

**Intraspecific variation in symbiont density in an insect-microbe symbiosis**

Benjamin J. Parker,<sup>1,2,3\*</sup> Jan Hrček,<sup>1,4</sup> Ailsa H.C. McLean,<sup>1</sup> Jennifer A. Brisson,<sup>3</sup> & H. Charles J. Godfray<sup>1</sup>

<sup>1</sup> Department of Zoology, University of Oxford, South Parks Road, Oxford OX13PS, UK

<sup>2</sup> Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA

<sup>3</sup> Department of Biology, University of Rochester, Rochester, NY 14627, USA

<sup>4</sup> Institute of Entomology, Biology Centre of the Czech Academy of Sciences, Branisovska 31, Ceske Budejovice 37005, Czech Republic

ORCIDs: BJP: 0000-0002-0679-4732; JH: 0000-0003-0711-6447; AHCM: 0000-0002-8195-2593; JAB: 0000-0001-7000-0709; HCJG: 0000-0001-8859-7232

\* Author for correspondence: [bjp@utk.edu](mailto:bjp@utk.edu), phone: +1 865 974 9818

22 ABSTRACT

23  
24 Many insects host vertically-transmitted microbes, which can confer benefits to their hosts but  
25 are costly to maintain and regulate. A key feature of these symbioses is variation: for example,  
26 symbiont density can vary among host and symbiont genotypes. However, the evolutionary  
27 forces maintaining this variation remain unclear. We studied variation in symbiont density using  
28 the pea aphid (*Acyrtosiphon pisum*) and the bacterium *Regiella insecticola*, a symbiont that can  
29 protect its host against fungal pathogens. We found that relative symbiont density varies both  
30 between two *Regiella* phylogenetic clades and among aphid 'biotypes'. Higher-density symbiont  
31 infections are correlated with stronger survival costs, but variation in density has little effect on  
32 the protection *Regiella* provides against fungi. Instead, we found that in some aphid genotypes,  
33 a dramatic decline in symbiont density precedes the loss of a symbiont infection. Together, our  
34 data suggest that the optimal density of a symbiont infection is likely different from the  
35 perspective of aphid and microbial fitness. *Regiella* might prevent loss by maintaining high  
36 within-host densities, but hosts do not appear to benefit from higher symbiont numbers and may  
37 be advantaged by losing costly symbionts in certain environments. The standing variation in  
38 symbiont density observed in natural populations could therefore be maintained by antagonistic  
39 coevolutionary interactions between hosts and their symbiotic microbes.

40  
41  
42 KEYWORDS

43  
44 Symbiosis, intraspecific variation, density, pea aphid, facultative symbiont

## INTRODUCTION

Animals vary in the composition and abundance of the microbial communities that they harbor. Many microbial symbionts associated with insects, for example, are not found in all individuals within a species (Ferrari & Vavre, 2011). These facultative microbes can benefit their hosts (Brownlie & Johnson, 2009) but also impose metabolic costs and can require regulation through immunological control (Douglas, 2014; Herren & Lemaitre, 2011; Login et al., 2011; Pan et al., 2018). In light of these trade-offs, theoretical (Jaenike, 2012) and empirical studies (Hansen, Jeong, Paine, & Stouthamer, 2007; Jaenike, Unckless, Cockburn, Boelio, & Perlman, 2010) have considered how the prevalence of facultative symbionts across hosts changes in response to ecological factors. In addition, the mechanisms underlying these partnerships might evolve, for example leading to changes in the density of symbiont infections within hosts. For natural selection to shape host-symbiont associations there must be heritable variation in the traits affecting the partnership and this variation must be linked to host and/or symbiont fitness.

Recent studies have found that variation among microbial strains and host genotypes influences symbiont density (Unckless, Boelio, Herren, & Jaenike, 2009) and have identified specific molecular mechanisms controlling symbiont numbers. For example, different strains of the heritable bacteria *Wolbachia* transferred into a single *Drosophila simulans* line establish at different densities, and density is correlated with the protection conferred by *Wolbachia* against viruses (Martinez et al., 2017). Within the *Wolbachia* strain wMelPop, a genomic region termed Octomom was shown to influence density—artificial selection for higher Octomom copy number led to increased *Wolbachia* numbers in *Drosophila melanogaster* and to higher costs for its host (Chrostek & Teixeira, 2015). In the *Wolbachia*–*Drosophila* system, symbiont strain seems to be a more important factor than host genotype driving variation in within-host bacterial density, but other studies have emphasized the importance of host genetic variation in determining symbiont density. A gene called *Wolbachia density suppressor* (*Wds*) in the parasitic wasp *Nasonia* controls the density of *Wolbachia* through suppression of vertical transmission of the bacteria, and explains much of the variation in density between two closely related *Nasonia* species (Funkhouser-Jones, van Opstal, Sharma, & Bordenstein, 2018). This strain of *Wolbachia* is a clear reproductive parasite, and *Wds* was shown to be under positive selection. Another example comes from the pea aphid (*Acyrtosiphon pisum*) that we study here, where host

genotypes vary in the density at which they maintain their obligate primary symbiont *Buchnera aphidicola* (Chong & Moran, 2016; Vogel & Moran, 2011; Wilkinson & Douglas, 1998).

What remains unclear, however, is what the existence of standing variation in symbiont density implies about the evolution of beneficial host-symbiont interactions. Is the optimal density of symbionts within a host the same from the perspective of the host and microbe? One possibility is host and microbial fitnesses are maximized at the same density, but we find variation in symbiont density because of mismatches between host and symbiont genotypes—either in natural combinations or in those made artificially in the laboratory—that would be selected against in natural populations. Alternatively, hosts and symbionts might have different optimal densities, for example when higher symbiont density lowers host fitness but increases the probability of symbiont transmission. Variation in natural populations could then reflect an antagonistic process as hosts and symbionts struggle for control of symbiont density. To understand the evolutionary forces shaping a host-microbe association, we need to determine which partner is controlling the symbiosis and how variation among hosts and symbionts is linked with the fitness of both partners.

Here we study genetic variation in symbiont density using a tractable model host-microbe interaction: pea aphids and the facultative bacterial symbiont *Regiella insecticola*. In addition to their obligate symbiont *Buchnera*, aphids host a variety of bacteria that are found at intermediate frequencies in natural populations. These facultative symbionts are costly to maintain (Oliver, Campos, Moran, & Hunter, 2008; Vorburger & Gouskov, 2011) but also provide conditional benefits to their hosts. *Hamiltonella defensa*, for example, confers resistance against parasitoid wasps (Oliver, Russell, Moran, & Hunter, 2003) but decreases in frequency in the absence of infection pressure (Ives et al., 2020; Oliver et al., 2008). Similarly, *Regiella* makes aphids more resistant to natural fungal pathogens from the order Entomophthorales (Parker, Spragg, Altincicek, & Gerardo, 2013; Scarborough, Ferrari, & Godfray, 2005). Within pea aphids, *Regiella* forms two main phylogenetic clades (Henry et al., 2013), and the clades provide stronger protection against specific genotypes of the fungal pathogen *Pandora* in a genotype by genotype ( $G_{\text{Symbiont}} \times G_{\text{Pathogen}}$ ) dependent manner (Parker, Hrcek, McLean, & Godfray, 2017).

The pea aphid species complex is comprised of multiple populations, termed ‘biotypes’, that are to some degree reproductively isolated (Peccoud et al., 2014; Peccoud, Ollivier, Plantegenest, &

Simon, 2009) and are adapted to live on different species of plants in the Fabaceae family (Ferrari, Via, & Godfray, 2008; Via, 1991). Importantly, biotypes have been shown to differ in the frequencies at which they harbor facultative symbionts (Henry et al., 2013). Some biotypes are strongly associated with a particular symbiont species: for example, in collections from different continents *Regiella* is commonly found in aphids from *Trifolium spp.* (Ferrari, West, Via, & Godfray, 2012; Russell et al., 2013; Tsuchida, Koga, Shibao, Matsumoto, & Fukatsu, 2002). In contrast, some other biotypes rarely carry *Regiella*, or do not commonly harbor any facultative symbionts. To explain biotype-level patterns in symbiont frequencies, some studies have focused on the effects of symbionts on host-plant use with mixed success (McLean, van Asch, Ferrari, & Godfray, 2011; Tsuchida, Koga, & Fukatsu, 2004). Other studies have found that aphid and symbiont genotypes vary in the success with which they form artificial host-microbe pairings in the lab (Niepoth, Ellers, & Henry, 2018; Parker, McLean, Hrcek, Gerardo, & Godfray, 2017), suggesting that genotype may influence facultative symbiont frequencies in natural populations.

In this study, we tested a series of hypotheses about variation in relative density across both host and symbiont genotypes and the relationship between density and fitness. Our overall expectation was that *Regiella* density would be correlated with stronger protection against a fungal pathogen but higher survival costs to aphids. **(1)** We first tested the hypothesis that *Regiella* strains from two phylogenetic clades vary in within-host density, and **(2)** that higher density symbiont strains impose stronger survival costs on their hosts than do lower density strains. **(3)** We then tested whether the variation in density we uncovered among *Regiella* strains is dependent on host genotype (a  $G_{\text{Host}} \times G_{\text{Symbiont}}$  interaction). This experiment used aphid lines that were naturally collected with a *Regiella* infection, and we next tested the hypothesis **(4)** that a single strain of *Regiella* would establish at different densities across a diverse panel of 28 host genotypes representing eight host-plant associated biotypes, some of which rarely harbor *Regiella*. **(5)** We then measured the effects of host genotype on the strength of protection conferred by *Regiella* against fungus and on aphid survival and **(6)** if these costs and/or benefits are correlated with symbiont density. **(7)** Last, we tested the hypothesis that low density lines are in the process of losing their symbiont infections. We found a high degree of variation in density across both hosts and microbes. However, we did not find strong evidence that density is associated with the fitness effects of harboring *Regiella*. We instead found that symbiont loss is

preceded by a drop in density and varies among genotypes. We discuss these findings in the context of host-symbiont coevolution.

## METHODS

Pea aphids are cyclically parthenogenetic allowing us to maintain asexual genotypes in the laboratory under a light regime of 16L:8D. We maintain aphid stock lines at 14°C, and we conducted all of the experiments in this study at 20°C. Aphids were housed in Petri dishes with a *Vicia faba* leaf (cultivar 'The Sutton'), a 'universal' host plant on which all aphid biotypes are able to feed (Sandström & Pettersson, 1994), with petioles inserted in 2% tap water agar. When new lines were collected from the wild, we screened them for seven pea aphid facultative symbionts using PCR and symbiont-specific primers (94°C 2m, 12 cycles x (94°C 30s, 56°C (declining 1°C each cycle) 50s, 72°C 90s), 26 cycles x (94°C 30s, 45°C 50s, 72°C 90s) and a final extension of 72°C 5m) (Henry et al., 2013). We cleared lines of any facultative symbionts by immersing the petioles of *V. faba* leaves in an antibiotic solution (100 mg ml<sup>-1</sup> ampicillin, 50 mg ml<sup>-1</sup> cefotaxime and 50 mg ml<sup>-1</sup> gentomycin), and feeding first-instar aphids on the leaves for 48 hours (McLean et al., 2011); note that lines collected with facultative symbionts that we are currently unable to cure (*Spiroplasma* & *Rickettsia*) were not included in our experiments. After antibiotic exposure, we moved aphids to fresh leaves until they became adults. We collected the late offspring (>10 in birth order) of antibiotic-fed aphids, maintained them until they became adults, tested them for symbionts using PCR as above, and retained the offspring of symbiont-free individuals. We kept these lines for at least 15 generations to ensure there were no maternal effects of antibiotic treatment on experimental aphids, while periodically re-testing to ensure the lines were free of secondary symbionts.

Facultative symbionts can be introduced into aphid lines via microinjection, which allows us to create different combinations of host and symbiont genotypes. We established symbionts in new host lines by injecting a small volume of hemolymph (approximately 0.25µl) from an infected adult donor aphid into a one-day-old 1<sup>st</sup> instar recipient using a capillary needle (Lukasik et al., 2015). Injected aphids were then reared on leaves of *V. faba* in Petri dishes at 14°C until they became adults when each aphid was moved into its own dish. These adults were allowed to reproduce, and approximately the 10<sup>th</sup> offspring was reared to maturity, allowed to reproduce,

and then frozen at -20°C. DNA extraction and PCR were used as described above to screen the lines for symbiont establishment. Lines that successfully acquired a *Regiella* infection were then reared individually and re-tested for infection before use in experiments. All of the aphids from a particular line (a combination of a host genotype and *Regiella* strain) are descendants of a single injected aphid that have been propagated for use in experiments and molecular work.

**(1) Density of *Regiella* strains in a common host genotype.** We tested the hypothesis that different *Regiella* strains in genetically identical hosts vary in their relative densities. As part of a different project, symbiont strains were collected from the United Kingdom and the United States, and six housekeeping genes were sequenced to build a phylogeny (Figure 1A) (Parker, Hrcek, et al., 2017). The 13 *Regiella* strains included representatives of the two major phylogenetic clades found in pea aphids (which separated ~500,000 years ago (Henry et al., 2013)). In this previous study, we transferred each *Regiella* strain into an uninfected aphid recipient (genotype 145) using the protocol described above.

In this study, we measured relative symbiont density using quantitative PCR (qPCR). We reared lines for eight generations and froze adult aphids (11 days old) individually from each line at -80°C. Note that aphids frozen for this study were sampled from the same generation as those used to collect the experimental data in Parker, Hrcek, et al. (2017). We extracted DNA from each aphid (N=2 aphids per *Regiella* strain) using the Qiagen DNEasy kit under recommended conditions. We designed qPCR primers based on a conserved region of the *hrpA* gene (F: CGCATTGGGAGAAAAGCCAAG; R: CCTTCCACCAAGCCATGACG). We compared amplification of *Regiella hrpA* to Glyceraldehyde-3-phosphatedehydrogenase (*g3PDH*) in the aphid genome (F: CGGGAATTTTCATTGAACGAC; R: TCCACAACACGGTTGGAGTA). It is important to note, therefore, that our measurements of symbiont density are relative to the number of host cells in a sample. PCR conditions for both sets of primers were optimized against a 1:10 serial dilution of gDNA (200ng – 0.2ng gDNA per reaction) to 99.2% for *hrpA* and 100.0% for *g3PDH*, and were ultimately run at concentrations of 400nM F/350nM R and 300nM for *g3PDH* and *hrpA*, respectively. qPCR reactions were run on a Bio-RAD CFX96 Real-Time System machine, with an initial step of 95°C for 3 minutes and 40 cycles of 95°C for 10s and 60°C for 30s. Each 20µL reaction included a 1X PCR buffer, Mg<sup>+2</sup> at 2mM, dNTPs at 0.2mM, EvaGreen at 1X, 0.025 units/µL of Invitrogen taq, and 40ng of gDNA. We ran three technical replicates for each reaction.

We calculated a  $-\Delta C_T$  value for each sample by subtracting the mean  $C_T$  value of the three technical replicates for *g3PDH* from the mean  $C_T$  value of *hrpA*. We analyzed  $-\Delta C_T$  values using a 2-factor ANOVA after testing for model assumptions using R v.3.4.1. *Regiella* strain was nested within clade, and we analyzed the main effects of clade and strain.

**(2) Correlation between *Regiella* density and aphid survival.** In a previous study, we quantified the survival costs of carrying *Regiella* in these same 13 strains in a single aphid host genetic background (Parker, Hrcek, et al., 2017). Briefly, adult aphids were housed on *V. fava* leaves in agar in Petri dishes at 20°C, and survival was recorded after 8 days. We calculated the percentage of aphids harboring each strain of *Regiella* that survived during the experiment, and we subtracted this value from the survival of symbiont-free aphids of the same genotype to calculate survival cost. Here, we used this survival data together with the symbiont density data to test the hypothesis that higher density symbiont strains impose stronger costs on their hosts. We calculated a non-parametric (Kendall's rank-order) correlation coefficient to look for an association between symbiont density and survival cost using the statistical package R v.3.4.1. We first analyzed all of the strains together in a single analysis, and then repeated the analysis for the strains in the two major clades separately.

**(3) GxG effects on *Regiella* density.** Using a second panel of lines generated previously (Parker, Hrcek, et al., 2017), we tested the hypothesis that variation in density among *Regiella* strains varies across host genotypes (a  $G_{\text{Host}} \times G_{\text{Symbiont}}$  interaction). Each of the six aphid genotypes included in this panel had a 'natural' *Regiella* infection when originally collected, and the aphid lines belonged to two biotypes (associated with the host plants *Trifolium pretense* or *Medicago sativa*). Six *Regiella* strains were established in each of the six host genotypes as described above (except for three combinations which failed). We focused our study of symbiont density only on the novel pairings of host and symbiont genotypes (which were generated using injection) and did not include the 'natural' pairings. Samples were collected and *Regiella* densities were measured using qPCR as above (N=3-8 aphids/line). We analyzed the  $-\Delta C_T$  values using a generalized linear model with a quasi-likelihood distribution in R v.3.4.1. The model included terms for *Regiella* genotype, host genotype, and the interaction between these terms. We performed model comparisons using ANOVA, by first removing the interaction term, then host genotype, then symbiont genotype.



**(4) *Regiella* density across diverse aphid biotypes.** We found a significant but small effect of host genotype on *Regiella* density in the experiment above; we then tested the hypothesis that the density of a single strain of *Regiella* would vary across a more diverse panel of genotypes from multiple host-plant associated biotypes, some of which rarely harbor *Regiella*. We collected genotypes from 8 different host plants from field sites in southern England (collection information can be found in Table S1). Aphids were assigned to biotypes using microsatellite sequencing (Hrcek et al., 2018). Lines were tested for symbionts and these were then removed as described above. We then established a single strain of *Regiella* (strain 313, collected in Gloucestershire, UK in 2007) in each aphid genotype using hemolymph transfers as above, and we confirmed successful *Regiella* transfer by PCR detection. We reared aphids in the lab for eight generations, at which point we screened the lines for *Regiella* again. This procedure often required multiple replicate injected lines of an aphid genotype because it was difficult to establish *Regiella* in many of the lines and some replicates lost their symbiont infections by generation eight. This protocol was successful for 28 aphid genotypes, from which we then collected three adult aphids each, extracted DNA, and measured *Regiella* density using qPCR as above.

**(5) Host genotype effects on fungal protection and survival costs.** Using the panel of 28 aphid lines with and without *Regiella*, we tested the hypothesis that the strength of symbiont mediated protection against *Pandora* and the survival costs of harboring *Regiella* varied across host genotypes. We obtained strains of the fungal pathogen *Pandora neoaphidis* from the USDA Agricultural Research Service's Collection of Entomopathogenic Fungal Cultures. We maintained fungal cultures in the laboratory by infecting aphids from a symbiont-free aphid clone that is a genomic mixture of multiple aphid biotypes every two weeks (genotype 145), and we stored dried cadavers at 4°C. Fungal cadavers were hydrated to induce sporulation by placing them on tap water agar for 12h overnight. We used spores from *Pandora* strain ARSEF2588 for the main fungal infection.

We infected aphids from the 28 host genotypes with and without *Regiella* strain 313 in a single temporal block. Aphids were reared on *V. faba* plants in cup-cages at 20°C for several generations. We then subjected adult aphids (11 days old) to a 'spore shower' from two fungal cadavers in an infection chamber, with all of the aphids from an aphid line together in a single chamber. Fungal cadavers were rotated among the infection chambers so that each cadaver was present for the same amount of time above each chamber to ensure equal spore dose and

infection risk across aphids. We have used this protocol previously to control for fungal cadaver age, pathogen dose, and aphid cadaver quality (Hrcek et al., 2018; McLean et al., 2018; Parker, Hrcek, et al., 2017). Control aphids were handled similarly but were not exposed to fungal spores. Between 23 and 106 aphids (mean 90.6) were used per line depending on availability. After infection, we maintained aphids in dishes as above in groups of five. Each dish was wrapped in parafilm to keep the humidity high (>95%) and kept at 20°C for 48 hours. We then moved aphids to new dishes without parafilm and with fresh plant material. The experiment was blinded (i.e. observers were unaware of aphid genotype, symbiont status or fungal exposure treatment), and pathogen infection frequencies and survival were recorded on day 7 after infection.

We analyzed sporulation status (whether an aphid produced visible signs of sporulation) using binomial-error, linear mixed-effects models implemented in the lme4 package in R v.3.4.1 (Bates, Mächler, Bolker, & Walker, 2015). Symbiont status (no secondary symbionts or *Regiella* strain 313) and aphid biotype were treated as fixed effects, with aphid genotype nested within biotype and modeled as a random effect. Minimal models were derived by first removing the interaction term followed by biotype and then symbiont, with the statistical importance of each term assessed using ANOVA. Post-hoc comparisons of aphid biotypes were made using Tukey's HSD pair-wise comparisons in the multcomp package (Hothorn, Bretz, & Westfall, 2008) in R v.3.4.1. We analyzed aphid survival to assess the costs of harboring *Regiella* using the same methods.

We repeated this experiment using six host genotypes, again with and without *Regiella* strain 313, this time using a different genotype of fungal pathogen (ARSEF6631). Sporulation status was recorded as above on day seven after infection as above. These data were modeled using binomial-error linear models (with aphid genotype and symbiont status as fixed effects) and Tukey's HSD pair-wise comparisons.

**(6) Correlations between *Regiella* density and host fitness across aphid biotypes.** We then tested the hypothesis that variation in density across diverse aphid genotypes is correlated with the protective benefits and survival costs of harboring *Regiella*. We calculated Kendall's rank correlation coefficients (in R v.3.4.1) between symbiont density and fitness effects for the 28 genotypes in the fungal infection experiment.  $-\Delta C_T$  values were used as the measurement of density as above. The strength of protection was calculated by (percent sporulated aphids with

*Regiella*) / (percent sporulated aphids without *Regiella*) for each genotype. Survival costs were calculated with (percent survival with *Regiella*) / (percent survival without *Regiella*), with larger values indicating higher survival costs.

**(7) *Regiella* density and loss across aphid generations.** In the experiments above we found several aphid lines with very low *Regiella* densities, and we noticed that many aphid lines lost their *Regiella* infections over the eight generations after symbiont establishment. We tested the hypothesis that these low-density lines are in the process of losing their symbiont infections by tracking symbiont density over multiple generations after injection. We selected two aphid genotypes from the previous experiments, one from the *Trifolium* biotype (genotype 317) and one from *Lotus corniculatus* (genotype 663). We selected these lines for further examination because in a previous study we found that aphids from *Trifolium* retain their symbionts at a higher frequency than do aphids from *Lotus corniculatus*. In the experiment above, genotype 317 from *Trifolium* had a much higher *Regiella* density ( $-\Delta C_T = 2.86$ ) than genotype 663 from *Lotus corniculatus* ( $-\Delta C_T = -6.02$ ) after eight generations, and we expected genotype 663 to have a low density of *Regiella* because it was in the process of losing its symbiont. We conducted the experiment in two blocks that were several months apart. We injected replicates of both genotypes (genotype 663: block 1, n = 9; block 2, n = 10; genotype 317: block 1, n = 3; block 2, n = 4) with *Regiella* strain 313 using the injection protocols described above. We then screened adult aphids from each replicate at the first generation after injection (generation 1) and each of the next four generations (generations 2-5). We collected 2-5 adult aphids from each line on the first day of reproduction at each generation and stored the samples at  $-20^{\circ}\text{C}$ . We extracted DNA and screened each sample for *Regiella* using PCR protocols as described above. For samples with a positive PCR result indicating the presence of *Regiella*, we used qPCR to measure symbiont density at generation 2 and at generation 4. We compared the ratio of symbionts lost and retained between the two genotypes with a 2 x 2 contingency table and a Yates corrected Chi-squared test.

## RESULTS

**(1) *Regiella* strains vary in density in a common host background.** Using the panel of 13 *Regiella* genotypes in a common host background, we found that two *Regiella* clades significantly differed in density (ANOVA; clade:  $F = 212$ , 1DF,  $p < 0.0001$ ; Figure 1A), and *Regiella* strains significantly varied in density independently from the effect of clade (Strain:  $F = 4.31$ , 11DF,  $p = 0.01$ ; Figure 1A). Densities of Clade 2 *Regiella* were on average 13.3 times higher than those in Clade 1, and phylogenetic clade explained a much larger portion of the variance in our model (78.1% total variance) than did strain (17.5%).

**(2) Higher density is correlated with lower host survival.** We found that higher *Regiella* densities are correlated with higher survival costs to hosts ( $\tau = 0.45$ ;  $z = 2.12$ ,  $p = 0.032$ ; Figure 1B). An important caveat, however, is that differences between the two clades other than density might impact host survival, and we did not find significant correlations between *Regiella* densities and survival cost among strains within each clade (Clade 1:  $\tau = 0.33$ ,  $z = 0.94$ ,  $p = 0.35$ ; Clade 2:  $\tau = 0.20$ ,  $z = 0.61$ ,  $p = 0.54$ ).

**(3) Variation in density due to symbiont genotype is not strongly dependent on host genotype (a  $G_{\text{Host}} \times G_{\text{Symbiont}}$  interaction).** Using the panel of six *Regiella* genotypes in six host genotypes, we again found that Clade 1 *Regiella* established at lower densities than Clade 2 symbionts (Figure 1C). In our model, symbiont genotype ( $F = 1084$ , 5DF,  $p < 0.0001$ ), host genotype ( $F = 57.0$ , 5DF,  $p < 0.0001$ ), and their interaction ( $F = 15.4$ , 16DF,  $p < 0.0001$ ) had statistically significant effects on density. Symbiont genotype explained 93.6% of the deviance accounted for by the full statistical model, while host genotype (2.1%) and the host genotype by symbiont genotype interaction (4.3%) explained less of the deviance in the model.

**(4) Aphid genotypes from diverse biotypes vary in *Regiella* density.** Using 28 different host genotypes in which we established *Regiella* (strain 313), we found extensive variation in symbiont density due to host genotype (Kruskal-Wallis  $\chi^2 = 71.7$ , DF = 26,  $p < 0.0001$ ; Figure 2A) and among aphid biotypes (Kruskal-Wallis  $\chi^2 = 26.8$ , DF = 7,  $p < 0.001$ ; Figure 2A). Some lines were found to have very low *Regiella* densities that were several orders of magnitude lower than average (Figure 2A).

**(5) Host genotype influences the protective benefits and survival costs of harboring**

***Regiella*.** Using the host genotype panel, we found that *Regiella* protected aphids from *Pandora* infection as expected (Figure 2B; Symbiont:  $\chi^2 = 169.32$ , 1DF,  $p < 0.0001$ ). However, biotypes differed significantly in the extent of this protection (Symbiont \* Biotype:  $\chi^2 = 89.9$ , 7DF,  $p < 0.0001$ ; post-hoc comparisons Figure 2B; Table S2). Specifically, we found that harboring *Regiella* provided no protection against fungal infection in aphids from the *Medicago lupulina* and *Lathyrus pratensis* biotypes (post-hoc comparisons, Figure 2B; Table S2). A replicate experiment using a second fungal isolate (ARSEF6631, which produced a higher rate of sporulation than the main infection) mirrored these results: protection against fungus was again strongly dependent on aphid host genotype, with the two genotypes from *Medicago lupulina* experiencing no protection from *Regiella* (Figure S1; Table S2). Host genotypes also differed in the severity with which they experienced survival costs due to *Regiella*, with this pattern differing significantly across aphid biotypes (Biotype x Symbiont:  $\chi^2 = 26.3$ , 7DF,  $p < 0.001$ ; Figure 2C; post-hoc comparisons, Table S2).

**(6) The costs but not the benefits of symbiosis are correlated with density.** Across the host biotype panel, we found that lines with higher *Regiella* densities experienced greater fitness costs of harboring *Regiella* (stronger decrease in survival of control aphids when harboring *Regiella*; tau = 0.31;  $z = 2.33$ ;  $p = 0.020$ , Figure 2D), but the correlation between density and protection against fungal pathogens was not statistically significant (tau = 0.27;  $z = 1.93$ ;  $p = 0.053$ , Figure 2E).

**(7) Symbiont loss is preceded by a drop in density and varies among genotypes.** We observed that some aphid lines lost their *Regiella* infections after injection, and in a previous study we found that aphid genotypes varied in their ability to maintain *Regiella* across generations (Parker, McLean, et al., 2017). We hypothesized that the very low densities we measured in some aphid lines in the experiment described above could reflect lines that are in the process of losing their symbiont infections. In the experiment above, genotype 317 from *Trifolium* had a much higher *Regiella* density ( $-\Delta C_T = 2.86$ ) than genotype 663 from *Lotus corniculatus* ( $-\Delta C_T = -6.02$ ) after 8 generations. In this experiment, we found that no lines from genotype 317 (0 out of 7) lost their *Regiella* infection over 5 generations. Among lines from genotype 663, 11 out of 19 lines lost their symbiont. In 10 lines from genotype 663, loss was

preceded by a sharp drop in symbiont density (Figure 3). The rate of symbiont loss differed significantly between the two genotypes ( $\chi^2$  with Yates' correction = 4.37, 1DF,  $p = 0.037$ ).

## DISCUSSION

We confirmed that symbiont density varies extensively across both host and microbial genotypes. Strains from the two major *Regiella* clades, when introduced into a single host genetic background, were found to be present at average densities that varied by ~13X (with clade explaining a large part of this variation; Figure 1A). A similar pattern was found when *Regiella* strains from the two clades were introduced into multiple host genotypes (Figure 1C). *Regiella* densities were quite similar across those aphid genotypes which when collected in the field naturally harbored this symbiont (Figure 1C). But we found much more extensive variation in symbiont density across aphid genotypes from multiple biotypes that had not all originally carried *Regiella* (Figure 2A). We found that low densities of symbionts are a prelude to the complete loss of the association, and that the likelihood this occurs is influenced by host genotype (Figure 3).

Across host genotypes, the strength of symbiont-mediated resistance to *Pandora* was not significantly correlated with *Regiella* density: some aphid lines with a high symbiont density experienced no protection against *Pandora* infection, while other lines with fewer symbionts were well-protected (Figures 2B; 2D). Variation in density across *Regiella* strains may not be correlated with fungal protection because, in this system, there is a symbiont strain by fungal genotype (GxG) interaction (Parker, Hreck, et al., 2017) with particular *Regiella* strains conferring strong protection only against specific *Pandora* genotypes. Other studies have found correlations between symbiont density and protection from pathogens across symbiont genotypes (e.g. Chrostek et al., 2013; Martinez et al., 2015; Martinez et al., 2017). A better understanding of when density influences symbiont-mediated protection will be aided by more examples of the molecular mechanisms underlying protection (Gerardo & Parker, 2014). One possibility is that some types of mechanisms, like the production of anti-pathogen toxins (e.g. (Ballinger & Perlman, 2017; Oliver, Degnan, Hunter, & Moran, 2009)), are density-dependent while others, like protection that is mediated by the host's immune system (e.g. Hughes, Koga, Xue, Fukatsu, & Rasgon, 2011; Kambris et al., 2010), are not influenced by symbiont density.

Evidence for our hypothesis that symbiont density is correlated with fitness costs to hosts was mixed. Survival costs were significantly correlated with *Regiella* density across host genotypes, but there were a number of lines with a high density *Regiella* infection that experienced no measurable costs of symbiont infection (Figures 2C; 2E). Across *Regiella* strains, higher density infections were correlated with higher survival costs to aphids (Figure 1B). But other genetic differences between the two *Regiella* clades might be driving both costs to hosts and differences in density, and we found no correlation between density and survival within the symbiont clades. These results might suggest that *Regiella* from the two clades, which are separated by a half million years (Henry et al., 2013), have adopted different strategies for surviving within hosts, i.e. Clade 2 *Regiella* establish at higher densities in hosts but are more costly to hosts. It is possible that higher density *Regiella* infections might protect symbionts against failure to be transmitted to offspring or increase the probability of horizontal transmission (Henry et al., 2013), but come at the expense of host survival.

By following infections across multiple generations, we found that low density infections tend to precede symbiont loss, and that the rate of low-density followed by symbiont loss varies between two genotypes (Figure 3). Aphid facultative symbionts are thought to be transmitted with high fidelity from mothers to offspring, but recent studies have shown how host biotype (Parker, McLean, et al.) and donor and recipient relatedness (Lukasik et al., 2015; Niepoth et al., 2018) affect the probability that symbionts are lost from novel infections created in the laboratory. Aphids living on different species of host plant are likely to vary in their risk of exposure to natural enemies, and aphids in some environments might be more strongly selected to maintain their symbionts than others (Hrcek et al., 2018).

Overall, we find considerable variation in the density of symbionts in this system that is driven by both symbiont strain and host biotype. In addition, other abiotic and biotic factors that we did not explore here could also impact symbiont density. For example, aphid symbionts are frequently found in co-infections of multiple bacterial species (Russell et al., 2013), and some studies suggest that symbiont density (and the phenotypic effects a symbiont has on a host) might be influenced by the presence of other facultative endosymbiont species within a host (Guay, Boudreault, Michaud, & Cloutier, 2009; McLean et al., 2018; Oliver, Moran, & Hunter, 2006). Our data suggest that the main relevance of density for this symbiosis is not through effects on the

strength of symbiont mediated protection, but rather through its potential effects on symbiont transmission. Hosts will benefit from losing protective symbionts in environments where the risk of pathogen infection is low and keeping symbionts in other environments. Symbionts, in contrast, will be under pressure to avoid being lost from a host in all environments, but symbiont fitness is tied to that of its host, and trade-offs between density and host fitness might shape symbiont mechanisms that determine density. Together our data suggest that the optimal density of a *Regiella* infection is different from the perspective of aphid and symbiont fitness, and that the standing variation in density found in natural populations might be maintained in some systems by antagonistic evolutionary interactions between hosts and their microbes.

ACKNOWLEDGEMENTS: We would like to thank Ciara Mann and Holly Nichols for valuable assistance with the experiments, Julia Ferrari for providing aphid genotypes, and Richard Humber and the USDA ARSEF collection of entomopathogenic fungi for providing the strains of *Pandora neoaphidis*. This work was funded by US NSF postdoctoral fellowship DBI-1306387 to BJP and UK NERC grant NE/K004972/1 to HCJG. BJP is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trusts. JH was supported by Czech Science Foundation grant no. 17-27184Y.

DATA ACCESSIBILITY AND BENEFIT-SHARING STATEMENT: All experimental data have been uploaded to the Dryad Digital Repository: <https://doi.org/10.5061/dryad.j3tx95xcv>

STATEMENT OF AUTHORSHIP: Designed research: BJP, JH, ACHM, HCJG. Performed research: BJP, JH, ACHM. Analyzed data: BJP, JAB. Wrote the paper: BJP, HCJG. Revised the manuscript critically for important intellectual content and approval of the version of the manuscript to be submitted: all authors.



## REFERENCES

- Ballinger, M. J., & Perlman, S. J. (2017). Generality of toxins in defensive symbiosis: Ribosome-inactivating proteins and defense against parasitic wasps in *Drosophila*. *PLoS Pathog*, 13(7), e1006431. doi:10.1371/journal.ppat.1006431
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). doi:10.18637/jss.v067.i01
- Brownlie, J. C., & Johnson, K. N. (2009). Symbiont-mediated protection in insect hosts. *Trends Microbiol*, 17(8), 348-354. doi:10.1016/j.tim.2009.05.005
- Chong, R. A., & Moran, N. A. (2016). Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts. *Proc Natl Acad Sci U S A*, 113(46), 13114-13119. doi:10.1073/pnas.1610749113
- Chrostek, E., Marialva, M. S., Esteves, S. S., Weinert, L. A., Martinez, J., Jiggins, F. M., & Teixeira, L. (2013). Wolbachia variants induce differential protection to viruses in *Drosophila melanogaster*: a phenotypic and phylogenomic analysis. *PLoS Genet*, 9(12), e1003896. doi:10.1371/journal.pgen.1003896
- Chrostek, E., & Teixeira, L. (2015). Mutualism breakdown by amplification of Wolbachia genes. *PLoS Biol*, 13(2), e1002065. doi:10.1371/journal.pbio.1002065
- Douglas, A. E. (2014). The molecular basis of bacterial-insect symbiosis. *J Mol Biol*, 426(23), 3830-3837. doi:10.1016/j.jmb.2014.04.005
- Ferrari, J., & Vavre, F. (2011). Bacterial symbionts in insects or the story of communities affecting communities. *Philos Trans R Soc Lond B Biol Sci*, 366(1569), 1389-1400. doi:10.1098/rstb.2010.0226
- Ferrari, J., Via, S., & Godfray, H. C. (2008). Population differentiation and genetic variation in performance on eight hosts in the pea aphid complex. *Evolution*, 62(10), 2508-2524. doi:10.1111/j.1558-5646.2008.00468.x
- Ferrari, J., West, J. A., Via, S., & Godfray, H. C. (2012). Population genetic structure and secondary symbionts in host-associated populations of the pea aphid complex. *Evolution*, 66(2), 375-390. doi:10.1111/j.1558-5646.2011.01436.x
- Funkhouser-Jones, L. J., van Opstal, E. J., Sharma, A., & Bordenstein, S. R. (2018). The Maternal Effect Gene Wds Controls Wolbachia Titer in *Nasonia*. *Curr Biol*, 28(11), 1692-1702 e1696. doi:10.1016/j.cub.2018.04.010
- Gerardo, N. M., & Parker, B. J. (2014). Mechanisms of symbiont-conferred protection against natural enemies: an ecological and evolutionary framework. *Curr Opin Insect Sci*, 4, 8-14. doi:10.1016/j.cois.2014.08.002
- Guay, J. F., Boudreault, S., Michaud, D., & Cloutier, C. (2009). Impact of environmental stress on aphid clonal resistance to parasitoids: Role of *Hamiltonella defensa* bacterial symbiosis in association with a new facultative symbiont of the pea aphid. *J Insect Physiol*, 55(10), 919-926. doi:10.1016/j.jinsphys.2009.06.006
- Hansen, A. K., Jeong, G., Paine, T. D., & Stouthamer, R. (2007). Frequency of secondary symbiont infection in an invasive psyllid relates to parasitism pressure on a geographic scale in California. *Appl Environ Microbiol*, 73(23), 7531-7535. doi:10.1128/AEM.01672-07
- Henry, L. M., Peccoud, J., Simon, J. C., Hadfield, J. D., Maiden, M. J., Ferrari, J., & Godfray, H. C. (2013). Horizontally transmitted symbionts and host colonization of ecological niches. *Curr Biol*, 23(17), 1713-1717. doi:10.1016/j.cub.2013.07.029
- Herren, J. K., & Lemaitre, B. (2011). Spiroplasma and host immunity: activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-

negative bacterial pathogens in *Drosophila melanogaster*. *Cell Microbiol*, 13(9), 1385-1396. doi:10.1111/j.1462-5822.2011.01627.x

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom J*, 50(3), 346-363. doi:10.1002/bimj.200810425

Hrcek, J., Parker, B. J., McLean, A. H. C., Simon, J. C., Mann, C. M., & Godfray, H. C. J. (2018). Hosts do not simply outsource pathogen resistance to protective symbionts. *Evolution*. doi:10.1111/evo.13512

Hughes, G. L., Koga, R., Xue, P., Fukatsu, T., & Rasgon, J. L. (2011). Wolbachia infections are virulent and inhibit the human malaria parasite *Plasmodium falciparum* in *Anopheles gambiae*. *PLoS Pathog*, 7(5), e1002043. doi:10.1371/journal.ppat.1002043

Ives, A. R., Barton, B. T., Penczykowski, R. M., Harmon, J. P., Kim, K. L., Oliver, K., & Radeloff, V. C. (2020). Self-perpetuating ecological-evolutionary dynamics in an agricultural host-parasite system. *Nat Ecol Evol*, 4(5), 702-711. doi:10.1038/s41559-020-1155-0

Jaenike, J. (2012). Population genetics of beneficial heritable symbionts. *Trends Ecol Evol*, 27(4), 226-232. doi:10.1016/j.tree.2011.10.005

Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M., & Perlman, S. J. (2010). Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science*, 329(5988), 212-215. doi:10.1126/science.1188235

Kambris, Z., Blagborough, A. M., Pinto, S. B., Blagrove, M. S., Godfray, H. C., Sinden, R. E., & Sinkins, S. P. (2010). Wolbachia stimulates immune gene expression and inhibits plasmodium development in *Anopheles gambiae*. *PLoS Pathog*, 6(10), e1001143. doi:10.1371/journal.ppat.1001143

Login, F. H., Balmand, S., Vallier, A., Vincent-Monegat, C., Vigneron, A., Weiss-Gayet, M., . . . Heddi, A. (2011). Antimicrobial peptides keep insect endosymbionts under control. *Science*, 334(6054), 362-365. doi:10.1126/science.1209728

Lukasik, P., Guo, H., van Asch, M., Henry, L. M., Godfray, H. C., & Ferrari, J. (2015). Horizontal transfer of facultative endosymbionts is limited by host relatedness. *Evolution*, 69(10), 2757-2766. doi:10.1111/evo.12767

Martinez, J., Ok, S., Smith, S., Snoeck, K., Day, J. P., & Jiggins, F. M. (2015). Should Symbionts Be Nice or Selfish? Antiviral Effects of Wolbachia Are Costly but Reproductive Parasitism Is Not. *PLoS Pathog*, 11(7), e1005021. doi:10.1371/journal.ppat.1005021

Martinez, J., Tolosana, I., Ok, S., Smith, S., Snoeck, K., Day, J. P., & Jiggins, F. M. (2017). Symbiont strain is the main determinant of variation in Wolbachia-mediated protection against viruses across *Drosophila* species. *Mol Ecol*, 26(15), 4072-4084. doi:10.1111/mec.14164

McLean, A. H., van Asch, M., Ferrari, J., & Godfray, H. C. (2011). Effects of bacterial secondary symbionts on host plant use in pea aphids. *Proc Biol Sci*, 278(1706), 760-766. doi:10.1098/rspb.2010.1654

McLean, A. H. C., Parker, B. J., Hrcek, J., Kavanagh, J. C., Wellham, P. A. D., & Godfray, H. C. J. (2018). Consequences of symbiont co-infections for insect host phenotypes. *J Anim Ecol*, 87(2), 478-488. doi:10.1111/1365-2656.12705

Niepoth, N., Ellers, J., & Henry, L. M. (2018). Symbiont interactions with non-native hosts limit the formation of new symbioses. *BMC Evol Biol*, 18(1), 27. doi:10.1186/s12862-018-1143-z

Oliver, K. M., Campos, J., Moran, N. A., & Hunter, M. S. (2008). Population dynamics of defensive symbionts in aphids. *Proc Biol Sci*, 275(1632), 293-299. doi:10.1098/rspb.2007.1192

- Oliver, K. M., Degnan, P. H., Hunter, M. S., & Moran, N. A. (2009). Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science*, 325(5943), 992-994. doi:10.1126/science.1174463
- Oliver, K. M., Moran, N. A., & Hunter, M. S. (2006). Costs and benefits of a superinfection of facultative symbionts in aphids. *Proc Biol Sci*, 273(1591), 1273-1280. doi:10.1098/rspb.2005.3436
- Oliver, K. M., Russell, J. A., Moran, N. A., & Hunter, M. S. (2003). Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc Natl Acad Sci U S A*, 100(4), 1803-1807. doi:10.1073/pnas.0335320100
- Pan, X., Pike, A., Joshi, D., Bian, G., McFadden, M. J., Lu, P., . . . Xi, Z. (2018). The bacterium *Wolbachia* exploits host innate immunity to establish a symbiotic relationship with the dengue vector mosquito *Aedes aegypti*. *ISME J*, 12(1), 277-288. doi:10.1038/ismej.2017.174
- Parker, B. J., Hrcek, J., McLean, A. H. C., & Godfray, H. C. J. (2017). Genotype specificity among hosts, pathogens, and beneficial microbes influences the strength of symbiont-mediated protection. *Evolution*, 71(5), 1222-1231. doi:10.1111/evo.13216
- Parker, B. J., McLean, A. H. C., Hrcek, J., Gerardo, N. M., & Godfray, H. C. J. (2017). Establishment and maintenance of aphid endosymbionts after horizontal transfer is dependent on host genotype. *Biol Lett*, 13(5). doi:10.1098/rsbl.2017.0016
- Parker, B. J., Spragg, C. J., Altincicek, B., & Gerardo, N. M. (2013). Symbiont-mediated protection against fungal pathogens in pea aphids: a role for pathogen specificity? *Appl Environ Microbiol*, 79(7), 2455-2458. doi:10.1128/AEM.03193-12
- Peccoud, J., de la Huerta, M., Bonhomme, J., Laurence, C., Outreman, Y., Smadja, C. M., & Simon, J. C. (2014). Widespread host-dependent hybrid unfitness in the pea aphid species complex. *Evolution*, 68(10), 2983-2995. doi:10.1111/evo.12478
- Peccoud, J., Ollivier, A., Plantegenest, M., & Simon, J. C. (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proc Natl Acad Sci U S A*, 106(18), 7495-7500. doi:10.1073/pnas.0811117106
- Russell, J. A., Weldon, S., Smith, A. H., Kim, K. L., Hu, Y., Lukasik, P., . . . Oliver, K. M. (2013). Uncovering symbiont-driven genetic diversity across North American pea aphids. *Mol Ecol*, 22(7), 2045-2059. doi:10.1111/mec.12211
- Sandström, J., & Pettersson, J. (1994). Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrtosiphon pisum*) performance. *Journal of Insect Physiology*, 40(11), 947-955. doi:10.1016/0022-1910(94)90133-3
- Scarborough, C. L., Ferrari, J., & Godfray, H. C. (2005). Aphid protected from pathogen by endosymbiont. *Science*, 310(5755), 1781. doi:10.1126/science.1120180
- Tsuchida, T., Koga, R., & Fukatsu, T. (2004). Host plant specialization governed by facultative symbiont. *Science*, 303(5666), 1989. doi:10.1126/science.1094611
- Tsuchida, T., Koga, R., Shibao, H., Matsumoto, T., & Fukatsu, T. (2002). Diversity and geographic distribution of secondary endosymbiotic bacteria in natural populations of the pea aphid, *Acyrtosiphon pisum*. *Mol Ecol*, 11(10), 2123-2135. doi:10.1046/j.1365-294x.2002.01606.x
- Unckless, R. L., Boelio, L. M., Herren, J. K., & Jaenike, J. (2009). *Wolbachia* as populations within individual insects: causes and consequences of density variation in natural populations. *Proc Biol Sci*, 276(1668), 2805-2811. doi:10.1098/rspb.2009.0287
- Via, S. (1991). The Genetic Structure of Host Plant Adaptation in a Spatial Patchwork: Demographic Variability among Reciprocally Transplanted Pea Aphid Clones. *Evolution*, 45(4), 827-852. doi:10.1111/j.1558-5646.1991.tb04353.x

- Vogel, K. J., & Moran, N. A. (2011). Effect of Host Genotype on Symbiont Titer in the Aphid-Buchnera Symbiosis. *Insects*, 2(3), 423-434. doi:10.3390/insects2030423
- Vorburger, C., & Gouskov, A. (2011). Only helpful when required: a longevity cost of harbouring defensive symbionts. *J Evol Biol*, 24(7), 1611-1617. doi:10.1111/j.1420-9101.2011.02292.x
- Wilkinson, T. L., & Douglas, A. E. (1998). Host cell allometry and regulation of the symbiosis between pea aphids, *Acyrtosiphon pisum*, and bacteria, Buchnera. *Journal of Insect Physiology*, 44, 629-635.
- Data citation:
- Parker, Benjamin et al. (2021), Intraspecific variation in symbiont density in an insect-microbe symbiosis, Dryad, Dataset, <https://doi.org/10.5061/dryad.j3tx95xcv>

## FIGURE AND TABLE LEGENDS

**Figure 1: *Regiella* strains vary in the density at which they establish in hosts.** **A:** The phylogeny of *Regiella* strains is shown along the Y-axis, along with the strain names. A *Regiella* strain from another aphid species (*Diuraphis noxia*) was used to root the phylogenetic tree.  $-\Delta C_T$  (critical threshold) values measured using qPCR are shown along the X-axis, which can be interpreted as the relative symbiont densities of each sample on a  $\log_2$  scale. Each of the biological replicates (a DNA extraction from a single adult aphid) for each strain are shown as points on the plot. Clade 1 and 2 genotypes shown in light and dark grey points, respectively. **B:** This plot shows the relationship between symbiont density (as shown by  $-\Delta C_T$ , on the X-axis) and survival cost (fraction of aphid without *Regiella* surviving the experiment - fraction with *Regiella* surviving on the Y-axis). The mean value for each genotype is shown by a single point, with the Clade 1 and 2 genotypes shown in light and dark grey, respectively. **C:** This graph shows the density of multiple *Regiella* strains established in different host genotypes. The Y-axis shows symbiont density as above, with each point representing a biological replicate. Symbiont and host genotype names are shown along the bottom of the figure as indicated.

**Figure 2: Host genotypes influence *Regiella* density and loss.** The host biotype and genotype names for each line used in these experiments is shown along the top of the figure. **A:** The Y-axis represents *Regiella* strain 313 density as  $-\Delta C_T$  values, which can be interpreted as the relative symbiont density on a  $\log_2$  scale. Each replicate is shown by a grey point. **B:** This plot shows sporulation data (the appearance of a sporulating fungal cadaver after 7 days) of aphid genotypes with (filled circles) and without (empty circles) *Regiella*. The strength of symbiont-mediated protection is reflected by the relative sporulation of lines with and without *Regiella* and is represented by a vertical grey line. Post-hoc tests comparing these differences across biotypes (interaction terms between biotype and *Regiella* presence) are shown along the top of the figure. **C:** The Y-axis shows the percent survival data of control aphids (not exposed to fungus) after 7 days. Circles and lines are as above. **D:** This plot shows the relationship between symbiont density and survival costs (percent survival without *Regiella* – percent survival with *Regiella*). **E:** This plot shows the relationship between symbiont density and protective benefit (percent sporulated aphids with *Regiella* / percent sporulated aphids without *Regiella*).

**Figure 3: Symbiont loss over several generations.** The y-axes show symbiont density as above, with samples taken at multiple generations shown along the top of the figure. Each point represents a different aphid line established with a *Regiella* infection from host genotype 663 (blue) or 317 (red) connected by a grey line. Lines that lost their symbiont infection as determined by PCR are shown along the bottom of the figure.

**Supplementary Information:**

**Supplementary Figure 1:** Results of fungal infection using a second *Pandora* fungal genotype (ARSEF6631). Aphid genotypes are a subset of the lines used in the first infection. Error bars show standard error, and post-hoc tests comparing the interaction between genotype and symbiont status are shown along the top of the figure. Aphid biotypes are indicated by the colors of the genotype labels along the x-axis.

**Table S1:** Information on the collection location, year, symbiont background, and biotype of aphid lines used in this study. Symbiont abbreviations: *Ham*: *Hamiltonella defensa*, *Reg*: *Regiella insecticola*, *X*: X-type (also known as *Candidatus Fukatsuia symbiotica*), *Rick*: *Rickettsia* sp., *Serr*: *Serratia symbiotica*.

**Table S2:** Results of post-hoc tests from the fungal infections. Statistically significant differences ( $P < 0.05$ ) are indicated in bold.

Figure 1

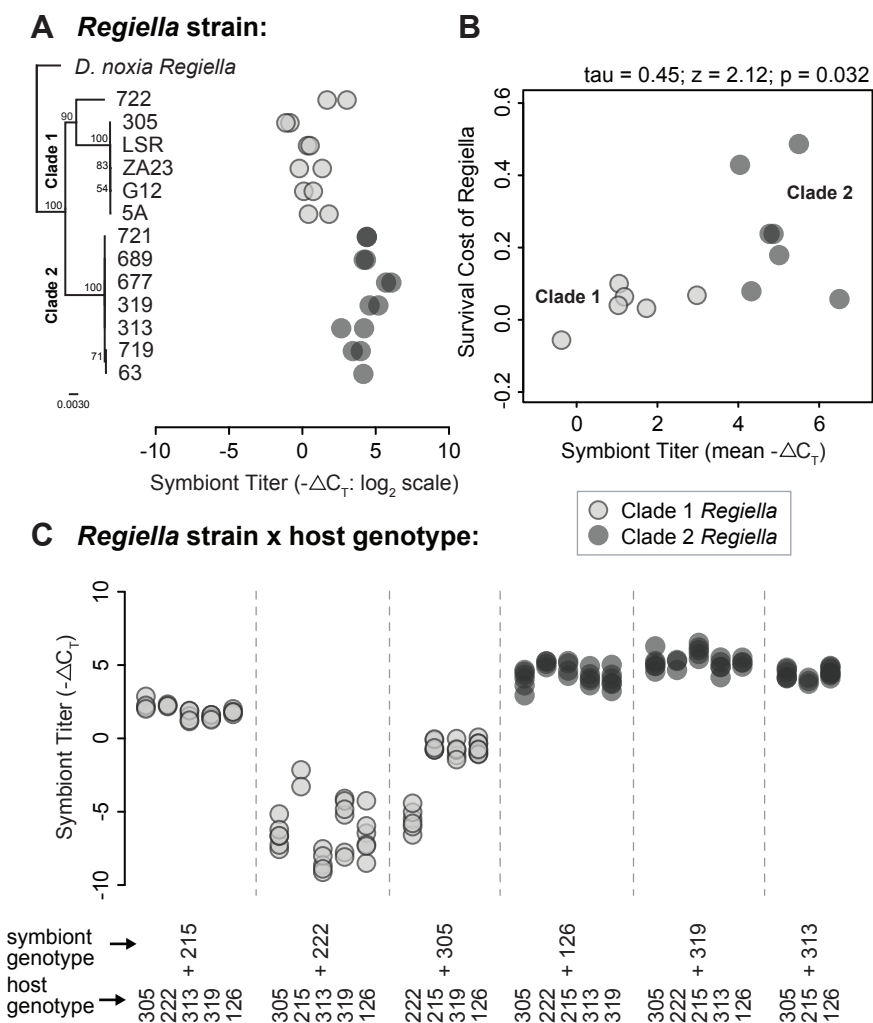


Figure 2

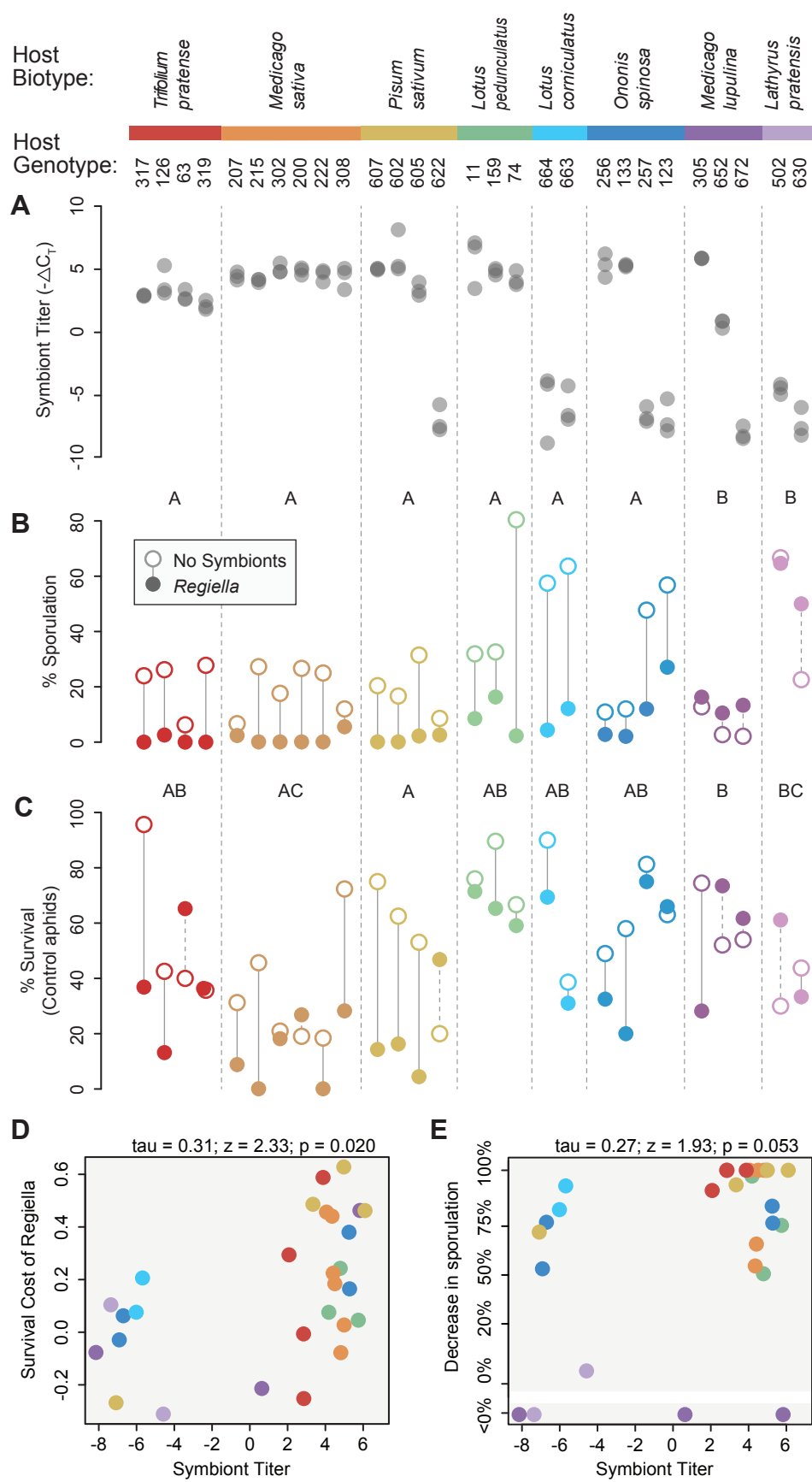




Figure 3

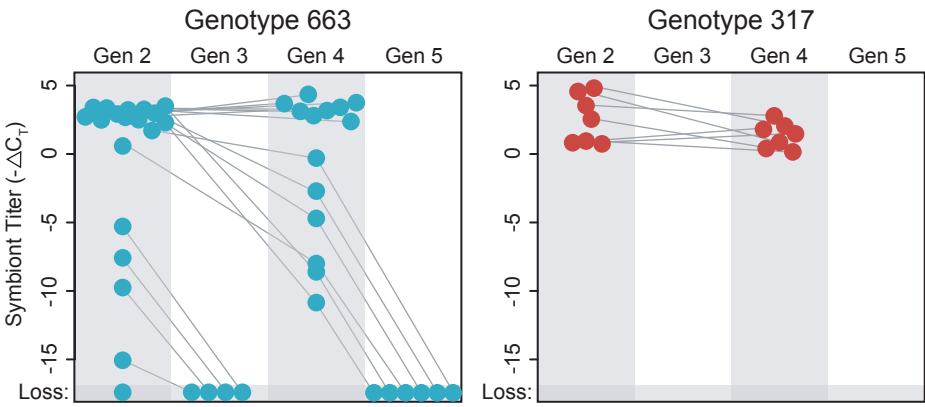
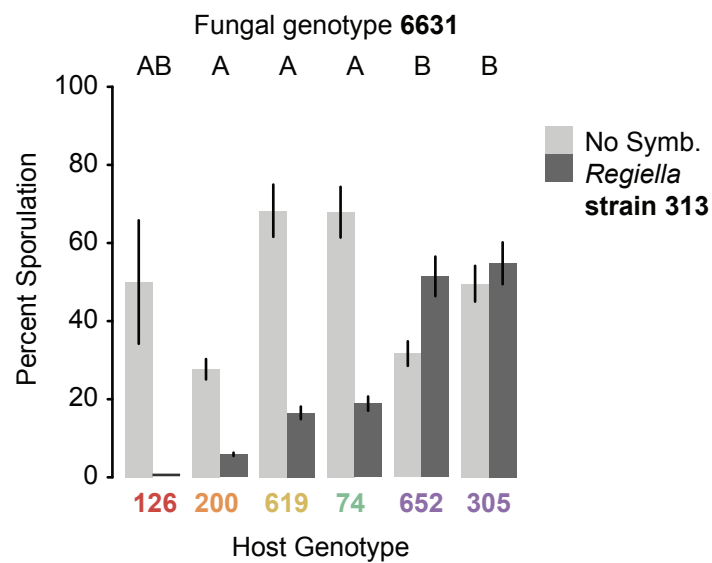


Figure S1



**Table S1:** Information on the collection location, year, symbiont background, and biotype of aphid lines used in this study. Symbiont abbreviations: *Ham*: *Hamiltonella defensa*, *Reg*: *Regiella insecticola*, *X*: X-type (also known as *Candidatus Fukatsuia symbiotica*), *Rick*: *Rickettsia* sp., *Serr*: *Serratia symbiotica*.

Host Genotype (lab code)	Location Collected	Year Collected	Original Symbionts	Biotype
502	Maidens Green	2003	None	<i>Lathyrus pratensis</i>
630	Bernwood meadows	2013	None	<i>Lathyrus pratensis</i>
663	Bernwood meadows	2014	None	<i>Lotus corniculatus</i>
664	Bernwood meadows	2014	None	<i>Lotus corniculatus</i>
11	Whitby	2012	<i>Ham</i>	<i>Lotus pedunculatus</i>
159	Windsor	2012	<i>Ham</i>	<i>Lotus pedunculatus</i>
74	Windsor	2012	<i>Ham</i>	<i>Lotus pedunculatus</i>
652	Bourton-on-the-water	2014	None	<i>Medicago lupulina</i>
672	Chilswell (nr. Oxford)	2014	None	<i>Medicago lupulina</i>
305	Warburg Reserve	2003	<i>Reg, Ham</i>	<i>Medicago lupulina</i>
207	Lincoln	2012	<i>Ham, X</i>	<i>Medicago sativa</i>
215	Lincoln	2012	<i>Reg</i>	<i>Medicago sativa</i>
302	Lincoln	2012	<i>Ham, X</i>	<i>Medicago sativa</i>
200	Lincoln	2012	None	<i>Medicago sativa</i>
222	Whitby field	2003	<i>Reg, Ham, Rick</i>	<i>Medicago sativa</i>
308	Lincoln	2012	<i>Ham</i>	<i>Medicago sativa</i>
123	Wernershoehe	2012	<i>Ham</i>	<i>Ononis spinosa</i>
256	Speen	2012	<i>Ham</i>	<i>Ononis spinosa</i>
257	Speen	2012	<i>Ham</i>	<i>Ononis spinosa</i>
133	Cock Marsh	2012	<i>Ham</i>	<i>Ononis spinosa</i>
607	Wytham Wood	2013	<i>Serr</i>	<i>Pisum sativum</i>
602	Oxford Botanical Garden	2014	<i>Serr</i>	<i>Pisum sativum</i>
622	Barracks Lane, East Oxford	2014	None	<i>Pisum sativum</i>
605	Meadow Lane, East Oxford	2013	<i>Serr</i>	<i>Pisum sativum</i>
619	Barracks Lane, East Oxford	2014	<i>Serr</i>	<i>Pisum sativum</i>
126	Legoland	2003	<i>Reg</i>	<i>Trifolium spp.</i>
319	Oddington	2012	<i>Reg</i>	<i>Trifolium spp.</i>
63	Windsor Ranger's Gate	2010	<i>Reg</i>	<i>Trifolium spp.</i>
317	Oxfordshire	2012	<i>Reg</i>	<i>Trifolium spp.</i>
145	Berkshire	2003	None	Mix

**Table S2:** Results of post-hoc tests from the fungal infections. Statistically significant differences ( $P < 0.05$ ) are indicated in bold.

Sporulation	Test Statistic	p
<i>Lotus corniculatus</i> vs <i>Lotus pedunculatus</i>	$z = 1.090$	0.95471
<i>Lotus corniculatus</i> vs <i>Lathyrus pratensis</i>	$z = 5.515$	<b>&lt; 0.001</b>
<i>Lotus corniculatus</i> vs <i>Medicago lupulina</i>	$z = 5.845$	<b>&lt; 0.001</b>
<i>Lotus corniculatus</i> vs <i>Medicago sativa</i>	$z = 0.624$	0.99839
<i>Lotus corniculatus</i> vs <i>Ononis spinosa</i>	$z = 2.598$	0.14447
<i>Lotus corniculatus</i> vs <i>Pisum sativum</i>	$z = -0.041$	1.00000
<i>Lotus corniculatus</i> vs <i>Trifolium spp.</i>	$z = -0.610$	0.99860
<i>Lotus pedunculatus</i> vs <i>Lathyrus pratensis</i>	$z = 5.249$	<b>&lt; 0.001</b>
<i>Lotus pedunculatus</i> vs <i>Medicago lupulina</i>	$z = 5.660$	<b>&lt; 0.001</b>
<i>Lotus pedunculatus</i> vs <i>Medicago sativa</i>	$z = -0.321$	0.99998
<i>Lotus pedunculatus</i> vs <i>Ononis spinosa</i>	$z = 1.814$	0.58962
<i>Lotus pedunculatus</i> vs <i>Pisum sativum</i>	$z = -0.847$	0.98919
<i>Lotus pedunculatus</i> vs <i>Trifolium spp.</i>	$z = -1.240$	0.91211
<i>Lathyrus pratensis</i> vs <i>Medicago lupulina</i>	$z = -0.071$	1.00000
<i>Lathyrus pratensis</i> vs <i>Medicago sativa</i>	$z = -4.655$	<b>&lt; 0.001</b>
<i>Lathyrus pratensis</i> vs <i>Ononis spinosa</i>	$z = -4.043$	<b>0.00120</b>
<i>Lathyrus pratensis</i> vs <i>Pisum sativum</i>	$z = -4.361$	<b>&lt; 0.001</b>
<i>Lathyrus pratensis</i> vs <i>Trifolium spp.</i>	$z = -3.960$	<b>0.00165</b>
<i>Medicago lupulina</i> vs <i>Medicago sativa</i>	$z = -4.921$	<b>&lt; 0.001</b>
<i>Medicago lupulina</i> vs <i>Ononis spinosa</i>	$z = -4.404$	<b>&lt; 0.001</b>
<i>Medicago lupulina</i> vs <i>Pisum sativum</i>	$z = -4.538$	<b>&lt; 0.001</b>
<i>Medicago lupulina</i> vs <i>Trifolium spp.</i>	$z = -3.087$	<b>0.03789</b>
<i>Medicago sativa</i> vs <i>Ononis spinosa</i>	$z = 1.700$	0.66743
<i>Medicago sativa</i> vs <i>Pisum sativum</i>	$z = -0.535$	0.99940
<i>Medicago sativa</i> vs <i>Trifolium spp.</i>	$z = -0.988$	0.97356
<i>Ononis spinosa</i> vs <i>Pisum sativum</i>	$z = -1.919$	0.51651
<i>Ononis spinosa</i> vs <i>Trifolium spp.</i>	$z = -2.036$	0.43541
<i>Pisum sativum</i> vs <i>Trifolium spp.</i>	$z = -0.521$	0.99950
Survival	Test Statistic	p
<i>Lotus corniculatus</i> vs <i>Lotus pedunculatus</i>	$z = 0.262$	1.00000
<i>Lotus corniculatus</i> vs <i>Lathyrus pratensis</i>	$z = 1.502$	0.79900
<i>Lotus corniculatus</i> vs <i>Medicago lupulina</i>	$z = 1.330$	0.88260
<i>Lotus corniculatus</i> vs <i>Medicago sativa</i>	$z = -1.191$	0.93139
<i>Lotus corniculatus</i> vs <i>Ononis spinosa</i>	$z = 0.223$	1.00000
<i>Lotus corniculatus</i> vs <i>Pisum sativum</i>	$z = -1.813$	0.60083
<i>Lotus corniculatus</i> vs <i>Trifolium spp.</i>	$z = 0.659$	0.99783
<i>Lotus pedunculatus</i> vs <i>Lathyrus pratensis</i>	$z = 1.414$	0.84457
<i>Lotus pedunculatus</i> vs <i>Medicago lupulina</i>	$z = 1.247$	0.91352
<i>Lotus pedunculatus</i> vs <i>Medicago sativa</i>	$z = -1.716$	0.66633
<i>Lotus pedunculatus</i> vs <i>Ononis spinosa</i>	$z = -0.067$	1.00000
<i>Lotus pedunculatus</i> vs <i>Pisum sativum</i>	$z = -2.457$	0.20593
<i>Lotus pedunculatus</i> vs <i>Trifolium spp.</i>	$z = 0.459$	0.99980
<i>Lathyrus pratensis</i> vs <i>Medicago lupulina</i>	$z = -0.554$	0.99930
<i>Lathyrus pratensis</i> vs <i>Medicago sativa</i>	$z = -2.683$	0.12220
<i>Lathyrus pratensis</i> vs <i>Ononis spinosa</i>	$z = -1.527$	0.78465
<i>Lathyrus pratensis</i> vs <i>Pisum sativum</i>	$z = -3.215$	<b>0.02729</b>
<i>Lathyrus pratensis</i> vs <i>Trifolium spp.</i>	$z = -1.097$	0.95533
<i>Medicago lupulina</i> vs <i>Medicago sativa</i>	$z = -3.113$	<b>0.03691</b>
<i>Medicago lupulina</i> vs <i>Ononis spinosa</i>	$z = -1.446$	0.82921
<i>Medicago lupulina</i> vs <i>Pisum sativum</i>	$z = -3.908$	<b>0.00207</b>
<i>Medicago lupulina</i> vs <i>Trifolium spp.</i>	$z = -1.425$	0.83954

<i>Medicago sativa</i> vs <i>Ononis spinosa</i>	z = 1.819	0.59546
<i>Medicago sativa</i> vs <i>Pisum sativum</i>	z = -0.761	0.99463
<i>Medicago sativa</i> vs <i>Trifolium spp.</i>	z = 2.238	0.31852
<i>Ononis spinose</i> vs <i>Pisum sativum</i>	z = -2.638	0.13613
<i>Ononis spinose</i> vs <i>Trifolium spp.</i>	z = 0.573	0.99912
<i>Pisum sativum</i> vs <i>Trifolium spp.</i>	z = 3.003	0.05171

<b>Fungal infection with <i>Pandora</i> strain ARSEF6631</b>	<b>Test Statistic</b>	<b>p</b>
Genotype 126 vs Genotype 200	z = 0.043	1.00000
Genotype 126 vs Genotype 619	z = 0.042	1.00000
Genotype 126 vs Genotype 74	z = 0.042	1.00000
Genotype 126 vs Genotype 652	z = 0.049	1.00000
Genotype 126 vs Genotype 305	z = 0.048	1.00000
Genotype 200 vs Genotype 619	z = -0.381	0.99865
Genotype 200 vs Genotype 74	z = -0.048	1.00000
Genotype 200 vs Genotype 652	z = 5.320	<b>&lt; 0.001</b>
Genotype 200 vs Genotype 305	z = 4.406	<b>&lt; 0.001</b>
Genotype 619 vs Genotype 74	z = 0.381	0.99865
Genotype 619 vs Genotype 652	z = 6.434	<b>&lt; 0.001</b>
Genotype 619 vs Genotype 305	z = 5.479	<b>&lt; 0.001</b>
Genotype 74 vs Genotype 652	z = 5.237	<b>&lt; 0.001</b>
Genotype 74 vs Genotype 305	z = 6.225	<b>&lt; 0.001</b>
Genotype 652 vs Genotype 305	z = 1.384	0.69533