Open access • Posted Content • DOI:10.1101/2020.12.17.420083

Intraspecific variation in immune gene expression and heritable symbiont density — Source link

Holly L. Nichols, Elliott B. Goldstein, Omid Saleh Ziabari, Benjamin J. Parker

Institutions: University of Tennessee, University of Rochester

Published on: 18 Dec 2020 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Acyrthosiphon pisum and Host (biology)

Related papers:

- Intraspecific variation in symbiont density in an insect-microbe symbiosis
- Genotype specificity among hosts, pathogens, and beneficial microbes influences the strength of symbiont-mediated protection.
- Aphid Heritable Symbiont Exploits Defensive Mutualism.
- Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts.
- Diversity in symbiont consortia in the pea aphid complex is associated with large phenotypic variation in the insect host

Share this paper:

View more about this paper here: https://typeset.io/papers/intraspecific-variation-in-immune-gene-expression-and-212v4fyng2

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 # Intraspecific variation in immune gene expression and heritable symbiont density
2
3 Holly L. Nichols,[1] Elliott B. Goldstein,[1] Omid Saleh Ziabari,[2] Benjamin J. Parker [1,*]
4
5 [1] Department of Microbiology, University of Tennessee, Knoxville, TN, 37996, USA
6
7 [2] Department of Biology, University of Rochester, Rochester, NY 14627, USA
8
9 ORCIDs: HLN: 0000-0003-2158-1767, EBG: 0000-0001-7651-0486, OSZ: 0000-0003-1949-9035, BJP:
10 0000-0002-0679-4732
11
12 *author for correspondence: 1311 Cumberland Avenue, Knoxville, TN, USA 37996, bjp@utk.edu, +1 865
13 974 9818
14
15 Classification: Biology > Evolution
16
17 Keywords: Symbiosis, Genetic variation, Gene expression, Host-microbe interactions, Immunity, Pea
18 aphid

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

19 ## ABSTRACT

20

21 Host genetic variation plays an important role in the structure and function of heritable microbial
22 communities. Recent studies have demonstrated that insects use immune mechanisms to regulate
23 heritable symbionts. Here we test the hypothesis that variation in symbiont density within hosts is linked
24 to intraspecific differences in the immune response to harboring symbionts. We show that pea aphids
25 (*Acyrthosiphon pisum*) harboring the bacterial endosymbiont *Regiella insecticola* (but not all other
26 species of symbionts) suppress expression of key immune genes. We then functionally link immune
27 suppression with symbiont density using RNAi. The pea aphid species complex is comprised of multiple
28 reproductively-isolated host plant-adapted populations. These 'biotypes' have distinct patterns of
29 heritable symbiont infections: for example, aphids from the *Trifolium* biotype are strongly associated with
30 *Regiella*. Using RNAseq, we compare patterns of gene expression in response to *Regiella* in aphid
31 genotypes from multiple biotypes, and we show that *Trifolium* aphids experience no immune gene
32 suppression from *Regiella* and host symbionts at lower densities. We then generated F1 hybrids
33 between two biotypes and found that symbiont density and immune suppression are both intermediate in
34 hybrids. We suggest that in this system, *Regiella* symbionts are suppressing aphid immune mechanisms
35 to increase their density, but that some hosts have adapted to prevent immune suppression in order to
36 control symbiont numbers. The specific immune mechanisms suppressed by *Regiella* have been
37 previously demonstrated to combat pathogens in aphids, and thus this work highlights the immune
38 system's complex dual role in interacting with both beneficial and harmful microbes.

39

40

41 ## AUTHOR SUMMARY

42

43 Heritable microbes are found in most insects including agriculturally and medically relevant pests.
44 Explaining the variation in the distribution and abundance of symbionts in natural populations is critical to
45 understanding these interactions. This work contributes to our mechanistic understanding of an important
46 model of host-microbe symbiosis and suggests more broadly that variation in insect immune responses
47 plays a role in intraspecific variation in host-symbiont interactions. Our work also suggests that
48 antagonistic coevolution can play a role in host-microbe interactions even when microbes are transmitted
49 vertically and provide a clear benefit to their hosts.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

50 INTRODUCTION:

51

52 Most insects harbor heritable microbes that have important effects on host fitness (1-3). A key aspect of
53 these symbioses is variation. Across species, host taxonomy has been shown to play a role in structuring
54 heritable microbial communities (4-6). Within species, microbes referred to as 'facultative' symbionts are
55 not found in all individuals, and symbiont frequencies are subject to selection on the relative costs and
56 benefits of harboring microbes (7, 8). In addition to microbiome composition, hosts vary in other aspects
57 of symbioses like the density of microbial infections (9). For example, two closely-related species of
58 *Nasonia* wasps vary in the density at which they harbor *Wolbachia* bacteria, and this variation is due to a
59 single gene that somehow suppresses maternal transmission of bacteria (10). Except for a few
60 examples, little is known about the mechanisms that underlie variation in heritable symbioses or the
61 evolutionary genetics of these interactions (11).

62

63 Invertebrate immune systems have been shown to play a direct role in mediating interactions with
64 heritable microbes. In grain weevils, for example, an antimicrobial peptide acts to confine mutualistic
65 symbionts to specialized cells called bacteriocytes (12), and silencing expression of immune pathways
66 allows symbionts to escape bacteriocytes (13). Other studies have found more complex interactions
67 between pathogens, the immune system, and vertically-transmitted symbionts: In *Drosophila*
68 *melanogaster*, for example, activation of the Toll and IMD immune pathways results in an increase in
69 density of *Spiroplasma* symbionts (14) (and see similar examples in mosquitos (15) and tsetse flies (16)),
70 suggesting in some systems the immune system can promote beneficial microbes by inhibiting other
71 microbes.

72

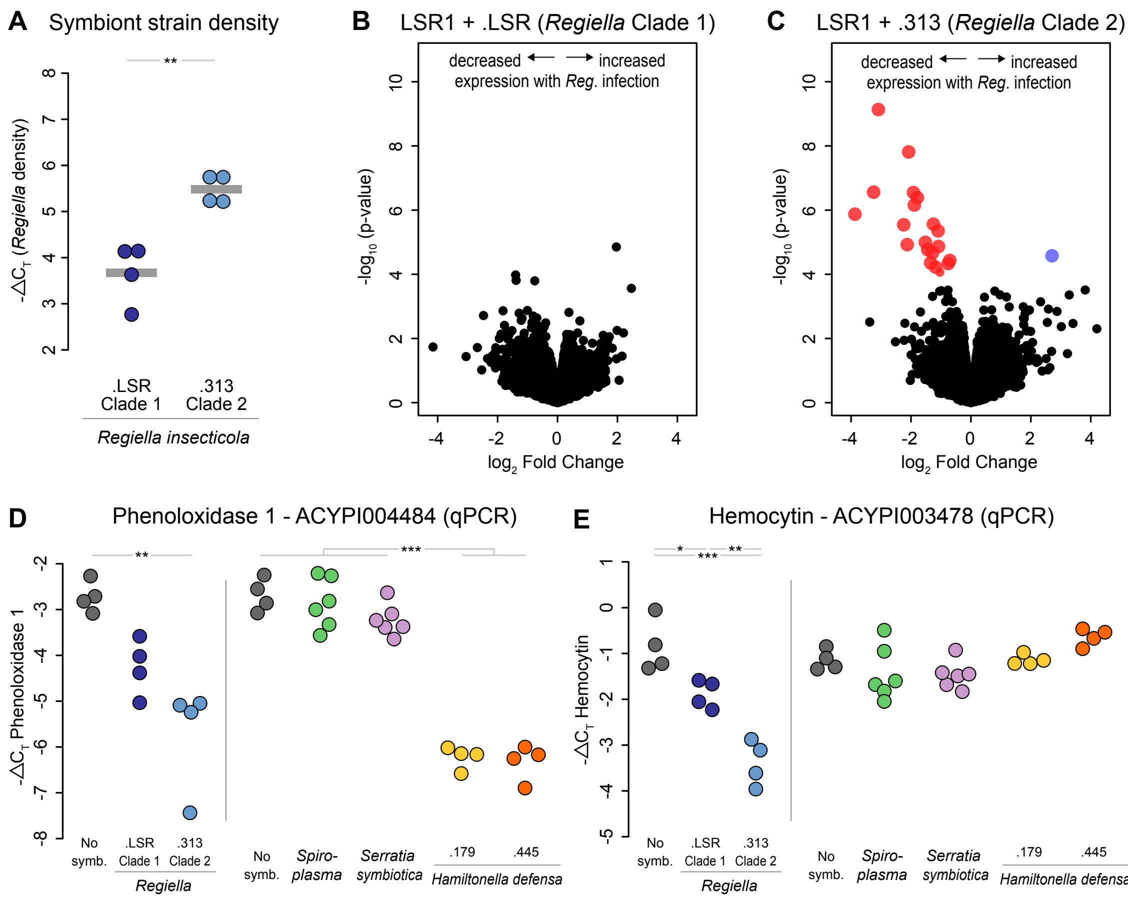
73 Immune genes are among the fastest evolving genes in eukaryotic genomes (17-19), and natural
74 populations harbor genetic variation in immune mechanisms (20). Given the importance of the immune
75 system in regulating insect symbioses, it seems likely that variation in immune mechanisms contributes
76 to variation in symbiont density among hosts. This hypothesis is complicated, however, because our
77 models for the maintenance of genetic variation in immune systems are based on antagonistic
78 coevolution between hosts and pathogenic microbes (21, 22). Heritable symbiont infections are thought
79 to spread through host populations because the fitness interests of host and microbe are generally
80 aligned, and many symbionts have been shown to benefit their hosts for example by providing protection
81 from pathogens (reviewed in (23)). But symbionts can impose costs on their hosts (e.g. (24)), and
82 selection may favor the loss of symbionts in certain contexts (25). In addition, within-host selection might
83 lead to a separation of the fitness interests of hosts and microbes. For example, a mutation in a symbiont
84 genome that increases symbiont density might increase the likelihood of symbiont transmission but come
85 at the expense of host fitness. Hosts, in turn, could evolve greater control over symbiont numbers in an

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

86 ongoing arms-race for control over a symbiosis. It is unclear, however, whether the 'arms-race' dynamics
87 underlying host-pathogen coevolution also govern the evolutionary interactions between immune
88 systems and beneficial microbes.
89
90 The pea aphid (*Acyrthosiphon pisum*) is an important insect-symbiont model system (26). The pea aphid
91 species complex is composed of multiple reproductively-isolated populations adapted to live on different
92 host plants within the family Fabaceae. These 'biotypes' are genetically differentiated and are estimated
93 to have radiated onto different host plants ~500,000 years ago (27) (but see (28)). In addition to obligate
94 intracellular bacteria called *Buchnera aphidicola*, aphids can harbor several species of facultative
95 symbionts. Multiple studies have found that facultative symbionts are non-randomly distributed across
96 aphid biotypes (29-32). For example, *Regiella insecticola* (which confers protection against fungal
97 pathogens to its host (33-35)) is strongly associated with aphids from the *Trifolium spp.* (clover) biotype
98 across continents. A number of studies have explored whether the strong association between *Trifolium*
99 biotype aphids and *Regiella* is due to improved host plant use with mixed results (36-39). Alternatively,
100 this association could be due to the risk of exposure to fungal pathogens (though see (40)), to historical
101 contingency (though see (41)), or to host and/or symbiont genetic mechanisms. This system therefore
102 provides a useful natural laboratory to study host-microbe adaptation across multiple environments within
103 a single species.
104
105 In this study, we show that immune mechanisms play a role in intraspecific variation in the density of a
106 heritable bacterial symbiont. We first demonstrate that pea aphids that harbor *Regiella* (but not all other
107 species of symbionts) sharply downregulate key innate immune genes, and that experimental
108 suppression of the immune gene phenoloxidase via RNAi increases symbiont density. We then measure
109 immune suppression across aphids from multiple biotypes and find that aphids from *Trifolium spp.* do not
110 experience immune suppression and harbor symbionts at relatively low density. By performing an F1
111 cross between genotypes from two biotypes we find that hybrid aphids show intermediate symbiont
112 densities and immune suppression, shedding light on the role of host genetic variation and the genomic
113 architecture of this variation. We discuss these findings in light of the biology of this system and suggest
114 that antagonistic coevolution between 'beneficial' microbes and their hosts can shape host-symbiont
115 associations.
116
117
118 RESULTS:
119
120 **Hosting some symbiont species leads to suppressed host immune gene expression.** Aphid lines
121 reproduce parthenogenetically under summer conditions, and facultative bacteria can be introduced into

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

122 or removed from host lines. We established lines that have the same aphid host genotype (LSR1,
123 collected from *Medicago sativa* (42)) with two different strains of *Regiella*: one from each of the two main
124 phylogenetic clades of *Regiella* found in natural populations of pea aphids (43). *Regiella* strain .LSR was
125 originally collected with the LSR1 aphid genotype and is a representative of *Regiella* 'Clade 1,' and
126 *Regiella* strain .313 is from 'Clade 2' (44). We maintained aphids in the lab for four generations after
127 symbiont establishment and then measured *Regiella* densities using quantitative PCR (qPCR). We have
128 found previously that 'Clade 2' *Regiella* establish at higher densities in hosts than do 'Clade 1' strains,
129 independent of host genotype (45). Consistent with this previous work, *Regiella* strain .313 established in
130 aphid genotype LSR1 at a significantly higher density (3.5X) than *Regiella* strain .LSR ($t = 5.1$, $p = 0.006$,
131 Figure 1A).
132



133
134

135 **Figure 1: Effects of hosting secondary symbionts on aphid gene expression. A:** *Regiella* density of
136 lines harboring clade 1 and clade 2 *Regiella*. The y-axis shows the -ΔC_T values which can be interpreted
137 on a $\log_2$ scale. **B&C:** Volcano plots of RNAseq data comparing expression of each expressed gene in
138 the aphid genome, represented by a point in each figure, between aphids with and without *Regiella*. A
139 and B show this analysis for aphids with a *Regiella* strain from clade 1 (strain .LSR) and clade 2 (strain
140 .313), respectively. The x-axes show the $\log_2$ fold change for each gene, and the y-axis shows the
141 significance of expression. Blue and red dots represent genes that were significantly up- or down-
142 regulated, respectively, at an FDR < 0.05. **D:** qPCR analysis of gene expression of Phenoloxidase 1
143 (*PO1*; ACYPI004484) in response to different species of facultative symbiont. grey dots represent aphids

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

144 without symbionts, and colored points show those with symbiont infections 4 generations after symbiont
145 establishment. The different symbiont species and strains are shown along the bottom of the figure. The
146 y-axis shows the -ΔC_T values of expression, which can be interpreted on a log_2 scale. Statistical
147 significance among species or strains is shown along the top of the figure; the two experiments,
148 separated by a grey line, were analyzed separately. **E:** Same as panel D but for the gene hemocytin
149 (ACYPI003478).
150

151 We then used RNAseq to measure how harboring *Regiella* influences aphid gene expression by
152 comparing these lines with symbiont-free aphids that were sham-injected. We sequenced cDNA made
153 from mRNA for 4 biological replicates of each line (where a biological replicate represents an
154 independent aphid line injected with symbionts or sham-injected). Overall, harboring *Regiella* strain .LSR
155 did not significantly alter expression of any genes in the aphid genome (FDR < 0.05; Figure 1B), while
156 strain .313 led to significantly decreased expression of 19 genes and upregulation of 1 (FDR < 0.05;
157 Figure 1C).

158 The 20 genes with altered expression included key innate immune system genes (Table S1). In
159 particular, the two copies of phenoloxidase in the pea aphid genome (referred to here as *PO1* and *PO2*)
160 were downregulated in the presence of *Regiella*. Also downregulated was a gene called hemocytin,
161 which encodes a protein released by immune cells that plays a role in immune cell aggregation (46, 47).
162 Other differentially expressed genes included a toll-like receptor and a putative lipopolysaccharide
163 recognition protein (Table S1).

164 We used qPCR to directly compare expression of two immune genes between lines harboring the
165 two *Regiella* strains and to confirm our RNAseq results. *PO1* was significantly downregulated in lines
166 harboring *Regiella* strain .313 (Figures 1D & 1E, left panels). Hemocytin was significantly downregulated
167 in lines harboring either symbiont, and the magnitude of this change was significantly stronger for aphids
168 harboring strain .313 than those with strain .LSR (Figures 1D & 1E, left panels).

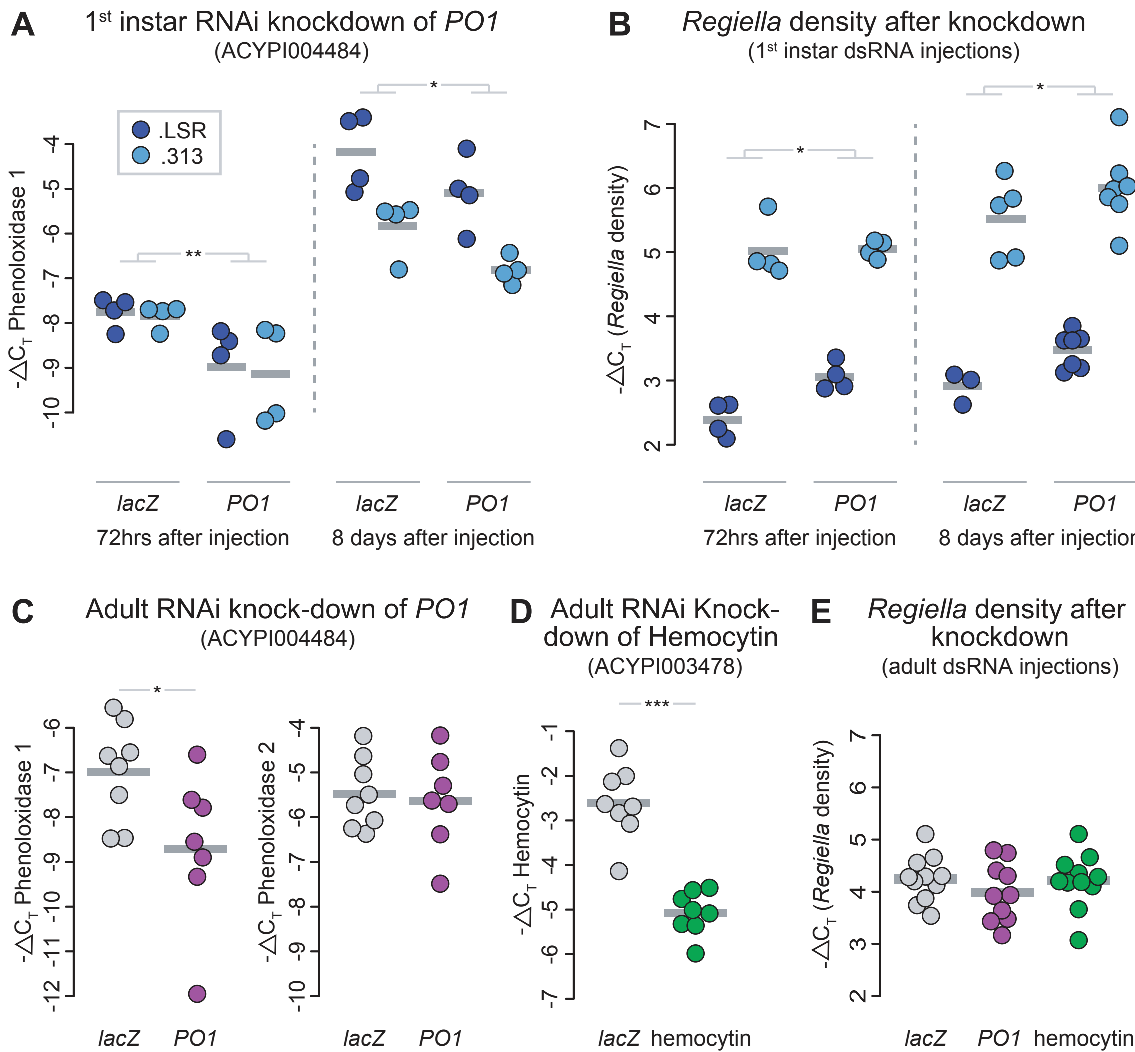
169 Next, we established aphid lines that harbored several additional species of aphid facultative
170 endosymbionts as above, and we looked for changes in *PO1* and hemocytin expression. *Spiroplasma*
171 and *Serratia symbiotica* did not alter expression of either gene, but two strains of *Hamiltonella defensa*
172 significantly downregulated *PO1* expression but not hemocytin (post-hoc tests, Table S2; Figures 1D &
173 1E, right panels). Like *Regiella*, the specific strain of *Spiroplasma* (.161) we used was found in previous
174 work to protect against fungal pathogens (48), but *Serratia* and *Hamiltonella* have been found not to
175 influence fungal resistance (35, 49, 50). These results therefore suggest that the changes we identify in
176 immune gene expression do not reflect the mechanism by which *Regiella* confers protection to aphids
177 against fungal pathogens, which is currently unknown.
178

179 **Immune gene expression influences symbiont density during development.** We studied the
180 function of immune gene expression on *Regiella* densities using RNA interference (RNAi) (51). We
181 knocked-down expression of *PO1* early in development and measured the effects on symbiont density.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

182 We synthesized dsRNA for *PO1* and a *lacZ* control, and we injected ~100 ng of dsRNA in salt buffer into
183 1-day-old aphids. Aphids harbored either *Regiella* strain .LSR or .313. We then sampled aphids at two
184 timepoints: at 72hrs after injection and after aphids had become adults (8 days after injection).
185 Injection with *PO1* dsRNA reduced *PO1* expression, on average, by ~60% at 72hrs (2-way ANOVA;
186 Treatment: F = 9.8, p = 0.009; Figure 2A). At this early timepoint, we found no significant difference
187 between aphids harboring *Regiella* strain .LSR vs strain .313 in *PO1* expression (Strain: F = 0.10, p =
188 0.75; Figure 2A). By the time aphids became adults (8 days after injection), *PO1* expression in *PO1*
189 dsRNA injected aphids was still reduced by ~60% compared with controls (Treatment: F = 3.6, p = 0.02;
190 Figure 2A). By this later timepoint, aphids harboring the two symbiont strains had diverged in expression
191 as found above (Strain: F = 11.5, p = 0.004; Figure 2A). Note that we dissected out and removed
192 developing embryos only from the adult samples before nucleic acid extraction, so we do not directly
193 compare expression in the 72hrs vs. adult samples, but qualitatively *PO1* expression increased during
194 development (Figure 2A).
195 *PO1* knockdown led to a 59% and 2% increase in *Regiella* strain .LSR and .313 density at 72hrs,
196 respectively (2-way ANOVA; Treatment: F = 4.7, p = 0.05; Figure 2B). *Regiella* density in aphids
197 harboring strain .LSR vs .313 also differed significantly (Strain: F = 245, p < 0.0001; Figure 2B),
198 suggesting that strain-level differences in symbiont density are present even at this early developmental
199 timepoint. The increase in *Regiella* density due to *PO1* knock-down persisted to adulthood (Treatment: F
200 = 5.7, p = 0.03, Figure 2B), with *PO1* dsRNA injection increasing *Regiella* density by 48% and 40% in
201 aphids harboring *Regiella* strains .LSR and .313, respectively. As we found above, the density of strain
202 .313 was higher than strain .LSR in adult aphids (Strain: F = 145, p < 0.0001; Figure 2B). Together,
203 these results show that knockdown of *PO1* increases *Regiella* density over development.
204
205 **Symbiont density is not impacted by immune gene knock-down later in development.** We
206 performed a similar experiment studying the effect of RNAi on symbiont density, but injected dsRNA into
207 adult aphids rather than 1st instars. We injected ~1μg of dsRNA synthesized from *PO1* or hemocytin into
208 adult (9 day old) aphid genotype LSR1 aphids infected with *Regiella* strain .LSR. We measured gene
209 expression and *Regiella* density at 72hrs after injection. This led to a ~69% and ~82% reduction in
210 expression of *PO1* and hemocytin, respectively (t-tests; *PO1*: t = -2.3, p = 0.05; hemocytin: t = -7.3, p <
211 0.0001; Figures 1C and 1D). We note that injection with dsRNA from *PO1* had no effect on expression of
212 the other copy of phenoloxidase in the aphid genome (*PO2*: t = -0.91, p = 0.38; Figure 1C),
213 demonstrating that our phenoloxidase RNAi assay is specific to *PO1* as designed. Knockdowns had no
214 effect on *Regiella* density in aphids injected as adults (ANOVA; Treatment: F = 0.80, p = 0.46; Figure
215 1E).
216

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



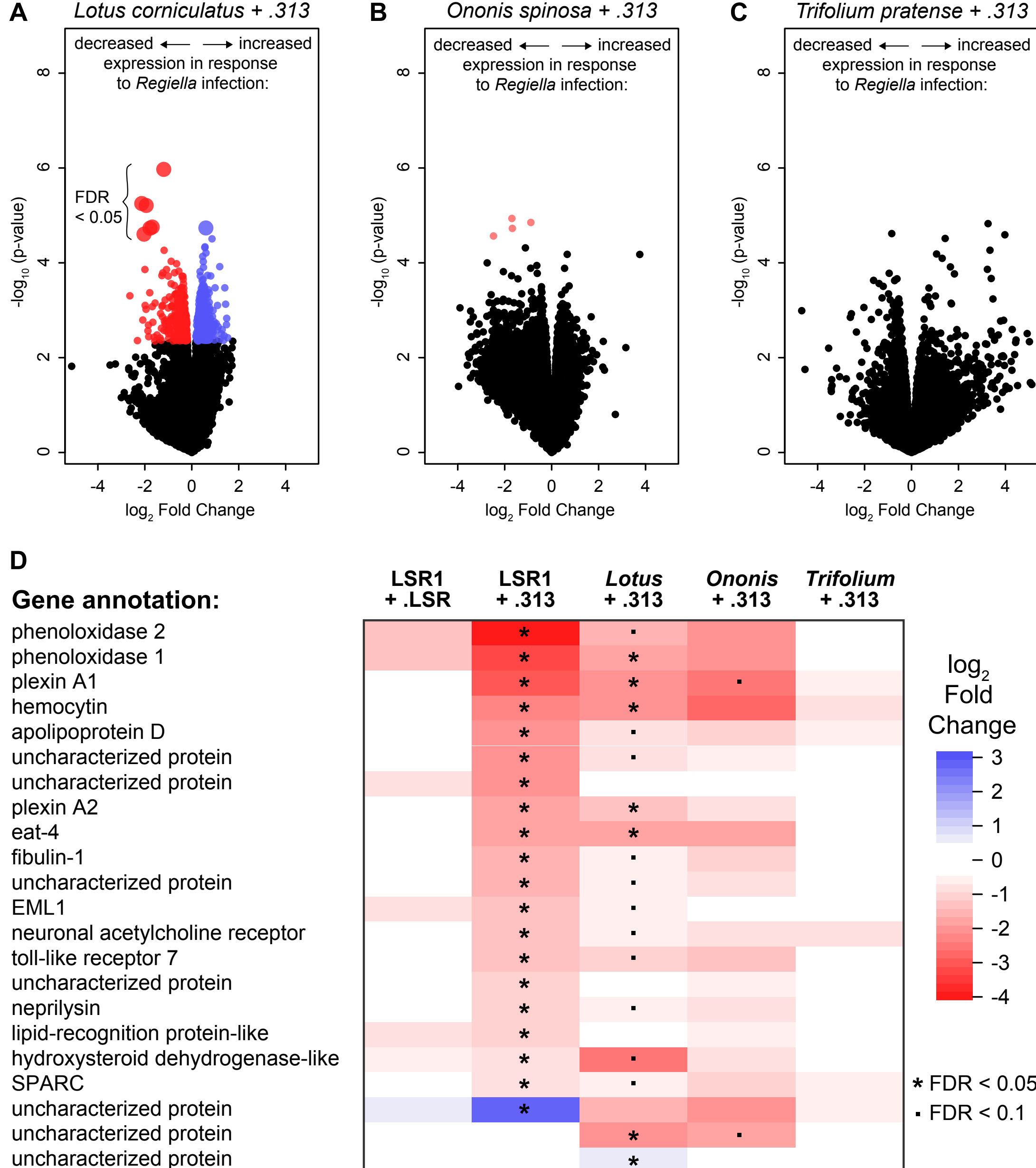
217

218 **Figure 2: RNAi knockdowns and *Regiella* density. A:** Validation of the RNAi knockdown for *PO1* in 1st
219 instar aphids. The y-axis shows -ΔC_T values of expression, which can be interpreted on a $\log_2$ scale.
220 Collection time-points (72hrs or adults) and treatment are shown along the bottom of the figure; the two
221 *Regiella* strains are represented by different colors as shown in the key. Bars show the mean of each
222 treatment group, with biological replicates shown as points. **B:** *Regiella* density after knockdown of
223 aphids injected as 1st instars. The y-axis shows -ΔC_T values of symbiont density, which can be
224 interpreted on a $\log_2$ scale. Time-points, treatment, and *Regiella* strain are indicated as in A. **C:**
225 Validation of the RNAi knockdown for *PO1* in adult aphids. Y-axes are as above, treatment is shown
226 along the bottom of the figures, with lacZ and *PO1* dsRNA-injected aphids represented by grey and
227 purple dots, respectively. Grey gars shown mean expression for a treatment. The left panel shows
228 expression of *PO1*, and the right panel shows expression of the other copy of phenoloxidase in the aphid
229 genotype (*PO2*). **D:** Knock-down validation of hemocytin. **E:** *Regiella* density after knockdown of aphids
230 injected as adults. Grey, purple, and green points represent different treatments (*lacZ*, *PO1*, and
231 hemocytin dsRNA injections, respectively). The y-axis shows *Regiella* density measured by -ΔC_T values
232 as above.
233

234 **Immune gene suppression differs across aphid biotypes.** As noted above the pea aphid species
235 complex includes multiple reproductively isolated 'biotypes' associated with different species of host
236 plants. We repeated the RNAseq experiment to study the effects of *Regiella* on gene expression across

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

237 multiple aphid biotypes. We used a genotype from the *Lotus corniculatus* biotype (663), a genotype from
238 *Ononis spinosa* (C133), and a genotype from *Trifolium pratense* (C317). For each aphid genotype, we
239 compared replicate lines that had each been infected with an independent *Regiella* Clade 2 (.313)
240 infection or had been sham injected as above after 4 generations (after verifying that the symbiont
241 infection had not been lost using PCR).
242



243

244 **Figure 3: Gene expression across aphid biotypes. A-C:** Volcano plots of expression data comparing
245 control vs. *Regiella*-infected aphids. Each expressed gene in the aphid genome is represented by a
246 point. The X-axes show the $\log_2$ fold change of each gene, with points to the right side of each plot
247 indicating increased expression in the presence of symbionts, and points to the left showing decreased
248 expression. The y-axes show the $-\log_{10}$ of the p-values indicating statistical significance of each gene's
249 expression change. Colored points are those where the expression change was found to be statistically
250 significant at an FDR < 0.1, with expression at < 0.05 shown by larger points as indicated. Panels A, B,

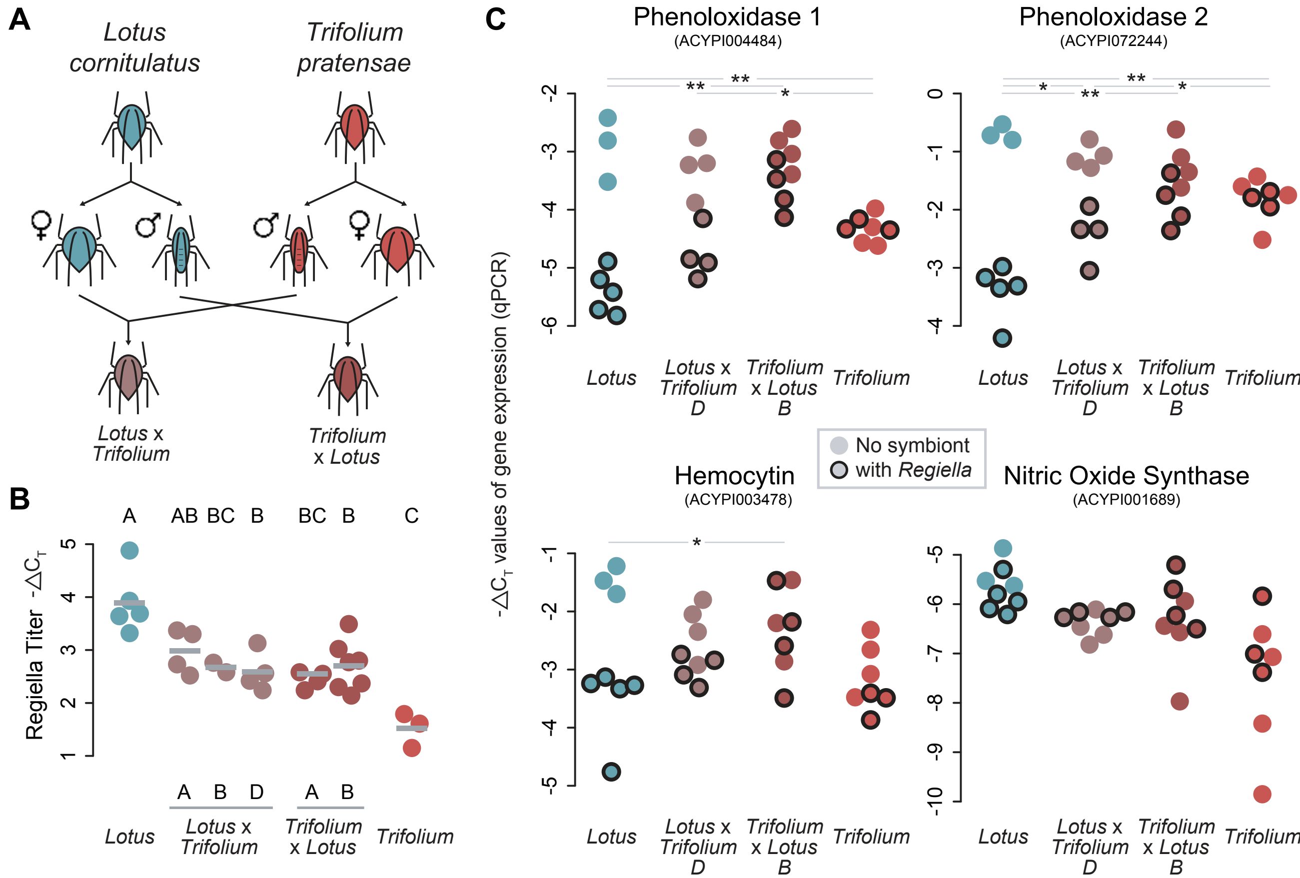
bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

251 and C show plots for the *Lotus, Ononis,* and *Trifolium* genotypes respectively, as shown along the top of
252 the figures. **D:** A heat-map comparing gene expression in response to *Regiella* strain .313 infection
253 across host genotypes. The 22 differentially expressed genes identified in the LSR1 transcriptome,
254 above, are listed to the left of the figure. Colors represent the $\log_2$ fold change of these genes in
255 response to *Regiella* as indicated in the key to the right of the figure (with red panels representing a
256 decrease in expression, and blue indicating an increase in expression). The five transcriptomes
257 generated in this study are shown in each column, as indicated at the top of the figure. Statistical
258 significance of each gene is indicated by * for FDR < 0.05, and . for FDR < 0.1.
259

260 In aphids from the *Lotus corniculatus* genotype, harboring *Regiella* strain .313 had a significant effect
261 on the expression of aphid genes (7 genes differentially expressed at an FDR < 0.05, and 612 genes at
262 an FDR < 0.1; Figure 3A). Of these 7 genes, five were also downregulated by the experiment described
263 above using genotype LSR1, including *PO1* and hemocytin (Figure 3D). There is therefore some degree
264 of conservation in the response to *Regiella* across genetically distinct aphid lines (Figure 3D). In contrast,
265 zero genes were differentially expressed in response to *Regiella* in the *Ononis* biotype line at an FDR of
266 < 0.05, and only 4 genes were differentially expressed at < 0.1 (Figure 3B). Similarly, zero genes differed
267 in expression in response to *Regiella* from the *Trifolium* line (Figure 3C) at either significance level.
268

269 **F1 genotypes have an intermediate phenotype in symbiont density and immune suppression.** We
270 performed an F1 cross (Figure 4A) between two of the biotype lines in order to better understand the role
271 of host genetic variation in *Regiella* density and immune gene expression. We crossed the *Lotus* (663)
272 and *Trifolium* (C317) lines (52), and generated multiple replicate infections with *Regiella* strain .313 in
273 each line as above. After four generations, we measured *Regiella* density using qPCR. *Regiella* density
274 differed between the parental lines (Figure 4B; post-hoc tests Table S3), with the *Lotus* line harboring a
275 significantly higher density of *Regiella* strain .313 (5.2X higher) than the *Trifolium* line. Further, the F1
276 lines harbored *Regiella* at densities intermediate to the parental lines (Figure 4B; post-hoc tests, Table
277 S3).
278 We then sampled aphids from this same generation to compare changes in immune gene expression
279 due to *Regiella* in parental and F1 lines. We selected two F1 lines for this assay with each aphid
280 genotype serving as the maternal line. We used qPCR to measure expression of both copies of
281 phenoloxidase, hemocytin, and also nitric oxide synthase (*NOS*; an important innate immune mechanism
282 that was not significantly differentially expressed in any of our RNAseq studies). Confirming our RNAseq
283 findings, harboring *Regiella* led to a decrease in expression of *PO1*, *PO2,* and Hemocytin in the *Lotus*
284 genotype, but *Regiella* did not change gene expression in the *Trifolium* genotype (Figure 3C). Further,
285 the F1 lines showed significant differences in gene expression that were intermediate to the two parental
286 lines: three immune genes were suppressed in response to *Regiella,* but to a significantly lesser extent
287 than in the lotus genotype (Figure 3C; post-hoc tests, Table S4).
288

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



289

290

291 **Figure 4: Immune suppression in F1 hybrid lines. A:** Diagram of the crossing scheme. **B:** *Regiella*
292 density in parental and F1 lines. The y-axis shows the -ΔC_T values reflecting *Regiella* density, which can
293 be interpreted on a $\log_2$ scale. The genotypes are shown along the bottom of the figure. Each biological
294 replicate, representing an independently injected aphid + *Regiella* line, is shown by a colored point, with
295 the means for each genotype shown in grey bars. Significance groups are shown along the top of the
296 figure at $p < 0.05$. **C:** Immune gene expression in lines with and without *Regiella*. The y-axis of each plot
297 shows -ΔC_T values (qPCR output) of gene expression. Aphid genotype is shown along the bottom of
298 each figure. Each biological replicate (an independently injected line) is shown by a colored point, with
299 lines harboring *Regiella* indicated by a dark black outline as indicated in the legend. The results of a
300 post-hoc analysis of the interaction term between *Regiella* presence/absence and host genotype is
301 indicated at the top of the figure, with an * or ** indicating that two genotypes differ in the extent to which
302 *Regiella* altered gene expression at $p < 0.05$ and $p < 0.01$, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

303 DISCUSSION
304
305 We show that some aphids harboring the facultative bacterial symbiont *Regiella insecticola* suppress
306 expression of key immune genes, and we link immune suppression with increases in *Regiella* density.
307 We further find that this mechanism is influenced by host genetic factors: some genotypes harbor
308 *Regiella* at lower densities and do not experience altered gene expression when infected, where other
309 aphid genotypes experience suppressed immune gene expression and have higher-density *Regiella*
310 infections. This study shows that intraspecific variation in the immune system affects a heritable
311 symbiosis.
312
313 We found no significantly differentially expressed genes in the transcriptome of aphids harboring a Clade
314 1 (.LSR) *Regiella* strain. Using qPCR we confirmed that the immune suppression we uncover occurs
315 more strongly in aphids harboring the Clade 2 *Regiella* than the Clade 1 strain. An important question is
316 whether symbionts are suppressing host immune mechanisms in order to reach higher densities in hosts,
317 or if hosts are modifying immune mechanisms in order to accommodate symbionts. Fitness costs to
318 aphids of harboring symbionts (including *Regiella*) have been measured in the laboratory and field (34,
319 53), and we have found previously that higher density Clade 2 *Regiella* strains impose stronger survival
320 costs on hosts than the lower density Clade 1 strains (45). In addition, the two *Regiella* clades confer
321 protection against specific genotypes of the fungal pathogen *Pandora neoaphidis* (44), and therefore
322 *Regiella* density will be positively correlated with symbiont-mediated protection for some fungal
323 genotypes and negatively correlated for others. Together these results suggest that immune suppression
324 is not an adaptation on the part of the host in order to accommodate symbionts, but instead some
325 *Regiella* strains suppress immune mechanisms in order to establish at higher densities in hosts.
326 Establishing at a higher density could benefit *Regiella* through improved competitive outcomes with other
327 strains and species of symbionts, or through increased horizontal transmission which occurs on
328 evolutionary (43) and even ecological (54) timescales.
329
330 Aphid biotypes harbor facultative symbionts at different frequencies that are to some degree conserved
331 across continents (29-32), and decades of research have gone into explaining these patterns in order to
332 better understand the ecological and evolutionary forces shaping beneficial host-microbe interactions.
333 One particularly clear association is between aphids from the *Trifolium* biotype and *Regiella*—and
334 specifically *Regiella* from Clade 2 (43). Studies attempting to explain this pattern have considered factors
335 like the potential effects of *Regiella* on host plant use (36-39) and pressure from fungal pathogens on
336 different host plants (40). We suggest that host genetic effects represent an additional factor shaping the
337 aphid-symbiont frequencies in natural populations. Whether a cause or consequence of the strong
338 association between *Regiella* and *Trifolium*, it seems likely that that the immune systems of *Trifolium*

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

339 aphids are better adapted to harboring this symbiont than biotypes that rarely interact with Clade 2
340 *Regiella*. We found no evidence of immune suppression in an aphid genotype from the *Trifolium* biotype:
341 zero genes were differentially expressed in response to harboring *Regiella* strain .313 (which was
342 confirmed using qPCR on three immune genes). Together, these results are consistent with a scenario
343 where Clade 2 *Regiella* have evolved increased within-host density that harms host survival, and aphids
344 from the *Trifolium* biotype have adapted to prevent immune suppression to control symbiont numbers.
345
346 The phenoloxidase enzyme is required for the activation of melanogenesis in invertebrates. Against
347 multicellular parasites, melanin is deposited around a foreign object via immune cells (hemocytes), and
348 the melanin capsule prevents the growth and reproduction of parasites (55). *PO* is also upregulated in
349 response to microbial pathogens in many studies (56, 57) and it is thought that because phenoloxidase
350 is cytotoxic it helps immune cells kill phagocytosed microbes (58). Aphid immune cells express
351 phenoloxidase (59), are known to have phagocytic properties (59, 60), and have been shown using
352 microscopy to contain secondary symbionts including *Regiella* (60). Harboring *Regiella* (and *Hamiltonella*
353 but not other species of symbionts) leads to a sharp decrease in the numbers of circulating immune cells
354 (called granulocytes) (60). One possibility is that *PO* knockdowns via RNAi in our study disrupted the
355 cellular immune responses aphids use to regulate symbionts during development, but the natural
356 mechanisms *Regiella* are using to suppress *PO* and other immune genes are unknown.
357
358 In addition to the effects on symbionts we have uncovered in this study, phenoloxidase has been shown
359 to be an important part of the pea aphid's response to pathogens (e.g. fungal pathogens (61, 62)).
360 Functionally, a recent study found that silencing of *PO1* and *PO2* via RNAi leads to decreased resistance
361 of pea aphids against pathogenic bacteria and a generalist fungal pathogen (63). Together, these results
362 show that the same molecular mechanisms are influencing interactions with both beneficial and
363 pathogenic microbes in this system. The protection against specialist fungal pathogens conferred by
364 *Regiella* might benefit hosts, but immune suppression by some strains of *Regiella* could trade-off with an
365 increased risk of infection with other pathogens. Recent work on animal immune systems has
366 emphasized the role of immune mechanisms in regulating mutualistic interactions between microbes and
367 hosts. In general, how immune systems evolve to manage the complex task of interacting with distinct
368 microbes with different effects on host fitness is an important question.
369
370 Finally, our findings emphasize the importance of host genetic variation in associations with beneficial
371 microbes (64). We found that hybrids between aphid biotypes harbor symbionts at intermediate densities
372 to their parental lines and only partially suppress immune gene expression. The extent to which a host
373 responds to symbiont infection is therefore likely be a quantitative trait, much like resistance against
374 pathogens, that is influenced by variation at multiple to-be-determined loci, and subject to natural

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

375 selection on the relative costs and benefits of symbiosis. This work thus contributes to a growing view of
376 animal microbiomes as complex phenotypes that influence animal fitness, and that are to some extent
377 under host control.
378
379

380 METHODS

381

382 **Pea aphids and symbiont establishment**: Pea aphids reproduce parthenogenetically under certain
383 light and temperature conditions (16L:8D at 20°C), allowing us to rear large numbers of genetically
384 identical and developmentally synchronized individuals for use in experiments. Wild-collected lines were
385 cured of their original symbiont infections using antibiotics (36) and maintained asexually in the lab for
386 several years before use in experiments (Tables S5 and S6 specify collection information for aphid
387 genotypes and symbiont strains).

388 Throughout these experiments we used established protocols to infect aphids with facultative
389 symbionts (65-67). We inject a small volume of hemolymph from an infected donor aphid into a 1st instar
390 recipient using a glass capillary needle. We then rear these aphids to adulthood and then collect an
391 offspring from late in the birth-order to establish the infected line. When this aphid produces offspring, we
392 extract DNA (using a lysis buffer with proteinase K and an ethanol precipitation (68)) and screen the line
393 for symbionts using PCR with symbiont-specific primers (43) (Table S7): (94°C 2 min, 11 cycles of 94°C
394 20s, 56°C (declining 1°C each cycle) 50 s, 72°C 30 s, 25 cycles of 94°C 2 min, 45°C 50 s, 72°C 2 min
395 and a final extension of 72°C 5 min). Each biological replicate in these experiments (an "aphid line")
396 originated from a separate symbiont-injection and screening except where noted.
397

398 **Measurements of symbiont density using qPCR:** We established aphid lines from the LSR1 genotype
399 with two strains of *Regiella*: "Clade 1" *Regiella* (strain .LSR), and "Clade 2" *Regiella* (strain .313). These
400 two strains of *Regiella* each come from one of the two main clades of *Regiella* found among pea aphids
401 as determined using an established protocol for MLST sequence typing (44, 69). We reared lines that
402 had successfully acquired a *Regiella* infection under asexual conditions for four generations, at which
403 point we re-screened lines for *Regiella* infection. We then used qPCR to compare symbiont density
404 between these strains. We removed embryos from groups of 7 adult aphids, and extracted DNA using
405 the Qiagen DNEasy kit under recommended conditions. We used qPCR primers that amplify a
406 conserved region of the *Regiella hrpA* gene (Table S7). Amplification of *g3PDH* was used as an
407 endogenous reference gene that controlled for the relative abundance of host DNA in each sample.
408 Primer concentrations were optimized against a serial dilution of gDNA (400/350nM F/R and 300nM for
409 *g3PDH* and *hrpA*, respectively). We calculated $-\Delta C_T$ values by $-(C_{T_hrpA} - C_{T\ g3PDH})$ and analyzed these

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

410 values with a t-test. Note that this approach reveals the relative density of symbionts relative to host
411 tissue across different samples, but does not measure the absolute abundance of symbionts.
412
413 **Effects of *Regiella* on host gene expression using RNAseq**: We then measured the effects of
414 harboring *Regiella* on gene expression using RNAseq. We used the lines established above with either
415 *Regiella* strain .LSR or .313, and symbiont-free aphids of the same host genotype. For the 'no symbiont'
416 treatment, we sham injected aphids (injected aphids with a small volume of hemolymph (0.25μl) from an
417 uninfected adult donor aphid) and handled aphids in the same way as with symbiont-injected aphids.
418 For transcriptome sequencing, we collected adult, fourth generation aphids on the first day that each
419 line produced offspring and dissected and removed developing embryos (in order to measure gene
420 expression of the mother without including RNA from her embryos). We stored carcasses in TRIzol
421 (Invitrogen) at -80°C. Each sample contained ~14 adult carcasses collected from multiple host plants.
422 We extracted RNA using TRIzol-chloroform and an isopropanol precipitation with an ethanol wash. We
423 digested genomic DNA and cleaned the RNA using the Zymo Clean & Concentrate–5 kit with the DNAse
424 I enzyme. RNA quality control was conducted on a bioanalyzer chip, and 12 sequencing libraries (4
425 biological replicates x 3 treatments) were constructed using the NEBNext Ultra II RNA Library Prep Kit
426 for Illumina (including poly-A selection and 15 rounds of PCR amplification). Libraries were sequenced
427 across one lane of Illumina PE150 sequencing (approximately 20 million reads per library) with a 250-
428 300bp insert per library.
429
430 **RNAseq analysis:** We estimated the average insert size of paired-end libraries using Picard Tools
431 v.2.21.3 in java 1.8.0, and mapped reads to the pea aphid reference genome v.2.1 (42) using tophat
432 v.2.1.1 (70). We counted reads mapped to each annotated gene (using a modified version of pea aphid
433 genome annotation v.2.1 (https://bipaa.genouest.org/sp/acyrthosiphon_pisum/) with several duplicated
434 genes removed from the file) using the count function in htseq v.0.9.1 (71) and the 'union' overlap mode
435 (Table S8). We analyzed read counts using EdgeR v.3.22.3 in R v.3.5.0. Genes with a minimum
436 threshold of aligned reads, determined by the filterByExpr function in edgeR, were retained in the
437 analysis. We fit a quasi-likelihood model to the data using the glmQLFit function, and we tested for
438 statistically significant differential expression of each gene using a quasi-likelihood F-test, interpreting
439 genes with a false discovery rate (FDR) of < 0.05 as differentially expressed in response to *Regiella*
440 infection.
441
442 **Immune gene expression across facultative symbiont species via qPCR:** We used qPCR to verify
443 our RNAseq results, and to explore how differences in gene expression due to *Regiella* in key innate
444 immunity genes varied across facultative symbiont species. We used qPCR primers (Table S7) that
445 amplified 80-120bp of two target genes of interest (*PO1*: ACYPI004484 and hemocytin: ACYPI003478)

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

446 and four endogenous control genes (Glyceraldehyde 3-phosphate dehydrogenase (*g3PDH*):
447 ACYPI009769, *NADH dehydrogenase*: ACYPI009382, *β-tubulin*: ACYPI001007, and *Rpl32*:
448 ACYPI000074). Primer concentrations were optimized against a 1:10 serial dilution of gDNA (200ng –
449 0.2ng gDNA per reaction) to an efficiency of 100 +/- 10% (*PO1*: 100nM; hemocytin: 100nM; *g3PDH*:
450 400/350nM F/R; *NADH*: 350/300nM F/R; *β-tubulin*: 400nM; and *rpl32*: 200nM). Reactions were run on a
451 Bio-RAD CFX96 Real-Time System machine, with an initial step of 95°C for 3 minutes and 40 cycles of
452 95°C for 10s and 60°C for 30s. Each 20μL reaction included a 1X PCR buffer, Mg^{+2} at 2mM, dNTPs at
453 0.2mM, EvaGreen at 1X, 0.025 units/μL of Invitrogen taq, and 40ng of cDNA. Three technical replicates
454 were run for each reaction.

455 We measured expression of these genes in lines with and without symbionts in two separate
456 experiments. First, we collected aphids from the *Regiella*-infected lines used for the RNAseq above (no
457 symbionts, Clade 1 .LSR *Regiella*, and Clade 2 .313 *Regiella*). We dissected out and removed embryos,
458 pooled adult carcasses, extracted and cleaned RNA, and DNAse treated samples as above. We
459 synthesized cDNA using the BioRad iScript cDNA synthesis kit under recommended conditions. For
460 each sample we averaged the C_T values from the endogenous control genes, and calculated $-\Delta C_T$ values
461 by $-(C_{T\ target} - C_{T\ mean\ endogenous\ control})$. We analyzed differences in gene expression between symbiont-free,
462 clade 1, and clade 2 lines with one-way ANOVAs on the $-\Delta C_T$ values, and used Tukey HSD tests for pair-
463 wise comparisons among different symbiont backgrounds. We performed separate analyses for the two
464 genes.

465 In a second experiment, we injected three additional symbiont species into aphids and measured
466 *PO1* and hemocytin expression as above. For donor aphids, we used an aphid line harboring *Serratia*
467 *symbiotica*, a line harboring *Spiroplasma* sp. (strain .161), and two strains of *Hamiltonella defensa* (Table
468 S6). We only successfully established *Serratia* from two injection events after multiple attempts, and so
469 the biological replicates of this assay were generated by splitting the lines onto multiple plants after 2
470 generations before sampling at generation 5; the other lines represent independently injected lines. We
471 maintained sham-injected (symbiont-free) aphids under identical conditions as above. Gene expression
472 was measured and analyzed as above. We note that the *Spiroplasma sp.* strain used in this experiment
473 has been shown, like *Regiella*, to be protective against fungal pathogens while the other symbiont
474 species used have not found to confer fungal protection. The two *Hamiltonella* strains used were
475 collected in the same field and may not represent distinct symbiont genotypes from each other. Data
476 were analyzed as above.
477

478 **Expression knock-down via RNAi**: We designed primers that amplify regions of two target genes
479 (531bp of *PO1* and 483bp of Hemocytin) with the T7 promoter sequence
480 (TAATACGACTCACTATAGGG) on the 5' end of each primer using the e-RNAi Webservice
481 (https://www.dkfz.de/signaling/e-rnai3/). Primer sequences can be found in Table S7. We PCR amplified

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

482 these regions from cDNA under recommended conditions. PCR products were sequenced using Sanger
483 sequencing primed with the T7 promoter sequence to confirm target identity. We then purified 160μL of
484 PCR product (using NaOAc and EtOH precipitation) and concentrated it to 500ng/μL. We used the
485 MEGAScript RNAi kit to synthesize dsRNA from PCR amplicons under recommended conditions and a
486 15hr transcription incubation. We ran dsRNA (at a 1/400 dilution) on a 2% agarose gel to verify that a
487 single band was obtained of the correct size, and we then concentrated the dsRNA product to
488 approximately 3300ng/μL using LiCl and an ethanol precipitation, and eluted the final dsRNA product in
489 MEGAScript buffer. We repeated these protocols to generate dsRNA from lacZ as a control as in (72).
490 In a first experiment, we injected 1st instar (1-day-old) aphids with approximately 100μg dsRNA from
491 either *PO1* or *lacZ* as a control using a glass capillary needle attached to a syringe on the underside of
492 thorax of each aphid. We used two aphid lines that harbored either *Regiella* strain .LSR or .313. We
493 collected injected aphids at two time-points after injection: at 72hrs and at day 9 when they had
494 undergone their final molt to the adult stage. For the 72hr samples, groups of 3 whole aphids were
495 pooled in TRIzol and stored at -80°C for RNA extraction, or were stored in tubes at -20°C for gDNA
496 extraction. For the adult samples, we dissected out developing embryos and stored adult carcasses
497 individually for RNA or DNA extraction, and we also pooled embryos from three adults and stored them
498 in tubes for DNA extraction.
499 From the samples stored in TRIzol, we extracted RNA, synthesized cDNA, and measured the
500 expression of *PO1* and four endogenous control genes using qPCR as above. We calculated $-\Delta C_T$
501 values as above and analyzed these using two-way ANOVAs (with treatment (*lacZ* vs. *PO1*) and *Regiella*
502 strain (.LSR vs. .313) as factors). We conducted post-hoc analyses using Tukey's HSD tests. The two
503 time-points were analyzed separately. We extracted DNA from the remaining samples using the Qiagen
504 DNEasy kit and measured *Regiella* densities using qPCR amplification of the *hrpA* gene as above. We
505 analyzed these data using two-way ANOVAs and Tukey's HSD tests as above in R v.3.5.1 after testing
506 for model assumptions.
507 For adult injections, we reared LSR1 aphids with *Regiella* strains .LSR and .313 at low densities as
508 above. We then injected adult aphids (9 days old) with 0.3μL of dsRNA (approximately 1μg total). In this
509 experiment we performed knock-downs of two genes: *PO1* and Hemocytin. We first collected aphids
510 harboring strain .LSR at 72 hours after injection to validate our knock-downs. For each gene, we pooled
511 groups of three aphids into 8 samples, removed developing embryos, extracted RNA from adult
512 carcasses, synthesized cDNA, and measured gene expression as above. We analyzed $-\Delta C_T$ values for
513 each gene using t-tests with treatment as the independent variable. In this experiment we also measured
514 expression of *PO2* in aphids that had been injected with dsRNA from *PO1* to verify that our knock-down
515 was specific to *PO1*. We then injected adult aphids harboring either strain .LSR or strain .313 with
516 approximately 1μg dsRNA, removed and discarded developing embryos after 72hrs, pooled samples into
517 groups of 4 dissected aphids, extracted DNA as above, and measured *Regiella* densities using qPCR as

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

518 above. We analyzed -ΔC_T values from this experiment using a two-way ANOVA, and included symbiont
519 strain (.LSR or .313) and treatment (*lacZ* (control), *PO1*, and Hemocytin) as factors in the analysis. We
520 conducted post-hoc analyses using a Tukey's HSD test to compare levels within treatment.
521
522 **RNAseq on aphid genotypes from multiple host-plant associated biotypes:** We selected three
523 aphid genotypes each from a different host-plant associated biotype (*Lotus corniculatus, Ononis spinosa,*
524 *and Trifolium pratense*); information on collection location can be found in Table S5. We established
525 *Regiella* strain .313 infections in each line as above. Each *Regiella*-infected line was established from a
526 different symbiont injection and maintained separately on *V. faba* plants. In parallel, control aphids were
527 sham injected as above. After 4 generations, we froze seven adult aphids from each line in liquid
528 nitrogen and stored at -80°C. RNA was extracted and purified as above. For the *Trifolium pratense*
529 biotype, RNAseq libraries and sequencing was conducted as described for the LSR genotype used
530 above. For the *Lotus corniculatus* and *Ononis spinosa* biotypes, dual-indexed stranded sequencing
531 libraries were constructed using the NEBNext polyA selection and Ultra Directional RNA library
532 preparation kits. Libraries were sequenced on one lane of Illumina HiSeq 4000 (Paired-end 150bp)
533 generating a target of > 2x 280M reads. We analyzed each genotype separately, comparing libraries with
534 and without a *Regiella* infection, as above.
535
536 **F1 crosses:** We used two of these biotype lines (C317 from *Trifolium* and 663 from *Lotus*) for the F1
537 genetic cross because we found from RNAseq data that these lines responded differently to *Regiella*
538 infection. To induce male and female aphids for genetic crosses, we transferred stocks to 'autumn'
539 conditions (short day, 13L:11D at 18°C). After 30 days we moved third and fourth instar nymphs onto
540 leaf-plates (a fava bean leaf in 2% agar in Petri dishes) to isolate virgin egg-laying sexual females
541 (oviparae) and males. Oviparae have a characteristic thicker hind tibia, and this feature was used to
542 isolate probable oviparae from males. The male screening was less stringent because virgin males were
543 not needed. We setup each cross by placing the corresponding genotype oviparae and males onto a
544 fava bean seedling, replenishing breeding stocks as they became available. After 24 hours, we treated
545 melanized eggs with 10% calcium propionate to clean off the surface and then transferred eggs using
546 fine-tipped forceps to a small petri dish with Whatman filter paper moistened with sterile water. We
547 sealed the plates with parafilm wrap and left them in autumn conditions for a further 24 hours, after which
548 dishes were transferred to a 2°C incubator to diapause. After 3 months, eggs were removed from the
549 diapause conditions and with fine-tip forceps, rolled against a Kimwipe to reduce any microbial growth.
550 We then transferred diapaused eggs to a new leaf-plate and placed them in 'autumn' conditions (as
551 above) until a fundatrix hatched. Each fundatrix was separated and a line was considered stable after
552 two generations. We used this protocol to generate five F1 lines: three with line 663 as the maternal
553 genotype and two with C317 as the maternal genotype.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

554
555 **qPCR measures of gene expression in the F1 panel:** We established replicate *Regiella* strain .313
556 infections in parental and the five F1 hybrid lines as above. After 4 generations, we collected three adults
557 from each biological replicate, removed embryos via dissection, extracted DNA, and measured symbiont
558 density all as above. We grouped the replicate F1 lines from each direction of the cross together to
559 analyze these data, and used a one-way ANOVA in R version 3.5.0, comparing density among
560 genotypes using a Tukey's HSD post-hoc test.
561 We then measured expression of four immune genes aphids with and without *Regiella* using qPCR in
562 the parental and two of the F1 lines. Four generations after injection, four aphids from each biological
563 replicate from each genotype were removed from plants, embryos were dissected out of adult carcasses
564 and stored in TRIzol at -80C. We extracted RNA and synthesized cDNA as above. For each sample, we
565 measured expression against the four endogenous control genes used above. Here we measured
566 expression of four target genes: two copies of Phenoloxidase (*PO1:* ACYPI04484 and *PO2:*
567 ACYPI072244), Hemocytin (ACYPI003478), and Nitric Oxide Synthase (ACYPI001689). $-\Delta C_T$ values
568 were analyzed using an ANOVA after testing for model assumptions; post-hoc tests using the multcomp
569 package in R v.3.5.0 were conducted to compare the interaction terms between *Regiella*
570 presence/absence and host genotype. Analysis of expression of each gene was conducted separately.
571
572 ACKNOWLEDGEMENTS: BJP is a Pew Scholar in the Biomedical Sciences, supported by The Pew
573 Charitable Trusts. OSZ is supported by an NSF Graduate Research Fellowship.
574
575 STATEMENT OF AUTHORSHIP: HLN, EBG, OSZ, and BJP carried out the molecular and experimental work.
576 BJP conceived of the study, analyzed the data, wrote the manuscript, and funded the work. All authors
577 edited and approved the manuscript before submission.
578
579 DATA ACCESSIBILITY STATEMENT: All experimental data is included in a supplementary file. Sequence data
580 has been uploaded to the NCBI Sequence Read Archive (SRA) with accession number PRJNA684046.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

581 REFERENCES
582

583 1. K. Hilgenboecker, P. Hammerstein, P. Schlattmann, A. Telschow, J. H. Werren, How many
584 species are infected with Wolbachia?--A statistical analysis of current data. *FEMS Microbiol Lett*
585 **281**, 215-220 (2008).
586 2. N. A. Moran, J. P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial
587 symbionts. *Annu Rev Genet* **42**, 165-190 (2008).
588 3. J. Ferrari, F. Vavre, Bacterial symbionts in insects or the story of communities affecting
589 communities. *Philos Trans R Soc Lond B Biol Sci* **366**, 1389-1400 (2011).
590 4. H. Hawlena *et al.*, The arthropod, but not the vertebrate host or its environment, dictates bacterial
591 community composition of fleas and ticks. *ISME J* **7**, 221-223 (2013).
592 5. J. A. Chandler, J. M. Lang, S. Bhatnagar, J. A. Eisen, A. Kopp, Bacterial communities of diverse
593 Drosophila species: ecological context of a host-microbe model system. *PLoS Genet* **7**,
594 e1002272 (2011).
595 6. A. H. C. McLean, H. C. J. Godfray, J. Ellers, L. M. Henry, Host relatedness influences the
596 composition of aphid microbiomes. *Environmental Microbiology Reports* **11**, 808-816 (2019).
597 7. J. Jaenike, R. Unckless, S. N. Cockburn, L. M. Boelio, S. J. Perlman, Adaptation via symbiosis:
598 recent spread of a Drosophila defensive symbiont. *Science* **329**, 212-215 (2010).
599 8. A. K. Hansen, G. Jeong, T. D. Paine, R. Stouthamer, Frequency of secondary symbiont infection
600 in an invasive psyllid relates to parasitism pressure on a geographic scale in California. *Appl*
601 *Environ Microbiol* **73**, 7531-7535 (2007).
602 9. R. A. Chong, N. A. Moran, Intraspecific genetic variation in hosts affects regulation of obligate
603 heritable symbionts. *Proc Natl Acad Sci U S A* **113**, 13114-13119 (2016).
604 10. L. J. Funkhouser-Jones, E. J. van Opstal, A. Sharma, S. R. Bordenstein, The Maternal Effect
605 Gene Wds Controls Wolbachia Titer in Nasonia. *Curr Biol* **28**, 1692-1702 e1696 (2018).
606 11. K. S. Stoy, A. K. Gibson, N. M. Gerardo, L. T. Morran, A need to consider the evolutionary
607 genetics of host-symbiont mutualisms. *J Evol Biol* 10.1111/jeb.13715 (2020).
608 12. F. H. Login *et al.*, Antimicrobial peptides keep insect endosymbionts under control. *Science* **334**,
609 362-365 (2011).
610 13. J. Maire, C. Vincent-Monegat, F. Masson, A. Zaidman-Remy, A. Heddi, An IMD-like pathway
611 mediates both endosymbiont control and host immunity in the cereal weevil Sitophilus spp.
612 *Microbiome* **6**, 6 (2018).
613 14. J. K. Herren, B. Lemaitre, Spiroplasma and host immunity: activation of humoral immune
614 responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial
615 pathogens in Drosophila melanogaster. *Cell Microbiol* **13**, 1385-1396 (2011).
616 15. X. Pan *et al.*, The bacterium Wolbachia exploits host innate immunity to establish a symbiotic
617 relationship with the dengue vector mosquito Aedes aegypti. *ISME J* **12**, 277-288 (2018).
618 16. R. V. Rio, Y. N. Wu, G. Filardo, S. Aksoy, Dynamics of multiple symbiont density regulation
619 during host development: tsetse fly and its microbial flora. *Proc Biol Sci* **273**, 805-814 (2006).
620 17. A. J. Shultz, T. B. Sackton, Immune genes are hotspots of shared positive selection across birds
621 and mammals. *Elife* **8** (2019).
622 18. T. A. Schlenke, D. J. Begun, Natural selection drives Drosophila immune system evolution.
623 *Genetics* **164**, 1471-1480 (2003).
624 19. C. Drosophila 12 Genomes *et al.*, Evolution of genes and genomes on the Drosophila phylogeny.
625 *Nature* **450**, 203-218 (2007).
626 20. B. P. Lazzaro, T. J. Little, Immunity in a variable world. *Philos Trans R Soc Lond B Biol Sci* **364**,
627 15-26 (2009).
628 21. J. Rolff, M. T. Siva-Jothy, Invertebrate ecological immunology. *Science* **301**, 472-475 (2003).
629 22. R. L. Unckless, B. P. Lazzaro, The potential for adaptive maintenance of diversity in insect
630 antimicrobial peptides. *Philos Trans R Soc Lond B Biol Sci* **371** (2016).
631 23. J. C. Brownlie, K. N. Johnson, Symbiont-mediated protection in insect hosts. *Trends Microbiol* **17**,
632 348-354 (2009).
633 24. J. Martinez *et al.*, Should Symbionts Be Nice or Selfish? Antiviral Effects of Wolbachia Are Costly
634 but Reproductive Parasitism Is Not. *PLoS Pathog* **11**, e1005021 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

635 25. K. M. Oliver, J. Campos, N. A. Moran, M. S. Hunter, Population dynamics of defensive symbionts
636 in aphids. *Proc Biol Sci* **275**, 293-299 (2008).
637 26. J. A. Brisson, D. L. Stern, The pea aphid, Acyrthosiphon pisum: an emerging genomic model
638 system for ecological, developmental and evolutionary studies. *Bioessays* **28**, 747-755 (2006).
639 27. V. Fazalova, B. Nevado, Low Spontaneous Mutation Rate and Pleistocene Radiation of Pea
640 Aphids. *Mol Biol Evol* **37**, 2045-2051 (2020).
641 28. J. Peccoud, A. Ollivier, M. Plantegenest, J. C. Simon, A continuum of genetic divergence from
642 sympatric host races to species in the pea aphid complex. *Proc Natl Acad Sci U S A* **106**, 7495-
643 7500 (2009).
644 29. J. Ferrari, J. A. West, S. Via, H. C. Godfray, Population genetic structure and secondary
645 symbionts in host-associated populations of the pea aphid complex. *Evolution* **66**, 375-390
646 (2012).
647 30. S. H. Chung, B. J. Parker, F. Blow, J. A. Brisson, A. E. Douglas, Host and symbiont genetic
648 determinants of nutritional phenotype in a natural population of the pea aphid. *Mol Ecol*
649 10.1111/mec.15355 (2020).
650 31. J. A. Russell *et al.*, Uncovering symbiont-driven genetic diversity across North American pea
651 aphids. *Mol Ecol* **22**, 2045-2059 (2013).
652 32. T. Tsuchida, R. Koga, H. Shibao, T. Matsumoto, T. Fukatsu, Diversity and geographic distribution
653 of secondary endosymbiotic bacteria in natural populations of the pea aphid, Acyrthosiphon
654 pisum. *Mol Ecol* **11**, 2123-2135 (2002).
655 33. C. L. Scarborough, J. Ferrari, H. C. Godfray, Aphid protected from pathogen by endosymbiont.
656 *Science* **310**, 1781 (2005).
657 34. B. J. Parker, C. J. Spragg, B. Altincicek, N. M. Gerardo, Symbiont-mediated protection against
658 fungal pathogens in pea aphids: a role for pathogen specificity? *Appl Environ Microbiol* **79**, 2455-
659 2458 (2013).
660 35. P. Lukasik, M. van Asch, H. Guo, J. Ferrari, H. C. Godfray, Unrelated facultative endosymbionts
661 protect aphids against a fungal pathogen. *Ecol Lett* **16**, 214-218 (2013).
662 36. A. H. McLean, M. van Asch, J. Ferrari, H. C. Godfray, Effects of bacterial secondary symbionts on
663 host plant use in pea aphids. *Proc Biol Sci* **278**, 760-766 (2011).
664 37. J. Ferrari, C. L. Scarborough, H. C. Godfray, Genetic variation in the effect of a facultative
665 symbiont on host-plant use by pea aphids. *Oecologia* **153**, 323-329 (2007).
666 38. T. E. Leonardo, Removal of a specialization-associated symbiont does not affect aphid fitness.
667 *Ecology Letters* **7**, 461-468 (2004).
668 39. T. Tsuchida, R. Koga, T. Fukatsu, Host plant specialization governed by facultative symbiont.
669 *Science* **303**, 1989 (2004).
670 40. J. Hrcek *et al.*, Hosts do not simply outsource pathogen resistance to protective symbionts.
671 *Evolution* 10.1111/evo.13512 (2018).
672 41. H. Mathe-Hubert, H. Kaech, C. Hertaeg, J. Jaenike, C. Vorburger, Nonrandom associations of
673 maternally transmitted symbionts in insects: The roles of drift versus biased cotransmission and
674 selection. *Mol Ecol* **28**, 5330-5346 (2019).
675 42. C. International Aphid Genomics, Genome sequence of the pea aphid Acyrthosiphon pisum.
676 *PLoS Biol* **8**, e1000313 (2010).
677 43. L. M. Henry *et al.*, Horizontally transmitted symbionts and host colonization of ecological niches.
678 *Curr Biol* **23**, 1713-1717 (2013).
679 44. B. J. Parker, J. Hrcek, A. H. C. McLean, H. C. J. Godfray, Genotype specificity among hosts,
680 pathogens, and beneficial microbes influences the strength of symbiont-mediated protection.
681 *Evolution* **71**, 1222-1231 (2017).
682 45. B. J. Parker, J. Hrcek, A. H. C. McLean, J. A. Brisson, H. C. J. Godfray, Intraspecific variation and
683 within-host density in an insect-microbe symbiosis. *bioRxiv* 10.1101/2020.11.03.365353 (2020).
684 46. I. Arai *et al.*, Immunohistochemical analysis of the role of hemocytin in nodule formation in the
685 larvae of the silkworm, Bombyx mori. *J Insect Sci* **13**, 125 (2013).
686 47. W. Ni *et al.*, Hemocytin facilitates host immune responses against Nosema bombycis. *Dev Comp*
687 *Immunol* **103**, 103495 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

688 48. A. H. C. McLean *et al.*, Multiple phenotypes conferred by a single insect symbiont are
689 independent. *Proc Biol Sci* **287**, 20200562 (2020).
690 49. P. Lukasik, H. Guo, M. van Asch, J. Ferrari, H. C. Godfray, Protection against a fungal pathogen
691 conferred by the aphid facultative endosymbionts Rickettsia and Spiroplasma is expressed in
692 multiple host genotypes and species and is not influenced by co-infection with another symbiont.
693 *J Evol Biol* **26**, 2654-2661 (2013).
694 50. K. M. Oliver, P. H. Degnan, G. R. Burke, N. A. Moran, Facultative symbionts in aphids and the
695 horizontal transfer of ecologically important traits. *Annu Rev Entomol* **55**, 247-266 (2010).
696 51. C. Ye *et al.*, Induction of RNAi Core Machinery's Gene Expression by Exogenous dsRNA and the
697 Effects of Pre-exposure to dsRNA on the Gene Silencing Efficiency in the Pea Aphid
698 (Acyrthosiphon pisum). *Front Physiol* **9**, 1906 (2018).
699 52. C. Braendle, M. C. Caillaud, D. L. Stern, Genetic mapping of aphicarus -- a sex-linked locus
700 controlling a wing polymorphism in the pea aphid (Acyrthosiphon pisum). *Heredity (Edinb)* **94**,
701 435-442 (2005).
702 53. J. Hrcek, A. H. McLean, H. C. Godfray, Symbionts modify interactions between insects and
703 natural enemies in the field. *J Anim Ecol* **85**, 1605-1612 (2016).
704 54. O. Duron, T. E. Wilkes, G. D. Hurst, Interspecific transmission of a male-killing bacterium on an
705 ecological timescale. *Ecol Lett* **13**, 1139-1148 (2010).
706 55. I. González-Santoyo, A. Córdoba-Aguilar, Phenoloxidase: a key component of the insect immune
707 system. *Entomologia Experimentalis et Applicata* **142**, 1-16 (2012).
708 56. Y. Wang *et al.*, Activation of Aedes aegypti prophenoloxidase-3 and its role in the immune
709 response against entomopathogenic fungi. *Insect Mol Biol* **26**, 552-563 (2017).
710 57. O. Binggeli, C. Neyen, M. Poidevin, B. Lemaitre, Prophenoloxidase activation is required for
711 survival to microbial infections in Drosophila. *PLoS Pathog* **10**, e1004067 (2014).
712 58. L. Cerenius, B. L. Lee, K. Soderhall, The proPO-system: pros and cons for its role in invertebrate
713 immunity. *Trends Immunol* **29**, 263-271 (2008).
714 59. A. M. Laughton, J. R. Garcia, B. Altincicek, M. R. Strand, N. M. Gerardo, Characterisation of
715 immune responses in the pea aphid, Acyrthosiphon pisum. *J Insect Physiol* **57**, 830-839 (2011).
716 60. A. Schmitz *et al.*, The cellular immune response of the pea aphid to foreign intrusion and
717 symbiotic challenge. *PLoS One* **7**, e42114 (2012).
718 61. B. J. Parker, S. M. Barribeau, A. M. Laughton, L. H. Griffin, N. M. Gerardo, Life-history strategy
719 determines constraints on immune function. *J Anim Ecol* **86**, 473-483 (2017).
720 62. M. N. Grell, A. B. Jensen, P. B. Olsen, J. Eilenberg, L. Lange, Secretome of fungus-infected
721 aphids documents high pathogen activity and weak host response. *Fungal Genet Biol* **48**, 343-
722 352 (2011).
723 63. L. Xu, L. Ma, W. Wang, L. Li, Z. Lu, Phenoloxidases are required for the pea aphid's defence
724 against bacterial and fungal infection. *Insect Mol Biol* **28**, 176-186 (2019).
725 64. E. R. Davenport, Elucidating the role of the host genome in shaping microbiome composition. *Gut*
726 *Microbes* **7**, 178-184 (2016).
727 65. R. Koga, T. Tsuchida, T. Fukatsu, Changing partners in an obligate symbiosis: a facultative
728 endosymbiont can compensate for loss of the essential endosymbiont Buchnera in an aphid. *Proc*
729 *Biol Sci* **270**, 2543-2550 (2003).
730 66. K. M. Oliver, J. A. Russell, N. A. Moran, M. S. Hunter, Facultative bacterial symbionts in aphids
731 confer resistance to parasitic wasps. *Proc Natl Acad Sci U S A* **100**, 1803-1807 (2003).
732 67. D. Q. Chen, A. H. Purcell, Occurrence and transmission of facultative endosymbionts in aphids.
733 *Curr Microbiol* **34**, 220-225 (1997).
734 68. W. Bender, P. Spierer, D. S. Hogness, Chromosomal Walking and Jumping to Isolate DNA from
735 the Ace and rosy Loci and the Bithorax Complex in Drosophila melanogaster. *J. Mol. Biol.* **168**,
736 17-33 (1983).
737 69. L. M. Henry, M. C. Maiden, J. Ferrari, H. C. Godfray, Insect life history and the evolution of
738 bacterial mutualism. *Ecol Lett* **18**, 516-525 (2015).
739 70. D. Kim *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions,
740 deletions and gene fusions. *Genome Biol* **14**, R36 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

741 71. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput
742 sequencing data. *Bioinformatics* **31**, 166-169 (2015).
743 72. B. J. Parker, J. A. Brisson, A Laterally Transferred Viral Gene Modifies Aphid Wing Plasticity.
744 *Curr Biol* **29**, 2098-2103 e2095 (2019).
745 73. B. Li *et al.*, A large genomic insertion containing a duplicated follistatin gene is linked to the pea
746 aphid male wing dimorphism. *eLife* **9** (2020).
747
748

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

749 Table S1: Significantly differentially expressed genes from the RNAseq experiments. Numbers
750 show log_2 fold changes of expression in aphids with vs. without *Regiella.*
751

Gene ID	Annotation	LSR + .LSR	LSR + .313	Lotus + .313	Ononis + .313	Trifolium + .313
ACYPI061678	plexin A	0.03	-1.79 **	-1.19 **	-0.86	-0.41
ACYPI44738	plexin A1-like †	-0.19	-3.09 **	-2.13 **	-2.48 *	-0.46
ACYPI003478	Hemocytin	-0.37	-2.24 **	-1.94 **	-2.75	-0.78
ACYPI004484	Phenoloxidase 1 (subunit A3)	-1.40	-3.25 **	-1.68 **	-2.06	-0.14
ACYPI009767	uncharacterized protein	--	--	0.60 **	0.16	--
ACYPI008883	probable vesicular glutamate transporter eat-4 †	-0.13	-1.89 **	-1.78 **	-1.71	-0.05
ACYPI53900	Uncharacterized protein	-0.34	0.09	-2.04 **	-1.67 *	-0.45
ACYPI008487	Apolipoprotein D	0.03	-2.08 **	-0.83 *	-1.04	-0.50
ACYPI006183	Uncharacterized protein	0.17	-1.92 **	-0.88 *	-0.69	-0.22
ACYPI072244	Phenoloxidase 2 (subunit 2)	-1.38	-3.87 **	-1.56 *	-2.08	-0.20
ACYPI001483	Echinoderm microtubule-associated protein-like 1	-0.76	-1.25 **	-0.49 *	-0.45	-0.18
ACYPI001736	Uncharacterized protein	0.09	-1.09 **	-0.33	-0.64	-0.15
ACYPI50923	Fibulin-1 †	-0.22	-1.52 **	-0.62 *	-1.12	-0.37
ACYPI009930	Uncharacterized protein	-0.77	-2.13 **	0.09	--	-0.12
ACYPI007618	Neprilysin-11-like †	0.37	-1.08 **	-0.64 *	-0.88	-0.21
ACYPI007421	Uncharacterized protein	0.31	-1.44 **	-0.57 *	-0.88	-0.36
ACYPI001380	Cys-loop ligand-gated ion channel subunit-like / Neuronal acetylcholine receptor subunit	0.06	-1.28 **	-0.60 *	-0.79	-0.71
ACYPI061541	Uncharacterized protein	0.74	2.71 **	-1.63	-2.10	-0.51
ACYPI001359	SPARC (Secreted protein acidic and rich in cysteine)	-0.42	-0.71 **	-0.69 *	-1.05	-0.54
ACYPI24889	Toll-like receptor 7 †	0.01	-1.35 **	-1.08 *	-1.26	-0.20
ACYPI000953	Hydroxysteroid dehydrogenase-like protein 2	-0.69	-0.76 **	-2.63 *	-0.76	-0.13
ACYPI47960	MD-2 related lipid-recognition protein-like †	-0.76	-1.18 **	-0.39	-0.52	-0.21

752
753 ** FDR < 0.05
754 * FDR < 0.1
755
756 † Uncharacterized in aphid genome v.2; annotation based on blast results.
757
758
759

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

760 Table S2: Post-hoc tests (Tukey's HSD) corresponding with Figures 1D and 1E.
761
762 *Regiella*:
763

PO1	**difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
.313 vs No Symb	2.98	1.44	4.53	0.0011
.LSR vs No Symb	1.53	-0.01	3.08	0.0513
.LSR vs .313	-1.45	-2.99	0.09	0.0652

764

Hemocytin	**difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
.313 vs No Symb	2.54	1.61	3.47	< 0.001
.LSR vs No Symb	1.03	0.10	1.96	0.031
.LSR vs .313	-1.51	-2.44	-0.57	0.004

765
766
767 Other Symbiont Species:
768

PO1	**difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
Hamiltonella .179 vs *Spiroplasma*	3.36	2.57	4.15	< 0.001
Hamiltonella .445 vs *Spiroplasma*	3.47	2.68	4.26	< 0.001
Serratia vs *Spiroplasma*	0.37	-0.34	1.07	0.54
No Symb. vs *Spiroplasma*	-0.18	-0.97	0.61	0.96
Hamiltonella .445 vs *Ham.* 179	0.10	-0.76	0.97	0.99
Serratia vs. *Hamiltonella* .179	-3.00	-3.79	-2.21	< 0.001
No Symb. vs *Hamiltonella* .179	-3.54	-4.41	-2.68	< 0.001
Serratia vs *Hamiltonella* .445	-3.10	-3.89	-2.31	< 0.001
No Symb. vs *Hamiltonella* .445	-3.65	-4.51	-2.78	< 0.001
No Symb. vs *Serratia*	-0.54	-1.34	0.24	0.27

769

Hemocytin	**difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
Hamiltonella .179 vs *Spiroplasma*	-0.29	-0.99	0.41	0.73
Hamiltonella .445 vs *Spiroplasma*	-0.79	-1.50	-0.09	0.02
Serratia vs *Spiroplasma*	0.03	-0.59	0.66	0.99
No Symb. vs *Spiroplasma*	-0.29	-0.99	0.42	0.74
Hamiltonella .445 vs *Ham.* 179	-0.50	-1.27	0.27	0.33
Serratia vs. *Hamiltonella* .179	0.33	-0.38	1.03	0.64
No Symb. vs *Hamiltonella* .179	0.00	-0.77	0.78	1.00
Serratia vs *Hamiltonella* .445	0.83	0.12	1.53	0.02
No Symb. vs *Hamiltonella* .445	0.50	-0.27	1.28	0.32
No Symb. vs *Serratia*	-0.32	-1.02	0.39	0.65

770
771

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

772 Table S3: Post-hoc tests (Tukey's HSD) of *Regiella* densities in F1 lines
773
774

Comparison	difference	Lower bound	Upper bound	Adjusted p-value
663 vs (663x317.A)	0.91	0.02	1.84	0.056
663 vs (663x317.B)	1.22	0.07	2.38	**0.033 ***
663 vs (663x317.D)	1.30	0.38	2.23	**0.003 ****
663 vs (317x663.A)	1.45	0.52	2.37	**< 0.001 *****
663 vs (317x663.B)	1.19	0.38	2.00	**0.002 ****
663 vs 317	2.38	1.37	3.38	**< 0.001 *****
(663x317.A) vs (663x317.B)	0.31	-0.89	1.51	0.98
(663x317.A) vs (663x317.D)	0.39	-0.58	1.37	0.84
(663x317.A) vs (317x663.A)	0.54	-0.44	1.51	0.58
(663x317.A) vs (317x663.B)	0.28	-0.58	1.15	0.94
(663x317.A) vs 317	1.46	0.41	2.52	**0.003 ****
(663x317.B) vs (663x317.D)	0.08	-1.11	1.28	> 0.99
(663x317.B) vs (317x663.A)	0.23	-0.97	1.42	> 0.99
(663x317.B) vs (317x663.B)	-0.03	-1.14	1.08	> 0.99
(663x317.B) vs 317	1.15	-0.11	2.41	0.088
(663x317.D) vs (317x663.A)	0.14	-0.83	1.12	> 0.99
(663x317.D) vs (317x663.B)	-0.11	-0.98	0.76	> 0.99
(663x317.D) vs 317	1.07	0.02	2.13	**0.045 ***
(317x663.A) vs (317x663.B)	-0.25	-1.11	0.61	> 0.99
(317x663.A) vs 317	0.93	-0.13	1.98	0.11
(317x663.B) vs 317	1.18	0.23	2.14	**0.009 ****

775

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

776 Table S4: Post-hoc tests (Tukey's HSD) of gene expression in F1 lines
777
778

Gene: *PO1*	**Difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
663 vs 317	2.58	0.42	6.12	**< 0.001 *** **
663 vs (663x317)	0.99	0.40	2.45	0.095
663 vs (317x663)	1.82	0.40	4.52	**< 0.001 *** **
317 vs (663x317)	-1.60	0.41	-3.88	**0.0039 ** **
317 vs (317x663)	-0.77	0.41	-1.86	0.26
(663x317) vs (317x663)	0.83	0.40	2.10	0.18

779
780

Gene: *PO2*	**Difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
663 vs 317	2.72	0.41	6.64	**< 0.001 *** **
663 vs (663x317)	1.38	0.39	3.50	**0.0097 ** **
663 vs (317x663)	2.00	0.39	5.06	**< 0.001 *** **
317 vs (663x317)	-1.34	0.40	-3.32	**0.014 * **
317 vs (317x663)	-0.72	0.40	-1.80	0.30
(663x317) vs (317x663)	0.61	0.39	1.59	0.41

781
782

Gene: *Hemocytin*	**Difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
663 vs 317	1.38	0.59	2.33	0.12
663 vs (663x317)	1.37	0.57	2.41	0.10
663 vs (317x663)	1.82	0.59	3.09	**0.026 * **
317 vs (663x317)	-0.01	0.58	-0.02	> 0.99
317 vs (317x663)	0.45	0.60	0.74	0.88
(663x317) vs (317x663)	0.46	0.58	0.78	0.86

783
784

Gene: *NOS*	**Difference**	**Lower bound**	**Upper bound**	**Adjusted p-value**
663 vs 317	1.77	0.77	2.31	0.12
663 vs (663x317)	0.81	0.74	1.11	0.69
663 vs (317x663)	1.35	0.74	1.83	0.29
317 vs (663x317)	-0.96	0.75	-1.27	0.59
317 vs (317x663)	-0.42	0.75	-0.56	0.94
(663x317) vs (317x663)	0.53	0.72	0.74	0.88

785
786

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

787 Table S5: Collection information for aphid genotypes.
788
789

Host Genotype (lab code)	Location Collected	Year Collected	Original Symbionts	Biotype
LSR	Ithaca, NY, USA	1998	*Reg*	*Medicago sativa*
317	Glouchestershire, UK	2003	*Reg*	*Trifolium pratensae*
133	Berkshire, UK	2012	*Ham*	*Ononis spinosa*
663	Oxfordshire, UK	2014	*None*	*Lotus corniculatus*

790
791
792

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

793 Table S6: Collection information for symbiont strains.
794

Symbiont Genotype (lab code)	Species	Location Collected	Year Collected
.LSR	*Regiella insecticola* (clade 1)	Ithaca, NY, USA	1998
.313	*Regiella insecticola* (clade 2)	Gloucestershire, UK	2007
.161	*Spiroplasma* sp.	Oxfordshire, UK	2006
.179	*Hamiltonella defensa*	Ithaca, NY, USA	2015
.445	*Hamiltonella defensa*	Ithaca, NY, USA	2015
.509	*Serratia symbiotica*	Knoxville, TN, USA	2019

795
796
797
798

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

799 Table S7: Primer Sequences
800

Primer Target	F sequence (5' to 3')	R sequence (5' to 3')	Citation
Screening for secondary symbiont infections:			
Regiella	AGTTTGATCATGGCTCAGATTG	GGTAACGTCAATCGATAAGCA	(43)
Serratia	AGAGTTTGATCMTGGCTCAG	TTTGAGTTCCCGACTTTATCG	(43)
Spiroplasma	ATTCTTCAGTAAAAATGCTTGGA	ACACATTTACTTCATGCTATTGA	(48)
Hamiltonella	AGTTTGATCATGGCTCAGATTG	AAATGGTATTSGCATTTATCG	(43)
***Regiella* MLST sequencing:**			
accD	CAYATGSGCATCTCTGCC	AATTCACTACTTTGAAAACCCGG	(43)
hrpA	AAAACATTGTCTTCCGGG	TTTTCAAARTTNAGCAARTCMGG	(43)
***Regiella* density via qPCR**			
Regiella hrpA	CGCATTGGGAGAAAAGCCAAG	CCTTCCACCAAGCCATGACG	This study
Gene knockdown via RNAi			
lacZ	**TAATACGACTCACTATAGGG** AGACCACACCATGATTACGCCAAGCTC	**TAATACGACTCACTATAGGG** AGACCACCATATCGGTGGTCATCATGC	(72)
PO1 (ACYPI004484)	**TAATACGACTCACTATAGGG** CGAGCTACTGCGGTATCCTT	**TAATACGACTCACTATAGGG** ACATTATTGGTGTTTGCGAATG	This study
Hemocytin (ACYPI003478)	**TAATACGACTCACTATAGGG** TCGATCTTCGTCAACAATCA	**TAATACGACTCACTATAGGG** AGGCCAACCTTGTTCTACTCC	This study
			***T7 promoter sequence shown in bold**
Immune gene expression using qPCR:			
G3PDH (ACYPI009769)	CGGGAATTTCATTGAACGAC	TCCACAACACGGTTGGAGTA	(73)
NADH (ACYPI009382)	CGAGGAGAACATGCTCTTAGAC	GATAGCTTGGGCTGGACATATAG	(73)
β-tubulin (ACYPI001007)	GGCCAAGGGTCATTACACTGA	TGCGAACCACGTCCAACA	(30)
Rpl32 (ACYPI000074)	CAAAGTGATCGTTATGACAAACTCAA	CGTCTTCGGACTCTGTTGTCAA	(30)
PO1 (ACYPI004484)	CACTGTCCGTAGCATTGAT	GGCAGAATAATCGTGAGGTA	(63)
PO2 (ACYPI072244)	ACGTGCGTATACGTTTCTCGAA	TGGCTTCCTATTCTGTTTTGCA	This study
Hemocytin (ACYPI003478)	ACAATTCGGCGTAAAGGAGGT	TGGCATGTAATCGACGGTGT	This study
NOS (ACYPI001689)	TAGTGCTATCGGCAAACGGT	CGGATACTGCGGGAAGACAG	This study

801
802

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.17.420083; this version posted December 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

803 Table S8: Sequencing and alignment results
804

Host Genotype	Treatment	Read Pairs (after QC)	Map Rate	Read pairs mapped to an exon
LSR1	Control A	26,834,362	L: 90.3% R: 89.7%	22,247,678
LSR1	Control B	23,898,256	L: 89.0% R: 88.4%	20,040,867
LSR1	Control C	20,567,298	L: 90.1% R:89.5%	17,513,945
LSR1	Control D	30,839,554	L: 90.5% R: 89.4%	26,232,845
LSR1	+ Clade 1 (.LSR) A	25,834,362	L: 90.3% R: 89.7%	22,247,678
LSR1	+ Clade 1 (.LSR) B	22,145,457	L: 89.1% R: 88.1%	18,466,505
LSR1	+ Clade 1 (.LSR) C	27,986,917	L: 89.4% R: 89.1%	23,737,329
LSR1	+ Clade 1 (.LSR) D	21,490,058	L: 89.8% R: 89.3%	18,248,263
LSR1	+ Clade 2 (.313) A	20,099,181	L: 89.2% R:88.5%	16,863,392
LSR1	+ Clade 2 (.313) B	19,020,667	L: 88.9% R: 88.5%	15,942,798
LSR1	+ Clade 2 (.313) C	21,612,263	L: 89.7% R: 89.0%	18,295,757
LSR1	+ Clade 2 (.313) D	23,324,084	L: 89.6% R: 88.6%	19,704,560
Lotus corniculatus	Control A	28,664,619	L: 86.7% R: 86.1%	23,297,559
Lotus corniculatus	Control B	32,207,455	L: 87.5% R: 86.8%	25,222,986
Lotus corniculatus	Control C	25,987,704	L: 86.7% R: 86.3%	21,067,089
Lotus corniculatus	*Regiella* A	24,420,268	L: 87.1% R: 86.7%	19,780,712
Lotus corniculatus	*Regiella* B	22,622,062	L: 87.7% R: 87.2%	18,485,841
Lotus corniculatus	*Regiella* C	30,883,446	L: 87.3% R: 86.8%	25,106,295
Ononis spinosa	Control A	25,428,345	L: 87.6% R: 86.9%	20,652,853
Ononis spinosa	Control B	24,322,991	L: 87.8% R: 87.2%	19,823,044
Ononis spinosa	Control C	30,277,729	L: 87.7% R: 87.2%	24,579,747
Ononis spinosa	*Regiella* A	26,677,905	L: 87.9% R: 87.3%	21,741,656
Ononis spinosa	*Regiella* B	29,108,115	L: 87.8% R: 87.3%	23,498,863
Ononis spinosa	*Regiella* C	25,669,486	L: 87.8% R: 86.4%	20,955,594
Trifolium pratense	Control A	22,440,188	L:86.7% R:85.9%	17,919,579
Trifolium pretense	Control B	20,932,628	L: 86.8% R: 85.9%	16,741,330
Trifolium pretense	Control C	22,905,550	L: 86.6% R: 85.6%	18,228,176
Trifolium pretense	*Regiella* A	27,502,459	L: 87.8% R: 87.3%	22,631,920
Trifolium pretense	*Regiella* B	20,253,813	L: 87.0% R: 86.7%	16,395,529
Trifolium pratense	*Regiella* C	23,107,988	L: 86.3% R: 85.2%	18,251,833

805
806