Combining phenotypic and genomic approaches reveals no evidence for adaptation to the local mutualist in *Medicago lupulina*

Tia L. Harrison[1]*, Corlett W. Wood[1]*, Isabela L. Borges[1], and John R. Stinchcombe[1,2]

[1] Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario M5S3B2, Canada

[2] Centre for Genome Evolution and Function, University of Toronto, Toronto, Ontario, M5S3B2, Canada

* These authors contributed equally to this work.

Running title: Local adaptation in *Medicago lupulina*

Corresponding author: Corlett W. Wood

Address:

Department of Ecology & Evolutionary Biology

University of Toronto

25 Willcocks Street, Room 3055

Toronto, ON

Canada M5S 3B2

Fax: 1 (416) 978-5878

Phone: 1 (647) 936-0565

Email: corlett.wood@utoronto.ca

Abstract

Local adaptation is a common but not ubiquitous feature of species interactions, and understanding the circumstances under which it evolves illuminates the factors that influence adaptive population divergence. Antagonistic species interactions dominate the local adaptation literature relative to mutualistic ones, preventing an overall assessment of adaptation within interspecific interactions. Here, we tested whether the legume *Medicago lupulina* is locally adapted to two species of mutualistic nitrogen-fixing rhizobial bacteria that vary in frequency across its eastern North American range. We reciprocally inoculated northern and southern *M. lupulina* genotypes with the northern (*Ensifer medicae*) or southern bacterium (*E. meliloti*) in a greenhouse experiment, and performed a genome scan for loci that showed elevated differentiation between field-collected plants that hosted different bacteria. Despite producing different numbers of root nodules (the structures in which the plants house the bacteria), neither northern nor southern plants produced more seeds, flowered earlier, or were more likely to flower when inoculated with their local rhizobia. None of the loci identified in our genomic analysis belonged to the well-characterized suite of legume-rhizobia symbiosis genes, suggesting that the rhizobia do not drive genetic divergence between *M. lupulina* populations. Our results demonstrate that local adaptation has not evolved in this mutualism despite large-scale geographic variation in the identity of the interacting species.

Keywords

Coevolution, legume, rhizobia, reciprocal transplant, genome scan, mutualism

Introduction

Characterizing the circumstances under which local adaptation evolves informs our understanding of the relative importance of gene flow and selection, and thereby the extent and limitations of adaptive evolution (Antonovics, 1976; Bridle & Vines, 2007; Hereford, 2009; Savolainen *et al.*, 2013; Whitlock, 2015). However, existing tests of local adaptation to the biotic environment focus disproportionately on antagonistic interactions (but see Anderson et al. 2004, Hoeksema and Thompson 2007, Barrett et al. 2012), limiting our understanding of adaptation within the broad suite of interspecific interactions that occur in nature. Here we combined a reciprocal transplant experiment with a genome scan to test for local adaptation in a classic mutualism: the symbiosis between legumes and nitrogen-fixing bacteria.

Local adaptation—when native genotypes outperform foreign genotypes in their home environment (Hereford 2009)—is driven by differences in selection in alternative environments, and is reflected in divergent phenotypes and genotypes between populations. The literature on local adaptation to the biotic environment remains dominated by antagonistic species interactions such as those between hosts and their parasites, pathogens, or prey (Brodie *et al.*, 2002; Kawecki & Ebert, 2004; Hoeksema & Forde, 2008; Koskella *et al.*, 2012). Tests for local adaptation in mutualisms are fairly rare (but see (Anderson *et al.*, 2004; Hoeksema & Thompson, 2007; Johnson *et al.*, 2010; Barrett *et al.*, 2012)). Bias in the type of interspecific interactions used to study local adaptation is potentially problematic because the nature of species interactions may influence the degree of local adaptation that evolves (Bergstrom & Lachmann, 2003; Anderson *et al.*, 2004; Barrett *et al.*, 2012). Several evolutionary processes are expected to differ between mutualisms and antagonisms, including the maintenance of variation within interactors (Kopp &

Gavrilets, 2006; Heath & Stinchcombe, 2014), the impact on species diversification (Yoder & Nuismer, 2010), and the rate of adaptation (Bergstrom & Lachmann, 2003; Damore & Gore, 2011). In particular, slower rates of coevolution in mutualisms may result in weaker patterns of adaptation to the local symbiont than those observed in antagonisms (Bergstrom & Lachmann, 2003; Barrett *et al.*, 2012)

Indirect evidence, however, suggests there is substantial potential for coevolution and local adaptation in mutualisms (Hoeksema & Thompson, 2007; Heath, 2010; Newman *et al.*, 2015; Rubin & Moreau, 2016). A recent study found that ants that form mutualistic relationships with acacia trees exhibit faster rates of genome evolution than their non-mutualistic relatives, suggesting that mutualism accelerates evolution in this system (Rubin & Moreau, 2016). In some mutualistic taxa, genotype-by-genotype interactions—which occur when an organism's fitness depends jointly on its genotype and that of its symbiont—account for a substantial proportion of genetic variation in fitness-related traits within populations (Heath, 2010; Heath *et al.*, 2012; Ehinger *et al.*, 2014). On a broad geographic scale, these interactions are predicted to manifest as local adaptation when coupled with population differences in symbiont genotype frequencies (Heath & Nuismer, 2014). Ultimately, though, directly testing for local adaptation in mutualisms requires assaying the fitness consequences of sympatric and allopatric symbionts in a mutualism with among-population variation in symbiont identity (Heath & Stinchcombe, 2014).

The economically and ecologically important mutualism between legumes in the genus *Medicago* and nitrogen-fixing bacteria ("rhizobia") is well suited to performing this test (Cook *et al.*, 1997; Cook, 1999; Young *et al.*, 2011). In the facultative *Medicago*-rhizobia symbiosis, soil bacteria in the genus *Ensifer* (formerly *Sinorhizobium*) (Young, 2010) fix atmospheric nitrogen

for their plant hosts in exchange for carbohydrates and housing in specialized root organs called nodules (Mylona *et al.*, 1995; van Rhijn & Vanderleyden, 1995). In eastern North America the relative frequencies of two principal symbionts (*Ensifer medicae* and *E. meliloti*) (Béna *et al.*, 2005) vary along a latitudinal cline (Figure S1) (Harrison, 2015), which may generate strong selection on *Medicago* populations to adapt to their local *Ensifer* species. The bacteria are essential for plant growth in nitrogen-poor edaphic environments (Simonsen & Stinchcombe, 2014a), and genes mediating the association experience strong selection in both *Medicago* and *Ensifer* (Bailly *et al.*, 2006; De Mita *et al.*, 2007; Epstein *et al.*, 2012; Bonhomme *et al.*, 2015). Finally, there is substantial evidence for genotype-by-genotype interactions for fitness traits between *Medicago* and its *Ensifer* symbionts (Heath, 2010; Gorton *et al.*, 2012; Heath *et al.*, 2012), and suggestive evidence for some degree of co-speciation in the two genera (Béna *et al.*, 2005).

A major challenge in testing for local adaptation in the legume-rhizobia mutualism is that the fitness benefit of the symbiosis depends on the biotic and abiotic environmental conditions in which it is expressed (Heath & Tiffin, 2007; Heath *et al.*, 2010; Porter *et al.*, 2011; Barrett *et al.*, 2012; Simonsen & Stinchcombe, 2014a). It is therefore essential to perform tests that are robust to ancillary environmental variation in this mutualism. Local adaptation is classically tested in reciprocal transplant experiments, in which its diagnostic signature is a genotype-by-environment interaction for fitness (Clausen *et al.*, 1940; Clausen & Hiesey, 1958; Nunez-Farfan & Schlichting, 2001; Kawecki & Ebert, 2004). Although such experiments are powerful because they reflect whole-organism performance in native and foreign environments, genotype-by-environment interactions are sensitive to experimental conditions (Kawecki & Ebert, 2004) and

103 null results from any single experiment could be due to experimental conditions not adequately
104 reflecting the typical natural environment.

105 Genomic scans for selection are a complementary tool for detecting local adaptation that
106 address this problem (Buehler *et al.*, 2014; de Villemereuil *et al.*, 2015; Jensen *et al.*, 2016).
107 Genome scans identify loci that exhibit heightened differentiation between populations
108 inhabiting alternative environments, which are presumed to constitute the genetic basis of local
109 adaptation (Coop *et al.*, 2010; Günther & Coop, 2013; Savolainen *et al.*, 2013; Tiffin & Ross-
110 Ibarra, 2014). Unlike reciprocal transplant experiments, these tests integrate across generations
111 and ancillary environmental variation, capturing the cumulative effects of long-term selection in
112 alternative environments (Tiffin & Ross-Ibarra, 2014). However, genome scans are vulnerable to
113 the criticism that the phenotypic effects of candidate loci are often unknown (Pavlidis *et al.*,
114 2012). Moreover, when the relevant phenotypes have a diffuse genetic basis, each of the many
115 underlying genes experiences weak selection and exhibits low levels of genetic differentiation
116 that are undetectable in outlier analyses (McKay & Latta, 2002; Tiffin & Ross-Ibarra, 2014).

117 Although genome scans and reciprocal transplant experiments are typically treated as
118 alternatives because they draw on fundamentally different data, together the two approaches
119 constitute an exceptionally rigorous test for local adaptation in environmentally sensitive
120 symbioses such as the legume-rhizobia mutualism. Combined, the two approaches integrate over
121 the effects of all loci in the genome (reciprocal transplant experiments) and across ancillary
122 environmental variation (genome scans), producing inferences that are less vulnerable to the
123 weaknesses of either method (Buehler *et al.*, 2014; de Villemereuil *et al.*, 2015; Jensen *et al.*,
124 2016). Both approaches are feasible to apply in the *Medicago*-rhizobia mutualism because

125 *Medicago* has a short generation time (Turkington & Cavers, 1979), its rhizobia are easily
126 manipulated (Heath & Tiffin, 2007), an annotated genome is available in the genus (Young *et*
127 *al.*, 2011), and the genes involved in the mutualism are extensively characterized (Mylona *et al.*,
128 1995; Cook *et al.*, 1997; Young *et al.*, 2011).

129 In the present study, we combined a reciprocal transplant experiment and genome scan to
130 test for adaptation to the local rhizobia species in the black medic (*Medicago lupulina*). We
131 tested the effect of sympatric and allopatric rhizobia on plant fitness in a greenhouse experiment,
132 and performed a genome scan to test for loci that exhibited elevated differentiation between
133 field-collected plants associated with different bacterial species in natural populations. Together,
134 these two experiments captured naturally occurring plant-rhizobia associations in the field and
135 tested the fitness consequences of those associations in a controlled laboratory environment.
136 Neither the phenotypic nor genomic approaches revealed strong evidence of adaptation to the
137 local rhizobia in *M. lupulina*, suggesting that local adaptation has not evolved in this mutualism's
138 North American range.

139

140 **Materials and Methods**

141 *Study system*

142 *Medicago lupulina* is an annual, highly self-fertilizing legume native to Eurasia
143 (Turkington & Cavers, 1979; Yan *et al.*, 2009). After its introduction to North America in the
144 1700s, *M. lupulina* expanded its range to occupy nitrogen-poor areas of the continent's temperate
145 and subtropical regions (Turkington & Cavers, 1979). In eastern North America, the relative

146 frequencies of *M. lupulina*'s two symbiotic rhizobia species (*Ensifer medicae* and *E. meliloti*)
147 vary along a northwest-to-southeast cline (Figure S1) (Harrison, 2015).
148
149 *Reciprocal transplant experiment*
150 To test for adaptation to the local rhizobia, we inoculated *M. lupulina* genotypes from the
151 northern and southern portions of the plant's eastern North American range with either the
152 locally abundant rhizobium species in the north (*E. medicae*) or in the south (*E. meliloti*). From a
153 total of 39 *M. lupulina* populations sampled between Delaware and Ontario in September-
154 October of 2013 (Harrison, 2015), we selected 7 southern and 7 northern plant populations in
155 which Harrison (2015) detected only a single *Ensifer* species (Figure 1, Table S1; see Figure S1
156 for a complete map with all 39 sampled populations). Within each population, seeds and root
157 nodules were collected from 2-10 randomly chosen *M. lupulina* individuals. All sampled plants
158 were at least 0.5m apart. Nodules were stored at 4°C in plastic bags until they were processed.
159 Field-collected seeds from these populations were grown in the greenhouse for one generation to
160 reduce maternal and environmental effects from the field, and we performed our experiments
161 using the progeny of these greenhouse-grown plants.
162 We planted F_1 greenhouse-derived seeds of 43 maternal families (27 from the north and
163 16 from the south) in a split-plot randomized complete block design in the greenhouse at the
164 University of Toronto. Each block was divided into two bacterial treatments, each containing 15
165 northern and 11 southern plants, the locations of which were randomized within blocks.
166 Populations were split across blocks. Due to seed limitations, not all families were represented in
167 every block, but within a block both bacterial treatments comprised the same 26 families. We

168 replicated this design across six blocks, for a total of 312 plants (6-13 replicates per family for 37
169 families; 1-4 replicates per family for 6 families). An additional block containing 42 plants (33
170 from the north and 9 from the south) served as an inoculation control, and a means for estimating
171 plant performance and fitness in the absence of either bacterial species. Prior to planting, seeds
172 were scarified with a razor blade, sterilized with ethanol and bleach, and stratified on 8% water
173 agar plates at 4˚C for 7 days to germinate. We planted with sterile forceps into cone-tainers filled
174 with sand (autoclaved twice at 121°C). We misted seedlings with water daily and fertilized with
175 5mL of nitrogen-free Fahraeus medium (noble.org/medicagohandbook) twice before inoculation
176 with rhizobia.

177 The *Ensifer* strains used for inoculation were recovered from frozen samples collected by
178 Harrison (2015) from two of the populations used in our experiment. The strains were originally
179 cultured from field-collected root nodules by sterilizing one nodule per plant in ethanol and
180 bleach, and crushing and plating it onto a 2% tryptone yeast (TY) agar plate. Strains were re-
181 streaked onto TY agar four times to reduce contamination and grown at 30°C for 48 hours, after
182 which they were transferred to liquid TY media and cultured for two days at 30°C. To identify
183 each strain to species (*E. medicae* or *E. meliloti*), DNA was extracted from liquid cultures (cell
184 density: 8 x 10^8 cells/ml) using the MoBio UltraClean Microbial DNA Isolation Kit, whole-
185 genome sequenced at SickKids Hospital (Toronto, Ontario), and genotyped using GATK
186 (McKenna *et al.*, 2010). We used alignment scores and the *Ensifer* 16S locus (Rome *et al.*, 1997)
187 to determine species identity of rhizobia strains associated with the sampled plants.

188 We selected one *E. medicae* strain from the northernmost population in Ontario and one
189 *E. meliloti* strain from the southernmost population in Delaware for our experiment ("SEG" and

190 "DE" in Figure 1). Genetic diversity is very low among strains within *Ensifer* species across
191 North America (Harrison 2015), so the specific strains used are not likely to influence our
192 results. Prior to inoculation, these strains were cultured as described above from samples stored
193 at -80°C. Liquid cultures were diluted with sterile TY media to an OD600 reading of 0.1 (a
194 concentration of ~10^6 cells per mL) (Simonsen & Stinchcombe, 2014b). Each plant received 1
195 mL of inoculate 13 days after planting, and 1 mL again 10 days later. Controls were also
196 inoculated twice with sterile TY media 10 days apart, and were used to assess rhizobia
197 contamination across treatments. Throughout the remainder of the experiment, all plants were
198 bottom-watered three times a week. We used two bottom-watering trays per block, such that all
199 plants in a given bacterial treatment had the same tray, while those from the alternative bacterial
200 treatment had a different tray.

201 We scored mortality weekly throughout the experiment, counted the number of leaves
202 every 4 weeks, recorded the date of first flower, and collected seeds. After five months, which
203 approximates the length of the April-October growing season in southern Ontario (Turkington &
204 Cavers, 1979), we harvested all plants and collected any remaining unripe seeds. We dried and
205 weighed aboveground tissue from each plant to the nearest 0.1 mg, and counted all seeds and
206 root nodules (symbiotic organs housing the rhizobia).

207 We analyzed five traits to test for local adaptation of northern and southern *M. lupulina*
208 plants to their local rhizobium: number of seeds, aboveground biomass, flowering time
209 (excluding plants that did not flower), probability of flowering, and number of nodules. All
210 analyzes were performed in R v.3.2.4 with sum-to-zero contrasts ("contr.sum") (R Core Team,
211 2016), and we tested significance using type III sums of squares in the function Anova in the *car*

212 package (Fox & Weisberg, 2011). Log-transformed aboveground biomass and flowering time
213 were analyzed with general linear mixed models using the function lmer in the *lme4* package
214 (Bates *et al.*, 2015). Probability of flowering and number of nodules were analyzed with
215 generalized linear mixed models with binomial and Poisson error distributions, respectively,
216 using the function glmer in the *lme4* package (Bates *et al.*, 2015). We verified that all dependent
217 variables met the assumptions of linearity, normality, and homoscedasticity through visual
218 inspection of quantile-quantile plots, plots of the residuals versus fitted values, and scale-location
219 plots. Seed number was severely zero-inflated (42% of plants did not produce seeds), so we
220 analyzed it using a mixture model (see below).

221 Each of the above models included rhizobia treatment (*E. medicae* or *E. meliloti*), region
222 (north or south), and the rhizobia-by-region interaction as fixed effects. A significant rhizobia-
223 by-region interaction, in which northern plants have higher fitness when inoculated with *E.*
224 *medicae* and southern plants have higher fitness with *E. meliloti*, would be evidence for local
225 adaptation. We included a fixed effect of researcher in our analysis of nodule counts. Block,
226 population, and family nested within population were included as random effects. We also
227 included the block-by-treatment interaction as a random effect because the rhizobia treatment
228 was applied at the half-block rather than at the plant level (Altman & Krzywinski, 2015). While
229 this design provides a weaker test of the rhizobia main effect, it is sensitive to the detection of
230 rhizobia-by-region interactions, the main goal of our experiment (Altman & Krzywinski, 2015).

231 We analyzed seed number with a zero-inflated Poisson model implemented with the
232 function MCMCglmm in the package *MCMCglmm* (Hadfield, 2010). Zero-inflated models are a
233 type of mixture model in which the zero class is modeled as the combined result of binomial and

count processes (Zuur *et al.*, 2009). In MCMCglmm, zero-inflated Poisson GLMMs are fit as multi-response models with one latent variable for the binomial zero-generating process and one for the Poisson count-generating process (Hadfield, 2015). We fit a model for seed number that included fixed effects of rhizobia, region, the rhizobia-by-region interaction, and the reserved MCMCglmm variable "trait" that indexes the binomial and Poisson latent variables. We omitted the interaction between trait and other fixed effects in order to estimate a single effect of rhizobia, region, and the rhizobia-by-region interaction across both the binomial and Poisson processes. Block, population, family, and the block-by-treatment effect were included as random effects. Different random effect variances were fit to the binomial and Poisson processes using the "idh" variance structure in MCMCglmm (Hadfield, 2015). We fit a residual variance (R) structure using the argument rcov = ~ us(trait):units, which allows a unique residual for all predictors in the model, used the default priors for the fixed effects (mean = 0, variance = 10^{10}) and specified parameter-expanded priors (alpha.mu = 0, alpha.v = 1000) for the random effects (Hadfield, 2010).

We ran the model for 1,300,000 iterations, discarded the first 300,000 iterations, and stored every 1,000th iterate. Model convergence was assessed with traceplots, running mean plots, and autocorrelation plots of the fixed and random effects using the *coda* (Plummer *et al.*, 2006) and *mcmcplots* (McKay Curtis, 2015) packages. Even though we used parameter-expanded priors on the random effects, the estimates of the population and block random effects remained close to zero, but omitting these terms from our model did not qualitatively change the results.

Finally, we calculated pairwise correlations between all traits using Spearman's correlation on the family means for each trait. We obtained family means for biomass, flowering time, and number of nodules by extracting the conditional modes (also known as the best linear unbiased predictors, or BLUPs) for each level of the family random effect from the models described above. For number of seeds, we used the marginal posterior modes of the family random effect as our family mean estimates.

Genomic outlier analysis

We used *M. lupulina* SNP data collected by Harrison (2015) to perform genomic scans of local adaptation. Field-collected seeds from 190 *M. lupulina* individuals were grown in the greenhouse as described in the "Reciprocal transplant experiment" section above. We extracted DNA from leaf tissue collected from one individual per maternal line using the Qiagen DNeasy Plant Tissue Mini Protocol. These samples were sequenced at Cornell University using genotyping-by-sequencing (GBS) in two Illumina flow cell lanes (Elshire *et al.*, 2011). Genomic libraries were prepared with the restriction enzyme EcoT22I, and SNPs were called using the program Stacks (Catchen *et al.*, 2011, 2013). We extracted and sequenced rhizobia DNA from one nodule from each field-sampled plants, and determined the species identity of each strain as described in the "Reciprocal transplant experiment" section above. We successfully determined the species identity of the rhizobia associated with 73 out of 190 *M. lupulina* plants, and performed all subsequent analyses on these 73 plants (or a subset thereof; see below). Our bioinformatics and SNP discovery pipelines for *Ensifer* and *Medicago* are described in detail in Appendix S1.

We searched for outlier loci between *M. lupulina* plants hosting *E. medicae* and *E. meliloti* to assess whether there is evidence for genetic divergence between plants associated with different *Ensifer* species. We used the program Bayenv2 to calculate X^TX statistics for each SNP in the *M. lupulina* sample (Coop *et al.*, 2010): X^TX is an F_{ST}-like statistic that controls for population variation and covariation in allele frequencies. We first estimated the covariance matrix using 100,000 iterations. Because we only wanted to calculate X^TX statistics and did not wish to calculate environmental correlations, we included an environmental file of dummy values to run Bayenv2 but avoid environmental analysis.

We ranked SNPs from highest to lowest X^TX values and identified the top 1% of SNPs to BLAST against the reference genome of *M. truncatula* to identify the outlier loci involved in rhizobia association in *M. truncatula* (taxonomy ID 3880) (Tang *et al.*, 2014). We used nucleotide BLAST (blastn) to search somewhat similar sequences in the unannotated *M. truncatula* genome in order to retrieve chromosome positions for our outlier loci. To identify the orthologous gene associated with each outlier locus, we then looked up the chromosome position of each outlier in the annotated *Medicago truncatula* genome (Mt. 4.0 http://jcvi.org/medicago/). We performed the BLAST test in two ways: first using the range-wide sample of plants that hosted different bacterial species (73 plant individuals), and second, focusing on southern Ontario samples (49 plant individuals). We performed the latter test because of the possibility that many loci unrelated to bacterial specificity (e.g., climatic adaptation) could be differentiated between southern Ontario and the mid-Atlantic United States due to environmental gradients that covary with bacterial species composition.

298 Outlier loci detected in genotyping-by-sequencing (GBS) data are rarely the actual loci
299 responsible for adaptation; instead, they are usually in linkage disequilibrium (LD) with the
300 causal genes. To account for this possibility, we searched for genes involved in the legume-
301 rhizobia symbiosis within either 5 or 10 kb of the *M. truncatula* orthologs of the outlier loci that
302 we detected in both the range-wide and Ontario samples. This approach assumes synteny
303 between *M. truncatula* and *M. lupulina*. We chose 5 and 10 kb based on the scale of LD in *M.*
304 *truncatula* (Branca *et al.*, 2011). While the scale of LD between even closely related species is
305 likely to differ based on mutation rates, recombination, population structure, and a host of other
306 demographic and evolutionary factors, we viewed this approach as superior to simply confining
307 our searches to the GBS loci without accounting for potential LD with causal genes.
308 Finally, we measured the distance between the *M. truncatula* orthologs of the outlier loci
309 that we detected in both the range-wide and Ontario samples and key *M. truncatula* genes
310 involved in the rhizobia symbiosis (again assuming synteny between *M. truncatula* and *M.*
311 *lupulina*). We considered genes involved in the initial signal exchange between the legume and
312 rhizobia (NSP, IPD3, and DMI1-DMI3); genes involved in infection thread development (LIN);
313 and genes involved in both rhizobia signaling and infection (NFP, LYK3, and NIN) (Jones *et al.*,
314 2007; Oldroyd *et al.*, 2009; Young *et al.*, 2011; Oldroyd, 2013; Tang *et al.*, 2014).
315
316 **Results**
317 *Reciprocal transplant experiment: Uninoculated plants*
318 Uninoculated *Medicago lupulina* plants performed extremely poorly without rhizobia.
319 None of our uninoculated control plants flowered or set seed, and the biomass of control plants

was approximately 20-fold smaller than inoculated plants (least squares mean ± SE (mg): controls: 21.01 ± 0.05; inoculated plants from both rhizobia treatments: 476.01 ± 0.03; $F_{1,14.808}$ = 610.7, $P < 0.001$). The performance of the control plants also demonstrates that cross-contamination between the two rhizobia treatments was likely minimal in our experiment. Only 1 of 42 uninoculated control plants produced nodules, and this anomalous individual was similar in size to the rest of the controls for the first several months, indicating that it probably did not nodulate until late in the experiment.

Reciprocal transplant experiment: Inoculated plants

In plants inoculated with *E. medicae* or *E. meliloti*, pairwise family mean correlations between all measured traits were generally low, indicating that the traits that we measured were largely independent of one another ($r \leq |0.10|$, $P \geq 0.54$). Only flowering time and aboveground biomass were significantly correlated ($r = 0.49$, $P = 0.002$); later-flowering plants had greater aboveground biomass.

Our analysis of seed number, probability of flowering, and flowering time revealed no evidence of adaptation to the local rhizobia. There was no significant rhizobia-by-region interaction for any of these reproductive traits (Figure 2, Table 1). There was a marginally significant effect of region on seed number; southern plants produced more seeds than northern plants in both rhizobia treatments (Figure 2A, Table 1). There was no significant effect of rhizobia treatment or region on either flowering trait (Figure 2C, Table 1).

The rhizobia-by-region interaction for aboveground biomass was marginally significant ($P_{\text{rhizobia-by-region interaction}} = 0.054$, Table 1). While the biomass of northern plants was unaffected by

rhizobia treatment, southern plants produced more aboveground biomass when inoculated with *E. meliloti* (Figure 2B), the locally abundant rhizobia in south.

We found a highly significant rhizobia-by-region interaction for nodule number (Table 1). Northern plants produced more nodules that southern plants when inoculated with *E. medicae*, the locally abundant rhizobia in the north. The difference between northern and southern plants decreased when inoculated with *E. meliloti*, an effect that was driven by both an increase in nodulation in southern plants and a decrease in nodulation in northern plants (Figure 2D). There was also a significant effect of region, indicating that northern plants produced more nodules across both rhizobia treatments, and a significant effect of researcher (Table 1).

Genomic outlier analysis

We identified a distribution of X^TX statistics around the null expectation of $X^TX = 2$, reflecting the 2 populations assigned in Bayenv2 (*M. lupulina* plants hosting *E. medicae* and plants hosting *E. meliloti*). In the range-wide sample, 16% (354 of 2209) of SNPs had X^TX scores greater than the null expectation of 2; in the Ontario sample, 29% (573 of 1977) of SNPs had X^TX scores greater than 2. We detected a range of alignment scores when we used BLAST to align outlier loci with top X^TX statistics from the whole sample and the Ontario sample to the *M. truncatula* reference genome (Table 2). The loci mapped to several different chromosomes in the *M. truncatula* reference genome.

Of the top 1% of SNPs detected in the range-wide sample (20 SNPs total), eight were associated with a specific *M. truncatula* gene (BLAST scores: 35.6 – 102; E value: 1.00e-19 – 0.31). Higher BLAST scores reflect higher-quality alignments; these scores indicate that our

364 sequences generally aligned moderately well to the *M. truncatula* genome. E (expectation)-
365 values reflect the number of hits expected by chance; lower E-values indicate better matches.
366 These 8 loci did not map to any genes known to be involved in the legume-rhizobia mutualism.
367 The remaining 12 loci did not map to a specific gene in the *M. truncatula* genome (BLAST
368 scores: 35.6 – 102; E-values: 3.00e-20 – 3.10e-1).
369 The results were qualitatively similar for the Ontario sample (Table 2). The BLAST
370 scores of the top 1% of outlier SNPs (20 SNPs total) ranged from 37.4 to 111 (E-values: 5.00e-
371 24 – 8.90e-2). Twelve of the top 1% of SNPs in the Ontario sample mapped to genes that are not
372 known to be involved in the legume-rhizobia mutualism. The remaining eight loci did not
373 associate with a specific gene in the *M. truncatula* annotated genome. The BLAST scores for
374 these loci were similar to the twelve loci that did map to specific *M. truncatula* genes (score:
375 35.6 – 95.1; E-value: 3.00e-20 – 3.10e-1).
376 There were only three outlier loci that appeared in the top 1% of SNPs in both the range-
377 wide sample and the Ontario sample (Table 3). These loci mapped to chromosomes 1, 5, and 7 in
378 the *M. truncatula* genome, but did not map to a specific gene. No genes found within 5 or 10kb
379 of the *M. truncatula* orthologs of these three outliers are known to be involved in the legume-
380 rhizobia symbiosis (assuming synteny between *M. truncatula* and *M. lupulina*). The *M.*
381 *truncatula* ortholog of the outlier on chromosome 5 had two genes within 5 kb, a phosphate
382 putative gene and a Ty3/Gypsy polyprotein/retrotransposon. The ortholog of the outlier on
383 chromosome 1 had no genes within a 5 kb window, and the ortholog of the outlier on
384 chromosome 7 had two genes within 5 kb, a DUF247 domain protein and a Gypsy-
385 likepolyprotein/retrotransposon putative gene. When we increased our window size to 10 kb we

386 found more genes, but none related to infection with rhizobia. For example, the *M. truncatula*
387 ortholog of the outlier on chromosome 5 was close to a DUF679 domain membrane protein and
388 an alpha/beta fold hydrolase putative gene. The ortholog of the outlier on chromosome 1 had a
389 reverse transcriptase zinc binding protein and a homeobox knotted-like protein in its 10 kb
390 window. The ortholog of the outlier on chromosome 7 had a phosphoenolpyruvate carboxylase
391 within its 10 kb window, along with several putative proteins.

392 Finally, we calculated the distance in base pairs between the *M. truncatula* orthologs of
393 the three outlier loci found in both the range-wide and Ontario analyses and several genes
394 involved in *Medicago*-rhizobia association. None of the symbiosis genes that we considered
395 were close to the orthologs of any of these three outliers. Most of the symbiosis genes are located
396 on chromosome 5, but none were close to the ortholog of the outlier locus on chromosome 5
397 (Table 4). The ortholog of the outlier locus on chromosome 1 was approximately 35,822 kb
398 away from the only symbiosis gene we considered that is located on chromosome 1 (LIN). The
399 remaining two symbiosis genes—DMI1 and DMI2—are located on chromosomes 2 and 8 (Ané
400 *et al.*, 2002), neither of which contained any outlier loci in our analysis.

401

402 **Discussion**

403 We combined phenotypic and genomic approaches to test for local adaptation of *M.*
404 *lupulina* to its mutualistic nitrogen-fixing bacteria across its eastern North American range.
405 Although our results confirm that *M. lupulina* performs poorly without any rhizobia, we found
406 no evidence for adaptation to the local rhizobia species in our reciprocal transplant experiment
407 for the majority of traits, including our best proxy for fitness (number of seeds). Our genomic

scan for outlier loci between field-collected *M. lupulina* plants associated with different rhizobia produced similar results, detecting no genes implicated in the legume-rhizobia mutualism. Our results suggest that local rhizobia do not have differential fitness consequences for their host plants, nor do they drive genetic divergence in known symbiosis genes, indicating that local adaptation is either absent or weak in this mutualism's eastern North American range despite the strong cline in the relative abundances of the two rhizobia species.

Reciprocal transplant experiment

Uninoculated plants performed extremely poorly without either *Ensifer* species, demonstrating that *M. lupulina* is adapted to symbiosis with rhizobia. Despite differential nodulation with local and foreign rhizobia ($P_{\text{rhizobia-by-region}} < 0.001$, Table 1), however, there was no strong evidence for adaptation to the local rhizobia in other plant traits. One explanation for this pattern is that plants modify their nodulation strategy to compensate for differences in symbiotic efficiency with local and foreign rhizobia. The congeneric species *M. truncatula* adjusts its nodulation strategy in response to the rhizobia nitrogen fixation efficiency (Heath & Tiffin 2009), which jointly depends on plant and rhizobia genotype (Mhadhbi *et al.*, 2005). If plants produce more nodules with less efficient symbionts, increased nodulation may not translate to greater nitrogen uptake, masking any effects of differential nodulation on biomass and seed production. The fact that seed number, a reasonable proxy for total fitness in a selfing annual or short-lived perennial like *M. lupulina* (Turkington & Cavers, 1979), was unaffected by the local rhizobia strongly suggests that adaptation to the local rhizobia was absent in our experiment at the whole-plant level.

Even in the traits that exhibited a rhizobia-by-region interaction—the statistical signature of local adaptation—the data are only weakly consistent with the canonical pattern of local adaptation. The strongest test of local adaptation is whether local genotypes outperform foreign genotypes in all environments (the "local-versus-foreign" criterion) (Kawecki & Ebert, 2004). Neither trait that exhibited any rhizobia-by-region interaction (number of nodules and aboveground biomass) satisfied this criterion. Instead, our results were more closely aligned with a weaker test of local adaptation, which diagnoses local adaptation when each genotype's fitness is greater in its native environment than in alternative environments (the "home-versus-away" criterion) (Kawecki & Ebert, 2004).

One potential weakness of reciprocal transplant experiments is that the conditions used (in our case, cone-tainers, sterilized greenhouse soil, artificial day length control, absence of other biotic interactors, etc.) may not adequately mimic the conditions under which local adaptation is manifested. For example, when we ended our experiment five months after planting, nearly half of our plants had not yet set seed. It is possible that extending the experiment would have uncovered local adaptation in seed number, although this is unlikely because flowering was not accelerated by inoculating plants with their local rhizobia (probability of flowering: $P_{\text{rhizobia-by-region}} = 0.631$; time to flowering: $P_{\text{rhizobia-by-region}} = 0.242$; Table 1).

Genomic outlier analysis

Our genomic outlier locus scan should circumvent these weaknesses inherent in our reciprocal transplant experiment, because it should detect allele frequency differences between plants hosting different rhizobia that integrate across many generations of selection. However,

452 we also found very weak evidence of local adaptation in this analysis. The loci that were highly
453 differentiated between plants hosting different *Ensifer* species (the top 1% of loci in the X^TX
454 outlier analysis) were not associated with any genes involved in the legume-rhizobia symbiosis
455 in either the range-wide or Ontario samples. Instead, the *M. truncatula* orthologs of these loci
456 were genes encoding proteins involved in cellular structure (transmembrane protein, TPR repeat
457 protein), or cellular chemical reactions such as DNA binding (TLD-domain nuclear protein),
458 RNA binding (CRS1/YhbY CRM domain protein), protein transport (transportin 1 protein), and
459 oxidation-reduction reactions (FAD/NAD(P)-binding oxidoreductase family protein) (Young *et*
460 *al.*, 2011; Tang *et al.*, 2014).

461 There was very little overlap between the outlier loci identified by comparing plants
462 hosting different rhizobia from across eastern North America and from Ontario. The outlier loci
463 identified in the range-wide comparison may be predominately involved in adaptation to local
464 conditions unrelated to the symbiosis that also vary clinally in eastern North America. However,
465 the fact that no symbiosis genes were found in the *M. truncatula* orthologs of the outlier loci in
466 the Ontario-only comparison suggests that the rhizobia are not a major agent of selection even at
467 smaller spatial scales.

468 It is improbable that the loci identified in our genome scan are novel symbiosis genes
469 underlying adaptation to the local bacteria, although our data are subject to caveats common to
470 genome scans for selection (Pavlidis *et al.*, 2012). Outliers identified in genome scans are rarely
471 the causal loci; they are in linkage disequilibrium with the actual genes underlying local
472 adaptation. However, none of the *M. truncatula* orthologs of our outlier loci were located within
473 the scale of linkage disequilibrium (5-10 kb in *M. truncatula*) (Branca *et al.*, 2011) from known

474 symbiosis genes. Second, a significant portion of our outlier loci did not match any annotated
475 gene in the *M. truncatula* annotated genome. It is possible that the relevant part of the *M.*
476 *truncatula* genome has not yet been annotated, or that the loci fall between annotated genes and
477 may perform unknown regulatory functions. Finally, a few outlier loci aligned poorly to the *M.*
478 *truncatula* genome. If these poorly aligned loci were symbiosis genes that are specific to *M.*
479 *lupulina* and divergent from *M. truncatula*, using the *M. truncatula* genome as the reference
480 would bias us against inferring local adaptation from our genomic data.
481 However, the existence of *M. lupulina*-specific symbiosis genes is unlikely. The plant
482 genes involved in symbiotic interactions with rhizobia in the *Medicago* system are well
483 characterized and highly conserved in legumes (Rostas *et al.*, 1986; van Rhijn & Vanderleyden,
484 1995; De Mita *et al.*, 2006; Branca *et al.*, 2011; Gorton *et al.*, 2012; Stanton-Geddes *et al.*,
485 2013). *Medicago lupulina* is a close relative of *M. truncatula* (Bena, 2001; Yoder *et al.*, 2013),
486 and both plants fix nitrogen with both *Ensifer* species tested in our experiment (Béna *et al.*,
487 2005). It is therefore unlikely that *M. lupulina*-specific mutualism genes underlie adaptation to
488 different *Ensifer* species.
489
490 *Local adaptation in the legume-rhizobia symbiosis*
491 Our phenotypic and genomic data indicate that *M. lupulina* is not adapted to the local
492 rhizobia across its eastern North American range. The absence of local adaptation in this
493 mutualism is surprising given that the system is characterized by several features that ordinarily
494 strongly favor its evolution. Genotype-by-genotype interactions commonly occur between a
495 congener, *M. truncatula*, and different strains of the same *Ensifer* species (Heath & Tiffin, 2007;

Heath, 2010; Heath *et al.*, 2012), suggesting that the genetically divergent rhizobia *species* (Bailly *et al.*, 2006) we assayed would have even greater effects on their plant host. Furthermore, there is a cline in the frequencies of the two rhizobia across a large geographic scale that coincides with plant population genetic structure (Harrison, 2015). What might account for the lack of local adaptation in this mutualism?

Gene flow may overwhelm the effects of local selection, leading to a low equilibrium level of genetic differentiation between plants associated with different rhizobia (McKay & Latta, 2002). Although there is a strong geographic cline in the frequencies of the two *Ensifer* species, Harrison (2015) did detect *E. meliloti* in some northern populations and *E. medicae* in some southern populations. Local adaptation within *M. lupulina* populations could be swamped by gene flow from neighboring populations that encounter the alternative mutualist, or by the invasion of the alternative mutualist itself. Horizontal gene transfer between the two rhizobia could similarly homogenize any signature of local selection (Lenormand, 2002; Bailly *et al.*, 2007). Bacteria that form nitrogen-fixing symbioses with legumes have been shown to horizontally transfer genes involved in forming and maintaining the mutualism (Suominen *et al.*, 2001; Aoki *et al.*, 2013; Lemaire *et al.*, 2015), which could largely eliminate among-symbiont differences from the perspective of the legume host. Finally, temporal variation in the biotic and abiotic environment may modify the costs and benefits of the mutualism (Heath *et al.*, 2010; Heath & McGhee, 2012; Simonsen & Stinchcombe, 2014a), weakening selection favoring local rhizobia.

Alternatively, local adaptation may generate relatively weak fitness tradeoffs in mutualisms. The fitness tradeoffs that are the hallmark of local adaptation evolve whenever

adaptation to one environment results in maladaptation to another (Kawecki & Ebert, 2004). It has been hypothesized that selection in coevolving mutualisms strongly favors general compatibility and the reduction of fitness tradeoffs (Law & Koptur, 1986; Parker, 1999; Barrett *et al.*, 2012). Selection to minimize fitness tradeoffs may be especially strong in the legume-rhizobia mutualism, which is disproportionately important for plants growing in nitrogen-poor soils (Heath *et al.*, 2010). Under nitrogen-limited conditions, the cost of maladaptation to a locally rare rhizobium may be severe enough to outweigh the selective advantage of a marginal increase in the benefits obtained from the locally abundant rhizobium (Barrett *et al.*, 2012). However, this process should minimize plant-rhizobia interactions for fitness *within* rhizobia species as well, inconsistent with the pervasive genotype-by-genotype interactions documented between *M. truncatula* and *E. meliloti* strains (Heath *et al.*, 2012).

Finally, local adaptation may be restricted to the half of the mutualism that we did not examine; the rhizobia may be adapted to their local *M. lupulina* genotype even though the plant does not appear to be adapted to its local rhizobium. The strongest signature of local adaptation in our reciprocal transplant experiment occurred in nodule traits, a pattern that has also been documented in congeneric *Medicago* species (Porter *et al.*, 2011). Differential nodulation may impact the rhizobia more than the plant, given that nodule number is correlated with rhizobia fitness in *Medicago* (Heath, 2010). Stronger local adaptation in one partner commonly occurs in host-parasite systems (Hoeksema & Forde, 2008), but the phenomenon has not been systematically explored in the context of mutualism even though asymmetrical evolutionary rates in coevolving species pairs are expected in both mutualistic and antagonistic systems (Bergstrom & Lachmann, 2003).

Complementarity of phenotypic and genotypic approaches

Our study demonstrates the value of combining phenotypic and genomic approaches to test for local adaptation in a single system. Basing our inferences on both methods suggests that our results are robust to the assumptions of each, and that reciprocal transplant experiments and genome scans for outlier loci produce similar biological inferences despite relying on different diagnostic signatures of local adaptation. Concordance between the two approaches is important because they are often applied in very different study systems. Reciprocal transplant experiments are infeasible in long-lived species or those that are difficult to maintain in the lab (de Villemereuil *et al.*, 2015), and although genomic approaches are becoming increasingly possible to apply in non-model taxa, interpreting the results remains challenging for traits with an incompletely characterized genetic basis (Pavlidis *et al.*, 2012). If phenotypic and genomic methods generally produce concordant inferences, as we found here, then these two approaches are unlikely to be a cryptic source of bias in the local adaptation literature. Studies of local adaptation should consider pairing phenotypic and genomic approaches to validate their results with independent line of evidences and exclude alternative interpretations of the data (de Villemereuil *et al.*, 2015; Jensen *et al.*, 2016).

Acknowledgements

Our work is supported by Discovery Grants and graduate fellowships from NSERC Canada, and the EEB Postdoctoral Fellowship at the University of Toronto. We thank Bruce Hall and Andrew

561 Petrie for greenhouse assistance, and Adriana Salcedo, Michelle Afkhami, and Rebecca Batstone
562 for advice on experimental design and analysis.

563

564 **References**

565 Altman, N. & Krzywinski, M. 2015. Points of Significance: Split plot design. *Nat. Methods* **12**:
566 165–166.
567 Anderson, B., Olivieri, I., Lourmas, M. & Stewart, B.A. 2004. Comparative population genetic
568 structures and local adaptation of two mutualists. *Evolution* **58**: 1730–1747.
569 Ané, J.-M., Lévy, J., Thoquet, P., Kulikova, O., de Billy, F., Penmetsa, V., *et al.* 2002. Genetic
570 and cytogenetic mapping of DMI1, DMI2, and DMI3 genes of *Medicago truncatula*
571 involved in Nod factor transduction, nodulation, and mycorrhization. *Mol. Plant. Microbe.*
572 *Interact.* **15**: 1108–1118.
573 Antonovics, J. 1976. The Nature of Limits to Natural Selection. *Ann. Missouri Bot. Gard.* **63**:
574 224–247.
575 Aoki, S., Ito, M. & Iwasaki, W. 2013. From alpha- to beta-proteobacteria: The origin and
576 evolution of rhizobial nodulation genes nodIJ. *Mol. Biol. Evol.* **30**: 2494–2508.
577 Bailly, X., Olivieri, I., Brunel, B., Cleyet-Marel, J.C. & Béna, G. 2007. Horizontal gene transfer
578 and homologous recombination drive the evolution of the nitrogen-fixing symbionts of
579 *Medicago* species. *J. Bacteriol.* **189**: 5223–5236.
580 Bailly, X., Olivieri, I., De Mita, S., Cleyet-Marel, J.C. & Béna, G. 2006. Recombination and
581 selection shape the molecular diversity pattern of nitrogen-fixing *Sinorhizobium* sp.
582 associated to *Medicago*. *Mol. Ecol.* **15**: 2719–2734.
583 Barrett, L.G., Broadhurst, L.M. & Thrall, P.H. 2012. Geographic adaptation in plant-soil
584 mutualisms: Tests using *Acacia* spp. and rhizobial bacteria. *Funct. Ecol.* **26**: 457–468.
585 Bates, D., Mächler, M., Bolker, B. & Walker, S. 2015. Fitting Linear Mixed-Effects Models

using lme4. *J. Stat. Softw.* **67**.

Bena, G. 2001. Molecular phylogeny supports the morphologically based taxonomic transfer of the "medicagoid" Trigonella species to the genus *Medicago* L. *Plant Syst. Evol.* **229**: 217–236.

Béna, G., Lyet, A., Huguet, T. & Olivieri, I. 2005. *Medicago-Sinorhizobium* symbiotic specificity evolution and the geographic expansion of Medicago. *J. Evol. Biol.* **18**: 1547–1558.

Bergstrom, C.T. & Lachmann, M. 2003. The Red King effect: when the slowest runner wins the coevolutionary race. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 593–598.

Bonhomme, M., Boitard, S., San Clemente, H., Dumas, B., Young, N. & Jacquet, C. 2015. Genomic Signature of Selective Sweeps Illuminates Adaptation of *Medicago truncatula* to Root-Associated Microorganisms. *Mol. Biol. Evol.* **32**: 2097–2110.

Branca, A., Paape, T.D., Zhou, P., Briskine, R., Farmer, A.D., Mudge, J., *et al.* 2011. Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula. Proc. Natl. Acad. Sci. U. S. A.* **108**: E864-70.

Bridle, J.R. & Vines, T.H. 2007. Limits to evolution at range margins: when and why does adaptation fail? *Trends Ecol. Evol.* **22**: 140–147.

Brodie, E.D.J., Ridenhour, B.J. & Brodie, E.D.I. 2002. The evolutionary response of predators to dangerous prey: hotspots and coldspots in the geographic mosaic of coevolution between garter snakes and newts. *Evolution* **56**: 2067–82.

Buehler, D., Holderegger, R., Brodbeck, S., Schnyder, E. & Gugerli, F. 2014. Validation of outlier loci through replication in independent data sets: A test on *Arabis alpina*. *Ecol. Evol.*

608 **4**: 4296–4306.

609 Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. 2013. Stacks: An
610 analysis tool set for population genomics. *Mol. Ecol.* **22**: 3124–3140.

611 Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H. & De Koning, D.-J.
612 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3* **1**:
613 171–182.

614 Clausen, J. & Hiesey, W.M. 1958. Experimental studies on the nature of species. IV. Genetic
615 structure of ecological races. *Carnegie Inst. Washingt. 615. Washington, DC*.

616 Clausen, J., Keck, D.D. & Hiesey, W.M. 1940. Experimental studies on the nature of species. I.
617 Effect of varied environments on North American plants. *Carnegie Inst. Washingt. 520.*
618 *Washington, DC*.

619 Cook, D.R. 1999. *Medicago truncatula* - A model in the making! *Curr. Opin. Plant Biol.* **2**: 301–
620 304.

621 Cook, D.R., VandenBosch, K., de Bruijn, F.J. & Huguet, T. 1997. Model legumes get the nod.
622 *Plant Cell* **9**: 275–280.

623 Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J.K. 2010. Using environmental correlations
624 to identify loci underlying local adaptation. *Genetics* **185**: 1411–1423.

625 Damore, J.A. & Gore, J. 2011. A slowly evolving host moves first in symbiotic interactions.
626 *Evolution* **65**: 2391–2398.

627 De Mita, S., Santoni, S., Hochu, I., Ronfort, J. & Bataillon, T. 2006. Molecular evolution and
628 positive selection of the symbiotic gene NORK in *Medicago truncatula*. *J. Mol. Evol.* **62**:
629 234–244.

630 De Mita, S., Santoni, S., Ronfort, J. & Bataillon, T. 2007. Adaptive evolution of the symbiotic
631 gene NORK is not correlated with shifts of rhizobial specificity in the genus *Medicago*.
632 *BMC Evol. Biol.* **7**: 210.
633 de Villemereuil, P., Gaggiotti, O.E., Mouterde, M. & Till-Bottraud, I. 2015. Common garden
634 experiments in the genomic era: new perspectives and opportunities. *Heredity* **116**: 249–
635 254.
636 Ehinger, M., Mohr, T.J., Starcevich, J.B., Sachs, J.L., Porter, S.S. & Simms, E.L. 2014.
637 Specialization-generalization trade-off in a *Bradyrhizobium* symbiosis with wild legume
638 hosts. *BMC Ecol.* **14**: 8. BMC Ecology.
639 Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., *et al.* 2011. A
640 robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS*
641 *One* **6**: e19379.
642 Epstein, B., Branca, A., Mudge, J., Bharti, A.K., Briskine, R., Farmer, A.D., *et al.* 2012.
643 Population Genomics of the Facultatively Mutualistic Bacteria *Sinorhizobium meliloti* and
644 *S. medicae*. *PLoS Genet.* **8**: 1–10.
645 Fox, J. & Weisberg, S. 2011. *An R Companion to Applied Regression*, Second ed. Sage,
646 Thousand Oaks, CA.
647 Gorton, A.J., Heath, K.D., Pilet-Nayel, M.-L., Baranger, A. & Stinchcombe, J.R. 2012. Mapping
648 the genetic basis of symbiotic variation in legume-rhizobium interactions in *Medicago*
649 *truncatula*. *G3* **2**: 1291–303.
650 Günther, T. & Coop, G. 2013. Robust identification of local adaptation from allele frequencies.
651 *Genetics* **195**: 205–220.

652 Hadfield, J. 2015. MCMCglmm Course Notes.

653 Hadfield, J.D. 2010. MCMC methods for multi-response generalized linear mixed models: the
654 MCMCglmm R package. *J. Stat. Softw.* **33**: 1–22.

655 Harrison, T.L. 2015. Population genomics of the *Medicago lupulina* and *Ensifer* mutualism in
656 North America. Master's thesis. University of Toronto.

657 Heath, K.D. 2010. Intergenomic epistasis and coevolutionary constraint in plants and rhizobia.
658 *Evolution* **64**: 1446–1458.

659 Heath, K.D., Burke, P. V & Stinchcombe, J.R. 2012. Coevolutionary genetic variation in the
660 legume-rhizobium transcriptome. *Mol. Ecol.* **21**: 4735–47.

661 Heath, K.D. & McGhee, K.E. 2012. Coevolutionary Constraints? The environment alters
662 tripartite interaction traits in a legume. *PLoS One* **7**: e41567.

663 Heath, K.D. & Nuismer, S.L. 2014. Connecting functional and statistical definitions of genotype
664 by genotype interactions in coevolutionary studies. *Front. Genet.* **5**: 1–7.

665 Heath, K.D. & Stinchcombe, J.R. 2014. Explaining mutualism variation: a new evolutionary
666 paradox? *Evolution* **68**: 309–317.

667 Heath, K.D., Stock, A.J. & Stinchcombe, J.R. 2010. Mutualism variation in the nodulation
668 response to nitrate. *J. Evol. Biol.* **23**: 2494–2500.

669 Heath, K.D. & Tiffin, P. 2007. Context dependence in the coevolution of plant and rhizobial
670 mutualists. *Proc. R. Soc. B Biol. Sci.* **274**: 1905–1912.

671 Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173**:
672 579–588.

673 Hoeksema, J.D. & Forde, S.E. 2008. A Meta-Analysis of Factors Affecting Local Adaptation.

674 *Am. Nat.* **171**: 275–290.

675 Hoeksema, J.D. & Thompson, J.N. 2007. Geographic structure in a widespread plant-
676 mycorrhizal interaction: Pines and false truffles. *J. Evol. Biol.* **20**: 1148–1163.

677 Jensen, J.D., Foll, M. & Bernatchez, L. 2016. The past , present and future of genomic scans for
678 selection. *Mol. Ecol.* **25**: 1–4.

679 Johnson, N.C., Wilson, G.W.T., Bowker, M.A., Wilson, J.A. & Miller, R.M. 2010. Resource
680 limitation is a driver of local adaptation in mycorrhizal symbioses. *Proc. Natl. Acad. Sci. U.*
681 *S. A.* **107**: 2093–8.

682 Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E. & Walker, G.C. 2007. How rhizobial
683 symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nat. Rev. Microbiol.* **5**: 619–
684 33.

685 Kawecki, T.J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.

686 Kopp, M. & Gavrilets, S. 2006. Multilocus genetics and the coevolution of quantitative traits.
687 *Evolution* **60**: 1523–1536.

688 Koskella, B., Lin, D.M., Buckling, A. & Thompson, J.N. 2012. The costs of evolving resistance
689 in heterogeneous parasite environments. *Proc. R. Soc. B* **279**: 1896–903.

690 Law, R. & Koptur, S. 1986. On the evolution of non-specific mutualism. *Biol. J. Linn. Soc.* **27**:
691 251–267.

692 Lemaire, B., Van Cauwenberghe, J., Chimphango, S., Stirton, C., Honnay, O., Smets, E., *et al.*
693 2015. Recombination and horizontal transfer of nodulation and ACC deaminase (acdS)
694 genes within Alpha- and Betaproteobacteria nodulating legumes of the Cape Fynbos biome.
695 *FEMS Microbiol. Ecol.* **91**: 1–11.

696 Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**: 183–
697 189.
698 McKay, J.K. & Latta, R.G. 2002. Adaptive population divergence: Markers, QTL and traits.
699 *Trends Ecol. Evol.* **17**: 285–291.
700 McKay Curtis, S. 2015. *mcmcplots: Create plots from MCMC output. R package version 0.4.2*.
701 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., *et al.* 2010.
702 The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA
703 sequencing data. *Genome Res.* **20**: 1297–1303.
704 Mhadhbi, H., Jebara, M., Limam, F., Huguet, T. & Aouani, M.E. 2005. Interaction between
705 *Medicago truncatula* lines and *Sinorhizobium meliloti* strains for symbiotic efficiency and
706 nodule antioxidant activities. *Physiol. Plant.* **124**: 4–11.
707 Mylona, P., Pawlowski, K. & Bisseling, T. 1995. Symbiotic Nitrogen Fixation. *Plant Cell* **7**:
708 869–885.
709 Newman, E., Manning, J. & Anderson, B. 2015. Local adaptation: Mechanical fit between floral
710 ecotypes of *Nerine humilis* (Amaryllidaceae) and pollinator communities. *Evolution* **69**:
711 2262–2275.
712 Nunez-Farfan, J. & Schlichting, C.D. 2001. Evolution in Changing Environments: The
713 "Synthetic" Work of Clausen , Keck , and Hiesey. *Q. Rev. Biol.* **76**: 433–457.
714 Oldroyd, G.E.D. 2013. Speak, friend, and enter: signalling systems that promote beneficial
715 symbiotic associations in plants. *Nat. Rev. Microbiol.* **11**: 252–63.
716 Oldroyd, G.E.D., Harrison, M.J. & Paszkowski, U. 2009. Reprogramming plant cells for
717 endosymbiosis. *Science* **324**: 753–755.

718 Parker, M.A. 1999. Mutualism in metapopulations of legumes and rhizobia. *Am. Nat.* **153**: S48–
719 S60.
720 Pavlidis, P., Jensen, J.D., Stephan, W. & Stamatakis, A. 2012. A critical assessment of
721 storytelling: Gene ontology categories and the importance of validating genomic scans.
722 *Mol. Biol. Evol.* **29**: 3237–3248.
723 Plummer, M., Best, N., Cowles, K. & Vinces, K. 2006. CODA: Converence Diagnostics and
724 Output Analysis for MCMC. *R News* **6**: 7–11.
725 Porter, S.S., Stanton, M.L. & Rice, K.J. 2011. Mutualism and adaptive divergence: Co-invasion
726 of a heterogeneous grassland by an exotic legume-rhizobium symbiosis. *PLoS One* **6**.
727 R Core Team. 2016. R: A language and environment for statistical computing. Vienna, Austria.
728 Rome, S., Cleyet-Marel, J.-C., Materon, L.A., Normand, P. & Brunel, B. 1997. Rapid
729 identification of *Medicago* nodulating strains by using two oligonucleotide probes
730 complementary to 16S rDNA sequences. *Can. J. Microbiol.* **43**: 854–861.
731 Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A. & Kondorosi, A. 1986. Conservation of
732 extended promoter regions of nodulation genes in *Rhizobium*. *Proc. Natl. Acad. Sci. U. S. A.*
733 **83**: 1757–1761.
734 Rubin, B.E.R. & Moreau, C.S. 2016. Comparative genomics reveals convergent rates of
735 evolution in ant–plant mutualisms. *Nat. Commun.* **7**: 12679.
736 Savolainen, O., Lascoux, M. & Merilä, J. 2013. Ecological genomics of local adaptation. *Nat.*
737 *Rev. Genet.* **14**: 807–20.
738 Simonsen, A.K. & Stinchcombe, J.R. 2014a. Herbivory eliminates fitness costs of mutualism
739 exploiters. *New Phytol.* **202**: 651–61.

740 Simonsen, A.K. & Stinchcombe, J.R. 2014b. Standing genetic variation in host preference for
741 mutualist microbial symbionts. *Proc. R. Soc. B Biol. Sci.* **281**: 20142036–20142036.
742 Stanton-Geddes, J., Paape, T., Epstein, B., Briskine, R., Yoder, J., Mudge, J., *et al.* 2013.
743 Candidate Genes and Genetic Architecture of Symbiotic and Agronomic Traits Revealed by
744 Whole-Genome, Sequence-Based Association Genetics in *Medicago truncatula*. *PLoS One*
745 **8**: 1–9.
746 Suominen, L., Roos, C., Lortet, G., Paulin, L. & Lindström, K. 2001. Identification and structure
747 of the *Rhizobium galegae* common nodulation genes: evidence for horizontal gene transfer.
748 *Mol. Biol. Evol.* **18**: 907–916.
749 Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., *et al.* 2014. An
750 improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC*
751 *Genomics* **15**: 312.
752 Tiffin, P. & Ross-Ibarra, J. 2014. Advances and limits of using population genetics to understand
753 local adaptation. *Trends Ecol. Evol.* **29**: 673–680. Elsevier Ltd.
754 Turkington, R. & Cavers, P.B. 1979. The biology of Canadian weeds. 33. *Medicago lupulina* L.
755 *Can. J. Plant Sci.* **59**: 99–110.
756 van Rhijn, P. & Vanderleyden, J. 1995. The Rhizobium-Plant Symbiosis. *Microbiol. Rev.* **59**:
757 124–142.
758 Whitlock, M.C. 2015. Modern Approaches to Local Adaptation. *Am. Nat.* **186**: S000–S000.
759 Yan, J., Chu, H.J., Wang, H.C., Li, J.Q. & Sang, T. 2009. Population genetic structure of two
760 *Medicago* species shaped by distinct life form, mating system and seed dispersal. *Ann. Bot.*
761 **103**: 825–834.

762 Yoder, J.B., Briskine, R., Mudge, J., Farmer, A., Paape, T., Steele, K., *et al.* 2013. Phylogenetic
763 signal variation in the genomes of *Medicago* (Fabaceae). *Syst. Biol.* **62**: 424–438.
764 Yoder, J.B. & Nuismer, S.L. 2010. When Does Coevolution Promote Diversification? *Am. Nat.*
765 **176**: 802–817.
766 Young, J.M. 2010. Sinorhizobium versus Ensifer: May a taxonomy subcommittee of the ICSP
767 contradict the Judicial Commission? *Int. J. Syst. Evol. Microbiol.* **60**: 1711–1713.
768 Young, N.D., Debellé, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., *et al.*
769 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses.
770 *Nature* **480**: 520–4.
771 Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. 2009. *Mixed Effects Models*
772 *and Extensions in Ecology with R*. Springer, New York, NY.
773

774 **Data accessibility**

775 Sequence data will be uploaded to NCBI. VCF files and data from the reciprocal transplant
776 experiment will be available on Dryad. GPS coordinates of sampled plant and rhizobia
777 populations are reported in Table S1.

778 **Tables and Figures**

779 Table 1. Results of general(ized) linear mixed models testing for local adaptation in the
780 reciprocal transplant experiment.

				pMCMC
Seeds	Rhizobia			0.204
(*MCMC GLMM*)	Region			0.070
	Rhizobia × region			0.350

		F	df	P
Biomass	Rhizobia	1.955	1, 5.097	0.220
(*LMM*)	Region	0.131	1, 12.782	0.723
	Rhizobia × region	3.747	1, 248.656	0.054

		F	df	P
Flowering time	Rhizobia	0.016	1, 5.436	0.903
(*LMM*)	Region	0.252	1, 12.896	0.624
	Rhizobia × region	1.378	1, 164.795	0.242

		Wald χ^2	df	P
Prob. of flowering	Rhizobia	0.012	1	0.912
(*GLMM*)	Region	0.047	1	0.829
	Rhizobia × region	0.231	1	0.631

		Wald χ^2	df	P
Nodules	Rhizobia	0.107	1	0.743
(*GLMM*)	**Region**	**5.581**	**1**	**0.018**
	Researcher	**95.079**	**1**	**<0.001**
	Rhizobia × region	**34.806**	**1**	**<0.001**

The type of model used is indicated below each trait. GLMM: generalized linear mixed model (see text for error distribution). LMM: Linear mixed model (Gaussian error).

785 Table 2. Summary statistics of Bayenv2 and BLAST results for the top 1% of SNPs in the X^TX

786 outlier analysis.

Ontario sample					
X^TX	**SNP identity**	**BLAST**	**Query cover**	**E value**	**Gene**
4.20	585117	37.4	56	0.089	transmembrane protein
3.83	1192907	93.3	100	1.00E-18	indole-3-glycerol phosphate lyase IGL1
3.82	187811	93.3	98	1.00E-18	toprim domain protein
3.75	1110167	107	100	6.00E-23	TLD-domain nuclear protein
3.73	1610082	71.6	81	4.00E-12	no result
3.65	229813	35.6	60	3.10E-01	no result
3.59	1959186	107	100	6.00E-23	TPR repeat protein
3.56	129152	41	59	7.00E-03	no result
3.55	884266	89.7	100	2.00E-17	no result
3.44	616912	95.1	98	4.00E-19	no result
3.43	97240	111	100	5.00E-24	transportin-1 protein
3.39	666854	87.8	100	6.00E-17	no result
3.39	713735	96.9	100	1.00E-19	indole-3-glycerol phosphate lyase IGL1
3.33	1294820	69.8	95	2.00E-11	copia-like polyprotein/retrotransposon
3.33	686219	69.8	100	2.00E-11	no result
3.33	1463455	91.5	100	5.00E-18	single-stranded nucleic acid-binding protein R3H
3.29	671122	37.4	85	8.90E-02	cysteinyl-tRNA synthetase
3.26	1071597	98.7	100	3.00E-20	no result
3.25	109965	no result	no result	no result	no result
3.25	734196	78.8	100	3.00E-14	no result

Range-wide sample					
3.25	825707	96.9	100	1.00E-19	FAD/NAD(P)-binding oxidoreductase family protein
3.06	1131532	89.7	100	2.00E-17	novel plant SNARE-like protein
3.02	1482582	35.6	39	0.31	ASCH domain protein
2.99	405789	89.7	93	2.00E-17	CAAX amino terminal protease family protein
2.99	511074	48.2	96	5.00E-05	no result
2.97	175058	55.4	96	3.00E-07	no result
2.95	485278	91.5	93	5.00E-18	no result
2.94	254373	59	100	3.00E-08	CRS1/YhbY (CRM) domain protein
2.92	1090953	55.4	95	3.00E-07	galactose oxidase
2.90	921907	93.3	100	1.00E-18	DUF223 domain protein
2.86	1071597	98.7	100	3.00E-20	no result
2.82	870953	35.6	51	3.10E-01	
2.81	1590352	102	92	3.00E-21	no result
2.75	686219	69.8	100	2.00E-11	no result
2.73	774291	69.8	95	2.00E-11	no result
2.71	313573	69.8	98	2.00E-11	no result
2.68	1342844	60.8	90	8.00E-09	
2.67	1217147	89.7	100	2.00E-17	novel plant SNARE-like protein
2.64	616912	84.2	100	7.00E-16	no result
2.64	1057909	66.2	70	2.00E-10	no result

787

788

789

790

791 Table 3. Outlier loci found in the top 1% of Bayenv2 results in both the *M. lupulina* range-wide
792 and Ontario samples.

X^TX	SNP identity	BLAST	Query cover	E value	Gene name
3.26	1071597	98.7	100	3.00E-20	no result
3.44	616912	95.1	98	4.00E-19	no result
3.33	686219	69.8	100	2.00E-11	no result

793
794
795
796

797 Table 4. Base pair distances between the *M. truncatula* ortholog of the outlier locus on

798 chromosome 5 and well-characterized nodulation and rhizobial infection genes in *M. truncatula*.

Gene name	**Distance (kb)**
NFP (Nod-factor receptor 5)	12397
NSP (Nodulation receptor kinase-like protein)	6319
LYK3 (LysM receptor kinase K1B)	18
NIN (Nodule inception protein)	24
IPD3 (Cyclops protein putative)	8469

799

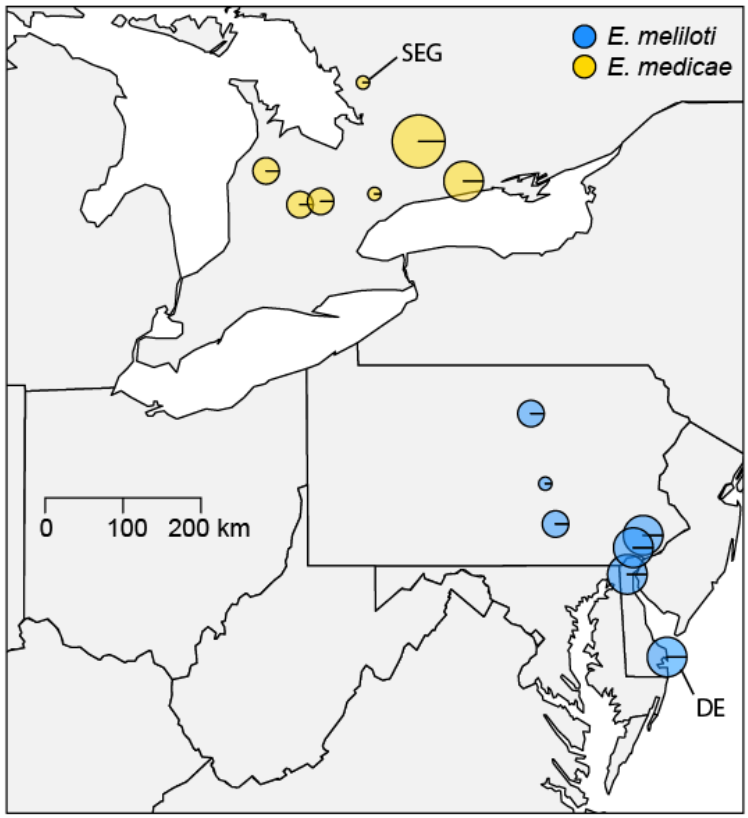
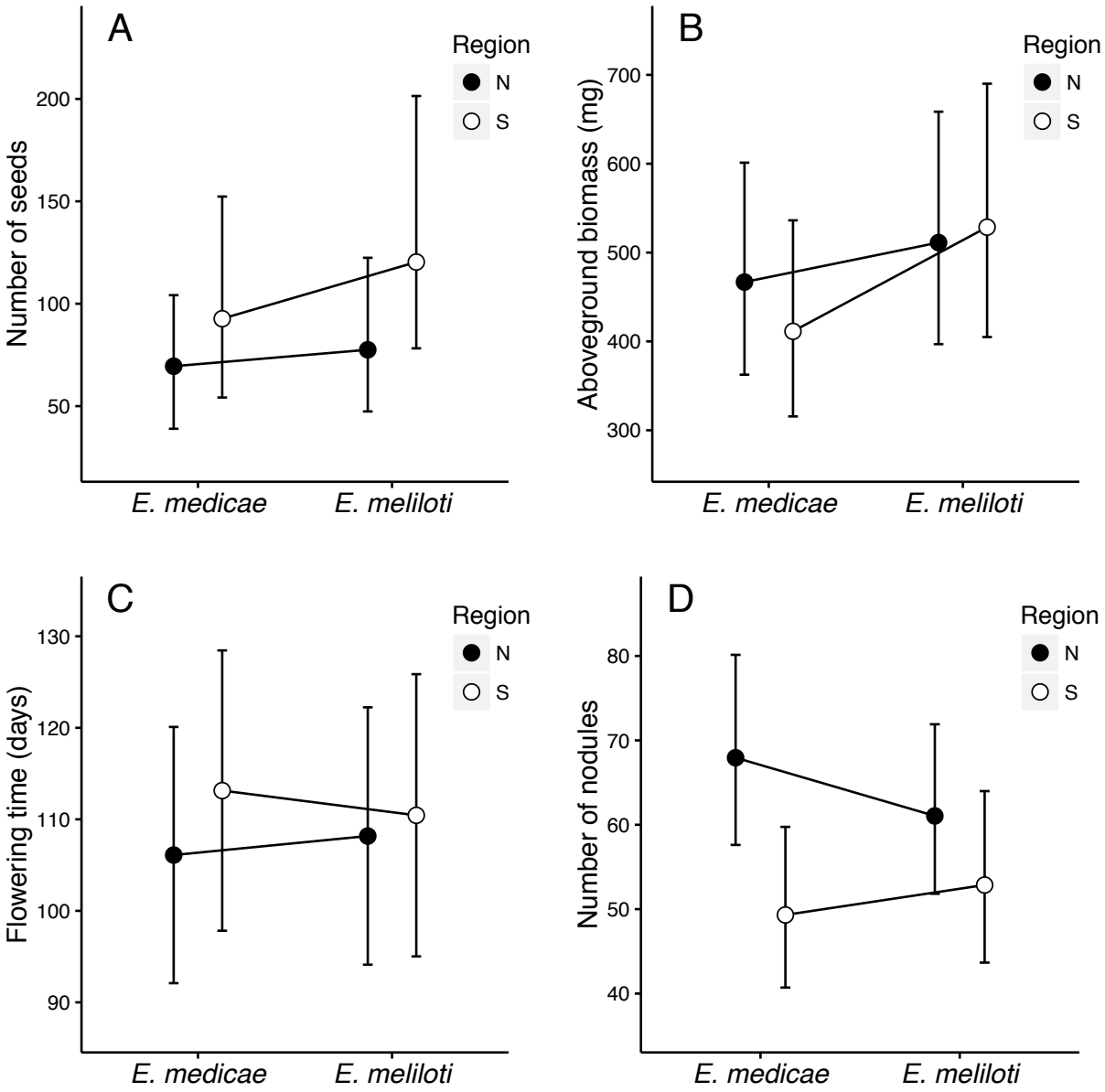


Figure 1. Locations of the 14 *M. lupulina* populations used in this study. The size of each circle corresponds to the number of plants sampled from the population, and the color indicates the rhizobia. The *E. medicae* strain used in the reciprocal transplant experiment was obtained from the northernmost population sampled ("SEG"); the *E. meliloti* strain was obtained from the southernmost population ("DE"). See Table S1 for GPS coordinates.



809

810 Figure 2. Least squares means and 95% confidence intervals for northern (black) and southern

811 (white) plants grown in the two rhizobia treatments. *Ensifer medicae* is the locally abundant

812 rhizobia in the north, and *E. meliloti* is the locally abundant rhizobia in the south. (A) Number of

813 seeds; (B) aboveground biomass; (C) flowering time; (D) number of nodules.

814