Update on Genomes of Nodule Bacteria

Genomes of the Symbiotic Nitrogen-Fixing Bacteria of Legumes

Allyson M. MacLean, Turlough M. Finan*, and Michael J. Sadowsky

Department of Biology, Center for Environmental Genomics, McMaster University, Hamilton, Ontario, Canada L8S 4K1 (A.M.M., T.M.F.); and Department of Soil, Water, and Climate, BioTechnology Institute and Microbial and Plant Genomics Institute, University of Minnesota, St. Paul, Minnesota 55108 (M.J.S.)

Over the last several decades, there have been a large number of studies done on the genetics, biochemistry, physiology, ecology, and agronomics of the bacteria forming nitrogen-fixing symbioses with legumes. These bacteria, collectively referred to as the rhizobia, are taxonomically and physiologically diverse members of the α and β subclasses of the Proteobacteria, and mostly comprise members of the genera Rhizobium, Bradyrhizobium, Mesorhizobium, Sino-rhizobium, and Azorhizobium (Fig. 1). Most studies have focused on mutational and biochemical analyses to define bacterial genes involved in root-nodule formation, symbiotic specificity, nitrogen fixation, and plant-microbe signal exchange. More recently, however, several genomic approaches have been used to define and understand the involvement of whole bacterial genomes in the symbiotic process. Genetic analyses of the model symbiotic bacterial species Sinorhizobium meliloti, Rhizobium leguminosarum, and Bradyrhizobium japonicum have revealed a few surprises concerning genome evolution and structure, how plant and microbes communicate, and physiological diversity among the microsymbionts of legumes. In this review we discuss what is currently known about the genomes of several rhizobia and how genome-enabled studies have provided insights into the symbiotic interaction of the rhizobia and their respective host legumes.

TRENDS IN RHIZOBIAL GENOMICS

The recent completion of the genomes of R. leguminosarum bv vicieae (Young et al., 2006), Rhizobium etli (González et al., 2006), and two photosynthetic Bradyrhizobium strains (Giraud et al., 2007) has increased the number of available complete rhizobial genome sequences to seven, including sequences obtained for B. japonicum (Kaneko et al., 2002), Mesorhizobium loti (Kaneko et al., 2000), and S. meliloti (Barnett et al., 2001; Capela et al., 2001; Finan et al., 2001; Galibert et al., 2001). The genomic architecture of these species varies considerably (Table I).

It is now apparent that multipartite genomes are typical among members of α-proteobacteria (Jumas-Bilak et al., 1998), although they are not restricted to members of this group. While multipartite genomes are often associated with species that interact with a host (Egan et al., 2005), the advantage afforded by a genome consisting of multiple replicons is not always obvious (Teyssier et al., 2004; Egan et al., 2005). However, it has been proposed that such an organization may increase the adaptive potential of a species (González et al., 2006), which, if valid, implies that the presence of multipartite genomes in many rhizobial species may reflect the complex lifestyle adopted by these plant-associated bacteria. The presence of additional replicons in the genome of the α-proteobacteria may enhance the bacterium’s metabolic or symbiotic capabilities, thus ensuring stable maintenance of the replicon within the population.

Genome size is influenced by environmental factors, and soil-dwelling species, such as the rhizobia, tend to have larger genomes (Bentley and Parkhill, 2004). B. japonicum has the largest chromosome size, approximately 9.2 Mb, and the difference in chromosome sizes among the rhizobia may in part be due to the presence of extrachromosomal (plasmid) DNA. While the bradyrhizobia generally lack plasmids, they are a common feature among the fast-growing microsymbionts. The complexity and heterogeneity of soil presumably necessitates a large inventory of genes to maximize survival of free-living cells (Bentley and Parkhill, 2004; Young et al., 2006), while the ability to establish a symbiotic relationship with a host plant imposes an additional genetic requirement upon rhizobia. Accordingly, it has been proposed that rhizobial genomes have evolved via expansion, primarily through lateral gene transfer and gene duplication, as a means of adjusting to the challenges imposed by this lifestyle (Batut et al., 2004; Boussau et al., 2004).

Symbiotically relevant genes in rhizobia are often clustered on large plasmids (pSym), or within genomic islands (referred to as symbiosis islands [SIs]), emphasizing
the accessory nature of the genes and their ability to be acquired via horizontal gene transfer. A SI present in *M. loti* strain ICMP3153 was found capable of transforming nonsymbiotic strains of *M. loti* into symbiotic counterparts (Sullivan and Ronson, 1998). While 611 and 681 kb SIs were also reported in *M. loti* strain MAFF303099 (Kaneko et al., 2000) and *B. japonicum* strain USDA110 (Kaneko et al., 2002), respectively, the transmissibility of these islands has yet to be confirmed. Intriguingly, the integration of SIs into the *M. loti* and *B. japonicum* genomes occurs within phe-tRNA and val-tRNA genes, respectively. The association of SIs in both species with a phage-related integrase implies that SIs may have originated from the ancient integration of a bacteriophage. In contrast, in *S. meliloti*, *R. etli*, *R. leguminosarum*, and *Sinorhizobium* sp. strain NGR234, the majority of genes involved in symbiosis (*nod*, *nif*, and *fix*) are plasmid borne (Freiberg et al., 1997; Barnett

---

**Table 1. Architecture of rhizobial genomes**

<table>
<thead>
<tr>
<th>Species</th>
<th>Genomic Architecture</th>
<th>Size</th>
<th>No. of Predicted ORFs</th>
<th>Accession No.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. japonicum</em></td>
<td>Chromosome</td>
<td>9,105,828</td>
<td>8,317</td>
<td>BA000040</td>
<td>Kaneko et al. (2002)</td>
</tr>
<tr>
<td><em>Bradyrhizobium</em> sp.</td>
<td>Chromosome</td>
<td>8,264, 689, 228,826</td>
<td>7,729</td>
<td>CP000494, CP000495</td>
<td>Giraud et al. (2007)</td>
</tr>
<tr>
<td>strain BTAi1</td>
<td>Chromosome, pBTAi1</td>
<td>8,264, 689, 228,826</td>
<td>7,729</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Bradyrhizobium</em> sp.</td>
<td>Chromosome</td>
<td>7,456,587</td>
<td>6,752</td>
<td>CU234118</td>
<td>Giraud et al. (2007)</td>
</tr>
<tr>
<td>strain ORS278</td>
<td>Chromosome</td>
<td>7,456,587</td>
<td>6,752</td>
<td>CU234118</td>
<td>Giraud et al. (2007)</td>
</tr>
<tr>
<td><em>M. loti</em></td>
<td>Chromosome, pMLa, pMLB</td>
<td>7,036,071, 351,911</td>
<td>6,752, 320, 209</td>
<td>AP002994–AP03017</td>
<td>Kaneko et al. (2000)</td>
</tr>
<tr>
<td><em>S. meliloti</em></td>
<td>Chromosome, pSymB, pSymA</td>
<td>3,654,135, 1,683,333</td>
<td>3,341, 1,570, 1,293</td>
<td>AL591688, AL591985,</td>
<td>Galibert et al. (2001)</td>
</tr>
<tr>
<td>bv viciae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 1.** Unrooted phylogenetic tree of 16S rDNA sequences from selected rhizobial strains in the α-proteobacteria. Sequences were aligned using ClustalW and the tree was constructed by using the neighbor-joining method. GenBank/EMBL accession numbers of the sequences used for analysis are as follows: CP000133.1, *R. etli* CFN 42; AM236080.1, *R. leguminosarum* bv viciae; EF054890.1, *R. tropici*; D11343.1, *Rhizobium galegae*; BA000012.4, *M. loti* MAFF303099; X77123.2, *Sinorhizobium fredii* strain LMG 8317; X67221.1, *Azorhizobium caulinodans*; EF428944.1, *S. meliloti*; Y17047.1, *Allorhizobium undicola*; AY904770.1, *Bradyrhizobium elkanii* strain SEMIA 6169; BA000040,2, *B. japonicum* USDA 110; CP000494, *Bradyrhizobium* sp. strain BTAi1; CU234118, *Bradyrhizobium* sp. strain ORS278; CP000463.1, *Rhodopseudomonas palustris* BioA53; AB298401.1, *Methylbacterium extorquens*; AF220762.1, *Methylbacterium nodulans*. Rhizobia with fully sequenced genomes are indicated in bold. Rhizobial genomes currently being sequenced are indicated by an asterisk (*).
et al., 2001; Galibert et al., 2001; González et al., 2003, 2006; Young et al., 2006) and several pSymS have the potential or documented ability to be transferred among bacteria via conjugation (Rao et al., 1994; Freiberg et al., 1997; Barnett et al., 2001; González et al., 2003; Brom et al., 2004; Pérez-Mendoza et al., 2004). The organization of symbiotic genes within mobile islands or plasmids permits the conversion of nonsymbiotic saprophytes into nitrogen-fixing plant endosymbionts (and vice versa) in a single step.

Rhizobial genomes appear to be highly dynamic entities and this is particularly reflected by the presence of many insertion sequence (IS) elements, transposases, and related genes, within regions encoding symbiotic functions. The ISs of M. loti and B. japonicum encode the majority of transposase genes and IS-related sequences in these genomes (Kaneko et al., 2000; Göttfert et al., 2001; Kaneko et al., 2002; Sullivan et al., 2002). In B. japonicum, an incredible 60% of all transposase genes (100 of 167) were located within the SI, which comprises a mere 7.5% of the entire genome. The islands of M. loti strains ICMP3153 and MAFF303099 differ in size by >110 kb; much of the interstrain variation is attributable to DNA of external origin (Sullivan et al., 2002). This is consistent with the dynamic nature of the islands (Finan, 2003). Similar trends for the accumulation of IS elements have also been reported for plasmid pSymA of S. meliloti (Galibert et al., 2001), pNGR234a from Sinorhizobium sp. strain NGR234 (Viprey et al., 2000), and plasmid p42d from R. etli (González et al., 2006), which is likely due to the nonessential nature of many of the genes within these genomic regions (Viprey et al., 2000). Reiterated sequences offer potential sites for recombination (Mavingui et al., 2002) and are likely involved in the movement of symbiotic genes within and among the genomes of the root- and stem-nodule bacteria. For example, in R. leguminosarum, almost identical symbiotic regions are present on the plasmids pRL10 and pRL1 (Young et al., 2006). Similarly, genes clustered within the SI of M. loti strain R7A are located on plasmid pMLa in strain MAFF303099 (Sullivan et al., 2002), and an approximately 42-kb region of pSymA was likely transferred to an accessory plasmid (pSymM11a) in S. meliloti (Stiens et al., 2006).

Palacios and colleagues have performed elegant studies of genome plasticity in rhizobia, demonstrating that extensive recombination occurs within the symbiotic plasmid of Sinorhizobium sp. NGR234 (Flores et al., 2000), as well as cointegration of this plasmid with the chromosome and/or megaplasmid of this species (Mavingui et al., 2002). They have also demonstrated cointegration of the S. meliloti megaplasmids with the chromosome (Guo et al., 2003). The plasticity and instability of rhizobial genomes are likely due to the presence of repeated DNA sequences (Flores et al., 1987), as well as IS elements and multiple replicons (Romero and Palacios, 1997). It has been postulated that large-scale recombinational events such as these offer a means of reshuffling genes among replicons over successive generations. It should be noted, however, that this potential source of genetic variation may be limited should excision events preferentially occur within cointegration sites, as has been demonstrated in S. meliloti (Guo et al., 2003).

The multiphasic lifestyle adopted by rhizobia has influenced the size, complexity, and also the content of their genomes (Bentley and Parkhill, 2004; Boussau et al., 2004). This appears to be a common trend among other members of the α-proteobacteria that have a varied lifestyle, including Rickettsia (an obligate intracellular animal pathogen), Bartonella and Brucella (facultative intracellular animal pathogens), Agrobacterium (an extracellular pathogen of plants), and the animal symbiont Wolbachia (Tsolis, 2002). Comparative genomic analyses of members of the α-proteobacteria have shown that these phylogenetically diverse bacteria share many similarities in their metabolism, physiology, overall genome architecture, and processes required for interaction with their cognate hosts (Paulsen et al., 2002; Tsolis, 2002). Interestingly, the obligate intracellular α-proteobacteria have genome structures indicative of prolonged gene loss, whereas the large size and complexity of the genomes of free-living species indicates expansion by both gene duplication and horizontal gene transfer events (Batut et al., 2004; Boussau et al., 2004; Giraud et al., 2007).

Rhizobial genomes are richly endowed with transport, regulatory, and stress-related systems, all of which are essential for interacting with the external environment and were acquired subsequent to the divergence of plant-associated Rhizobiales from related α-proteobacteria (Boussau et al., 2004). Some nodules bacteria, like the recently sequenced Bradyrhizobium sp. strains BTAI1 and ORS278 (Giraud et al., 2007) are amazingly metabolically and ecologically diverse and have very complex genomes. These strains can grow as heterotrophs, autotrophs, phototrophs, and symbiotically with members of the plant genus Aeschynomene. Analysis of the R. leguminosarum genome sequence indicates that this species shares more genes in common with two other plant symbionts, S. meliloti and M. loti, than with the more closely related plant pathogen Agrobacterium tumefaciens, consistent with the observation that genome content (as well as size) is directly influenced by a species’ lifestyle (Young et al., 2006).

In addition to the rhizobia discussed above, genome sequencing projects have also been initiated or completed for a number of plant-associated α-proteobacteria, including Agrobacterium vitis S4, and Agrobacterium radiobacter K84. Genome sequencing of Rhizobium tropici and Sinorhizobium medicae, and the nitrogen-fixing, plant-growth-promoting bacteria Azospirillum brasilense and Gluconacetobacter diazotrophicus, is also nearing completion.

PLASMIDOMICS: THE RHIZOBIA AND PLASMIDS

Since nodulation and nitrogen fixation genes in many rhizobia are plasmid borne, the sequences of
several large rhizobial plasmids were determined prior to the completion of the whole genome scale projects. The first such initiative involved the complete sequencing of the symbiotic plasmid pNGR234a (Freiberg et al., 1997). This work was among the first to report a high frequency of IS and related sequences within rhizobial genomes and led the authors to suggest that lateral gene transfer played a prominent role in the disparate evolution of a plant pathogen and endosymbiont from a common ancestor. The sequence of the symbiotic plasmid (p42d) of R. etli strain CFN42 (González et al., 2003) also reflects the mosaic nature of symbiotic regions in rhizobia. Intriguingly, IS sequences located on either side of a 125 kb symbiotic region and a putative integrase gene raised the possibility that this plasmid-borne symbiotic region may itself be motile. To date, there have been no published reports of plasmids in the slow-growing bradyrhizobia, which is likely due to the fact that no symbiotic genes have been localized to Bradyrhizobium plasmids.

Many rhizobial strains appear to carry one or more nonsymbiotic, accessory plasmids that may confer benefits with regard to overall fitness. One such plasmid, the 144 kb pSmeSM11a from S. meliloti, has been sequenced (Stiens et al., 2006). The majority of genes in pSmeSM11a were predicted to encode for proteins involved in DNA replication, recombination, and repair, as well as genes encoding various metabolic enzymes and transport systems. A gene encoding an 1-amino-cyclopropane-1-carboxylic acid deaminase (acdS) was also identified and it was postulated that AcdS indirectly promotes plant growth by decreasing ethylene levels (Glick, 2005). Consistent with this suggestion is the finding that the introduction of acdS and its regulatory gene into an S. meliloti strain lacking these genes resulted in a strain with considerably superior nodulating ability (Ma et al., 2004). This study demonstrates that genes localized to accessory plasmids may have a role in microbial fitness, and underscores the importance of undertaking similar projects in the future.

**TOWARDS FUNCTIONAL GENOMICS**

The transcriptome for S. meliloti has been examined under a variety of conditions, including in planta (Ampe et al., 2003; Bèrges et al., 2003; Rüberg et al., 2003; Barnett et al., 2004; Becker et al., 2004; Krol and Becker, 2004; Uchiumi et al., 2004; Capela et al., 2005; Barnett and Fisher, 2006; Capela et al., 2006; Domínguez-Ferreras et al., 2006). To simultaneously and systematically measure both plant and bacterial gene expression, Long and colleagues developed a dual-genome chip that carried a probe set specific for the entire S. meliloti genome, and an additional approximately 10,000 gene probe set corresponding to the host plant Medicago truncatula (Barnett et al., 2004). One of the more interesting results obtained from this study was the observation that the majority of tentative plant consensus sequences that were up-regulated in nitrogen-fixing nodules produced by wild-type S. meliloti were similarly expressed in Fix" nodules (i.e. unable to fix nitrogen) induced by a fix’ mutant strain. This result further supports the hypothesis that the nitrogen-fixing ability of colonizing rhizobia is likely not important with respect to plant root and nodule gene expression (Barnett et al., 2004; Barnett and Fisher, 2006). While there have been reports of host sanctions imposed upon rhizobia in nonfixing nodules in soybeans (Glycine max), and that host plants are sensitive to the nitrogen-fixing status of individual nodules (Kiers et al., 2003), this likely takes place long after the establishment of nodules. The functional genomics of rhizobia is ultimately dependent on the ability to rapidly develop mutants in interesting loci identified using transcriptional analyses. In S. meliloti, this issue was addressed by the development of an ORFeome functional genomics platform (Schroeder et al., 2005). In addition, in vivo expression technology has been used for the identification of genes with nodule-inducible expression (Oke and Long, 1999) and to define S. meliloti promoters that were actively expressed during the earliest stages of infection (Zhang and Cheng, 2006).

One of the difficulties inherent in working with plant hosts with indeterminate root nodules (such as M. truncatula and alfalfa [Medicago sativa]) is that colonizing bacteria (and plant cells) may be present at various developmental stages. Such a heterogeneous population makes the identification of genes relevant to a particular stage of symbiosis challenging, and may result in apparent discrepancies between different studies (Barnett et al., 2004). The use of plant and bacterial developmental mutants offers a means of analyzing gene expression at specific stages of infection, albeit with the caveat that the expression patterns of such mutants may not always be representative of that exhibited by the wild-type bacterium. In S. meliloti, a bacA mutant, which fails to survive upon entry into plant cells, permitted expression analysis of cells contained within infection threads (Ampe et al., 2003; Capela et al., 2006). Similar studies are required to tease apart gene expression patterns at each stage of nodulation.

Transcriptomics and functional analyses in other rhizobial species are less developed, although micro- and macroarray studies have been performed in B. japonicum and M. loti, respectively. In B. japonicum, an array was used to identify potential members of the NifA and RegR regulons using probes sets specific for >700 candidate genes (Hauser et al., 2006). Microarray analysis in this species has also been used to examine cell response to iron-limiting conditions, and to identify iron-regulated genes (Rudolph et al., 2006; Yang et al., 2006a, 2006b). Recently, full genome oligo-based microarrays for B. japonicum have been developed (see http://www.bradygenome.org/) and will likely facilitate other functional genomic studies in this bacterium.

Genome-wide macroarray experiments performed upon bacteroids from M. loti strain MAFF303099 resulted in the identification of expression islands scattered across the symbiotic island (Uchiumi et al., 2004).
In contrast, gene expression in areas outside of the island was strongly down-regulated during symbiosis, a pattern similar to that described by Barnett et al. (2004) for *S. meliloti*. The symbiotic islands of *M. loti* strains MAFF303099 and R7A differ by 109 kb, and several expression islands were located within MAFF303099 strain-specific regions. Intriguingly, in silico analyses indicated that many genes whose expression was up-regulated in bacteroids are likely not regulated by RpoN, a σ factor required for the transcription of nitrogen-fixation genes. Uchiumi et al. (2004) have suggested that chromosomal reconfiguration may be used by these bacteria as a means of elevating expression of genes within the symbiotic island.

**OTHER FUNCTIONAL GENOMIC APPROACHES**

While the compilation of genome sequence data represents a significant milestone in the characterization of a bacterial species, assigning gene function to open reading frames (ORFs) is essential to extract information encoded within the nucleotide sequences. One method of addressing this issue is through the construction of an ORFeome, as has been reported in *S. meliloti*. Schroeder et al. (2005) describes the amplification and cloning of all predicted ORFs within the *S. meliloti* genome (6,317) into an entry vector, which has been designed to facilitate the transfer via recombination of the cloned ORF into a destination plasmid(s) (Schroeder et al., 2005). Such destination plasmids may be used in the overexpression of a gene of interest, to monitor expression of the gene, or to generate mutants in which gene expression has been disrupted.

The conditions in which genes are expressed may offer valuable insight into the function of the encoded products. For that reason, Cowie et al. (2006) constructed a gene fusion library of *S. meliloti* that contains reporter genes fused at defined locations to 2,874 (47%) of the genes in the *S. meliloti* genome (see www.sinorhizobium.org). These fusions have been employed for the identification of genes whose expression is regulated by environmental inorganic phosphate (Yuan et al., 2006) and in studies directed to the identification of ligands transported by the 160 solute-binding-dependent uptake systems present in *Sinorhizobium*. Over 100 different inducing conditions were screened and remarkably specific inducers, and hence probable solutes, were detected for over 76 transport systems (Mauchline et al., 2006).

Systematic gene mutagenesis offers another means of determining gene function via loss of function. Luo et al. (2005) reported the targeted disruption of all putative LysR-type regulators in the *S. meliloti* genome, and identified two new regulators (LsrA and LsrB) as being essential for symbiosis. A signature-tagged transposon library, consisting of 12,000 *S. meliloti* mutant strains has also been constructed (Pobigaylo et al., 2006). Such libraries will prove useful in the identification of genes essential for symbiosis and those that are involved in competitive interactions in the early periods of legume nodulation.

**DIRECT GENOME FUNCTIONALITY REVEALED BY PROTEOMIC ANALYSES**

Several proteomic analyses have now provided additional insight into the symbiosis between *S. meliloti* and *B. japonicum* and their host legumes (Natera et al., 2000; Hoa et al., 2004; Sarma and Emerich, 2005, 2006). One trend that has emerged during proteomic studies of *B. japonicum* and *S. meliloti* is the preferential expression of a subset of ATP-binding cassette-type transporters in nodule bacteria (Djordjevic, 2004; Hoa et al., 2004; Sarma and Emerich, 2006). Free-living cells of both species express a large number of sugar transporters and expression of these systems is down-regulated in planta (Djordjevic, 2004; Sarma and Emerich, 2006). A different set of transport systems appears to be specific to bacteroids, with predicted substrates including amino acids, iron (in *S. meliloti*), and phosphate. Lodwig et al. (2003) suggested that amino acid cycling between plant cells and *R. leguminosarum* bacteroids occurs via ATP-binding cassette-type transporters (Aap and Bra) and is required for the assimilation of fixed nitrogen by pea (*Pisum sativum*) plants (Lodwig et al., 2003). While it is possible that similar systems are in place in the *B. japonicum* and *S. meliloti* symbioses, expression of the majority of these transport systems was not found to be similarly up-regulated in microarray analyses (Barnett et al., 2004). Since several studies have shown that there is often poor overlap between proteomic and transcriptomic data (Becker et al., 2004; Barnett and Fisher, 2006; Capela et al., 2006), both types of functional studies need to be done to obtain a clearer understanding of the role of specific genes and proteins in the symbiotic process.

The transcriptomes and proteomes of bacteroids differ considerably from those of free-living cells (Perret et al., 1999; Natera et al., 2000; Ampe et al., 2003; Barnett et al., 2004; Becker et al., 2004; Djordjevic, 2004; Hoa et al., 2004). Overall, gene expression is largely down-regulated in bacteroids (Barnett et al., 2004; Becker et al., 2004; Capela et al., 2006), possibly as a result of growth arrest and a stationary phase-like existence (Capela et al., 2006). Likewise, a considerably larger number of proteins in bacteroids are down-regulated than up-regulated, consistent with a general decrease in cell metabolism (Natera et al., 2000). Within the nodule, the plant supplies rhizobia with a carbon source in the form of dicarboxylic acids, which are then metabolized via the tricarboxylic acid cycle. In both *S. meliloti* and *B. japonicum*, enzymes involved in gluconeogenesis have been reported in bacteroid proteomes (Djordjevic, 2004; Sarma and Emerich, 2005, 2006), indicating that this pathway may be active in these cells. Although the enzyme involved in the first step of gluconeogenesis, phosphoenolpyruvate carboxykinase (PckA), was detected in *B. japonicum* (Sarma and
Emerich, 2005), it was not found in the proteome of S. meliloti bacteroids (Djordjevic, 2004) and expression of pckA was classified as repressed in microarray studies (Barnett et al., 2004; Becker et al., 2004). As well, PckA enzyme activity was not detected in alfalfa bacteroid extracts (Finan et al., 1991). Thus, bacteroid metabolism may differ in these two species. In exchange for reduced photosynthetic, nodule bacteria supply the host plant with a source of nitrogen. In S. meliloti bacteroids, assimilation of ammonia as a source of cellular nitrogen is repressed. This was evident in both gene expression (Barnett et al., 2004; Becker et al., 2004) and proteome analyses (Djordjevic, 2004). This result is consistent with the hypothesis that ammonia produced via nitrogen fixation is not readily utilized by bacteroids for their own nitrogen metabolism.

THE GENOMES OF OTHER SYMBIOTIC BACTERIA

The actinomycete Frankia forms nitrogen-fixing root nodules on nonlegumes and the genetic basis of the symbiotic interactions between Frankia strains and host plants is poorly understood. Accordingly, the recent publication of genomic sequences of three Frankia sp. strains is an exciting development (Normand et al., 2007). In fact, this work has already yielded valuable insight into the evolutionary histories of the sequenced strains, and has correlated differences in genome content and size with the biogeographic history of the host plants associated with each strain. This work documents the first example of both genome contraction in the narrow host range strain HFPCC13, and expansion in the broad host range strain EAN1pec, within closely related strains of bacteria. This further strengthens the idea that genome content is a reflection of a species’ lifestyle and that the microsymbiont genome is influenced by the evolutionary history and geographical distribution of their hosts.

CONCLUSION

Traditionally, studies involving rhizobia have focused upon a subset of genes involved in root nodulation, nitrogen fixation, and other symbiotic processes. However, the recent completion of genome sequencing projects in several rhizobial species permits a more comprehensive analysis of gene function and evolution. One trend that has emerged is the highly dynamic and complex nature of rhizobial genomes. As well, it is becoming increasingly evident that genome architecture and even content is significantly influenced by the multifaceted lifestyle adopted by legume endosymbionts.

Transcriptomic and proteomic analyses have provided valuable insight into rhizobial-legume symbioses, however, significant limitations still exist within these types of studies. This is perhaps best exemplified in studies involving indeterminate root nodules, where the issue of analyzing expression patterns at various stages of infection has been addressed through the use of bacterial and plant developmental mutants. The poor overlap that has been documented between transcriptomic and proteomic data emphasizes the importance of developing additional functional genomic approaches, such as studies focusing upon the metabolomics of stem- and root-nodule bacteria (Barsch et al., 2004; Colebatch et al., 2004). Such initiatives are necessary to extend our knowledge of the fascinating partnership that exists between legumes and their microsymbionts.

ACKNOWLEDGMENT

We thank Eddie Cytryn for help in the construction of the phylogenetic tree of 16S rDNA sequences.

Received April 26, 2007; accepted May 3, 2007; published June 6, 2007.

LITERATURE CITED

Capela D, Filipe C, Bobik C, Batut J, Bruand C (2006) Sinorhizobium meliloti differentiation during symbiosis with alfalfa: a transcriptomic dissec-
tion. Mol Plant Microbe Interact 19:363–372