What Makes the Rhizobia-Legume Symbiosis So Special?[1]

Ann M. Hirsch*, Michelle R. Lum, and J. Allan Downie

Department of Molecular, Cell, and Developmental Biology (A.M.H., M.R.L.) and Molecular Biology Institute (A.M.H.), University of California, Los Angeles, California 90095; and Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom (J.A.D.)

"We are too prone to give all credit to him who places the last piece in the puzzle and to forget that all his predecessors had prepared the way."

—Edwin W. Fred, Ira L. Baldwin, and Elizabeth McCoy (1932)

Ever since the identification by Hellriegel and Wilfarth (1888) of rhizobia as the source of fixed nitrogen in root nodules of legumes, people have wondered whether or not plants outside the Fabaceae could be manipulated to associate with rhizobia. The development of nodules, the keystone of Hellriegel and Wilfarth's findings, has since become the "Holy Grail" of the field of biological nitrogen fixation. It is well known that the rhizobia-legume interaction falls into cross inoculation groups, whereby certain rhizobial strains nodulate only certain legumes. For example, *Sinorhizobium meliloti* effectively nodulates species of *Medicago*, *Melilotus*, and *Trigonella*, whereas *Rhizobium leguminosarum* bv *viciae* induces nitrogen-fixing nodules on *Pisum*, *Vicia*, *Lens*, and *Lathyrus* spp. Closely related to the pea (*Pisum sativum*) strain is *R. leguminosarum* bv *trifolii,* which initiates nodules only on species of clover (*Trifolium*).

However, not all rhizobia strain-legume associations are this tight. For example, *Rhizobium* strain NGR234 nodulated 232 species of legumes from 112 genera tested and even nodulated the nonlegume *Parasponia andersonii*, a member of the elm family (Pueppke and Broughton, 1999). On the opposite end of the spectrum, not all members of the legume family nodulate. Of the three different subfamilies of legumes—Caesalpinoideae, Mimosoideae, and Papilionoideae—members of the basal subfamily, Caesalpinoideae, are mostly non-nodulating (Nod^-). Thus, nodulation and presumably nitrogen-fixing ability are not 100% correlated even within the legume family. Nodulation may have originated multiple times in the Fabaceae: once in the only caesalpinoid that is confirmed to be nodulated, *Chamaecrista* genus; once in the mimosoid line; and lastly, at the base of the papilionoid line (Doyle, 1998). Alternatively, there may have been a single origin of nodulation with multiple losses. In any case, more than 90% of the Papilionoideae and Mimosoideae are nodulated, whereas less than 25% of the Caesalpinoideae form nodules.

Moreover, other plant families can establish interactions with nitrogen-fixing bacteria exclusive of rhizobia. Members of eight different families, known as actinorhizal plants, are nodulated by *Frankia* spp. nitrogen-fixing actinomycetes. Various grasses, including such agronomically important ones as sugarcane (*Saccharum officinarum*), maize (*Zea mays*), and rice (*Oryza sativa*), associate with different nitrogen-fixing bacteria, among them species of *Glucoacetobacter*, *Azospirillum*, *Herbaspirillum*, and *Azoarcus*; these associations do not, however, result in nodule formation. Some of the positive responses on plant growth exhibited by these so-called "associative" nitrogen-fixing interactions are due to the production of phytohormones, but nitrogen fixation has also been demonstrated (Sevilla et al., 2001). Nevertheless, what makes rhizobia and *Frankia* spp. different from the associative nitrogen-fixing microbes is that most of the rhizobia or *Frankia* spp.-fixed nitrogen is transferred to and assimilated by the plant for the plant's growth. Is forming a new organ, the root nodule, essential for the evolution of this type of nitrogen assimilation?

The legumes and their association with *Rhizobium* spp. in the broad sense have always been extremely important agronomically. The use of crop rotations to enhance the productivity of nonlegume crops was vividly described by the Romans, who were probably aware of an even older tradition in Greece. Moreover, this nonpathogenic association between prokaryote and eukaryote is a fascinating phenomenon for investigation of basic biological principles. In this review, we address the fundamental question of this agriculturally and environmentally important symbiosis: "What makes this association so unique that only legumes form a symbiosis with rhizobia?"

[1] This work was supported by the U.S. Department of Agriculture, by the National Science Foundation, and by the University of California-Mexus program (to A.M.H.'s laboratory). J.A.D. was supported by the Biotechnology and Biological Sciences Research Council.

* Corresponding author; e-mail ahirsch@ucla.edu; fax 310–206–5413.

www.plantphysiol.org/cgi/doi/10.1104/pp.010866.

LEGUMES ARE A UNIQUE FAMILY

The evidence for the evolution of the legumes (Fabaceae), the third largest family of flowering plants, is fragmentary, at least based on fossil evidence. There are no obviously identifiable nodules associated with fossils that can be accurately described as legume roots. The mostly leaf fossils date from the Cretaceous era, which has been variably dated as 65 to 145 million years ago (MYA). Thus, we do not know how long ago the first legumes started to associate with rhizobia.

A phylogenetic analysis using a chloroplast gene sequence showed that the legumes and actinorhizal plants (nodulated by *Frankia* spp.) belong to the Rosid I clade, and suggested that there was a single origin for a predisposition for nodulation in this lineage (Soltis et al., 1995). However, it is unclear as to what this predisposition entails. Does it mean that the plants have unique receptors or unusual cell walls? Do these plants produce certain types of signal molecules to entice the symbiont or to repress various types of defense molecules, thus enabling the symbiosis to occur? Do they have different phytohormones or phytohormone sensitivities? The Soltis et al. (1995) study used *rbcL*, and other organellar sequences have been utilized as well to study the relationships of angiosperm genera. Would nuclear gene sequences generate the same results? If there was a predisposition for nodulation, then why do the vast majority of the plants in the Rosid I clade not associate with nitrogen-fixing organisms?

If all nodules are derived from a common progenitor, how do nodules of the legumes differ from those of other plant groups? Although the ontogeny of the various actinorhizal nodules is not identical, the nodules are developmentally and anatomically more related to lateral roots than are legume nodules. Nevertheless, the legume nodule shares more traits with a lateral root than with any other plant organ (Hirsch and LaRue, 1998). Legume and actinorhizal nodules can be indeterminate, growing by means of an apical meristem, but determinate nodules, those lacking a persistent apical meristem, are only found in legumes. Moreover, some legumes such as lupins or *Sesbania rostrata* develop nodules that fall into an intermediate category. Unlike the lateral root that is initiated from cell divisions in the pericycle, the legume nodule originates from cell divisions in the outer or inner cortex, depending on whether a determinate or indeterminate nodule is formed.

In the next section, we will address several questions. What are the features that enabled legumes to be predisposed to nodulation? What traits are unique to legumes and not found in other species within the Rosid I line? What are the genes in the two partners that enable the symbiosis to occur? Is there the remotest possibility that the "Holy Grail" can be attained?

Flavonoids: Signals and Modulators of Nodule Development

More than 4,000 different flavonoids have been identified in vascular plants, and a particular subset of them is involved in mediating host specificity in the legumes (Perret et al., 2000). All flavonoids consist of two benzene rings linked through a heterocyclic pyran or pyrone ring (Fig. 1). Specific substitutions on the ring produce flavonols, flavones, flavanones, as well as isoflavonoids, which are derived from a migration of the B ring from the 2 to the 3 position (Fig. 1A). Isoflavonoids are limited to the legume family. Daidzein and genistein (Fig. 1B), isoflavonoids produced by soybean (*Glycine max*), are effective inducers of *Bradyrhizobium japonicum nod* genes, but inhibit *S. meliloti nod* gene expression. *S. meliloti nod* genes can be induced by luteolin (Fig. 1C). This specificity enables rhizobia to distinguish their hosts from other legumes. The specific flavonoid not only induces *nod* gene expression, but also rhizobial chemotaxis. Nevertheless, other than the isoflavones, most flavonoids are not unique to legumes. How do soil rhizobia recognize their host

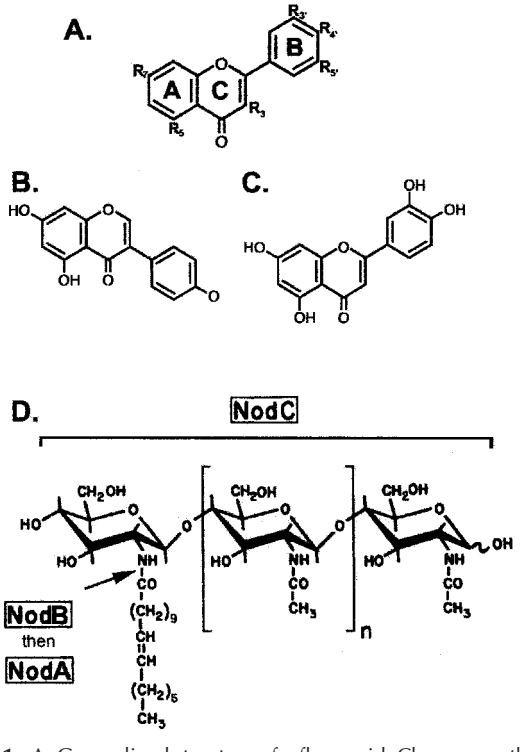

Figure 1. A, Generalized structure of a flavonoid. Changes on the ring or in the R groups result in flavonols, flavones, flavanones, and glycosylated flavonoids, among others. B, An isoflavonoid inducer, genistein. C, A flavone inducer, luteolin. D, Generalized structure of a Nod factor. NodC, a β-glucosaminyl transferase, links the UDP-*N*-acetyl glucosamine monomers into a chitin-like backbone. NodB removes an acetyl group from the terminal residue of the chitin oligomer. Then, NodA catalyzes the transfer of a fatty acyl chain onto the resulting free amino group, using acyl-ACP from fatty acid biosynthesis.

and initiate the symbiosis when nonlegume plant species growing in the same area are also sources of flavonoids? Apparently, it is the next stage, once the flavonoids are perceived, where another level of specificity comes into play.

Flavonoids are perceived as aglycones, which induce rhizobial *nod* genes by interacting with the gene product of *nodD*, a LysR-type regulator. This interaction results in a conformational change in the NodD protein such that it binds to *nod* box elements in the promoters of the *nod* genes (see Perret et al., 2000). The concerted expression of these genes leads to the synthesis of Nod factor molecules—lipochitooligosaccharides that usually consist of four or five *N*-acetylglucosamines, β-1–4 linked, with the terminal nonreducing sugar *N*-acylated with a fatty acid of 16 to 18 carbon residues (Fig. 1D). Nod factors can be chemically modified with acetate, sulfate, or carbamoyl groups, or can have different sugars, such as Ara, Man, Fuc, or substituted Fuc. The degree of saturation of the acyl tail can also vary (Perret et al., 2000). The assemblage of these substitutions result in a specific Nod factor that is recognized by a particular legume.

Nod Factor Responsiveness of Legumes

One of the key traits that differentiates the nodulating legumes from other plant species is their responsiveness to Nod factor. Early responses to Nod factor include ion flow across the plasma membrane and an associated depolarization of the membrane, followed by periodic oscillations in intracellular calcium, referred to as "calcium spiking"; these are followed by deformation of root hairs and initiation of cortical cell division (for review, see Downie and Walker, 1999). Root hair curling (which involves entrapment of the bacteria) and infection thread growth require the presence of the bacteria. Nodulation appears to have an absolute requirement for Nod factor because rhizobia that do not synthesize Nod factor do not nodulate, and legume mutants that are incapable of perceiving Nod factor or transducing it along a signal transduction pathway are Nod^- (Downie and Walker, 1999). The identity of a Nod factor receptor(s) in legumes is unknown, but research is well under way (see Cullimore et al., 2001). Its identity may be key to finding the "Holy Grail."

A biochemical approach has led to the characterization of high-affinity binding sites for Nod factors. One of these, NFBS2, is located in the plasma membrane and exhibits different selectivities for Nod factors in alfalfa (*Medicago sativa*) and bean (*Phaseolus vulgaris*; Gressent et al., 1999; J.J. Bono, personal communication). In *Dolichos biflorus*, an unusual lectin with Nod factor-binding activity has been characterized (Etzler et al., 1999). This protein has an apyrase activity and has been named lectin nucleotide phosphohydrolase. It is not clear, however, whether Nod factor-binding proteins are unique to legumes. Furthermore, to date, there is no evidence for linking a Nod factor-binding protein to a Nod^- mutation.

The study of Nod^- plant mutants has also yielded leads for identifying the proteins involved in Nod factor perception and signal transduction. Many Nod^- mutants have been identified in commercially and agronomically important legumes such as pea, bean, alfalfa, sweetclover, and others, and more recently, model legumes such as *Medicago truncatula* and *Lotus japonicus* have been used for genetic studies (Stougaard, 2001). Transposon tagging in *L. japonicus* led to identification of *NIN* (nodule inception), which encodes a transcription factor, and the first cloned gene that is directly involved in nodule development (Schauser et al., 1999). *nin* mutants showed root hair curling in response to rhizobia, but did not develop infection threads or nodules. In recent unpublished data, G.B. Kiss and colleagues positionally cloned a gene called *NORK* (nodule receptor kinase) from an alfalfa Nod^- mutant (MN1008) that shows neither Ca^{2+} spiking nor root hair deformation in response to rhizobia (Ehrhardt et al., 1996). Kiss and colleagues did a chromosome walk to the *Nn1* locus (mutated in MN1008) using a combination of bacterial artificial chromosome clones from *M. truncatula* (a diploid) and markers from alfalfa (a tetraploid). They identified the mutated *NORK* gene in a position equivalent to the previously mapped *Sym19* locus (pea) and also to the *Dmi2* locus (*M. truncatula*; G.B. Kiss, personal communication). *NORK* encodes a Leu-rich repeat kinase that could be a receptor, but so far it has not been shown whether or not it directly interacts with Nod factor. It is possible that NORK could interact with a Nod factor-binding protein via its Leu-rich repeats; a postulated position for this locus in nodulation signaling is shown in Figure 2.

The Mycorrhizal Connection: Not Specific to Legumes, But a Useful Correlation

Since the original coupling of Nod^- and Myc^- (inability to establish a mycorrhizal association) in pea and *Vicia faba* mutants by Duc et al. (1989), a number of papilionoid legumes that normally nodulate have been shown to be both Nod^- and Myc^-. Figure 2 illustrates the Nod^-Myc^- connection for *L. japonicus* (*Lj*), pea (*Ps*), alfalfa (*Ms*), *M. truncatula* (*Mt*; for references, see Marsh and Schultze, 2001; Stougaard, 2001) and white sweetclover (*Ma*; Lum et al., 2002). A Myc^- mutant allele has also been described for bean. The various mutants can be described as those that are blocked very early in the signaling cascade (before Ca^{2+} spiking) and those blocked later (after Ca^{2+} spiking). Those Nod^- Myc^- legume mutants blocked before Ca^{2+} spiking are presumed to be altered in a receptor that is common to the mycorrhizal and nodulation pathways (*Pssym8*, *Pssym19*, *Mtdmi1*, *Mtdmi2*, and MN1008). A group of

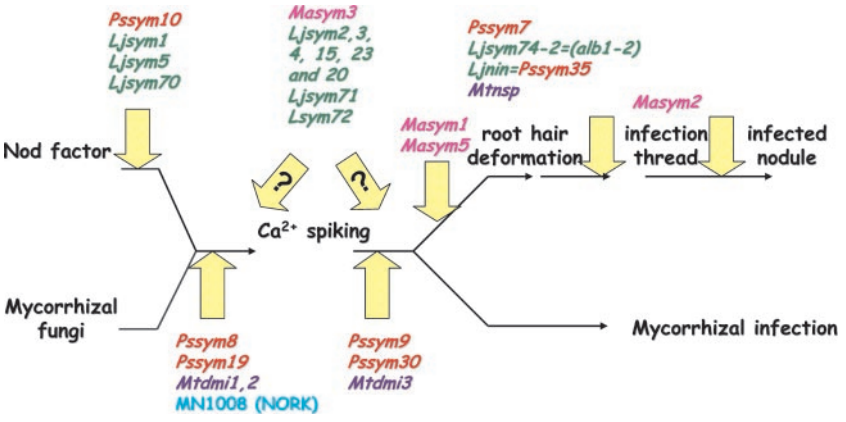

Figure 2. Analysis of various papilionoid legume non-nodulation mutants for their mycorrhizal phenotypes. A common pathway is observed where the two lines converge. The mutants are ordered on this pathway according to their Ca^{2+} spiking response, where known. Some of these mutants have not been tested for Ca^{2+} spiking response yet, nor have allelism tests been performed. The *Masym1* mutants are mostly Myc^{-}, but one mutant allele, BT62, which is leaky, is Myc^{+}. The one *Masym5* mutant is also Myc^{-} (Lum et al., 2002). The mycorrhizal phenotypes of *Pssym7* and *Mtnsp* have not been reported. Except for *Ljnin*, none of the genes have been cloned. The diagram is based on that of Walker et al. (2000). Green, *L. japonicus*; red, pea; pink, *Melilotus alba*; blue, *M. truncatula*; turquoise, alfalfa. The yellow arrows indicate the approximate stage where the mutants are blocked.

mutants, *Masym3* and the *L. japonicus sym* mutants, have not been tested yet for Ca^{2+} spiking, or if they have, not all mutant alleles have been examined. Therefore, it is not known whether they are blocked before or after Ca^{2+} spiking. In addition, the results on the *L. japonicus sym* mutants have been generated from several different laboratories, so it is not yet known whether some of the mutations are allelic.

Pssym10, *Ljsym1*, *Ljsym5*, and *Ljsym70* mutants are $Nod^{-}Myc^{+}$ and presumably blocked before Ca^{2+} spiking (to date, the *Ljsym* mutants have not been tested) and upstream of the NORK-type receptor (Fig. 2). This upstream gene(s) could encode a Nod factor-binding protein and/or a nodulation specific signaling protein. Also, many legume mutants that are very likely blocked after Ca^{2+} spiking, but before infection thread formation, have been found. They include those mutated in *NIN*, which has been characterized as a transcription factor (Schauser et al., 1999). Others blocked after infection thread development may have mutations in genes that encode elements of the signal transduction pathway leading to nodule morphogenesis.

The connection between mycorrhizal development and nodulation as well as the fact that the early nodulin (*ENOD*) genes are expressed in both symbioses suggest that nodulation may have evolved from the more ancient mycorrhizal condition (see later section). The legumes have given a special insight into the mycorrhizal association by enabling identification of genes required for this symbiosis. Of course, it is predicted that other mutations are likely to affect the mycorrhizal symbiosis, but not nodulation. However, such mutants have yet to be described. The mycorrhizal association is believed to have originated more than 400 MYA based on fossil evidence, and so it is possible that the nodulation symbiosis may have adapted some components of a much older symbiotic pathway.

The Specificity of Legume Lectins

For the infection thread to form, there has to be an intimate connection between the rhizobial cell surface and the plant cell wall. Based on the strong correlation between the inoculation specificity of bacteria of the family Rhizobiaceae on their legume hosts, and the ability of host-produced lectins to bind to *Rhizobium* sp. cells, the lectin recognition hypothesis was formulated to explain why alfalfa and *S. meliloti* or soybean and *B. japonicum* or any other legume and its nitrogen-fixing rhizobial species are symbiotic partners (for references, see Hirsch, 1999). Lectins frequently follow the various cross inoculation groups established by their host legumes due to their different carbohydrate-binding specificities. Soybean lectin (SBA or SBL), a galactosamine-binding protein, differs from pea lectin (PSA or PSL), a Glc-/Man-binding protein, and they both differ from other legume lectins.

Could the lectins that appear to be characteristic of their legume hosts be involved in infection thread formation or nodulation? Work on transgenic legume plants carrying a foreign lectin gene strongly suggested that the introduced lectin enhanced rhizobial attachment, infection thread formation, and nodulation in response to heterologous rhizobial strains (Kijne et al., 1997; van Rhijn et al., 1998, 2001). However, the heterologous rhizobia must produce the compatible Nod factor for the host legume; otherwise, no nodules develop. The requirement for the compatible Nod factor suggests that the introduced

lectin may be facilitating bacterial attachment, in so doing causing a localized increase in Nod factor concentration at the site of bacterial entry. However, lectins do not bind Nod factors, so the introduced lectin must be interacting with some other component(s) of the rhizobial cell surface. Different rhizobial strains have characteristic cell surfaces consisting of capsular polysaccharide, exopolysaccharide, and lipopolysaccharide. Neither *Bradyrhizobium elkanii* USDA31, which does not bind SBL, nor *exoB* mutants of *B. japonicum* attached to *Lotus corniculatus* roots or induced nodules over the non-transgenic and vector control levels (van Rhijn et al., 1998). Similarly, inoculation with an exopolysaccharide-deficient mutant of *R. leguminosarum* bv *viciae* did not result in infection threads or nodules on transgenic alfalfa plants carrying the PSL gene (van Rhijn et al., 2001). These data suggest that some component, which is missing or changed in the *B. japonicum* or *R. leguminosarum* bv *viciae exo* mutants, may be a ligand for the introduced lectin.

Nevertheless, a larger question that remains is whether the legume lectins are absolutely essential for nodulation. Would introduction of a legume lectin into a nonlegume result in significant rhizobial colonization such that a condition similar to associative nitrogen fixation arises? Thus far, we do not have an answer to this question.

Legume Genes and Gene Regulation: Unique Domains?

A number of early nodulin (*ENOD*) and nodulin (*NOD*) genes have been identified based on what was thought to be their exclusive expression in the nodule. However, it has now become clear that many of these genes are in fact expressed in nonsymbiotic tissues and/or during nonsymbiotic conditions. For example, *ENOD40* is an early nodulin gene induced within hours of *Rhizobium* sp. inoculation and its expression appears to be critical for proper nodule development (Charon et al., 1999). However, *ENOD40* transcripts are also found localized in the stele of the stem, lateral roots, and in other tissues. In addition, *ENOD40* homologs have now been identified in nonlegumes, including rice, a monocotyledon outside the Rosid I clade (Kouchi et al., 1999), although so far not in Arabidopsis (H. Kouchi, personal communication; M.R. Lum, N.A. Fujishige, and A.M. Hirsch, unpublished data). Similarly, plant hemoglobins were long thought to be nodule-specific proteins, but homologs have now been found in rice and Arabidopsis, among others (Arrendondo-Peter et al., 1997; Trevaskis et al., 1997).

There are some nodulin genes that appear to be novel, such as some of the peribacteroid membrane proteins, which may have originated due to gene duplications and/or recombination (Verma et al., 1991). However, it seems that many of the genes involved in nodule development and nitrogen fixation were recruited from their original task in plant growth and development to function in the nodule. Understanding how these genes are regulated may contribute to our understanding of what makes legumes unique. Recent data indicate that some of the regulatory genes have domains that may be found exclusively in legumes, such as the *Hy5* homolog, *LjBZF* (M. Kawaguchi, personal communication) and a DNA-binding protein, *VsENBP1* (E.Ø. Jensen, personal communication). Are there other genes with legume-specific regulatory domains and are these regions critical for nodulation?

WHERE DID RHIZOBIA EVOLVE FROM?

How did bacteria acquire the ability to establish a symbiosis with legumes? In the absence of a bacterial fossil record, it is difficult to date speciation within bacteria. However, analysis of evolutionary changes in highly conserved genes can be used as a "molecular clock." Such studies suggest that the fast-growing rhizobia (e.g. *Rhizobium* sp. and *Sinorhizobium)* diverged around 200 to 300 MYA, whereas divergence between fast-growing rhizobia and slow-growing bradyrhizobia occurred around 500 MYA. These times are earlier than the split between monocots and dicots (156–171 MYA) and the separation of brassicas and legumes (125–136 MYA). Therefore, rhizobia appear to have diverged well before the existence of legumes and probably before the appearance of angiosperms (Turner and Young, 2001). Therefore, nodulation capacity is thought to have been acquired after bacterial divergence and horizontally spread among different genera.

This concept is strongly supported by the recent finding that *Burkholderia* strain STM678 can nodulate legumes (Moulin et al., 2001). This genus is in a completely different subdivision (β) of the proteobacteria from the rhizobia (α-subdivision), and so these bacteria are essentially unrelated. Nevertheless, the nodulation genes are clearly similar to those from rhizobia (Moulin et al., 2001). The question as to where "rhizobia" evolved from can be restated as: "What are the unique elements that enable rhizobia to establish a symbiosis" or, perhaps more specifically, "where did the nodulation genes come from"?

Nodulation Genes Are Unique Qualities of Rhizobia

The ability to fix atmospheric N_2 is very widespread among bacteria and Archaea, although interestingly, this capacity is restricted to prokaryotes. Therefore, there are many different diazotrophs that, if equipped with the ability to invade plants, could theoretically evolve to establish a nitrogen-fixing symbiosis. One change, however, would be to uncouple the regulation of nitrogen fixation in planta from the microbial requirement for fixed nitrogen, something rhizobia have done very efficiently.

As mentioned earlier, there are many bacteria that grow endophytically within plants, but what distinguishes the rhizobia is their ability to make "Nod factors," molecules required to program the specialized infection process and nodule morphogenesis. The biosynthesis of Nod factors has been thoroughly reviewed (see Perret et al., 2000). Although Nod factors can carry many substituents, which are important for nodulating specific legumes, their basic structure requires the action of only three gene products, NodA, NodB, and NodC (Fig. 1D). NodC is an *N*-acetyl-glucosaminyl transferase that produces the chitin backbone from UDP-*N*-acetyl glucosamine. NodB removes an acetyl group from the terminal residue of the chitin oligomer, and NodA catalyzes the transfer of a fatty acyl chain onto the resulting free amino group, using acyl-ACP from fatty acid biosynthesis.

The origin of the *nodA*, *nodB*, and *nodC* genes therefore may be crucial. It is likely that they came from outside the Rhizobiaceae because, like most of the nodulation and nitrogen fixation genes, they have a G + C content that is significantly lower than the average G + C content of rhizobia; they also have a different codon usage from most chromosomal genes (Galibert et al., 2001). NodC is one of a large class of bacterial β-glucosyl transferases, many of which can incorporate *N*-acetyl glucosamine into cell wall polysaccharides. For example, *Streptococcus pyogenes* produces a polymer of alternating β-1,4-linked GlcNAc and GlcUA. Furthermore, the peptidoglycan of many bacteria is composed of a backbone of alternating β-1,4-linked GlcNAc and N-acetyl muramic acid (which is the lactic acid ether of GlcNAc). It is possible that a NodC-like protein could have evolved from such a bacterial enzyme. There are several NodB-like proteins in databases, and it is easy to imagine how a simple glucosamine-deacetylase like NodB could have been recruited.

The potential origin of NodA is an enigma and its function is unusual because it adds a fatty acyl chain to a preformed polysaccharide. Almost all bacterial fatty acylated polysaccharides studied are produced by incorporating acylated sugars during elongation of the polysaccharide. NodA-like proteins are special because thus far they have been found only in rhizobia and no related proteins are detected in database searches. Perhaps these nodulation genes came from some bacterial source that has yet to be sequenced. The unusual characteristics of NodA may enable us, in the future, to get an insight into what that source may have been.

An alternative view is that the key nodulation genes may have been acquired from fungi. Most fungi make chitin as part of their cell wall and therefore have chitin synthases, which are similar to NodC. Some fungi contain endosymbiotic bacteria. More significantly, one of the endomycorrhizal fungi, which can infect plant roots using a pathway that seems to share steps in common with nodulation, was found to contain a *Burkholderia* strain that harbored nitrogen fixation genes (Minerdi et al., 2001). This, taken together with the finding that a related *Burkholderia* strain can nodulate, may be a significant coincidence. However, *Burkholderia* spp. typically have a G + C content similar to rhizobia and so are unlikely to be the source of the low G + C symbiosis genes found in rhizobia.

What Do the Rhizobial Genomic Sequences Tell Us?

The complete sequences of *S. meliloti* (Galibert et al., 2001) and *Mesorhizobium loti* (Kaneko et al., 2000) have recently been completed and provide a wealth of data. Both genomes are large (6.7 and 7.6 Mb, respectively) and there is clustering of many genes known to be required for the symbiosis. In *M. loti*, many of the symbiosis genes are located on a chromosomal symbiosis island of 611 kb, whereas in *S. meliloti*, most of the symbiosis genes are located on either of two large plasmids, pSymA (1.35 Mb) or pSymB (1.7 Mbp). The location of symbiosis genes on "islands" or plasmids reinforces the idea that these regions have the potential to be horizontally transferred. Although the pSym plasmids of *S. meliloti* are not transmissible, they are clearly related to other highly transmissible plasmids. Earlier work on the symbiosis island of a *M. loti* strain demonstrated that this was an exceptionally efficient mechanism of transferring nodulation capacity to Nod^- bacteria in field experiments (Sullivan et al., 1995). The mechanism of excision and integration of the symbiosis island out of and into the chromosome has been established to occur via integration into a Phe-tRNA (Sullivan and Ronson, 1998).

It is surprising that 35% of *M. loti* genes have no orthologs in *S. meliloti*, and this diversity is further exemplified by the finding that over 50% of the genes on the 536-kb symbiosis plasmid of NGR234, a strain very closely related to *S. meliloti*, have no orthologs in *S. meliloti*. In fact, the most different region in the comparison of the predicted gene products of *M. loti* with *S. meliloti* (Galibert et al., 2001) corresponded to the symbiosis island! This suggests that although some very highly conserved nodulation and nitrogen fixation genes are required for symbiotic nitrogen fixation, many different genes are specifically required to optimize interactions with different legume hosts. Therefore, it is difficult to generalize, although it is evident that rhizobia have numerous solute transporters and are rich in catabolic genes, presumably enabling them to compete successfully in the rhizosphere and in soil.

The uneven distribution of insertion elements, intergenic mosaic elements, percent G + C, and altered codon usage on pSymA shed light on the evolution of *S. meliloti* (Galibert et al., 2001). Thus, a typical aerobic heterotrophic bacterium may have first greatly

extended its metabolic potential by acquisition of pSymB. The subsequent gain of pSymA conferred the ability to infect plants, form nodules, successfully colonize the low oxygen environment of the nodule, and thereafter fix nitrogen. However, both *S. meliloti* and *M. loti* seem to have acquired highly evolved symbiotic gene packages, and so we are still left with the conundrum about how the process originally started.

IS CELL DIVISION, I.E. MAKING A NODULE, ESSENTIAL?

Most of the research on rhizobia-legume symbioses has focused on papilionoid legumes and their symbionts, many of which have been selected for agronomic performance. Thus, several discoveries relevant to papilionoid legumes may not apply to all symbiotic interactions, particularly for understanding the evolution of nodulation. For example, extended infection threads are required for pea and alfalfa nodulation, but there are examples where infection threads are almost nonexistent and bacteria spread interstitially as in peanut (*Arachis hypogaea*; Chandler, 1978). In some tropical legumes and also in *Parasponia* sp., rhizobia are not released into membrane-bound symbiosomes; rather, they fix nitrogen within specialized fixation threads (for references, see Hirsch and LaRue, 1998). Is fixation thread development and nodule morphogenesis a prerequisite for this nitrogen-fixing symbiosis or could accumulations of bacteria between cells, such as what occurs in associative nitrogen-fixing interactions, have provided fixed nitrogen in primitive, evolving symbioses? There are reports of nitrogen-fixing bacteria, based on acetylene reduction assays, in more basal legumes (Bryan et al., 1996), but these studies remain preliminary.

Do any of the legume or rhizobial mutants characterized so far shed light on which genes can be dispensed with, yet allow the symbiosis to proceed in what might be considered akin to the primitive condition or ground state? Several of the mutations affecting host-specific modifications of Nod factors delay nodulation, but in many cases, the process continues normally (e.g. Ardourel et al., 1994). However, an interesting phenotype was described for a mutant of *R. leguminosarum* bv *viciae* lacking all of the host-specific nodulation genes but retaining the *nodABC* genes and their regulator (Walker and Downie, 2000). On vetch (*Vicia sativa*), many hundreds of root hairs were heavily infected, but infection threads and nodules were not formed. If similar levels of infection of root hairs on an evolutionarily more basal legume were to occur, and nitrogen fixation could take place within these infected cells, then we could postulate that there might be the potential to provide significant levels of nitrogen to the plant. The ability of many Nod^- legumes to accumulate high levels of nitrogen (McKey, 1994) could argue positively for some sort of non-nodular association with rhizobia. Alternatively, these plants may be efficient nitrogen scavengers. More studies are clearly needed.

The potential to induce cell division and create a nodule greatly enhances the efficiency of the symbiosis. It may be significant that minimalist Nod factor structures can induce early signaling events, whereas more highly substituted Nod factors are required to initiate cell division, nodule primordia, and infection thread structures (van Brussel et al., 1992; Ardourel et al., 1994; Walker and Downie, 2000). This has led to the idea that there may be different levels of recognition of Nod factors. Some plant genes involved in processes related to cell division, such as cell cycle control and nuclear endoreduplication, have been identified (Cebolla et al., 1999; Charon et al., 1999; Roudier et al., 2000), and these may act relatively late in relation to the developmental scheme briefly sketched in Figure 2. Pingret et al. (1998) suggested a role for a G protein-mediated signaling pathway for induction of legume early nodulin genes, based on inhibitor studies and the induction of gene expression by mastoparan, a G protein agonist. However, mastoparan did not induce calcium spiking in root hairs (Walker et al., 2000), and taken at face value, this would imply that a role for G protein-mediated signaling could be downstream of calcium spiking.

CONCLUSIONS AND PERSPECTIVES

"Take nature away and all your insight is in a biological vacuum."

—Fernando Nottebohm (2001), as quoted by Spector (2000)

By studying bird song, Nottebohm discovered that cells in the brain can be reactivated to produce new neurons. The implications of his research for treating the consequences of Parkinson's disease, stroke, Alzheimer's syndrome, and spinal cord injury are now being widely discussed (Specter, 2001). Whoever would have thought that the study of how birds learn to sing would have so much application for alleviating human suffering!

The diversity seen in the legumes and their interacting partners is as wide ranging as the difference between the brains of canaries and humans. For the past 20 years, rhizobial and legume biologists have pursued a scientific investigation based on this biodiversity for the purposes of understanding the complexities of the agriculturally and environmentally important nitrogen-fixing symbiosis epitomized by nodulation. Although model systems are valuable because they provide the tools for sophisticated and detailed analysis of one or two species, they cannot fully answer the fundamental questions; for example, the nuances of host specificity and whether or not nodulation/nitrogen fixation can be extended to nonrelated species, particularly plants outside the Rosid

I clade. Legumes are unique in their response to Nod factors in that they actively promote entry of bacteria into the root. However, as yet we do not know how legumes evolved the ability to recognize such signals or how entry is actually accomplished. Attention must be given to the broader aspects of the legume-rhizobia association. Darwin's revolution of biology could not have occurred without the unrestrained view he had of the organisms around him. Variation among individuals gave Darwin the insight to understand the origin of species. By recognizing that the evolution of the rhizobia-legume symbiosis is more akin to an interwoven tapestry than to a continuous thread, we may have a better chance of understanding the uniqueness of this association. Thus, genome projects and scientific pursuits that include a diversity of legumes and rhizobial species will better inform us as to which genes/proteins are conserved among all hosts and symbionts and help us determine whether the ability to fix N_2 into ammonia can be transferred to crops other than legumes. The "Holy Grail" awaits.

ACKNOWLEDGMENTS

We thank John G. Streeter for his insights, and we thank our colleagues who answered our e-mails and requests for information and unpublished data. We are also grateful to Stefan J. Kirchanski for his comments on the manuscript. We apologize to all those authors whose papers we have not cited because of space constraints. As stated in the first quote, we appreciate all the hard work that has gone before us.

Received September 24, 2001; accepted September 24, 2001.

LITERATURE CITED

Ardourel M, Demont N, Debelle F, Maillet F, de Billy F, Promé JC, Dénarié J, Truchet G (1994) *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. Plant Cell **6:** 1357–1374

Arrendondo-Peter R, Hargrove MS, Sarath G, Moran JF, Lohrman J, Olson JS, Klucas RV (1997) Rice hemoglobins: gene cloning, analysis, and O_2-binding kinetics of a recombinant protein synthesized in *Escherichia coli*. Plant Physiol **115:** 1259–1266

Bryan JA, Berlyn GP, Gordon JC (1996) Towards a new concept of the evolution of symbiotic nitrogen fixation in the Leguminosae. Plant Soil **186:** 151–159

Cebolla A, Vinardell JM, Kiss E, Olah B, Roudier F, Kondorosi A, Kondorosi E (1999) The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants. EMBO J **18:** 4476–4484

Chandler MR (1978) Some observations on the infection of *Arachis hypogaea* L. by *Rhizobium*. J Exp Bot **29:** 749–755

Charon C, Sousa C, Crespi M, Kondorosi A (1999) Alteration of *enod40* expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*. Plant Cell **11:** 1953–1966

Cullimore JV, Ranjeva R, Bono JJ (2001) Perception of lipo-chitooligosaccharidic Nod factors in legumes. Trends Plant Sci **6:** 24–30

Downie JA, Walker SA (1999) Plant responses to nodulation factors. Curr Opin Plant Biol **2:** 483–489

Doyle JJ (1998) Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. Trends Plant Sci **3:** 473–478

Duc G, Trouvelot A, Gianinazzi-Pearson V, Gianinazzi S (1989) First report of non-mycorrhizal plant mutants (Myc^-) obtained in pea (*Pisum sativum* L.) and fababean (*Vicia faba* L.). Plant Sci **60:** 215–222

Ehrhardt DW, Wais R, Long SR (1996) Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. Cell **85:** 673–681

Etzler ME, Kalsi G, Ewing NN, Roberts NJ, Day RB, Murphy JB (1999) A Nod factor binding lectin with apyrase activity from legume roots. Proc Natl Acad Sci USA **96:** 4704–4709

Fred EW, Baldwin IL, McCoy E (1932) Root Nodule Bacteria and Leguminous Plants. University of Wisconsin Studies, Madison, WI

Galibert F, Finan TM, Long SR, Pühler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P et al. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:** 668–672

Gressent F, Drouillard S, Mantegazza N, Samain E, Geremia RA, Canut H, Niebel A, Driguez H, Ranjeva R, Cullimore J et al. (1999) Ligand specificity of a high-affinity binding site for lipo-chitooligosaccharidic Nod factors in *Medicago* cell suspension cultures. Proc Natl Acad Sci USA **96:** 4704–4709

Hellriegel H, Wilfarth H (1888) Untersuchungen über die Stickstoffnahrung der Gramineon und Leguminosen. Beilageheft zu der Ztschr. Ver. Rübenzucker-Industrie Deutschen Reichs

Hirsch AM (1999) Role of lectins (and rhizobial exopolysaccharides) in legume nodulation. Curr Opin Plant Biol **2:** 320–326

Hirsch AM, LaRue TA (1998) Is the legume nodule a modified root or stem or an organ *sui generis*? Crit Rev Plant Sci **16:** 361–392

Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K et al. (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res **7:** 331–338

Kijne JW, Bauchrowitz MA, Díaz CL (1997) Root lectins and rhizobia. Plant Physiol **115:** 869–873

Kouchi H, Takane K, So RB, Ladha JK, Reddy PM (1999) Rice *ENOD40*: isolation and expression analysis in rice and transgenic soybean root nodules. Plant J **18:** 121–129

Lum M, Li Y, LaRue TA, Schwartz RD, Kapulnik Y, Hirsch AM (2002) Investigation of four non-nodulating mutants of white sweetclover (*Melilotus alba* annua Desr.)

mutants and their response to arbuscular-mycorrhizal fungi. Am Zool (in press)

Marsh JF, Schultze M (2001) Analysis of arbuscular mycorrhizas using symbiosis-defective plant mutants. New Phytol **150:** 525–532

McKey D (1994) Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. *In* JI Sprent, D McKey, eds, Advances in Legume Systematics, Part 5. Royal Botanic Gardens, Kew, UK, pp 211–228

Minerdi D, Fani R, Gallo R, Boarino A, Bonfante P (2001) Nitrogen fixation genes in an endosymbiotic *Burkholderia* strain. Appl Environ Microbiol **67:** 725–732

Moulin L, Munive A, Dreyfus B, Boivin-Masson C (2001) Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature **411:** 948–950

Perret X, Stahelin C, Broughton WJ (2000) Molecular basis of symbiotic promiscuity. Microbiol Mol Biol Rev **64:** 180–201

Pingret JL, Journet EP, Barker DG (1998) *Rhizobium* Nod factor signaling: evidence for a G protein-mediated transduction mechanism. Plant Cell **10:** 659–672

Pueppke SG, Broughton WJ (1999) *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. Mol Plant-Microbe Interact **12:** 293–318

Roudier F, Fedorova E, Gyorgyey J, Feher A, Brown S, Kondorosi A, Kondorosi E (2000) Cell cycle function of a *Medicago sativa* A2-type cyclin interacting with a PSTAIRE-type cyclin-dependent kinase and a retinoblastoma protein. Plant J **23:** 73–83

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature **402:** 191–195

Sevilla M, Burris RH, Gunapala N, Kennedy C (2001) Comparison of benefit to sugarcane plant growth and $^{15}N_2$ incorporation following inoculation of sterile plants with *Acetobacter diazotrophicus* wild-type and Nif$^-$ mutant strains. Mol Plant-Microbe Interact **14:** 358–366

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic fixation in angiosperms. Proc Natl Acad Sci USA **92:** 2647–2651

Specter M (2001) Rethinking the brain: how the songs of canaries upset a fundamental principle of science. New Yorker **23:** 42–53

Stougaard J (2001) Genetics and genomics of root symbiosis. Curr Opin Plant Biol **4:** 328–335

Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW (1995) Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc Natl Acad Sci USA **92:** 8985–8989

Sullivan JT, Ronson CW (1998) Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc Natl Acad Sci USA **95:** 5145–5149

Trevaskis B, Watts RA, Andersson CR, Llewellyn DJ, Hargrove MS, Olson JS, Dennis ES, Peacock WJ (1997) Two hemoglobin genes in *Arabidopsis thaliana*: the evolutionary origins of leghemoglobins. Proc Natl Acad Sci USA **94:** 12230–12234

Turner SL, Young JPW (2001) The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. Mol Biol Evol **17:** 309–319

van Brussel AAN, Bakuizen R, van Spronsen PC, Spaink HP, Tak T, Lugtenberg BJJ, Kijne JW (1992) Induction of preinfection thread structures in the leguminous host plant by mitogenic lipooligosaccharides of *Rhizobium*. Science **257:** 70–72

van Rhijn P, Fujishige NA, Lim PO, Hirsch AM (2001) Sugar-binding activity of pea lectin enhances heterologous infection of transgenic alfalfa plants by *Rhizobium leguminosarum* biovar *viciae.* Plant Physiol **126:** 133–144

van Rhijn P, Goldberg RB, Hirsch AM (1998) *Lotus corniculatus* nodulation specificity is changed by the presence of a soybean lectin gene. Plant Cell **10:** 1233–1249

Verma DPS, Hu C-A, Zhang M (1991) Root nodule development: origin, function and regulation of nodulin genes. Physiol Plant **85:** 253–265

Walker SA, Downie JA (2000) Entry of R*hizobium leguminosarum* bv. *viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. Mol Plant-Microbe Interact **13:** 754–762

Walker SA, Viprey V, Downie JA (2000) Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by nod factors and chitin oligomers. Proc Natl Acad Sci USA **97:** 13413–13418