how the predicted bASV abundance would permit detection given the detection threshold for each particular sample. This procedure is described in detail in the Supplemental Methods.

With these family-level abundances, we computed the difference in median abundances between damaged and undamaged leaf sets. We did not use means of predicted log CFU values because of the potentially large influence of positive outliers on the log scale on summary statistics calculated on the exponentiated scale (i.e., $\text{med}(\log x) = \log \text{med}(x)$, while $\log \text{E}[x] \neq \text{E}[\log x]$). We then expressed the median difference in predicted bacterial density between leaf types in units of $\log_2$ so that we could represent effects in terms of the number of bacterial doublings (i.e., cell divisions) that comprise the predicted abundance differences between sample types. We repeated this calculation for each of the 200 replicate posterior draws to generate the posterior predictive distribution of the difference in within-host density for bacteria under damaged versus undamaged leaves. This measure reflects the central tendency of the predicted effect sizes as well as parameter uncertainty at each stage of our modeling procedure.

Estimating compositional changes in leaf bacterial communities. We examined how differential growth under herbivory of distinct members of the bacterial community would impact overall signatures of ecological diversity and similarity. Using our predictive posterior distributions of bacterial abundances (i.e., median log CFU), we calculated standard ecological summary statistics of community-level compositional variation across herbivore-impacted versus non-impacted leaves (Shannon evenness J', and Shannon–Jensen divergence S–J; see Supplemental Methods for details). To do this, we used the vector of 200 simulated median predicted abundances in each sample class and transformed these values into relative abundances by dividing by the total abundance of the simulated communities. For the field studies, sample classes were damaged and undamaged leaves, and abundances were summed at the family level. We then calculated the posterior predictive distribution of J' using the 200 simulated damaged and undamaged communities, compared the distribution of differences between leaf classes, and finally examined evidence for community-level differentiation via S–J divergence.

Population-level analysis of herbivore–bacteria co-aggregation. We examined spatial variation in the extent of herbivore damage across plant patches and used this information to calculate predicted bacterial abundance at the whole-patch level. To do this, we first considered how inducible defenses applied at the patch level (see Field Studies, above) impacted subsequent levels of *S. nigrita* damage that accumulated over the remainder of the growing season. This analysis was formerly conducted and published for site EL (23), and we implemented a similar approach for site NP in the present study (see Supplementary Methods for details).

We next assessed how patch-level herbivore variation, given that it occurred, was correlated with bacterial infection intensity at the field-scale by focusing on the most highly abundant bASV in each sample site (Pseudomonas3). Leaf-level abundance of this bASV was estimated using the two-stage log-normal–abundance-model approach (as above) for each leaf in the herbivory survey data for site NP. We then summed the predicted abundance (on the linear scale) of bacteria across leaves within each plant patch and used these predicted bacterial sums to calculate the cumulative proportion of the total Pseudomonas3 population harbored by plant patches with differing levels of herbivory as measured by our study, which portrays the extent of co-aggregation of herbivores and bacteria across the plant population. Pseudomonas3 was mapped to the putatively phytopathogenic *P. syringae* species complex via maximum likelihood estimation of a phylogenetic tree using RAxML 8.2, (34) comprising all *Pseudomonas* bASVs ($n = 49$) as well as the non-redundant species-annotated sequences within the combined set of the top ten BLAST hits for each individual *Pseudomonas* bASVs in our dataset.

Experimental infections *in planta*. We directly evaluated within-host performance of bacteria in the presence of anti-herbivore inducible defenses using a greenhouse experiment

PRE-PRINT

with field-derived accessions of bittercress plants and their bacterial colonizers. Specifically, six *Pseudomonas* spp. strains were randomly selected from within each of the two dominant *Pseudomonas* clades found within bittercress (*P. syringae*, and *P. fluorescens*; (23); $n = 12$ strains total) and were infected in replicate into a series of bittercress plants that had been pre-treated 3 days prior with JA (1 mM; Sigma) or a mock carrier solution. We used 32 distinct bittercress clones grown in the greenhouse at University of Arizona for up to 12 months which were initially isolated as rhizomes from several sites within 2 km of the RMBL in 2012. Each of 15 leaves per plant clone were randomized to be infiltrated with bacteria ($n = 1$ per strain, per plant) or a sterile mock solution ($n = 3$ per plant). After 2 days of within-host growth, we sampled, sterilized, and weighed leaf discs from all plants. Samples were then flash frozen in liquid nitrogen and held at -80C. Subsequently, all samples were homogenized and dilution-plated on non-selective rich King's B media, following (23). We counted colonies on all visible dilutions after 3 days of growth at 28 °C, averaged counts across replicate dilutions, and calculated the density of CFU per gram of input leaf tissue. Densities of each bacterial strain were compared between sample classes (treated versus un-treated) using Gaussian Bayesian regression models in which the response variable was calculated as the $\log_2$ of the observed CFU per leaf sample.

To evaluate how anti-herbivore defenses would be expected to impact the community-level composition of this set of *Pseudomonas* spp. strains, we used the posterior abundances from the models above to generate replicate vectors of relative abundance values under treated versus untreated conditions. We then considered how the observed strain-level differences would translate into patterns at higher taxonomic levels. To do this, we summed strain-level abundances at the clade-level (*P. syringae* or *P. fluorescens*), or simply at the community-level (i.e., all 12 strains), for treated and untreated leaf classes separately.

Results

Host-derived 16S read counts are an index of bacterial load in host tissues. In order to study within-host abundance variation of leaf-colonizing bacteria *en masse*, we first linked marker gene sequence data to bacterial density within tissue samples. By taking advantage of the fact that most bittercress-associated bacteria can be cultured on rich media in the laboratory (23), we paired 16S amplicon sequencing and bacterial culturing on a subset of our samples ($n = 101$) in order to test for a relationship between the proportionality between bacteria- and host-derived 16S reads (γ; see Methods) and the density of colony forming units (log CFU g^{-1}) grown from the same leaf (figure 1a). We found a clear positive association between γ and log CFU (figure 1b; table S1; intercept 4.92 [4.74; 5.11], slope 0.29 [0.23; 0.35], 95% posterior credible intervals). This result validated our usage of γ as an estimator of absolute bacterial abundance using 16S sequence data and allowed us to examine how the population size of distinct bacterial groups varied across leaves with and

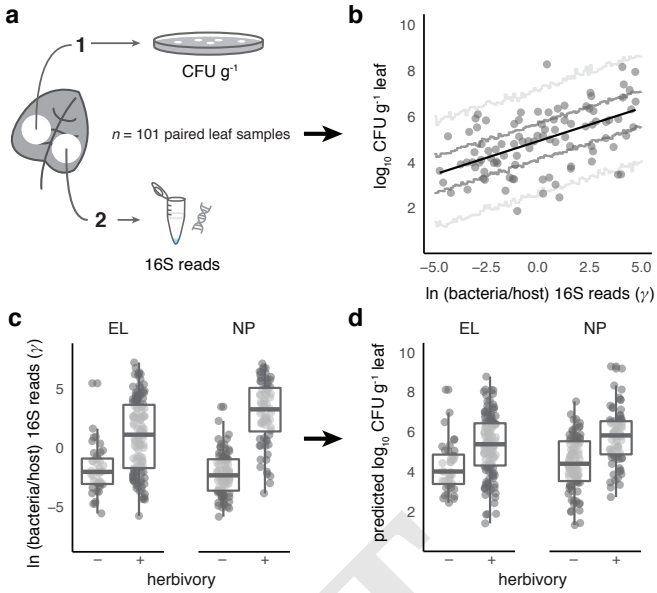

Fig. 1. Relating bacterial counts to 16S sequence counts reveals elevated bacterial loads in insect-damaged plant tissues. a. Paired sampling design allows direct comparison between two methods of detecting bacteria within plant tissues, using a subset of samples from one site (EL). **b**. Proportionality of log bacteria-to-host 16S read counts (γ) positively correlates with culturable bacterial load in paired leaf samples. Plot shows observed data with median (black), as well as 95% (light gray) and 50% (dark gray) quantiles of posterior predicted distribution of response variable from a Bayesian log-normal regression (see Methods). **c.** Validated response variable γ is elevated under insect herbivory, and this metric reflects higher bacterial loads when predicted based on the relationship between γ and log CFU (**d**).

without damage from insect herbivores.

Bacterial loads are amplified in insect-damaged leaf tissues. By linking information from 16S counts to absolute bacterial abundance using the estimation procedure above, we found that bacteria that infect herbivore-damaged leaves exhibit local population sizes several doublings greater compared to the bacteria found in undamaged leaves (figure 1c,d; figure 2a,b; median ± 95% credible interval of posterior predicted additional doublings: 2.5 [2.1; 3.9] site EL; 4.5 [3.6; 5.3] site NP). This result, rooted in sequence data, is further validated by the parallel observation that damaged leaves showed higher bacterial loads than undamaged leaves via culturing of the $n = 101$ calibration set (figure S4; mean difference 3.7 of bacterial doublings [1.8–5.6, 95% c.i. on mean difference]; Welch's unequal variance two-sample t-test, $t = 3.86$, $p < 0.001$), which is quantitatively consistent with a prior result from a parallel and independent purely culture-based experiment in this study system (23).

Herbivore-associated bacterial amplification is both community-wide and taxon-specific. When then capitalized on the high taxonomic resolution and sampling depth afforded by amplicon sequencing to examine shifts in abundance and distribution of diverse bacteria within the bittercress leaf microbiome. We found that the within-host density of bacterial bASVs from a wide range of bacterial families was substantially elevated in herbivore-damaged leaves compared to undamaged leaves (figure 2). For most families, the

PRE-PRINT

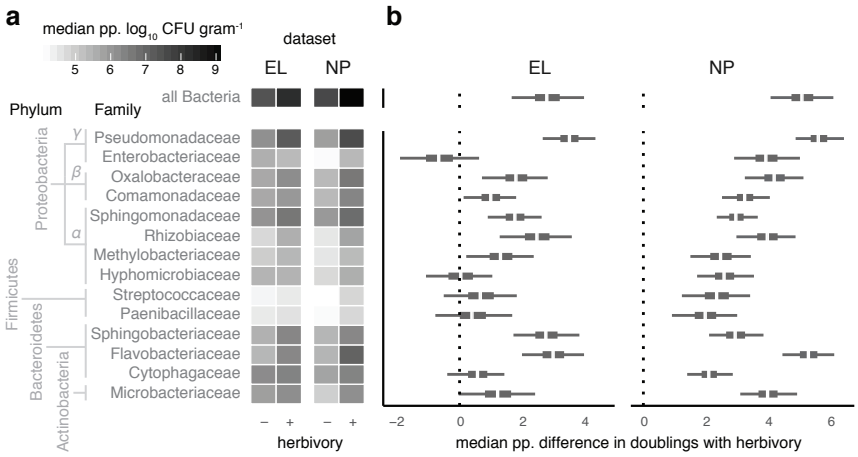

Fig. 2. Pervasive increases in within-host bacterial load in herbivore-damaged leaves. a–b. Posterior predicted ('pp.') infection intensity of bacterial amplicon sequence variants (bASV) from the 14 most prevalence bacterial families show variation in the extent of elevated growth in herbivore-damaged leaf tissue. **a.** Heatmap shows median predicted $\log_{10}$ bacterial abundance (colony-forming units, CFU) per g starting leaf material) from 200 posterior simulations of the best-fitting model of each bacterial family separately (see Methods). **b.** Median (white), 95% and 50% quantiles of the median difference in the number of predicted bacterial cell divisions (i.e., doublings) achieved in herbivore-damaged leaves compared to undamaged leaves from 200 posterior simulations for sites EL and NP separately.

relative increase in within-host density with herbivory was greater at site NP than at site EL (figure 2b), but this was largely because several taxa showed lower baseline loads in undamaged leaves at site NP compared to site EL (figure S5a). In contrast, bacterial loads for all families were similar for damaged leaves at both site (figure S5a,b). Pseudomonadaceae was the most abundant taxon across all leaves and also showed the greatest fold increase under herbivory (figure 2). Thus, a large fraction of the response seen at the level of all bacteria can be attributed to this abundant taxon. Taxa at lower abundance nonetheless achieved substantially higher within-host growth in damaged leaf tissue, indicating that the impact of insects on their host plants can alter within-host growth of a majority of the endophytic phyllosphere community.

At finer taxonomic scales, several family-level γ models showed support for bASV-level differences in intercept and slope values (table S2), including the abundant taxa Pseudomonadaceae and Sphingomonadaceae. Two individual bASVs in particular drove family-level patterns in each of these clades (figure S6) which together comprised $\sim 20\%$ of all sequencing reads across the two sample sets. The majority of the response at the level of family for Pseudomonadaceae and Sphingomonadaceae can thus be ascribed to these highly abundant individual bASVs. In particular, Pseudomonas3 was the single most abundant bASV across our datasets. This individual bASV falls within the putatively phytopathogenic *P. syringae* clade (figure S8), members of which we have previously shown to be pathogenic, induce chlorosis, and reduce leaf photosynthetic function in bittercress (23). Thus, a major component of the signal of elevated bacterial load in the presence of insect herbivory comes from putatively phyopathogenic genotypes within the group *P. syringae*.

Herbivory produces compositional shifts in leaf bacterial communities. When the absolute abundance patterns described above were analyzed in a compositional framework, we detected differences in overall community structure and ecological diversity between damaged and undamaged leaves. Specifically, we found lower evenness (J'; figure 3b) in damaged leaves, suggesting a stronger skew towards a smaller number of bacterial taxa: taxa within the Pseudomonadaceae comprise an even greater proportion of the population in damaged leaves owing to their already-high average abundance in undamaged leaves. Taken as a whole, the distribution of family-level relative abundances differed in terms of Shannon–Jensen divergence (i.e., β-diversity) between damaged versus un-damaged leaves (figure 3c) at both study sites, indicating that amplification of bacteria in herbivore-damaged leaves can produce community-wide signatures of reduced within-host diversity and elevated between-host diversity at broad taxonomic scales (family).

Using absolute abundances allowed us to infer the direction of ecological effects and their magnitudes, as expressed in units of numbers of cell divisions (doublings), and community compositional shifts are only interpretable in light of these data. For example, the relative abundance of several taxa declined under herbivory even while we know that their absolute abundance actually increased (figure 3a; e.g., Sphingomonadaceae at site NP, and many other rarer taxa). These effects are driven by the higher overall absolute abundances and the larger relative increase of Pseudomonadaceae species in damaged leaves, which drives down relative abundance of other taxa even while they also achieved higher densities in herbivore-damaged leaves. This makes clear the importance of accounting for bacterial load in host tissues before interpreting biological signal from relative compositional analyses.

Anti-herbivore plant defenses enhance growth of phytopathogens *in planta*. We next conducted experimental infections in the greenhouse to more directly test whether herbivore-inducible plant phenotypes increase within-host fitness of strains from the most dominant taxon, *Pseudomonas* spp., as observed in our field studies. Three days prior to infection, we pre-treated randomized subsets of greenhouse-grown bittercress accessions with JA or a mock solution and measured the growth after 3 d of twelve strains of *P. syringae* or *P. fluorescens* that represent the diversity of *Pseudomonas* spp. found within bittercress at our study sites (23). JA treatment caused statistically clear alterations in within-host growth of five of the twelve strains tested (fig-

PRE-PRINT

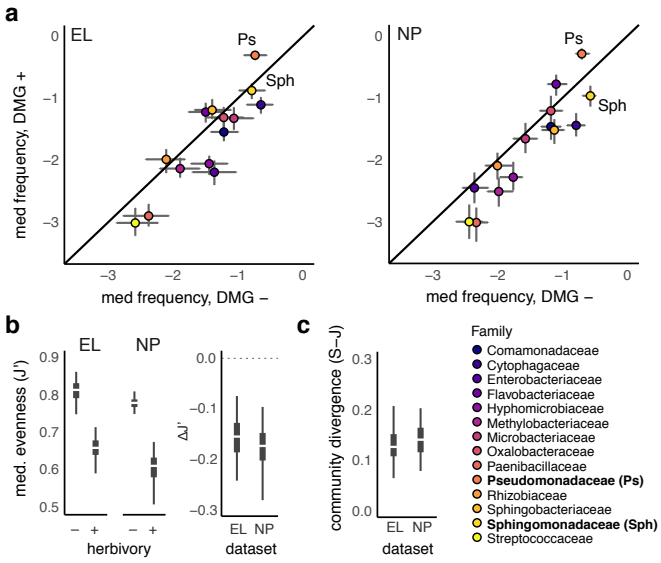

Fig. 3. Insect herbivory induces microbiome compositional shifts towards Pseudomonadaceae, decreasing ecological diversity and inducing community-level divergence. a. $\log_{10}$ relative abundance of each family in undamaged (x-axis) versus damaged (y-axis) samples shows skew towards Pseudomonadaceae (Ps) and relative reductions of abundance among most other taxa at both study sites, including Sphingomonas (Sph) which shows a ~ 2-fold increase in number of doublings in herbivore damaged leaves. **b.** Compositional changes caused by amplification of already abundant taxa (e.g., Pseudomonadaceae) causes reduced community-level evenness (J') and leads to compositional divergence (i.e., beta-diversity) between damaged and undamaged leaves at both study sites (**c**).

ure 4a), with the most pronounced changes resulting in 2.5–5 additional doublings of two phylogenetically distinct *P. syringae* isolates (RM20A and RM02A; table S3). Increased within-host density of these two strains can account for differences in total *Pseudomonas* abundance and the difference between changes at the level of bacterial clade (*P. syringae* versus *P. fluorescens*; figure 4b). This recapitulates patterns of elevated *P. syringae* and family-wide increased abundance under herbivory in the field. Notably, two strains (RM22B and RM20B) exhibited decreased within-host fitness in JA-treated leaves (figure 4a–b). Such herbivore-driven fitness variation among *P. syringae* strains results in overall compositional shifts within this experimental *Pseudomonas* community (figure 4c), including a decrease in Shannon evenness (figure 4d) and community-wide divergence between treated and untreated leaves (figure 4e). These effects at genotype-level resolution can account for patterns among bacterial species (e.g., *P. syringae*) and at the family level seen in the field at larger taxonomic scales (e.g., figure 2).

Phytopathogens are aggregated in highly herbivore–damaged plant patches. We next analyzed how Pseudomonas3, a bASV within the *P. syringae* group, varied across bittercress plant patches in relation to the level of herbivory on those plant patches. In our field study, we counted the number of damaged and undamaged leaves on each of ~ 1900 individual stems ($n = 16$ sampled per patch) at site NP and found a highly aggregated (i.e., right-skewed) distribution of herbivore loads across plant patches (figure 5a, top marginal density plot). This aggregated herbivore distributions across bittercress patches results in a predicted 50-fold enrichment of local density of *P. syringae* (i.e., Pseudomonas3) in the most-damaged compared to the least-damaged plant patch (figure 5). Analyzed in a more general framework, the extent of herbivore aggregation among discrete plant patches predicts the degree to which putative phytopathogens are concentrated in a subset of the plant population (figure S8): over half the predicted Pseudomonas3 population is harbored in just one fifth of the plant patches in the bittercress population at our study site NP.

Herbivore–bacteria co-aggregation is associated with lower plant fitness. By the end of the growing season, when herbivory and bacterial infection had run their course, we counted the number of fruits produced by each of our 16 focal stems across each of the 60 plant patches at site NP. We found that patches with higher herbivore intensity had lower reproductive success, with the most damaged patches estimated to produce half as many fruits as patches with no herbivore damage (figure 5b, table S4). Because herbivory amplifies bacterial loads within leaf tissues, plants with more insect damage tend to have higher levels of bacterial co-aggregation. Thus,

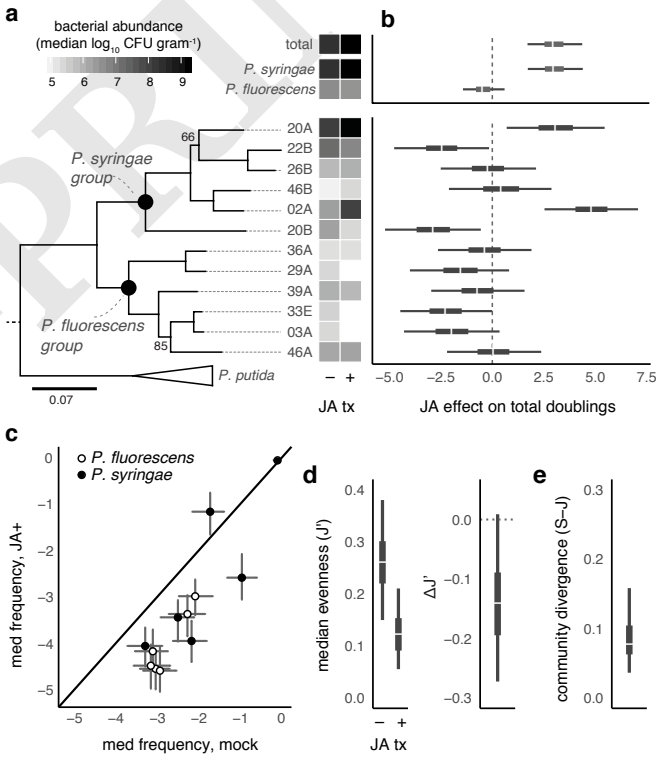

Fig. 4. Eliciting anti-herbivore plant defenses alters within-host performance of phytopathogens under laboratory conditions. a. Experimental infections with 12 *Pseudomonas* spp. strains, concurrently isolated from EL study site (23), show extensive strain-to-strain variation in growth under mock-treated (M) and jasmonic acid (JA) induced plants. Heatmap shows median $\log_{10}$CFU g^{-1} surface sterilized plant tissue 2 d post inoculation. Maximum likelihood phylogeny of strains estimated with four housekeeping loci (2951 bp) from (23). **b.** Median, 95%, and 50% quantiles of the posterior difference between the number of bacterial doublings attained by bacteria growing in JA- versus mock-treated leaves (see Supplemental Methods for model specifications). **c.** Compositional analysis of patterns deduced from bacterial abundances depicted in (**a**) reflect decreased evenness (J'; **d**) in JA-treated plant tissues, leading to overall community-level divergence (**e**). Median, 95%, and 50% quantiles from 200 posterior simulations of abundance (**c–e**) from the best-fitting model fit to data in (**b**).

PREPRINT

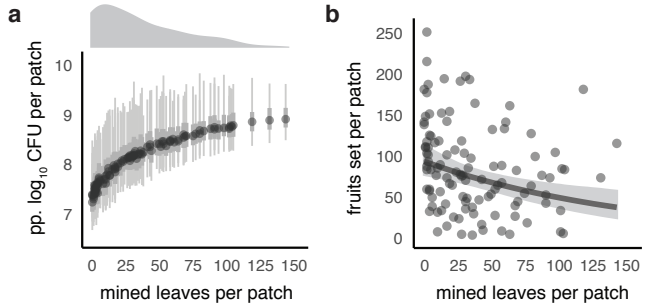

Fig. 5. Co-infection by herbivores and phytopathogenic bacteria is highly aggregated across plant populations and is associated with decreased plant reproduction. a. Median, 95%, and 50% quantiles of 200 posterior simulations of predicted ('pp.') bacterial infection intensity across plant patches ($n = 110$ at site NP; $n = 16$ stems sampled per patch). Density plots above x-axes exhibit right-skewed (i.e., aggregated) distributions of herbivore damage at the plant patch level. **b.** Patch-level herbivory (and thus co-infection intensity) is associated with decreased fruit-set in this native plant population. Plotted are raw fruit-set data summed at the patch level ($n = 16$ stems per patch), with marginal effects slope (and its 95% credible interval) plotted after accounting for average plant height (see table S4).

levels of co-aggregation of plant natural enemies may explain standing variation in reproductive success among bittercress patches across the landscape.

Causes of herbivore aggregation in nature plant populations. The degree of herbivore aggregation among host plants at site NP varied extensively across plant patches in our transect-based field study (figure S9a–b), and the marginal effects on herbivore damage arising from the prior plant treatments with defense hormones SA or JA were small (figure S9c). Estimates for both SA and JA treatment coefficients were elevated above the mock/control condition, but the posterior distribution for both hormone effects overlapped zero (lower 4th %-ile < 0 for JA; lower 15th %-ile < 0 for SA; table S5). Thus, while prior plant exposure to JA, and possible also SA, may cause elevated *S.nigrita* herbivory at the patch scale, standing variation in *S.nigrita* herbivory arising stochastically or from unmeasured factors at site NP dominates over any causal effects of our population-level defense hormone manipulation. We then examined the potential for plant hormone pre-treatment to impact abundance of bacterial colonizers. When examining damaged or undamaged leaf classes separately, plant-level hormone treatments with either JA or SA five weeks prior to leaf sampling showed no discernible effect on distributions of γ for overall or family-wise bacterial abundance, especially in comparison to the clear and large effects of herbivore damage (figure S10).

Discussion

Overview. Here we show that insect herbivory plays a major role in shaping bacterial abundance and diversity within plant microbiomes using field and greenhouse experiments. At two sub-alpine field sites, the growth of a majority of bacterial taxa found in the leaf microbiome of native bittercress was amplified in plant tissues damaged by the specialist herbivore *S. nigrita* (figure 2). These ecological effects were only detectable by linking sequence-based bacterial quantification to an external standard of absolute abundance (figure 1), rather than relying on compositional analysis as is commonly done with studies of both plant and animal microbiomes (but see (35)). The bacterial clades most altered under herbivory are well-known for causing plant disease (*P. syringae*), and our follow-up experimental work in the greenhouse showed that inducing anti-herbivore defenses in bittercress was sufficient to cause similar degrees of amplification of *P. syringae* genotypes in leaf tissue. Amplification of specific *P. syringae* genotypes can largely account for species- and family-level patterns seen in our field studies, which has coarser taxonomic resolution. Overall, these experiments suggest that *S. nigrita* herbivores may play a larger role than previously appreciated in promoting the growth and subsequent transmission of particular bacterial genotypes or pathovars in bittercress. Given that the majority of plants face herbivore attack to some degree (16, 36), it is possible that our results generalize across plant–microbe systems.

Herbivore-inducible plant defenses can amplify phytopathogens. The mechanisms governing growth-promoting effects of insect herbivory on leaf-colonizing bacteria are potentially numerous. Several possibilities stand out as plausible. First, leaf damage itself can release nutrients that may alleviate resource constraints for colonizing bacteria or create routes for colonization of the leaf interior by bacteria initially present on the leaf surface (37, 38). Simultaneously, leaf damage can trigger local and systemic signaling cascades which select for or against different bacteria. For example, *Pseudomonas* isolates exhibit quantitative variation in their tolerance to isothiocyanates, a toxic hydrolysis product of herbivore-inducible glucosinolates activated upon leaf tissue disruption in mustards such as bittercress (39) (and Humphrey, unpublished data). The immediate physical and chemical environment of herbivore-damaged leaves may thus allow for a subset of the resident microbes commonly found among the leaf microbiota to dominate.

Second, the architecture of plant defense provides numerous avenues through which anti-herbivore defenses could indirectly alter interactions with bacteria independent from, or in combination with, the more direct effects noted above. JA-dependent anti-herbivore defenses can suppress the subsequent activity of signaling pathways responsive to bacterial infection (reviewed in (20)), allowing bacteria (including strains of *P. syringae*) to reach higher densities within JA-affected leaf tissues (40). Insects such as *S. nigrita* trigger such JA-dependent host defenses in bittercress (22), and this form of defense signaling is widely conserved among diverse plant groups (20). Released from top-down control, a diversity of resident microbes can then proliferate as defenses are more strongly directed against herbivory, which may manifest in community-wide patterns of abundance changes as noted in our study (figure 2).

The hypothesis that anti-herbivore defenses pleiotropically increase bacterial growth is also consistent with our greenhouse experiment. We found that the within-host fitness of several strains of putatively phytopathogenic *P. syringae* was amplified within undamaged host tissues pre-treated with

PRE-PRINT

JA compared to healthy mock-treated leaves (figure 3). Although our experiment is consistent with this plausible mechanism by which herbivores can facilitate bacterial growth within plants, it does not identify the proximal mechanism(s) responsible for these effects. JA-induced defenses may have instead stressed the plants, decreasing basal tolerance to infection, or shunted resources towards investments which reduce the ability of plants to resist or tolerate bacterial infection (41). Under field conditions, herbivore attack may cause direct costs via loss of plant tissue as well as indirect costs associated with mounting an anti-herbivore defense response, and such resource constraints may also decrease tolerance to both herbivores and bacterial colonizers. Specifically, at site NP, more heavily damaged plants had lower reproductive success (figure 5c). Whether these plants were more stressed to begin with or achieved lower fruiting success because of their infestation with herbivores and phytopathogens cannot be resolved without future controlled studies that isolate the causal effects of single and multiple infection on plant fitness.

Herbivore distributions can alter the spatial patterning of plant disease. The impact of insect herbivory on phyllosphere bacteria can be observed at several spatial scales. The effects of herbivory were observed between individual leaves on single host plants and at larger spatial scales, because herbivore damage was highly clustered on a subset of hosts (figure 5). This pattern of aggregated herbivore damage across hosts is consistent with reports of high levels of aggregation across other plant–herbivore systems (42) and across a range of host–macroparasite systems more generally (43). The accompanying effects on bacterial growth may occur most strongly on a subset of plant patches or plant genotypes. We observed that plant patches with the highest levels of herbivore damage (and, thus, predicted bacterial abundance) had substantially ($\sim 50\%$) lower reproductive success compared to minimally-damaged plant patches (figure 5b). Although we cannot identify the cause of lower plant fitness, the co-aggregation of these distinct groups of plant colonizers may at least partly explain it.

What drives the highly skewed pattern of herbivory among host plants? Plant populations often display a patchwork of defensive phenotypes, influenced both by variation in the underlying genes and by the spatial and temporal pattern of defense induction caused by herbivore or microbial attack. While plant defensive chemistry is known to shape herbivore attack rates in the laboratory and in the field over wide spatial extents (44), less is known about how patterns of defense induction impact the population dynamics of insect herbivores (45, 46). Many insects are deterred by anti-herbivore defenses, but some specialist herbivores use these cues as attractants owing to detoxification mechanisms which confer resistance to such defenses (47). *S. nigrita* uses anti-herbivore plant defenses to locate host but also avoids high concentrations of these compounds when given the option (22). Thus, the joint expression of positive chemo-taxis towards JA-inducible compounds, coupled to aversion of high levels of JA responsive defensive chemistry, may influence which hosts are most susceptible to herbivory and where herbivory becomes concentrated among plants in natural populations.

Results from our field hormone treatments using JA and SA both showed elevated patch-level *S. nigrita* leaf miner damage compared to mock-treated patches (figure S9). However, the statistical signature of these treatment effects was not clear enough to confidently conclude that our field trials substantially altered natural patterns of herbivory by this specialist, given the high degree of stochastic or unexplained variation in herbivore damage we observed across plant patches (figure S9, table S5). Discovering the biotic and abiotic factors structuring herbivory patterns in natural host populations thus remains a challenge in this system (48) as well as many others (46), and we have not attempted to solve this problem in the present study.

Nonetheless, our study suggests that predicting herbivore distributions will be key for understanding population-level distributions of plant-associated bacteria, given that many bacteria exhibited strong responses to insects either directly and/or indirectly via herbivore inducible plant phenotypes. Regardless of its causes, chewing herbivore damage can be readily measured and incorporated into plant microbiome studies. Future work incorporating field-scale population dynamics of insect herbivores may help us better understand the major drivers of variation in microbiome abundance and diversity within and among plant communities.

Quantifying bacterial loads is crucial for understanding the ecology of the microbiome. By analyzing our abundance variation data in a compositional framework, our analyses revealed ambiguities inherent to sequence-based enumeration of bacterial communities. The patterns of abundance variation among bacterial taxa across leaf types, when distilled into a community-level compositional metric, revealed decreased ecological diversity (i.e., evenness) and compositional divergence between sets of damaged versus undamaged leaf samples (figures 3–4). We interpret this community-wide loss of ecological diversity under herbivory as being due to particular taxa undergoing larger absolute changes in abundance than other taxa, which leads to stronger biases in the composition of the community. However, the compositional analysis on its own would preclude inference of the direction or magnitude of changes in bacterial abundance (51), even though this is of primary interest to researchers exploring the population biology of host-colonizing microbes and the impacts of the microbiome on host fitness (13, 35, 52). Notably, however, even if all bacterial taxa experienced the same relative change in abundance, ecological diversity metrics calculated on the relative abundance scale would still differ between damaged and undamaged communities given different starting abundances of those taxa. Thus, compositional microbiome analysis is poorly suited for analyses seeking to understand the population biology of microbiome-dwelling bacterial taxa in the presence of abundance variation.

Rather than being discarded, 16S read counts derived from host organelles—once curated—can provide an internal reference population against which the proportionality of

PRE-PRINT

other taxa can be measured (49). Much of the validity of our approach comes from the fact that the majority of abundant bacterial taxa in bittercress can be cultivated in the laboratory (23), which permitted our high-throughput quantitative enumeration of many bacterial lineages at once via 16S sequencing. Direct quantification (35), as well as controlled DNA spike-ins (50), can also correct the biases and ambiguities inherent to purely compositional analyses of microbiome data. Our study provides one such a framework for enabling standard high-throughput amplicon sequencing data to quantitatively describe features of the population biology of phyllosphere bacteria.

Conclusions. Biotic interactions such as herbivory are absent from the classic 'disease triangle' of plant pathology. Instead, abiotic conditions are thought to govern whether the scales are tipped towards pathogenicity. More recently, microbe–microbe interactions have received attention as potential mediators of both plant immunity and *P. syringae* pathogenicity. Our study emphasizes that large effects on the population biology of *P. syringae*, and many other lineages of leaf-colonizing bacteria, stem from the action of insect herbivores. Their role in *P. syringae* epidemiology—and plant–microbiome relations in general—has been underappreciated. Variation in bacterial abundance across samples, and the implications of relative abundance changes for bacterial fitness, are not easily detectable via compositional analyses applied to 16S data, which typically do not utilize external or internal standards. Thus, studies aiming to decipher why plant microbiomes differ in structure or function should endeavor to quantify bacterial loads in order to retain this important axis of variation as a focal response variable.

DATA AVAILABILITY

See `https://github.com/phumph/coinfection` for all scripts and raw input data. Saved model objects can be found archived in a Dryad repository (pending). Sequence data can be found at the NIH Short Read Archive under project ID (pending).

ACKNOWLEDGEMENTS

PTH and NKW gratefully acknowledge funding from the National Science Foundation (DEB-1256758 to NKW, DEB-1309493 to PTH), as well as the Rocky Mountain Biological Laboratory. We are indebted to field assistance provided by Heather Briggs, Kara Cromwell, Aaron Koning, Lucy Anderson, Kyle Niezgoda, Devon Picklum, and Nicolas Alexandre, as well as laboratory assistance from Hoon Pyon and Amir Abidov.

References

1. James O Lloyd-Smith, Mary Poss, and Bryan T Grenfell. Hiv-1/parasite co-infection and the emergence of new parasite strains. *Parasitology*, 135(7):795–806, Jun 2008. doi: 10.1017/S0031182008000292.
2. Anna-Liisa Laine. Context-dependent effects of induced resistance under co-infection in a plant-pathogen interaction. *Evol Appl*, 4(5):696–707, Sep 2011. doi: 10.1111/j.1752-4571.2011.00194.x.
3. Charlotte Tollenaere, Hanna Susi, and Anna-Liisa Laine. Evolutionary and epidemiological implications of multiple infection in plants. *Trends Plant Sci*, 21(1):80–90, Jan 2016.
4. Anssi Karvonen, Jukka Jokela, and Anna-Liisa Laine. Importance of sequence and timing in parasite coinfections. *Trends Parasitol*, Dec 2018. doi: 10.1016/j.pt.2018.11.007.
5. Fletcher W Halliday, James Umbanhowar, and Charles E Mitchell. Interactions among symbionts operate across scales to influence parasite epidemics. *Ecol Lett*, 20:1285–1294, Sep 2017.
6. Hanna Susi, Benoit Barrès, Pedro F Vale, and Anna-Liisa Laine. Co-infection alters population dynamics of infectious disease. *Nat Commun*, 6:5975, Jan 2015. doi: 10.1038/ncomms6975.
7. Matthew W Horton, Natacha Bodenhausen, Kathleen Beilsmith, Dazhe Meng, Brian D Muegge, Sathish Subramanian, M Madlen Vetter, Bjarni J Vilhjálmsson, Magnus Nordborg, Jeffrey I Gordon, and Joy Bergelson. Genome-wide association study of arabidopsis thaliana leaf microbial community. *Nat Commun*, 5:5320, 2014.
8. Davide Bulgarelli, Matthias Rott, Klaus Schlaeppi, Emiel Ver Loren van Themaat, Nahal Ahmadinejad, Federica Assenza, Philipp Rauf, Bruno Huettel, Richard Reinhardt, Elmon Schmelzer, Joerg Peplies, Frank Oliver Gloeckner, Rudolf Amann, Thilo Eickhorst, and Paul Schulze-Lefert. Revealing structure and assembly cues for arabidopsis root-inhabiting bacterial microbiota. *Nature*, 488(7409):91–5, Aug 2012.
9. Natacha Bodenhausen, Miriam Bortfeld-Miller, Martin Ackermann, and Julia A Vorholt. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet*, 10(4):e1004283, Apr 2014.
10. Joseph Edwards, Cameron Johnson, Christian Santos-Medellín, Eugene Lurie, Natraj Kumar Podishetty, Srijak Bhatnagar, Jonathan A Eisen, and Venkatesan Sundaresan. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A*, 112(8):E911–20, Feb 2015.
11. Jason A Peiffer, Aymé Spor, Omry Koren, Zhao Jin, Susannah Green Tringe, Jeffery L Dangl, Edward S Buckler, and Ruth E Ley. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A*, 110(16):6548–53, Apr 2013.
12. Maggie R Wagner, Derek S Lundberg, Tijana G Del Rio, Susannah G Tringe, Jeffery L Dangl, and Thomas Mitchell-Olds. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun*, 7:12151, 2016.
13. Gwen Falony, Marie Joossens, Sara Vieira-Silva, Jun Wang, Youssef Darzi, Karoline Faust, Alexander Kurilshikov, Marc Jan Bonder, Mireia Valles-Colomer, Doris Vandeputte, Raul Y Tito, Samuel Chaffron, Leen Rymenans, Chloë Verspecht, Lise De Sutter, Gipsi Lima-Mendez, Kevin D'hoe, Karl Jonckheere, Daniel Homola, Roberto Garcia, Ettje F Tigchelaar, Linda Eeckhaudt, Jingyuan Fu, Liesbet Henckaerts, Alexandra Zhernakova, Cisca Wijmenga, and Jeroen Raes. Population-level analysis of gut microbiome variation. *Science*, 352(6285):560–4, Apr 2016. doi: 10.1126/science.aad3503.
14. Omri M Finkel, Gabriel Castrillo, Sur Herrera Paredes, Isai Salas González, and Jeffery L Dangl. Understanding and exploiting plant beneficial microbes. *Curr Opin Plant Biol*, 38: 155–163, 08 2017. doi: 10.1016/j.pbi.2017.04.018.
15. Ma Del Carmen Orozco-Mosqueda, Ma Del Carmen Rocha-Granados, Bernard R Glick, and Gustavo Santoyo. Microbiome engineering to improve biocontrol and plant growth-promoting mechanisms. *Microbiol Res*, 208:25–31, Mar 2018. doi: 10.1016/j.micres.2018.01.005.
16. John L Maron and Elizabeth Crone. Herbivory: effects on plant abundance, distribution and population growth. *Proc Biol Sci*, 273(1601):2575–84, Oct 2006. doi: 10.1098/rspb.2006.3587.
17. Agrawal. Induced responses to herbivory and increased plant performance. *Science*, 279 (5354):1201–2, Feb 1998.
18. Mélanie Bressan, Marie-Anne Roncato, Floriant Bellvert, Gilles Comte, Feth Zahar Haichar, Wafa Achouak, and Odile Berge. Exogenous glucosinolate produced by arabidopsis thaliana has an impact on microbes in the rhizosphere and plant roots. *ISME J*, 3(11): 1243–57, Nov 2009.
19. Simon C Groen, Parris T Humphrey, Daniela Chevasco, Frederick M Ausubel, Naomi E Pierce, and Noah K Whiteman. Pseudomonas syringae enhances herbivory by suppressing the reactive oxygen burst in arabidopsis. *J Insect Physiol*, 84:90–102, 01 2016. doi: 10.1016/j.jinsphys.2015.07.011.
20. Jennifer S Thaler, Parris T Humphrey, and Noah K Whiteman. Evolution of jasmonate and salicylate signal crosstalk. *Trends Plant Sci*, 17(5):260–70, May 2012. doi: 10.1016/j.tplants.2012.02.010.
21. Maggie R Wagner, Derek S Lundberg, Devin Coleman-Derr, Susannah G Tringe, Jeffery L Dangl, and Thomas Mitchell-Olds. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild arabidopsis relative. *Ecol Lett*, 17(6):717–26, Jun 2014. doi: 10.1111/ele.12276.
22. Parris T Humphrey, Andrew D Gloss, Nicolas M Alexandre, Martha M Villalobos, Marcella R Fremgen, Simon C Groen, Lisa N Meihls, Georg Jander, and Noah K Whiteman. Aversion and attraction to harmful plant secondary compounds jointly shape the foraging ecology of a specialist herbivore. *Ecol Evol*, 6(10):3256–68, 05 2016. doi: 10.1002/ece3.2082.
23. Parris T Humphrey, Trang T Nguyen, Martha M Villalobos, and Noah K Whiteman. Diversity and abundance of phyllosphere bacteria are linked to insect herbivory. *Mol Ecol*, 23(6): 1497–515, Mar 2014. doi: 10.1111/mec.12657.
24. Luke R Thompson, Jon G Sanders, Daniel McDonald, Amnon Amir, Joshua Ladau, Kenneth J Locey, Robert J Prill, Anupriya Tripathi, Sean M Gibbons, Gail Ackermann, Jose A Navas-Molina, Stefan Janssen, Evguenia Kopylova, Yoshiki Vázquez-Baeza, Antonio González, James T Morton, Siavash Mirarab, Zhenjiang Zech Xu, Lingjing Jiang, Mohamed F Haroon, Jad Kanbar, Qiyun Zhu, Se Jin Song, Tomasz Kosciolek, Nicholas A Bokulich, Joshua Lefler, Colin J Brislawn, Gregory Humphrey, Sarah M Owens, Jarrad Hampton-Marcell, Donna Berg-Lyons, Valerie McKenzie, Noah Fierer, Jed A Fuhrman, Aaron Clauset, Rick L Stevens, Ashley Shade, Katherine S Pollard, Kelly D Goodwin, Janet K Jansson, Jack A Gilbert, Rob Knight, and Earth Microbiome Project Consortium. A communal catalogue reveals earth's multiscale microbial diversity. *Nature*, 551(7681): 457–463, 11 2017. doi: 10.1038/nature24621.
25. Derek S Lundberg, Scott Yourstone, Piotr Mieczkowski, Corbin D Jones, and Jeffery L Dangl. Practical innovations for high-throughput amplicon sequencing. *Nat Methods*, 10 (10):999–1002, Oct 2013. doi: 10.1038/nmeth.2634.
26. Benjamin J Callahan, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. Dada2: High-resolution sample inference from illumina amplicon data. *Nat Methods*, 13(7):581–3, 07 2016. doi: 10.1038/nmeth.3869.
27. Qiong Wang, George M Garrity, James M Tiedje, and James R Cole. Naive bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy. *Appl Environ Microbiol*, 73(16):5261–7, Aug 2007. doi: 10.1128/AEM.00062-07.
28. Sarah L Lebeis, Sur Herrera Paredes, Derek S Lundberg, Natalie Breakfield, Jase Gehring, Meredith McDonald, Stephanie Malfatti, Tijana Glavina del Rio, Corbin D Jones, Susannah G Tringe, and Jeffery L Dangl. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science*, 349(6250):860–4, Aug 2015.
29. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.
30. Stan Development Team. RStan: the R interface to Stan, 2018. R package version 2.18.1.

PRE-PRINT

31. Paul-Christian Bürkner. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1):395–411, 2018.
32. Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.
33. Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432, Sep 2017. ISSN 1573-1375. doi: 10.1007/s11222-016-9696-4.
34. Alexandros Stamatakis. Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–3, May 2014. doi: 10.1093/bioinformatics/btu033.
35. Doris Vandeputte, Gunter Kathagen, Kevin D'hoe, Sara Vieira-Silva, Mireia Valles-Colomer, João Sabino, Jun Wang, Raul Y Tito, Lindsey De Commer, Youssef Darzi, Séverine Vermeire, Gwen Falony, and Jeroen Raes. Quantitative microbiome profiling links gut community variation to microbial load. *Nature*, 551(7681):507–511, 11 2017. doi: 10.1038/nature24460.
36. Martin M Turcotte, T Jonathan Davies, Christina J M Thomsen, and Marc T J Johnson. Macroecological and macroevolutionary patterns of leaf herbivory across vascular plants. *Proc Biol Sci*, 281(1787), Jul 2014. doi: 10.1098/rspb.2014.0555.
37. S S Hirano and C D Upper. Bacteria in the leaf ecosystem with emphasis on pseudomonas syringae-a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev*, 64(3):624–53, Sep 2000.
38. Steven E Lindow and Maria T Brandl. Microbiology of the phyllosphere. *Appl Environ Microbiol*, 69(4):1875–83, Apr 2003.
39. Jun Fan, Casey Crooks, Gary Creissen, Lionel Hill, Shirley Fairhurst, Peter Doerner, and Chris Lamb. Pseudomonas sax genes overcome aliphatic isothiocyanate-mediated non-host resistance in arabidopsis. *Science*, 331(6021):1185–8, Mar 2011. doi: 10.1126/science.1199707.
40. Jianping Cui, Adam K Bahrami, Elizabeth G Pringle, Gustavo Hernandez-Guzman, Carol L Bender, Naomi E Pierce, and Frederick M Ausubel. Pseudomonas syringae manipulates systemic plant defenses against pathogens and herbivores. *Proc Natl Acad Sci U S A*, 102(5):1791–6, Feb 2005. doi: 10.1073/pnas.0409450102.
41. Huang Huang, Bei Liu, Liangyu Liu, and Susheng Song. Jasmonate action in plant growth and development. *J Exp Bot*, 68(6):1349–1359, 03 2017. doi: 10.1093/jxb/erw495.
42. Nora Underwood, Kurt Anderson, and Brian D. Inouye. Induced vs. constitutive resistance and the spatial distribution of insect herbivores among plants. *Ecology*, 86(3):594–602, 2005. doi: 10.1890/03-0290.
43. D J Shaw and A P Dobson. Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. *Parasitology*, 111 Suppl:S111–27, 1995.
44. Tobias Züst, Christian Heichinger, Ueli Grossniklaus, Richard Harrington, Daniel J Kliebenstein, and Lindsay A Turnbull. Natural enemies drive geographic variation in plant defenses. *Science*, 338(6103):116–9, Oct 2012. doi: 10.1126/science.1226397.
45. Nora Underwood, Kurt Anderson, and Brian D. Inouye. Induced vs. constitutive resistance and the spatial distribution of insect herbivores among plants. *Ecology*, 86(3):594–602, 2005.
46. Richard Karban and Ian T Baldwin. *Induced Responses to Herbivory.* University of Chicago Press, 1997.
47. Ute Wittstock, Niels Agerbirk, Einar J Stauber, Carl Erik Olsen, Michael Hippler, Thomas Mitchell-Olds, Jonathan Gershenzon, and Heiko Vogel. Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proc Natl Acad Sci U S A*, 101(14):4859–64, Apr 2004. doi: 10.1073/pnas.0308007101.
48. Nicolas M Alexandre, Parris T Humphrey, Andrew D Gloss, Jimmy Lee, Joseph Frazier, Henry A Affeldt, 3rd, and Noah K Whiteman. Habitat preference of an herbivore shapes the habitat distribution of its host plant. *Ecosphere*, 9(9), Sep 2018. doi: 10.1002/ecs2.2372.
49. David Lovell, Vera Pawlowsky-Glahn, Juan José Egozcue, Samuel Marguerat, and Jürg Bähler. Proportionality: a valid alternative to correlation for relative data. *PLoS Comput Biol*, 11(3):e1004075, Mar 2015.
50. Frank Stämmler, Joachim Gläsner, Andreas Hiergeist, Ernst Holler, Daniela Weber, Peter J. Oefner, André Gessner, and Rainer Spang. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome*, 4(1):28, Jun 2016. ISSN 2049-2618. doi: 10.1186/s40168-016-0175-0.
51. Gregory B. Gloor, Jean M. Macklaim, Vera Pawlowsky-Glahn, and Juan J. Egozcue. Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, 8:2224, 2017. ISSN 1664-302X. doi: 10.3389/fmicb.2017.02224.
52. Jeroen Raes. Editorial overview: It's the ecology, stupid: microbiome research in the post-stamp collecting age. *Curr Opin Microbiol*, 44:iv–v, Aug 2018. doi: 10.1016/j.mib.2018.07.008.

PRE-PRINT