Tansley insight

Evolving together, evolving apart: measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants

Liana T. Burghardt

Department of Plant and Microbial Biology, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, St Paul, MN 55108, USA

Summary

Most plant–microbe interactions are facultative, with microbes experiencing temporally and spatially variable selection. How this variation affects microbial evolution is poorly understood. Given its tractability and ecological and agricultural importance, the legume–rhizobia nitrogen-fixing symbiosis is a powerful model for identifying traits and genes underlying bacterial fitness. New technologies allow high-throughput measurement of the relative fitness of bacterial mutants, strains and species in mixed inocula in the host, rhizosphere and soil environments. I consider how host genetic variation (G × G), other environmental factors (G × E), and host life-cycle variation may contribute to the maintenance of genetic variation and adaptive trajectories of rhizobia—and, potentially, other facultative symbionts. Lastly, I place these findings in the context of developing beneficial inoculants in a changing climate.

I. Introduction

Bacteria are a ubiquitous part of life and play an essential, and often beneficial, role in plant communities. Until recently, however, it has been easier to measure selection on plants than selection on bacteria. This is particularly true for bacteria that form facultative associations with plants—living both with and without hosts. Selection outside hosts can be crucial, as fitness in variable environments is a product of fitness in and frequency of exposure to each environment. High-throughput sequencing allows efficient measurement of bacterial fitness in complex, ecologically realistic environments, transforming our understanding of the genetic basis and environmental dependency of bacterial adaptation and serving to estimate parameters for modeling plant–microbe interactions.

The mutualism between rhizobial bacteria and leguminous plants has long been used to study the functional genetic basis of biological nitrogen (N) fixation and the evolutionary ecology of mutualisms (Heath & Grillo, 2016). This relationship is vital for plant productivity and nutrient cycling in both agricultural and natural systems (Ciais et al., 2013). In nature, rhizobia experience a complex series of selective environments. The symbiotic interaction begins when seeds germinate. As legume roots grow,
the rhizobia living in the soil colonize the root surface; only a tiny fraction of these rhizobia will invade plant tissue via root hairs and initiate nodule development (Fig. 1a). As the nodule develops, the invading rhizobial population expands and begins to fix N₂ into NH₃ in exchange for resources, including photosynthate. The specificity and functional details (e.g. how rhizobia enter roots) of these interactions vary widely. The extent to which these details influence the evolution of the mutualism itself is an active area of research (Sachs et al., 2018). Regardless, evolutionary changes in rhizobial populations that have measurable effects on rhizobial and host traits can occur over the timescale of a few host generations either through shifts in the frequency of lineages or through de novo mutations (Tang et al., 2012; Hollowell et al., 2016).

Most assays of rhizobial adaptation are conducted with single strains of bacteria. However, in soils, several rhizobial genera and scores of strains coexist in a single location (Graham, 2008; Rangin et al., 2008; Ndungu et al., 2018). Here, I review techniques to measure relative rhizobial fitness and provide an overview of our current understanding of rhizobial adaptation to the soil, the rhizosphere and nodules. I describe the influence of additional sources of variable selection, including host genetic variation, environmental variation and variation in host life cycles. Understanding how this variation affects mutualism stability and bacterial evolution is vital for predicting larger-scale responses to climate change and harnessing beneficial plant–microbe interactions in conservation and agriculture. These methodological advances also open new lines of inquiry of broad interest to evolutionary ecologists, including adaptation to variable environments, phenotypic plasticity and the maintenance of genetic variation.

II. Legumes and rhizobia: new tools drive discoveries

Legume–rhizobia systems provide tractable models for studying plant–microbe interactions, from identifying genes to predicting nutrient cycling. The annual, diploid legume Medicago truncatula (a close relative of alfalfa) and its symbiont Ensifer meliloti were the first legume–rhizobia pair with reference genomes and resequenced strain panels, which enable the use of genome-wide association studies (GWAS) to identify associations between genomic variants (e.g. single nucleotide polymorphisms) and phenotypes of interest. Decades of functional genetic work provide a strong foundation for contextualizing results from studies of rhizobia in ecologically

![Diagram](image-url)

**Fig. 1** Rhizobial fitness in the legume–rhizobia mutualism. (a) The major environments experienced by rhizobia in nature. N, ammonia (NH₃); C, malate (C₄H₆O₅). (b) Examples of pairwise and more complex mixed-inoculation experiments. (c) Illustration of the difference between nodule occupancy (proportion of nodules inhabited by each rhizobial lineage) and composite fitness (relative representation of a rhizobial lineage across all pooled nodules). Nodule photo credit: Diana Trujillo.
realistic conditions. For instance, host $DMI$ and bacterial $Nod$ genes are important for nodule formation, and host $DNF$ and bacterial $Nif/Fix$ genes are necessary for N fixation (Kang et al., 2016). Development of molecular and genomic resources in other legume–rhizobia systems is proceeding rapidly. Genome assemblies are now available for > 10 legume species and hundreds of rhizobial strains. Transcriptomic, metabolomic and proteomic databases, tools for genome modification and microscopy techniques (reviewed in diCenzo, 2019) are expanding scales of inquiry.

Understandably, much research has focused on assessing rhizobial traits in laboratory cultures or when a single strain grows with a single host, but these assays may not measure the traits or identify the genes that influence fitness in strain mixtures. Measurement of relative rhizobial fitness in mixed inoculations is difficult because of the challenge of distinguishing specific bacteria (species, strains and mutants) after selection. Previously, this was done for simple mixtures (two to three) via marker genes (e.g. 16S–23S RFLP), fluorescent tags or antibiotic resistance (Fig. 1b). Sequencing technology has made scaling up to complex mixtures possible: the fitness effects of individual mutations can be assayed via insertion sequencing (In-Seq), also known as transposon sequencing (Tn-Seq) (e.g. Pobigaylo et al., 2008; for more details, see Poole et al., 2018); fitness differences between strains can be assayed via ‘select and resequence’ (S&R, overview in Box 1; Burghardt et al., 2018); and fitness differences between microbial species/genera can be quantified using curated markers or ribosomal sequences (e.g. Vuong et al., 2017).

One limitation of conducting scaled-up mixed-inoculation experiments with legume hosts is that the process of nodule infection itself creates a strong sampling bottleneck, owing to each nodule being founded by only one or, at most, a few lineages. If there are many possible nodule inhabitants in the mixed inoculum, it is necessary to sample hundreds or even thousands of nodules to overcome stochastic processes, but if there are only two or three possible nodule occupants, far fewer nodules need to be sampled. Researchers assay the outcome of mixed inoculations in hosts in two complementary ways. Most early mixed-inoculum experiments assayed each nodule individually to determine the proportion of nodules formed by each inoculum member (i.e. nodule occupancy). By contrast, pooling all nodules from a plant before rhizobial identification provides a composite measurement of fitness that includes both which rhizobia form nodules and differences in rhizobial population sizes within and between nodules (larger nodules often have more rhizobia; Fig. 1c). Pooling can thus be used to identify genes and gene functions that influence the composite trait of relative rhizobial fitness (Box 1). When combined with data on the individual benefit that each strain provides to a host, mixed-inoculum experiments can assess whether hosts increase the fitness of strains that are beneficial to them and vice versa. The ability to efficiently measure fitness alignment between species in different environmental contexts has the potential to greatly advance our understanding of mutualism evolution (Friesen, 2012).

### Box 1 Using ‘select and resequence’ to identify key traits and genes underlying rhizobial fitness

The *Medicago–Ensifer* system (also known as *Sinorhizobium*) is particularly suitable for identifying the genomic basis of rhizobial fitness. One reason is that *Medicago* almost exclusively associates with one of two *Ensifer* lineages allowing the development and use of genome-wide association study (GWAS) panels for both host and microbe.

**Tools and symbiotic details of the *Medicago–Ensifer* system.**

<table>
<thead>
<tr>
<th>Symbiotic partners</th>
<th>Genetic model</th>
<th>Genetic model: <em>M. truncatula</em> (barrel medic)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Medicago</em> species</td>
<td>Forage/cover crops: <em>M. sativa</em> (alfalfa), <em>M. lupulina</em> (black medic), <em>M. polymorpha</em> (burr medic)</td>
<td></td>
</tr>
<tr>
<td>Host life cycle</td>
<td>Primarily winter annual and perennial, some biennials and summer annuals</td>
<td></td>
</tr>
<tr>
<td>Rhizobial genera</td>
<td><em>Ensifer</em> (all hosts), <em>Neorhizobium</em> (three host species), <em>Rhizobium</em> (one host species)</td>
<td></td>
</tr>
<tr>
<td>Symbiosis type</td>
<td>Root hair infection, indeterminate nodules, symbiosomes, terminal differentiation</td>
<td></td>
</tr>
<tr>
<td>Model system tools</td>
<td>Both: multiple genome assemblies, expression atlases, transformation</td>
<td></td>
</tr>
</tbody>
</table>


**Select and resequence (S&R) experiment:** Burghardt et al. (2018) measured relative strain fitness of 101 strains of bacteria by extracting DNA from nodules of two host genotypes and applying a maximum likelihood-based, haplotype reconstruction method to high-coverage, whole-genome single nucleotide polymorphism frequencies (Kessner et al., 2016). They found a weak, positive relationship between strain fitness across host genotypes, suggesting a partially shared genetic basis for strain fitness inside nodules of closely related host accessions.

**Identifying the genomic basis of bacterial fitness in nodules with GWAS:** Burghardt et al. (2018) found that bacterial fitness in hosts was associated with variation in genes involved in bacterial motility, host recognition and N fixation. While some of these variants are tightly linked to variants in other genes, a benefit of bacterial GWAS is that it is relatively easy to validate candidates.

**Genomic basis of fitness alignment:** One host genotype (A17) associated with more beneficial bacterial strains than the other (Burghardt et al., 2018). By looking for overlap between top GWAS candidates for plant size in single-strain inoculations and strain fitness, Epstein et al. (2019) identified bacterial genes, including *NifA* and *QueC*, that may influence both traits.

**Moving forwards:** S&R studies should expand to multiple generations, additional ecological contexts and various community complexities with nodule sampling based on age, morphology or functionality. This approach can also be used to screen for strain specificity of known legume mutants with impairment in nodule formation or nitrogen fixation.
III. Relative measures matter: selection on rhizobia in the soil, the rhizosphere and host plants

Rhizobia are frequently studied in the context of mutualisms with legumes, but each lineage will also exist as free-living bacteria in the soil and the rhizosphere (Denison & Kiers, 2011; Masson-Boivin & Sachs, 2017) (Fig. 1a). This facultative lifestyle means that a rhizobial lineage is exposed to temporally and spatially varying environments (Heath & Stinchcombe, 2013). Researchers are beginning to use the mixed-inoculum techniques outlined earlier to integrate this variation into our understanding of natural selection on rhizobia.

Mixed-inoculation experiments conducted on diverse legume hosts suggest that a vast majority of nodules are occupied (or, at the very least, dominated) by a single rhizobial lineage. However, coinfection rates >10% have been reported in glasshouse and laboratory experiments (e.g. Sachs et al., 2010; Checcucci et al., 2016), with rates of coinfection tending to increase with inoculum density (Daubech et al., 2017). Proxies of strain fitness from single-strain assays (e.g. nodule number) are rarely predictive of nodule occupancy in mixed-inoculation experiments (Mellor et al., 1987), and industrial inoculated strains are often outcompeted by rhizobia indigenous to agricultural fields (Box 2; Ndungu et al., 2018). Rates of nodule occupancy can be altered by deleting genes involved in biofilm formation and motility (Caetano-Anolles et al., 1988; Frederix et al., 2014), altering the copy number of an rRNA operon that influences growth rate (Crook et al., 2012), and modifying the presence of plasmids (Cherni & Perret, 2019). Select and resequence assays (Box 1) of pooled nodules from mixed-inoculation experiments in Ensifer meliloti confirmed the lack of correlation between single-strain and mixed-inoculum fitness metrics (Burghardt et al., 2018) and, along with a Tn-Seq screen (Pobigaylo et al., 2008), identified genes involved in motility, host recognition and N fixation, as well as genes never before implicated in symbiosis as determinants of strain fitness.

Few mixed-inoculation experiments have been conducted in soil and rhizosphere environments. Studies of the rhizosphere reveal the importance of bacterial travel towards and attachment to root surfaces – implicating quorum sensing, biofilms, exopolysaccharides, chemotaxis and motility (Frederix et al., 2014; Salas et al., 2017). In a large-scale metagenomic study, Levy et al. (2018) found many of these same genes and gene functions to be enriched in plant-associated bacteria compared with bacteria that live in other environments. Manipulative mixed-inoculation soil experiments (e.g. Saeki et al., 2017) provide a path for confirmation of results from culture environments while generating and testing new hypotheses.

Cumulative fitness of a microbial lineage that experiences multiple environments is a product of fitness in each environment (e.g. host and soil), but most mixed-inoculation experiments are conducted in a single environment. Simultaneous assessment of fitness in nodules, the rhizosphere and the soil will determine the magnitude and range of fitness correlations across environments. For instance, Burghardt et al. (2018) found very weak correlations in strain fitness between host and soil environments, but a strong tradeoff between fitness in the soil and in culture. The strength and directionality (positive or negative) of fitness correlations across environments have important implications for adaptive trajectories and inoculum development for agriculture (see Box 2; Kaminsky et al., 2019). Together with measures of rhizobial population size in each environment, measures of relative fitness provide comprehensive insights into selection pressures on rhizobia (e.g. Daubech et al., 2017).

IV. Selection as a result of host genetic variation and environmental variation

Rhizobia also experience temporally variable selection as a result of changes in host identity, resource availability, temperature, moisture and the larger biotic community. Both interspecific (Vuong et al., 2017; Pahua et al., 2018) and intraspecific (Heath & Tiffin, 2007; Rangin et al., 2008; Crook et al., 2012) genetic variations of hosts affect selection on rhizobia and can cause rank-order shifts in mixed-inoculum experiments. If legume species or genotypes co-occur, then host-specific fitness could promote the coexistence of multiple rhizobial strains. Even variation at a single host gene can have dramatic effects on strain occupancy and fitness in nodules (Kim et al., 2015) and on the frequency of microbial genera in the rhizosphere (Zgadzaj et al., 2016). These effects often depend on the rhizobial lineage and can be a result of genetic variation in a single rhizobial gene (Wang et al., 2018).

Box 2 Applications for agriculture

Legumes have been used to improve soil fertility for thousands of years and actively managed for human benefit for over a century (O’Callaghan, 2018; Kaminsky et al., 2016), and yet the rate of biological nitrogen (N) fixation in agriculture remains inconsistent across years and fields. Thus, ecologically and economically costly N fertilizer is still applied to many legume crops to maximize yields. There is currently intense interest and investment in developing new microbial inoculants to improve crop yield and stress tolerance and manage disease and pests (Busby et al., 2017). One significant difference between natural and agricultural systems is that a single crop variety is planted each year and not allowed to reseed. Thus, crops cannot coevolve with the rhizobia in the soil. Maintaining a mutually beneficial relationship requires both partners to have high fitness. However, legume breeding programs often choose agricultural strains based primarily on plant benefit rather than rhizobial benefit. Below, I highlight how a better understanding of ecological and evolutionary drivers of rhizobial fitness in agricultural systems could help to surmount these challenges that emerge in the large-scale deployment of rhizobia.

Failure to establish: Strains are inoculated onto fields but never make it to high enough densities to have a phenotypic effect on plant traits (Kaminsky et al., 2016). One path forward is to determine which bacterial traits influence bacterial fitness in different soil types (instead of in culture).

Competition problem: Indigenous strains of bacteria in the soil outcompete inoculant strains in terms of associating with hosts (Triplett & Sadowsky, 2016). One path forward is to determine which genes underlie strain competition for nodule formation.

Mutualism breakdown: Over time, strains in agricultural fields become less beneficial. One path forward is to find genes that influence the fitness of both organisms and breed for plants that specifically increase the fitness of beneficial bacteria (i.e. for fitness alignment).
Rhizobia have a primarily positive effect on plant growth, but the strength of the effect depends on environmental factors, such as N availability (Frederickson, 2017; Gibert et al., 2018). Less clear is the extent to which rhizobial fitness depends on environmental factors. The emerging pattern from mixed-inoculation experiments suggests that nutrients, water and temperature can have measurable effects on the outcome of competition for occupancy and fitness in nodules, but these effects are relatively small compared with the influence of host genetic variation (Regus et al., 2014; Vuong et al., 2017; Wendlandt et al., 2019; but see Shiro et al., 2016). A general understanding of how strain fitness, plant fitness and N fixation change with climate and land-use changes will require comprehensive experiments spanning legume systems. Future research should examine how additional direct and indirect biotic players, such as phages (Hashem & Angle, 1990), other competing bacteria (Xiao et al., 2017), arbuscular mycorrhizal fungi (Ossler et al., 2015), nematodes (Wood et al., 2018) and herbivores (Paudel & Bede, 2015) modify interactions between rhizobia and legumes. Such a systems-level approach will allow functional characterization of the interactions of specific community members and advance understanding of the natural history of microbes (Busby et al., 2017; Martiny & Walters, 2018).

V. Selection as a result of variation in legume life cycles

Although genes that directly affect rhizobia inside hosts have garnered most of the attention, other plant traits can profoundly alter selection on rhizobia (Fig. 2). For instance, changes in length and timing of plant life cycles alter the timing of bacterial exposure to selection outside hosts. These life-cycle shifts also alter the timing of nodule formation and bacterial release – understudied processes in both natural and agricultural environments. For instance, a perennial host that forms long-lived nodules creates a far more stable selection regime than one that forms new nodules and releases rhizobia back into the soil every year (Fig. 2c). In contrast to natural systems, the timing and frequency of legume presence in agricultural systems are determined by management practices such as planting times, crop rotations and cover cropping (Box 2). A handful of experimental evolution studies have examined bacterial adaptation after exposure to multiple generations of hosts (Guidot et al., 2014; Marchetti et al., 2017; Meaden & Koskella, 2017; Burghardt et al., 2019). Manipulative approaches like these could be adapted to address questions about the influence of host life cycles on rhizobial evolution.

Legumes exhibit wide variation in life cycles (Berger & Ludwig, 2014; Berger et al., 2017). Shifts in life-cycle expression can be a result of exposure to environmental variation (e.g. climate change) or of genetic changes in key transitions such as germination or flowering (Burghardt et al., 2015). Excitingly, many of the genes underlying flowering time and perenniality in Arabidopsis (Andrés & Coupland, 2012) underlie flowering-time variation in legumes (Weller & Ortega, 2015). Moving forward, legumes with mutations in flowering-time genes will enable the study of how gene networks underlying N fixation integrate with environmentally sensitive flowering-time pathways and above-ground processes more generally.

VI. Conclusions

I have argued that understanding and leveraging the legume–rhizobia mutualism requires an explicitly evolutionary approach, one that is widely applicable to other facultative plant–microbe
relationships – both pathogenic and beneficial. Methodological advances in the tracking of bacterial lineages in complex mixed inocula can be applied in other host–microbe systems, including plant pathogens (e.g. *Pseudomonas* Baltrus et al., 2017; Melnyk et al., 2019), bioluminescent bacteria in squid (e.g. *Vibrio Euprymna* Pankey et al., 2017), and members of the human gut microbiome (Garud et al., 2019). Recent advances in legume–rhizobium research can also be used as a model to understand the evolution and ecology of microbes involved in facultative symbiosis more generally. Explicit modeling of facultative relationships will allow prediction of how mutualistic interactions in nature will respond to environmental change (e.g. climate change). Similarly, the role played by host life cycles in shaping selection on bacteria will not be restricted to mutualists or plants (Barrett et al., 2008). Pursuing these lines of inquiry raises the possibility of scaling from genes and functional traits to ecosystem-level nutrient cycling relevant to maximizing yield and nutrient retention in croplands. In summary, taking an explicitly ecological and evolutionary perspective that incorporates the true variation in environments experienced by bacteria affords a fresh perspective on critical environmental and agricultural challenges of our time.

**Acknowledgements**

Thanks to P. Tiffin, L. D. Leverett, K. T. Burghardt, M. R. Wagner, J. Coughlan, E. Josephs, T. Bell and R. T. Batstone for helpful comments on early manuscript drafts. I apologize to those researchers whose thoughtful work was omitted as a result of space constraints. This work was supported by the National Science Foundation under grants IOS-1237993 and IOS-1724993. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

**ORCID**

Liana T. Burghardt  [https://orcid.org/0000-0002-0239-1071](https://orcid.org/0000-0002-0239-1071)

**References**


