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

Tia L. Harrison¹*, Corlett W. Wood¹*, Isabela L. Borges¹, and John R. Stinchcombe¹,²

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

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

Although genome scans and reciprocal transplant experiments are typically treated as alternatives because they draw on fundamentally different data, together the two approaches constitute an exceptionally rigorous test for local adaptation in environmentally sensitive symbioses such as the legume-rhizobia mutualism. Combined, the two approaches integrate over the effects of all loci in the genome (reciprocal transplant experiments) and across ancillary environmental variation (genome scans), producing inferences that are less vulnerable to the weaknesses of either method (Buehler *et al.*, 2014; de Villemereuil *et al.*, 2015; Jensen *et al.*, 2016). Both approaches are feasible to apply in the *Medicago*-rhizobia mutualism because
Medicago has a short generation time (Turkington & Cavers, 1979), its rhizobia are easily manipulated (Heath & Tiffin, 2007), an annotated genome is available in the genus (Young et al., 2011), and the genes involved in the mutualism are extensively characterized (Mylona et al., 1995; Cook et al., 1997; Young et al., 2011).

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

Materials and Methods

Study system

Medicago lupulina is an annual, highly self-fertilizing legume native to Eurasia (Turkington & Cavers, 1979; Yan et al., 2009). After its introduction to North America in the 1700s, M. lupulina expanded its range to occupy nitrogen-poor areas of the continent’s temperate and subtropical regions (Turkington & Cavers, 1979). In eastern North America, the relative
frequencies of *M. lupulina*’s two symbiotic rhizobia species (*Ensifer medicae* and *E. meliloti*) vary along a northwest-to-southeast cline (Figure S1) (Harrison, 2015).

**Reciprocal transplant experiment**

To test for adaptation to the local rhizobia, we inoculated *M. lupulina* genotypes from the northern and southern portions of the plant’s eastern North American range with either the locally abundant rhizobium species in the north (*E. medicae*) or in the south (*E. meliloti*). From a total of 39 *M. lupulina* populations sampled between Delaware and Ontario in September-October of 2013 (Harrison, 2015), we selected 7 southern and 7 northern plant populations in which Harrison (2015) detected only a single *Ensifer* species (Figure 1, Table S1; see Figure S1 for a complete map with all 39 sampled populations). Within each population, seeds and root nodules were collected from 2-10 randomly chosen *M. lupulina* individuals. All sampled plants were at least 0.5m apart. Nodules were stored at 4°C in plastic bags until they were processed. Field-collected seeds from these populations were grown in the greenhouse for one generation to reduce maternal and environmental effects from the field, and we performed our experiments using the progeny of these greenhouse-grown plants.

We planted F₁ greenhouse-derived seeds of 43 maternal families (27 from the north and 16 from the south) in a split-plot randomized complete block design in the greenhouse at the University of Toronto. Each block was divided into two bacterial treatments, each containing 15 northern and 11 southern plants, the locations of which were randomized within blocks. Populations were split across blocks. Due to seed limitations, not all families were represented in every block, but within a block both bacterial treatments comprised the same 26 families. We
replicated this design across six blocks, for a total of 312 plants (6-13 replicates per family for 37
families; 1-4 replicates per family for 6 families). An additional block containing 42 plants (33
from the north and 9 from the south) served as an inoculation control, and a means for estimating
plant performance and fitness in the absence of either bacterial species. Prior to planting, seeds
were scarified with a razor blade, sterilized with ethanol and bleach, and stratified on 8% water
agar plates at 4°C for 7 days to germinate. We planted with sterile forceps into cone-tainers filled
with sand (autoclaved twice at 121°C). We misted seedlings with water daily and fertilized with
5mL of nitrogen-free Fahraeus medium (noble.org/medicagohandbook) twice before inoculation
with rhizobia.

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

We selected one E. medicae strain from the northernmost population in Ontario and one
E. meliloti strain from the southernmost population in Delaware for our experiment ("SEG" and
Gene\footnote{DE in Figure 1}. Genetic diversity is very low among strains within \textit{Ensifer} species across North America (Harrison 2015), so the specific strains used are not likely to influence our results. Prior to inoculation, these strains were cultured as described above from samples stored at -80°C. Liquid cultures were diluted with sterile TY media to an OD600 reading of 0.1 (a concentration of \( \approx 10^6 \) cells per mL) (Simonsen & Stinchcombe, 2014b). Each plant received 1 mL of inoculate 13 days after planting, and 1 mL again 10 days later. Controls were also inoculated twice with sterile TY media 10 days apart, and were used to assess rhizobia contamination across treatments. Throughout the remainder of the experiment, all plants were bottom-watered three times a week. We used two bottom-watering trays per block, such that all plants in a given bacterial treatment had the same tray, while those from the alternative bacterial treatment had a different tray.

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

We analyzed five traits to test for local adaptation of northern and southern \textit{M. lupulina} plants to their local rhizobium: number of seeds, aboveground biomass, flowering time (excluding plants that did not flower), probability of flowering, and number of nodules. All analyzes were performed in R v.3.2.4 with sum-to-zero contrasts ("contr.sum") (R Core Team, 2016), and we tested significance using type III sums of squares in the function \texttt{Anova} in the \texttt{car} package.
package (Fox & Weisberg, 2011). Log-transformed aboveground biomass and flowering time were analyzed with general linear mixed models using the function lmer in the lme4 package (Bates et al., 2015). Probability of flowering and number of nodules were analyzed with generalized linear mixed models with binomial and Poisson error distributions, respectively, using the function glmer in the lme4 package (Bates et al., 2015). We verified that all dependent variables met the assumptions of linearity, normality, and homoscedasticity through visual inspection of quantile-quantile plots, plots of the residuals versus fitted values, and scale-location plots. Seed number was severely zero-inflated (42% of plants did not produce seeds), so we analyzed it using a mixture model (see below).

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

We analyzed seed number with a zero-inflated Poisson model implemented with the function MCMCglmm in the package MCMCglmm (Hadfield, 2010). Zero-inflated models are a 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 $X^T X$ statistics for each SNP in the *M. lupulina* sample (Coop et al., 2010): $X^T X$ 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 $X^T X$ 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 $X^T X$ 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.
Outlier loci detected in genotyping-by-sequencing (GBS) data are rarely the actual loci responsible for adaptation; instead, they are usually in linkage disequilibrium (LD) with the causal genes. To account for this possibility, we searched for genes involved in the legume-rhizobia symbiosis within either 5 or 10 kb of the *M. truncatula* orthologs of the outlier loci that we detected in both the range-wide and Ontario samples. This approach assumes synteny between *M. truncatula* and *M. lupulina*. We chose 5 and 10 kb based on the scale of LD in *M. truncatula* (Branca et al., 2011). While the scale of LD between even closely related species is likely to differ based on mutation rates, recombination, population structure, and a host of other demographic and evolutionary factors, we viewed this approach as superior to simply confining our searches to the GBS loci without accounting for potential LD with causal genes.

Finally, we measured the distance between the *M. truncatula* orthologs of the outlier loci that we detected in both the range-wide and Ontario samples and key *M. truncatula* genes involved in the rhizobia symbiosis (again assuming synteny between *M. truncatula* and *M. lupulina*). We considered genes involved in the initial signal exchange between the legume and rhizobia (NSP, IPD3, and DMI1-DMI3); genes involved in infection thread development (LIN); and genes involved in both rhizobia signaling and infection (NFP, LYK3, and NIN) (Jones et al., 2007; Oldroyd et al., 2009; Young et al., 2011; Oldroyd, 2013; Tang et al., 2014).

**Results**

**Reciprocal transplant experiment: Uninoculated plants**

Uninoculated *Medicago lupulina* plants performed extremely poorly without rhizobia. 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 $X^T X$ statistics around the null expectation of $X^T X = 2$, 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 $X^T X$ scores greater than the null expectation of 2; in the Ontario sample, 29% (573 of 1977) of SNPs had $X^T X$ scores greater than 2. We detected a range of alignment scores when we used BLAST to align outlier loci with top $X^T X$ 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
sequences generally aligned moderately well to the *M. truncatula* genome. E (expectation)-values reflect the number of hits expected by chance; lower E-values indicate better matches.

These 8 loci did not map to any genes known to be involved in the legume-rhizobia mutualism. The remaining 12 loci did not map to a specific gene in the *M. truncatula* genome (BLAST scores: 35.6 – 102; E-values: 3.00e-20 – 3.10e-1).

The results were qualitatively similar for the Ontario sample (Table 2). The BLAST scores of the top 1% of outlier SNPs (20 SNPs total) ranged from 37.4 to 111 (E-values: 5.00e-24 – 8.90e-2). Twelve of the top 1% of SNPs in the Ontario sample mapped to genes that are not known to be involved in the legume-rhizobia mutualism. The remaining eight loci did not associate with a specific gene in the *M. truncatula* annotated genome. The BLAST scores for these loci were similar to the twelve loci that did map to specific *M. truncatula* genes (score: 35.6 – 95.1; E-value: 3.00e-20 – 3.10e-1).

There were only three outlier loci that appeared in the top 1% of SNPs in both the range-wide sample and the Ontario sample (Table 3). These loci mapped to chromosomes 1, 5, and 7 in the *M. truncatula* genome, but did not map to a specific gene. No genes found within 5 or 10kb of the *M. truncatula* orthologs of these three outliers are known to be involved in the legume-rhizobia symbiosis (assuming synteny between *M. truncatula* and *M. lupulina*). The *M. truncatula* ortholog of the outlier on chromosome 5 had two genes within 5 kb, a phosphate putative gene and a Ty3/Gypsy polyprotein/retrotransposon. The ortholog of the outlier on chromosome 1 had no genes within a 5 kb window, and the ortholog of the outlier on chromosome 7 had two genes within 5 kb, a DUF247 domain protein and a Gypsy-likepolyprotein/retrotransposon putative gene. When we increased our window size to 10 kb we
found more genes, but none related to infection with rhizobia. For example, the *M. truncatula*
ortholog of the outlier on chromosome 5 was close to a DUF679 domain membrane protein and
an alpha/beta fold hydrolase putative gene. The ortholog of the outlier on chromosome 1 had a
reverse transcriptase zinc binding protein and a homeobox knotted-like protein in its 10 kb
window. The ortholog of the outlier on chromosome 7 had a phosphoenolpyruvate carboxylase
within its 10 kb window, along with several putative proteins.

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

**Discussion**

We combined phenotypic and genomic approaches to test for local adaptation of *M.
lupulina* to its mutualistic nitrogen-fixing bacteria across its eastern North American range.
Although our results confirm that *M. lupulina* performs poorly without any rhizobia, we found
no evidence for adaptation to the local rhizobia species in our reciprocal transplant experiment
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 strongcline 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,
we also found very weak evidence of local adaptation in this analysis. The loci that were highly
differentiated between plants hosting different *Ensifer* species (the top 1% of loci in the $X^T X$
outlier analysis) were not associated with any genes involved in the legume-rhizobia symbiosis
in either the range-wide or Ontario samples. Instead, the *M. truncatula* orthologs of these loci
were genes encoding proteins involved in cellular structure (transmembrane protein, TPR repeat
protein), or cellular chemical reactions such as DNA binding (TLD-domain nuclear protein),
RNA binding (CRS1/YhbY CRM domain protein), protein transport (transportin 1 protein), and
oxidation-reduction reactions (FAD/NAD(P)-binding oxidoreductase family protein) (Young *et

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

It is improbable that the loci identified in our genome scan are novel symbiosis genes
underlying adaptation to the local bacteria, although our data are subject to caveats common to
genome scans for selection (Pavlidis *et al.*, 2012). Outliers identified in genome scans are rarely
the causal loci; they are in linkage disequilibrium with the actual genes underlying local
adaptation. However, none of the *M. truncatula* orthologs of our outlier loci were located within
the scale of linkage disequilibrium (5-10 kb in *M. truncatula*) (Branca *et al.*, 2011) from known
symbiosis genes. Second, a significant portion of our outlier loci did not match any annotated
gene in the *M. truncatula* annotated genome. It is possible that the relevant part of the *M.*
*truncatula* genome has not yet been annotated, or that the loci fall between annotated genes and
may perform unknown regulatory functions. Finally, a few outlier loci aligned poorly to the *M.*
*truncatula* genome. If these poorly aligned loci were symbiosis genes that are specific to *M.*
*lupulina* and divergent from *M. truncatula*, using the *M. truncatula* genome as the reference
would bias us against inferring local adaptation from our genomic data.

However, the existence of *M. lupulina*-specific symbiosis genes is unlikely. The plant
genes involved in symbiotic interactions with rhizobia in the *Medicago* system are well
characterized and highly conserved in legumes (Rostas *et al.*, 1986; van Rhijn & Vanderleyden,
1995; De Mita *et al.*, 2006; Branca *et al.*, 2011; Gorton *et al.*, 2012; Stanton-Geddes *et al.*,
2013). *Medicago lupulina* is a close relative of *M. truncatula* (Bena, 2001; Yoder *et al.*, 2013),
and both plants fix nitrogen with both *Ensifer* species tested in our experiment (Béna *et al.*,
2005). It is therefore unlikely that *M. lupulina*-specific mutualism genes underlie adaptation to
different *Ensifer* species.

**Local adaptation in the legume-rhizobia symbiosis**

Our phenotypic and genomic data indicate that *M. lupulina* is not adapted to the local
rhizobia across its eastern North American range. The absence of local adaptation in this
mutualism is surprising given that the system is characterized by several features that ordinarily
strongly favor its evolution. Genotype-by-genotype interactions commonly occur between a
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
Petrie for greenhouse assistance, and Adriana Salcedo, Michelle Afkhami, and Rebecca Batstone for advice on experimental design and analysis.
References


Hadfield, J. 2015. MCMCglmm Course Notes.


Data accessibility

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

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

<table>
<thead>
<tr>
<th></th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeds</td>
<td>Seeds Rhizobia</td>
</tr>
<tr>
<td>(MCMC GLMM)</td>
<td>(MCMC GLMM) Region</td>
</tr>
<tr>
<td></td>
<td>(MCMC GLMM) Rhizobia × region</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass</td>
<td>Biomass Rhizobia</td>
<td>1.955</td>
<td>1, 5.097</td>
</tr>
<tr>
<td>(LMM)</td>
<td>(LMM) Region</td>
<td>0.131</td>
<td>1, 12.782</td>
</tr>
<tr>
<td></td>
<td>(LMM) Rhizobia × region</td>
<td>3.747</td>
<td>1, 248.656</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flowering time</td>
<td>Flowering time Rhizobia</td>
<td>0.016</td>
<td>1, 5.436</td>
</tr>
<tr>
<td>(LMM)</td>
<td>(LMM) Region</td>
<td>0.252</td>
<td>1, 12.896</td>
</tr>
<tr>
<td></td>
<td>(LMM) Rhizobia × region</td>
<td>1.378</td>
<td>1, 164.795</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Wald $\chi^2$</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of flowering</td>
<td>Prob. of flowering Rhizobia</td>
<td>0.012</td>
<td>1</td>
</tr>
<tr>
<td>(GLMM)</td>
<td>(GLMM) Region</td>
<td>0.047</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(GLMM) Rhizobia × region</td>
<td>0.231</td>
<td>1</td>
</tr>
<tr>
<td>Trait</td>
<td>Model</td>
<td>Wald $\chi^2$</td>
<td>df</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------</td>
<td>---------------</td>
<td>----</td>
</tr>
<tr>
<td>Nodules</td>
<td>Rhizobia</td>
<td>0.107</td>
<td>1</td>
</tr>
<tr>
<td>(GLMM)</td>
<td>Region</td>
<td>5.581</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Researcher</td>
<td>95.079</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rhizobia × region</td>
<td>34.806</td>
<td>1</td>
</tr>
</tbody>
</table>

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).
Table 2. Summary statistics of Bayenv2 and BLAST results for the top 1% of SNPs in the XT
outlier analysis.

<table>
<thead>
<tr>
<th>SNP identity</th>
<th>BLAST cover</th>
<th>E value</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.20</td>
<td>37.4</td>
<td>0.089</td>
<td>transmembrane protein</td>
</tr>
<tr>
<td>3.83</td>
<td>93.3</td>
<td>1.00E-18</td>
<td>indole-3-glycerol phosphate lyase IGL1</td>
</tr>
<tr>
<td>3.82</td>
<td>93.3</td>
<td>1.00E-18</td>
<td>toprim domain protein</td>
</tr>
<tr>
<td>3.75</td>
<td>107</td>
<td>6.00E-23</td>
<td>TLD-domain nuclear protein</td>
</tr>
<tr>
<td>3.73</td>
<td>71.6</td>
<td>4.00E-12</td>
<td>no result</td>
</tr>
<tr>
<td>3.65</td>
<td>35.6</td>
<td>3.10E-01</td>
<td>no result</td>
</tr>
<tr>
<td>3.59</td>
<td>107</td>
<td>6.00E-23</td>
<td>TPR repeat protein</td>
</tr>
<tr>
<td>3.56</td>
<td>41</td>
<td>7.00E-03</td>
<td>no result</td>
</tr>
<tr>
<td>3.55</td>
<td>89.7</td>
<td>2.00E-17</td>
<td>no result</td>
</tr>
<tr>
<td>3.44</td>
<td>95.1</td>
<td>4.00E-19</td>
<td>no result</td>
</tr>
<tr>
<td>3.43</td>
<td>111</td>
<td>5.00E-24</td>
<td>transportin-1 protein</td>
</tr>
<tr>
<td>3.39</td>
<td>87.8</td>
<td>6.00E-17</td>
<td>no result</td>
</tr>
<tr>
<td>3.39</td>
<td>96.9</td>
<td>1.00E-19</td>
<td>indole-3-glycerol phosphate lyase IGL1</td>
</tr>
<tr>
<td>3.33</td>
<td>69.8</td>
<td>2.00E-11</td>
<td>copia-like polyprotein/retrotransposon</td>
</tr>
<tr>
<td>3.33</td>
<td>69.8</td>
<td>2.00E-11</td>
<td>no result</td>
</tr>
<tr>
<td>3.33</td>
<td>91.5</td>
<td>5.00E-18</td>
<td>single-stranded nucleic acid-binding protein</td>
</tr>
<tr>
<td>3.29</td>
<td>37.4</td>
<td>8.90E-02</td>
<td>cysteinyl-tRNA synthetase</td>
</tr>
<tr>
<td>3.26</td>
<td>98.7</td>
<td>3.00E-20</td>
<td>no result</td>
</tr>
<tr>
<td>3.25</td>
<td>no result</td>
<td>no result</td>
<td>no result</td>
</tr>
<tr>
<td>3.25</td>
<td>78.8</td>
<td>3.00E-14</td>
<td>no result</td>
</tr>
</tbody>
</table>
## Range-wide sample

<table>
<thead>
<tr>
<th>Value</th>
<th>Accession</th>
<th>Confidence</th>
<th>Matches</th>
<th>E-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.25</td>
<td>825707</td>
<td>96.9</td>
<td>100</td>
<td>1.00E-19</td>
<td>FAD/NAD(P)-binding oxidoreductase family protein</td>
</tr>
<tr>
<td>3.06</td>
<td>1131532</td>
<td>89.7</td>
<td>100</td>
<td>2.00E-17</td>
<td>novel plant SNARE-like protein</td>
</tr>
<tr>
<td>3.02</td>
<td>1482582</td>
<td>35.6</td>
<td>39</td>
<td>0.31</td>
<td>ASCH domain protein</td>
</tr>
<tr>
<td>2.99</td>
<td>405789</td>
<td>89.7</td>
<td>93</td>
<td>2.00E-17</td>
<td>CAAX amino terminal protease family protein</td>
</tr>
<tr>
<td>2.99</td>
<td>511074</td>
<td>48.2</td>
<td>96</td>
<td>5.00E-05</td>
<td>no result</td>
</tr>
<tr>
<td>2.97</td>
<td>175058</td>
<td>55.4</td>
<td>96</td>
<td>3.00E-07</td>
<td>no result</td>
</tr>
<tr>
<td>2.95</td>
<td>485278</td>
<td>91.5</td>
<td>93</td>
<td>5.00E-18</td>
<td>no result</td>
</tr>
<tr>
<td>2.94</td>
<td>254373</td>
<td>59</td>
<td>100</td>
<td>3.00E-08</td>
<td>CRS1/YhbY (CRM) domain protein</td>
</tr>
<tr>
<td>2.92</td>
<td>1090953</td>
<td>55.4</td>
<td>95</td>
<td>3.00E-07</td>
<td>galactose oxidase</td>
</tr>
<tr>
<td>2.90</td>
<td>921907</td>
<td>93.3</td>
<td>100</td>
<td>1.00E-18</td>
<td>DUF223 domain protein</td>
</tr>
<tr>
<td>2.86</td>
<td>1071597</td>
<td>98.7</td>
<td>100</td>
<td>3.00E-20</td>
<td>no result</td>
</tr>
<tr>
<td>2.82</td>
<td>870953</td>
<td>35.6</td>
<td>51</td>
<td>3.10E-01</td>
<td></td>
</tr>
<tr>
<td>2.81</td>
<td>1590352</td>
<td>102</td>
<td>92</td>
<td>3.00E-21</td>
<td>no result</td>
</tr>
<tr>
<td>2.75</td>
<td>686219</td>
<td>69.8</td>
<td>100</td>
<td>2.00E-11</td>
<td>no result</td>
</tr>
<tr>
<td>2.73</td>
<td>774291</td>
<td>69.8</td>
<td>95</td>
<td>2.00E-11</td>
<td>no result</td>
</tr>
<tr>
<td>2.71</td>
<td>313573</td>
<td>69.8</td>
<td>98</td>
<td>2.00E-11</td>
<td>no result</td>
</tr>
<tr>
<td>2.68</td>
<td>1342844</td>
<td>60.8</td>
<td>90</td>
<td>8.00E-09</td>
<td></td>
</tr>
<tr>
<td>2.67</td>
<td>1217147</td>
<td>89.7</td>
<td>100</td>
<td>2.00E-17</td>
<td>novel plant SNARE-like protein</td>
</tr>
<tr>
<td>2.64</td>
<td>616912</td>
<td>84.2</td>
<td>100</td>
<td>7.00E-16</td>
<td>no result</td>
</tr>
<tr>
<td>2.64</td>
<td>1057909</td>
<td>66.2</td>
<td>70</td>
<td>2.00E-10</td>
<td>no result</td>
</tr>
</tbody>
</table>
Table 3. Outlier loci found in the top 1% of Bayenv2 results in both the *M. lupulina* range-wide and Ontario samples.

<table>
<thead>
<tr>
<th>X²X</th>
<th>SNP identity</th>
<th>BLAST identity</th>
<th>Query cover</th>
<th>E value</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.26</td>
<td>1071597</td>
<td>98.7</td>
<td>100</td>
<td>3.00E-20</td>
<td>no result</td>
</tr>
<tr>
<td>3.44</td>
<td>616912</td>
<td>95.1</td>
<td>98</td>
<td>4.00E-19</td>
<td>no result</td>
</tr>
<tr>
<td>3.33</td>
<td>686219</td>
<td>69.8</td>
<td>100</td>
<td>2.00E-11</td>
<td>no result</td>
</tr>
</tbody>
</table>
Table 4. Base pair distances between the *M. truncatula* ortholog of the outlier locus on chromosome 5 and well-characterized nodulation and rhizobial infection genes in *M. truncatula*.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Distance (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFP (Nod-factor receptor 5)</td>
<td>12397</td>
</tr>
<tr>
<td>NSP (Nodulation receptor kinase-like protein)</td>
<td>6319</td>
</tr>
<tr>
<td>LYK3 (LysM receptor kinase K1B)</td>
<td>18</td>
</tr>
<tr>
<td>NIN (Nodule inception protein)</td>
<td>24</td>
</tr>
<tr>
<td>IPD3 (Cyclops protein putative)</td>
<td>8469</td>
</tr>
</tbody>
</table>
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.
Figure 2. Least squares means and 95% confidence intervals for northern (black) and southern (white) plants grown in the two rhizobia treatments. *Ensifer medicae* is the locally abundant rhizobia in the north, and *E. meliloti* is the locally abundant rhizobia in the south. (A) Number of seeds; (B) aboveground biomass; (C) flowering time; (D) number of nodules.