1-2014

Functional Conservation of the Capacity for ent-Kaurene Biosynthesis and an Associated Operon in Certain Rhizobia

David M. Hershey
Iowa State University

Xuan Lu
Iowa State University, xuanlu@iastate.edu

Jiachen Zi
Iowa State University, jzi@iastate.edu

Reuben J. Peters
Iowa State University, rjpeters@iastate.edu

Functional Conservation of the Capacity for ent-Kaurene Biosynthesis and an Associated Operon in Certain Rhizobia

Abstract
Bacterial interactions with plants are accompanied by complex signal exchange processes. Previously, the nitrogen-fixing symbiotic (rhizo)bacterium Bradyrhizobium japonicum was found to carry adjacent genes encoding two sequentially acting diterpene cyclases that together transform geranylgeranyl diphosphate to ent-kaurene, the olefin precursor to the gibberellin plant hormones. Species from the three other major genera of rhizobia were found to have homologous terpene synthase genes. Cloning and functional characterization of a representative set of these enzymes confirmed the capacity of each genus to produce ent-kaurene. Moreover, comparison of their genomic context revealed that these diterpene synthases are found in a conserved operon which includes an adjacent isoprenyl diphosphate synthase, shown here to produce the geranylgeranyl diphosphate precursor, providing a critical link to central metabolism. In addition, the rest of the operon consists of enzymatic genes that presumably lead to a more elaborated diterpenoid, although the production of gibberellins was not observed. Nevertheless, it has previously been shown that the operon is selectively expressed during nodulation, and the scattered distribution of the operon via independent horizontal gene transfer within the symbiotic plasmid or genomic island shown here suggests that such diterpenoid production may modulate the interaction of these particular symbionts with their host plants.

Keywords
Biosynthetic pathways, genetics, diterpenes, mesorhizobium, synteny

Disciplines
Biochemistry, Biophysics, and Structural Biology | Genomics | Plant Biology

Comments
This article is from *Journal of Bateriology* 196 (2014): 100, doi:10.1128/JB.01031-13. Posted with permission.

JB
Journals.ASM.org

Functional Conservation of the Capacity for *ent*-Kaurene Biosynthesis and an Associated Operon in Certain Rhizobia

David M. Hershey,* Xuan Lu, Jiachen Zi, Reuben J. Peters

Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, Iowa, USA

Bacterial interactions with plants are accompanied by complex signal exchange processes. Previously, the nitrogen-fixing symbiotic (rhizo)bacterium *Bradyrhizobium japonicum* was found to carry adjacent genes encoding two sequentially acting diterpene cyclases that together transform geranylgeranyl diphosphate to *ent*-kaurene, the olefin precursor to the gibberellin plant hormones. Species from the three other major genera of rhizobia were found to have homologous terpene synthase genes. Cloning and functional characterization of a representative set of these enzymes confirmed the capacity of each genus to produce *ent*-kaurene. Moreover, comparison of their genomic context revealed that these diterpene synthases are found in a conserved operon which includes an adjacent isoprenyl diphosphate synthase, shown here to produce the geranylgeranyl diphosphate precursor, providing a critical link to central metabolism. In addition, the rest of the operon consists of enzymatic genes that presumably lead to a more elaborated diterpenoid, although the production of gibberellins was not observed. Nevertheless, it has previously been shown that the operon is selectively expressed during nodulation, and the scattered distribution of the operon via independent horizontal gene transfer within the symbiotic plasmid or genomic island shown here suggests that such diterpenoid production may modulate the interaction of these particular symbionts with their host plants.

Bacteria play critical roles in biogeochemical cycles, such as the fixation of nitrogen. Although nitrogen makes up approximately 80% of the Earth's atmosphere, its bioavailability remains a major limitation, e.g., to plant growth (1). This is due to the inability of plants to assimilate the diatomic nitrogen that occurs naturally in the atmosphere (2). Among plants, legumes uniquely host bacteria from the *Rhizobiales* order of the *Alphaproteobacteria* in nodules formed following invasion of their root cortical cells. Inside these nodules, the rhizobia develop into endosymbiont bacteroids, fixing nitrogen in exchange for carbon from their plant hosts (3). This agriculturally important collaboration is thought to be the main biological route for nitrogen fixation. Of particular relevance here, a number of the rhizobia have been shown to produce plant growth hormones, such as the gibberellins, which are thought to further promote growth of the host plant (4).

Both legume and rhizobial species exhibit a surprising amount of specificity with respect to symbiotic partners (5). Only rarely can a given rhizobial species nodulate more than a few closely related plants. This specificity is due to bacterial and plant factors (2). The host plant secretes flavonoid inducers that elicit rhizobial production of lipochitooligosaccharide Nod factors, which are recognized by the host plant, with subsequent steps in the nodulation process being dependent on recognition of bacterial cell surface chemistry and effector proteins as well. However, the host plant also applies the usual defense mechanisms—e.g., microbe-associated molecular pattern-triggered immunity and R-gene recognition of bacterial effectors—to restrict nodulation by unwanted strains (2, 3, 6). This complex signal exchange process exerts extreme evolutionary pressure on the rhizobia (7–9), which can be inferred, in part, from the presence of large plasmids or genomic islands with distinct G+C contents relative to the G+C content in the rest of the genome. These large plasmids or genomic islands contain the large set of genes required for nodulation, including nitrogen fixation (5, 10–13). This presumably reflects the ability of horizontal transfer to spread these distinct genetic elements, enabling nodulation by the recipient rhizobia (9, 14). Intriguingly, the prevalence of insertion sequences and phage integration is thought to promote rearrangement within these symbiotic modules (15).

Previously, we characterized two terpene synthases from *Bradyrhizobium japonicum* USDA110 (16). These proved to be diterpene cyclases capable of successively converting the general diterpenoid precursor (*E*,*E*,*E*)-geranylgeranyl diphosphate (GGPP) into *ent*-copalyl diphosphate (*ent*-CPP) and, hence, to *ent*-kaurene, a precursor to the gibberellin phytohormones (Fig. 1). The relevant genes, *blr2149* and *blr2150*, then encode an *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS), respectively. Notably, these two genes fall into a more extensive operon that was originally defined by Tully et al. (17) and suggested to be present in all rhizobia (18), although it did not appear to affect the ability of *B. japonicum* to nodulate soybean (*Glycine max*) or fix nitrogen in the resulting nodules (19). In the study described here, we investigated the functional conservation of the CPS and KS from this operon and demonstrate the production of the upstream GGPP by the isoprenyl diphosphate synthase encoded by the adjacent gene in the operon. In addition, we found that this operon exhibits a scattered distribution within the rhizobia. While examples are found in all four major genera, with conservation of the ability to produce *ent*-kaurene, the uneven distribution of the operon suggests that such diterpenoid production provides a selective advantage only under certain conditions.

Received 2 September 2013 **Accepted** 14 October 2013
Published ahead of print 18 October 2013
Address correspondence to Reuben J. Peters, rjpeters@iastate.edu.
* Present address: David M. Hershey, Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, California, USA.
D.M.H. and X.L. contributed equally to this article.

doi:10.1128/JB.01031-13

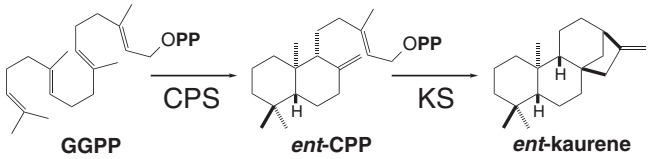

FIG 1 Production of *ent*-kaurene by *B. japonicum*. Shown are the steps catalyzed by the characterized *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (*ent*-KS).

MATERIALS AND METHODS

General. Unless otherwise noted, all molecular biology reagents were purchased from Invitrogen and all other chemicals were from Fisher Scientific. *B. japonicum* USDA110 was obtained from Michael Sandowsky (University of Minnesota), and *Mesorhizobium loti* MAFF303099, *Sinorhizobium fredii* NGR234, and *Rhizobium etli* CFN42 were all obtained from Philip Poole (John Innes Centre), while *Sinorhizobium meliloti* 1021 was obtained from Kathryn Jones (Florida State University). *Escherichia coli* was grown at 37 or 16°C on either NZY (for cloning) or TB medium (for expression). Rhizobia were cultured with YEM medium at 28°C. When necessary, 1.8% agar was added to the relevant medium to pour plates. Where applicable, antibiotics were used at the following concentrations: chloramphenicol, 30 μg/ml; carbenicillin, 50 μg/ml; spectinomycin, 50 μg/ml; and kanamycin, 50 μg/ml. Liquid cultures were grown with vigorous shaking (200 rpm), generally in 250-ml Erlenmeyer flasks with 50 ml medium. Microanaerobic cultures were grown in YEM medium with 10 mM KNO_3 under an atmosphere of nitrogen gas (N_2) and ~0.5% oxygen with moderate shaking (80 rpm) in rubber-stoppered flasks, with the atmosphere exchanged every 12 h (N_2 was blown into the flasks for 15 min).

Sequence retrieval and analysis. All sequences were retrieved from the National Center for Biotechnology Information (NCBI) website. The amino acid sequence of the previously characterized KS from *B. japonicum* (BjKS) was used as a query for BLAST searches against the *Rhizobiales* order (i.e., by restricting the search to this order, defined as taxid 356) on the NCBI website (20). This also was done using the amino acid sequence of CPS from *B. japonicum* (BjCPS) as the query sequence. Sequence analyses were carried out with the CLC Main Workbench program (version 6.8.4). Alignments used the following parameters: gap open cost, 10; gap extension cost, 10; and end gap cost, as any other. Trees were prepared using the neighbor-joining algorithm with a bootstrap analysis of 1,000 replicates. PAUP analysis was used to confirm the topology of the resulting trees. The phylogenetic analyses whose results are presented here were carried out using genes encoding biochemically analogous proteins from a bacterial species as distantly related as possible as the designated outgroup sequence. For CPS, this was from *Streptomyces* sp. strain KO-3988, which falls within the *Actinobacteria* phylum, yet this *Streptomyces* sp. CPS (SsCPS) (GenBank accession number AB183750) also produces *ent*-CPP (21). For NifK, this was from *Azotobacter vinelandii* (AvNifK; GenBank accession number Avin_01400), which falls within the *Proteobacteria* phylum but is in the distinct *Gammaproteobacteria* class.

Cloning and characterization of CPS and KS. Genomic DNA was isolated from rhizobia using a Generation capture kit (Qiagen) following the manufacturer's protocol. Each CPS and KS gene was amplified via PCR from genomic DNA using gene-specific primers and cloned into pENTR-SD-dTOPO (Invitrogen). Biochemical characterization of the CPS-KS pair from each species of rhizobia was carried out as described previously for *B. japonicum* (16). Briefly, the KS genes were subcloned into the plasmid pGG-DEST, which carries a plant GGPP synthase (GGPS), and the CPS genes were subcloned into pDEST14. This enabled use of a previously described metabolic engineering system, which included constructs analogous to the CPS and KS from *Arabidopsis thaliana* (22). Accordingly, the *E. coli* strain OverExpress C41 (Lucigen) was transformed with the various combinations of the pGG-DEST::CPS and pDEST14::KS plasmids described below, along with pIRS (i.e., to increase the isoprenoid precursor pool, as described previously [23]). Liquid cultures (50 ml) of the resulting recombinant *E. coli* strains were induced at an optical density at 600 nm of 0.6, the pH was adjusted to 7.1, and the bacteria were grown at 16°C for 72 h. The cultures were then extracted with an equal volume of hexanes. The organic extract was separated out and dried in a rotary evaporator, and the residue was resuspended in 100 μl hexanes. This concentrated extract was analyzed by gas chromatography (GC), carried out on a Varian (Palo Alto, CA) 3900 GC with a Saturn 2100 ion trap mass spectrometer (MS) in electron ionization (70 eV) mode. Samples (1 μl) were injected in splitless mode at 50°C, and after holding for 3 min at 50°C, the oven temperature was raised at a rate of 14°C/min to 300°C, where it was held for an additional 3 min. MS data from *m/z* 90 to 600 were collected starting 12 min after injection and were collected until the end of the run. The production of *ent*-kaurene was verified by comparison of the mass spectra and retention time to those of an authentic standard (enzymatically produced by the characterized CPS and KS from *A. thaliana*).

Mapping the diterpenoid biosynthesis operon. A 20-kb region surrounding each biochemically characterized KS was downloaded from the NCBI website and further analyzed. The predicted genes that either were homologous to those in the *B. japonicum* operon or had plausible predicted functions in (di)terpenoid biosynthesis were identified by alignment and open reading frame prediction. In each case, the boundaries of each operon were clear from the predicted functions of the adjacent genes (i.e., these have no plausible function in terpenoid biosynthesis). Putative RpoN and NifA binding sites were identified by searching for identical matches in the upstream region of each operon to previously defined 16-nucleotide motifs (24).

Characterization of GGPS. Fragments from the 3′ region of the operon from *S. fredii*, including genes for GGPS-CPS-KS or CPS-KS only, were amplified from genomic DNA via PCR. These were cloned into pZeroBlunt and then subcloned into a previously described *S. meliloti* overexpression vector (25), pstb-LAFR5 (obtained from Kathryn Jones, Florida State University), using BamHI and EcoRI restriction sites introduced by PCR, along with three upstream stop codons and an optimized ribosome binding site. The resulting constructs were transformed into *S. meliloti* 1021 by triparental mating using *E. coli* strain MM294A carrying the construct and *E. coli* strain MT616 as the helper, as described previously (25). These recombinant strains were grown for 5 days, and then the total culture was extracted with an equal volume of hexanes. This organic extract was separated out and dried under a gentle stream of N_2, with the residue then resuspended in 200 μl of hexanes for analysis by GC-MS, as described above.

Analysis of rhizobial diterpenoid production. Liquid cultures grown under aerobic or microanaerobic conditions were harvested 3, 6, or 9 days after inoculation, and the cells were separated from the spent medium by centrifugation (15 min at 10,000 × *g*). For analysis of the gibberellin content, the supernatant was acidified to pH 2.5 with acetic acid and then extracted with an equal volume of ethyl acetate saturated with acetic acid (1%, vol/vol). This organic extract was separated and passed over a 1-ml HP-20 resin column, which was then eluted with 3 ml each of 1% acetic acid in distilled H_2O (dH_2O) and 40% and 80% (vol/vol in dH_2O with 1% acetic acid) methanol. Each of these fractions was dried in a rotary evaporator, and the residue was resuspended in 100 μl acetic acid-saturated ethyl acetate for analysis by GC-MS as described above. For analysis of the *ent*-kaurene intermediate, the total culture was directly extracted with an equal volume of hexanes, which was separated out and passed over a 1-ml silica gel column to remove contaminating polar compounds. The resulting organic extract was dried under a gentle stream of N_2, and the residue was resuspended in 100 μl of hexanes for analysis by GC-MS, again, as described above.

RESULTS

Identification of KS and CPS homologs in rhizobia. As noted above, the BjKS that directly produces *ent*-kaurene exhibits dis-

TABLE 1 Rhizobial KS and CPS homologs

Organism	KS		CPS	
	GenBank accession no.	% identity to BjKS	GenBank accession no.	% identity to BjCPS
Bradyrhizobium japonicum	NP_768790		NP_768789	
Bradyrhizobium elkanii	WP_018270013	91	WP_016845990	92
Bradyrhizobium sp. strain WSM1253	WP_007600190	89	WP_007600189	91
Bradyrhizobium sp. strain WSM471	WP_007605894	88	WP_007605892	91
Mesorhizobium loti	NP_106894	93	NP_106893	93
Mesorhizobium alhagi	WP_008838313	94	WP_008838314	95
Mesorhizobium amorphae	WP_006204703	93	WP_006204702	93
Mesorhizobium ciceri	YP_004144785	92	YP_004144784	92
Mesorhizobium sp. strain STM 4661	WP_006329103	92	WP_006329109	93
Mesorhizobium sp. strain WSM4349	WP_018457688	92	WP_018457687	93
Sinorhizobium fredii	NP_443948	92	NP_443949	95
Sinorhizobium meliloti	WP_018098888	91	WP_018098887	94
Sinorhizobium medicae	WP_018009727	90	WP_018009726	92
Rhizobium etli	NP_659792	87	NP_659791	86
Rhizobium tropici	YP_007335933	87	YP_007335932	90
Rhizobium sp. strain CCGE 510	WP_007636919	88	WP_007636921	87
Rhizobium grahamii	WP_016558477	71	WP_016558476	72
Rhizobium mesoamericanum	WP_007539161	69	WP_007539159	72

tinct sequence homology relative to other characterized bacterial diterpene synthases. Accordingly, the BjKS sequence was used in initial BLAST searches of the NCBI database to identify bacteria from the *Rhizobiales* order that contain homologous diterpene synthases. Notably, homologs were found in species from all four major genera of rhizobia, i.e., *Rhizobium*, *Sinorhizobium*, and *Mesorhizobium*, in addition to *Bradyrhizobium*. In each case, immediately upstream of the BjKS homolog was a homolog to BjCPS (Table 1) with 2-nucleotide-overlapping open reading frames, just like the 2-nucleotide-overlapping open reading frame found in *B. japonicum*. Interestingly, these were not conserved by bacterial phylogeny; e.g., the KS from the various species of *Bradyrhizobium* shared less sequence identity than BjKS and the KS from *Mesorhizobium loti*, which not only is in a distinct genus but also falls into the separate *Phyllobacteriaceae* family. Accordingly, KS and CPS appear to have been distributed via horizontal gene transfer. Consistent with such an inheritance mechanism, the KS gene is not found in all rhizobia (e.g., no homolog is present in *Rhizobium leguminosarum*, whose genome has been fully sequenced [26], nor are homologs present in any of the genome sequences reported for various strains of *S. meliloti*, despite the presence of a homologous protein sequence annotated as being encoded by *Sinorhizobium meliloti* [27–29]).

Functional characterization of representative CPSs and KSs. To investigate the ability of the CPS and KS homologs found in our bioinformatics search to cooperatively produce *ent*-kaurene, we analyzed these from one species from each genus, specifically, examples of species for which complete genome sequences have been reported, *Mesorhizobium loti* MAFF303099 (12), *Sinorhizobium fredii* NGR234 (30), and *Rhizobium etli* CFN42 (10). We cloned and characterized the CPS and KS from each of these species much as previously described for those from *B. japonicum* (16). Briefly, each pair of CPS and KS homologs was coexpressed in recombinant *E. coli* also expressing a plant GGPP synthase (*Abies grandis* GGPS [AgGGPS]), which led to the production of kaurene (Fig. 2). To demonstrate stereospecificity, each CPS was

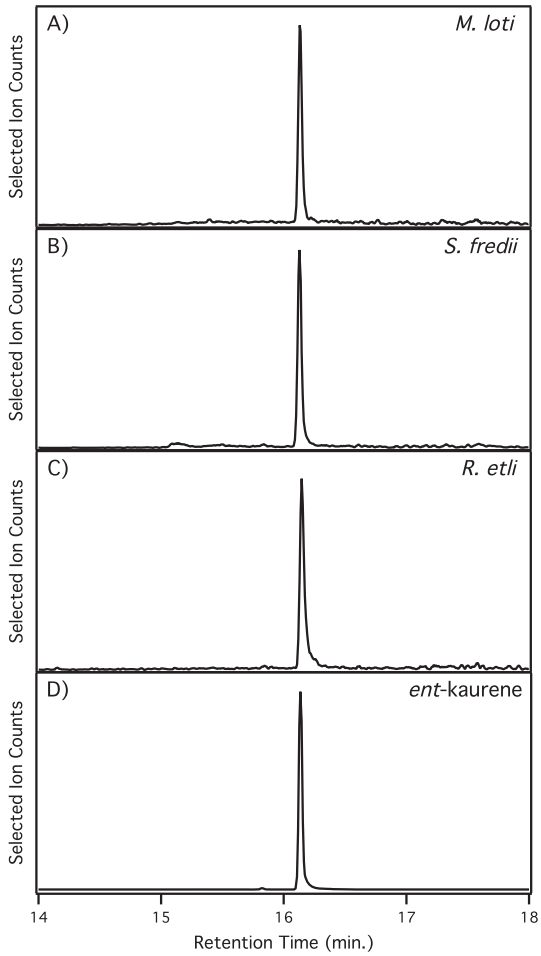

FIG 2 Selected ion (*m/z* 272) chromatograms obtained by GC-MS demonstrating production of *ent*-kaurene from GGPP by coexpressing CPS and KS from *M. loti* (A), *S. fredii* (B), and *R. etli* (C), along with an authentic standard (from coexpression of the CPS and KS from *Arabidopsis thaliana*) in *E. coli* (along with a GGPP synthase) (D).

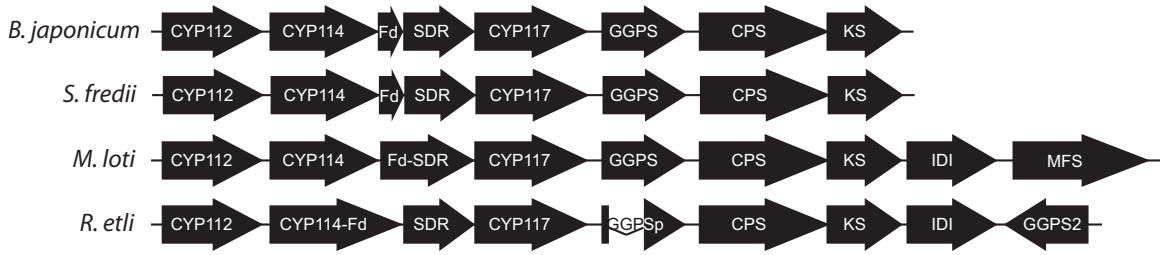

FIG 3 Schematic of diterpenoid biosynthesis operon from the designated rhizobia (with the gene designations described in the text).

expressed in recombinant *E. coli* with AgGGPS and the KS from *Arabidopsis thaliana* (AtKS), which is specific for *ent*-CPP. In addition, each KS was expressed in recombinant *E. coli* with AgGGPS and the *ent*-CPP-producing CPS from *Arabidopsis thaliana* (AtCPS). In each case, this led to the production of *ent*-kaurene (data not shown), demonstrating a stereochemistry consistent with that of the gibberellin plant hormones. These results confirmed a common catalytic activity for these distributed enzymatic genes and, importantly, that each genus identified here contains the capacity to produce *ent*-kaurene from GGPP.

Defining a rhizobial diterpenoid biosynthetic operon. Given that the genes for BjCPS and BjKS are neighboring genes in what has been proposed to be a more extensive operon (19), we examined the genomic context for each of the characterized CPSs and KSs to determine if these were similarly set in a more broadly conserved operon. Indeed, homologs to all of the other genes from the *B. japonicum* operon also were present, with retention of relative gene order. In particular, homologs to the cytochromes P450 CYP112 and CYP114, a ferredoxin (Fd), a short-chain alcohol dehydrogenase/reductase (SDR), another cytochrome P450 (CYP117), an isoprenyl diphosphate synthase that presumably makes GGPP (GGPS), as well as the orthologous CPS and KS were detected. Although it should be noted that some of these genes were fused together in certain cases, i.e., the CYP114 and Fd in *R. etli* and the Fd and SDR in *M. loti*, these still exhibited clear homology to the separate genes found elsewhere. Thus, these genes define a core diterpenoid biosynthetic gene cluster/operon that is conserved across all four of the major rhizobial genera, sharing 80 to 92% nucleotide sequence identity.

Notably, the GGPS gene in *R. etli* appears to be disrupted. While some homologous sequence is present, there is a large internal deletion, resulting in a clearly compromised open reading frame (accordingly, we suggest that this is a pseudogene and refer to it as GGPSp). However, *R. etli* contains another isoprenyl diphosphate synthase in close proximity to its core operon. Although this is not closely related to the GGPS found within the operon and is in the opposite orientation, we hypothesize that this might serve the same function (and refer to it here as GGPS2). There is an intervening gene. However, this gene appears to encode an isopentenyl diphosphate isomerase (IDI), which balances the isoprenoid precursor supply and, thus, similarly has a plausible role in (di)terpenoid biosynthesis as well. Further analysis demonstrated that a homologous IDI gene also occurs at the same position (3′ to the KS) in *M. loti*. Intriguingly, *M. loti* further has a gene encoding a major facilitator superfamily (MFS) member immediately downstream of its IDI, and we speculate that this might be involved in secretion of the diterpenoid natural product. Accordingly, in *R. etli* and *M. loti*, accessory genes appear to have been appended to the core diterpenoid biosynthesis operon (Fig. 3).

Upon sequence comparisons of the core operon, that from *R. etli* appeared to be the most divergent, sharing <82% identity, while the others were ≥90% identical to each other. Even when excluding the compromised GGPSp, comparison of the other genes from the *R. etli* operon revealed that these are only 86 to 89% identical to those from the other rhizobia, which is less than the level of identity shared by the other rhizobia. Accordingly, the *R. etli* operon is clearly the most divergent, consistent with distribution of the entire operon by horizontal gene transfer; e.g., despite their common phylogenetic origin in the *Rhizobiaceae* family, *R. etli* and *S. fredii* contain the most disparate rhizobial diterpenoid biosynthesis operons.

Demonstrating production of GGPP and capacity for *ent*-kaurene production. Diterpenoid biosynthesis generally proceeds via the initial formation of a hydrocarbon skeletal structure, followed by oxidative elaboration (31). In the organization of the rhizobial diterpenoid biosynthesis operons, it is notable that the genes predicted to be involved in oxidation are in the 5′ region, with all those predicted to be involved in the formation of the cyclized olefin *ent*-kaurene falling in the 3′ region. This includes the putative GGPP synthase (GGPS), as bacteria do not necessarily produce GGPP, leading to the presence of a GGPS in all of the identified bacterial diterpenoid biosynthetic gene clusters (21, 32–37). The observed organization of the rhizobial diterpenoid biosynthetic operon suggests that the 3′ and 5′ regions might form nominally independent subclusters, although no such subclusters appear in the currently available sequence information. Nevertheless, we took advantage of this gene arrangement to demonstrate both the production of GGPP by the isoprenyl diphosphate synthase and, hence, the ability of the operon to lead to the production of at least *ent*-kaurene. In particular, while initial efforts were directed at heterologous expression of the putative GGPP synthase in *E. coli* for use in coexpression studies such as those described above, that ultimately proved unsuccessful. We then turned to recombinant expression in a more closely related bacterium, specifically, the 1021 strain of *Sinorhizobium meliloti*, whose reported genome sequence does not contain the rhizobial diterpenoid biosynthesis operon (27). Accordingly, we overexpressed the 3′ region of the operon from the closely related *S. fredii*, either a fragment containing GGPS-CPS-KS or a fragment containing CPS-KS only. Consistent with the usual lack of GGPP production in bacteria, expression of the CPS-KS genes alone in *S. meliloti* 1021 did not result in the production of *ent*-kaurene, while expression of the GGPS-CPS-KS genes did (Fig. 4). These results, then, confirm that the associated isoprenyl diphosphate synthase pro-

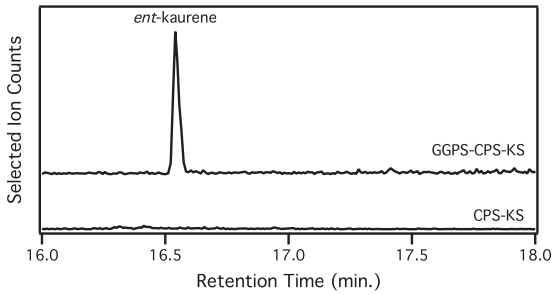

FIG 4 Selected ion (*m/z* 272) chromatograms obtained by GC-MS demonstrating production of *ent*-kaurene from *S. meliloti* 1021 expressing GGPS-CPS-KS, but not CPS-KS alone, from *S. fredii* NGR234 (as indicated).

duces GGPP, providing a critical link to central metabolism, and further demonstrate the capacity of the operon to confer the ability to produce at least *ent*-kaurene.

Evidence for a putative role in symbiosis. In each of the rhizobia examined here, the diterpenoid biosynthesis operon is located in the symbiotic module (plasmid or genomic island), suggesting that the resulting natural product may play a role in the symbiotic relationship of these rhizobia with their host plants. Consistent with such a role, it has been shown in *B. japonicum* that the genes in the diterpenoid biosynthesis operon are expressed upon bacterial differentiation in the nodules to the nitrogen-fixing bacteroid form and to a much lesser extent under microanaerobic culture conditions which mimic those found in the nodules (38). This expression has further been shown to be dependent on the symbiosis-specific sigma factor RpoN and associated transcription factor NifA (39). Similarly, expression of the genes from the diterpenoid biosynthesis operon in *R. etli* has also been shown to depend on RpoN and NifA and occur under similar conditions, i.e., in nodules and under microanaerobic conditions (24). In addition, sequences suggestive of regulation by the RpoN-NifA regulon also appear upstream of the diterpenoid biosynthesis operon in the other two genera. In particular, these are putative binding sites for both NifA and RpoN and, specifically, are the same sites identified upstream of the diterpenoid biosynthesis operon in *R. etli* (24). Accordingly, it seems likely that the diterpenoid biosynthesis operon in *M. loti* and *S. fredii* is also under the control of the RpoN-NifA regulon and expressed during symbiosis.

Given that *ent*-kaurene is the olefin intermediate in biosynthesis of the gibberellin phytohormones in both plants and fungi (40), coupled to previous reports of bacterial production of gibberellins (4), we have hypothesized that the capacity to produce this diterpene is indicative of the ability to produce gibberellins (16). However, although it has been reported that *B. japonicum* produces gibberellin A_3 (GA_3) under standard liquid culture conditions (41), we have been unable to detect this or *ent*-kaurene from *B. japonicum* grown under the previously described conditions or even under microanaerobic conditions. In addition, we were unable to detect production of *ent*-kaurene from any of the other rhizobia examined here under either aerobic or microanaerobic culture conditions.

While the rhizobial symbiotic modules (plasmid or genomic island) are known to be distributed via horizontal gene transfer, the diterpenoid biosynthetic operon seems to be separately distributed, albeit via selective integration into the symbiotic module. For example, it has been noted that this cytochrome P450-rich gene cluster exhibits a different G+C content relative to that in the rest of the symbiotic plasmid in *S. fredii* NGR234 (42). The phylogenetic relationship of the CPS genes from the analyzed rhizobial diterpenoid biosynthesis operons, for which homologs are evident in other bacteria, is distinct from that of the gene for the nitrogenase subunit NifK (Fig. 5). This suggests that the diterpenoid biosynthetic operon, which is not found in all such rhizobial symbiotic modules in any case, has been independently incorporated into these symbiotic modules. Although it has already been shown that polar disruption of the initial CYP112 gene in *B. japonicum* has no discernible effect on its ability to nodulate soybean or fix nitrogen (19), the resulting diterpenoid natural product may provide some selective advantage in the symbiotic growth phase of the rhizobia where it has been incorporated.

DISCUSSION

The results presented here demonstrate a scattered distribution of a diterpenoid biosynthetic operon within the rhizobia, with functional conservation of at least the capacity for the production of *ent*-kaurene (Fig. 2 and 4). Although the final diterpenoid end product remains unclear at this time, both plant and fungi produce *ent*-kaurene en route to the production of gibberellins (40), suggesting that this operon may also lead to such phytohormone production. The location of this operon in the symbiotic module of the relevant rhizobia, along with its previously demonstrated transcription in response to bacteroid differentiation in nodules, indicates a putative role for the resulting diterpenoid natural

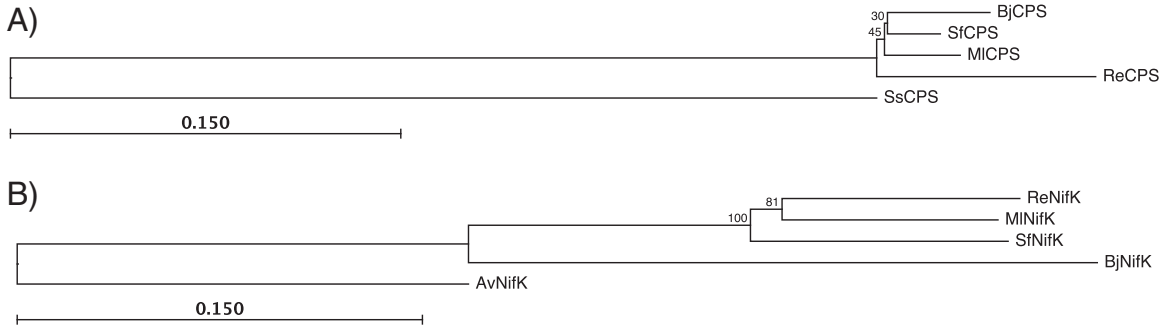

FIG 5 Molecular phylogenetic analysis of genes for the characterized rhizobial CPS (A) and genes for the nitrogenase subunit NifK (B) from the same rhizobia. SsCPS and AvNifK are the designated outgroup sequences, as described in the text. SfCPS, *S. fredii* CPS; MlCPS, *M. loti* CPS; ReCPS, *R. etli* CPS; ReNifK, *R. etli* NifK; MlNifK, *M. loti* NifK; SfNifK, *S. fredii* NifK; BjNifK, *B. japonicum* NifK.

product in the symbiotic relationship of these rhizobia with their host plants. While the exact physiological function of this diterpenoid remains unclear at this time as well, the scattered distribution of the operon, which appears to be a result of its apparently independent horizontal gene transfer between symbiotic modules, suggests that it provides a selective advantage only under certain conditions. Nevertheless, the striking conservation of this diterpenoid biosynthetic operon hints at its importance. Intriguingly, all the characterized operons are from rhizobia associated with determinate, rather than indeterminate, nodules. While nodule type is specified by the host plant species (43, 44), rhizobial specificity for plant host species indirectly controls bacterial nodulation phenotypes. Accordingly, it is tempting to speculate that the diterpenoid product of this operon specifically plays a role in rhizobial interactions within determinate nodules. While there are examples of the operon found in rhizobia usually isolated from plants that form indeterminate nodules (e.g., *S. meliloti*), it is unclear if these are complete. For example, while *Mesorhizobium* sp. strain WSM4349 was isolated from *Biserrula*, which forms indeterminate nodules, its operon has lost the gene for CYP112 and contains inactivating mutations in the SDR. Thus, the role of this operon and the resulting diterpenoid in determinate versus indeterminate nodules poses an interesting direction for future investigations.

ACKNOWLEDGMENTS

We thank Philip Poole (John Innes Centre) for helpful discussion and providing rhizobia and Kathryn Jones (Florida State University) for providing the pstb-LAFR5 expression vector. We also thank Matt Hillwig for productive discussions.

This work was supported by a grant from the National Science Foundation (MCB0919735) to R.J.P.

REFERENCES

1. **Vance CP.** 2001. Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. Plant Physiol. **127:**390–397. http://dx.doi.org//10.1104/pp.010331.
2. **Den Herder G, Parniske M.** 2009. The unbearable naivety of legumes in symbiosis. Curr. Opin. Plant Biol. **12:**491–499. http://dx.doi.org/10.1016/j.pbi.2009.05.010.
3. **Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC.** 2007. How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. Nat. Rev. Microbiol. **5:**619–633. http://dx.doi.org/10.1038/nrmicro1705.
4. **Bottini R, Cassan F, Piccoli P.** 2004. Gibberellin production by bacteria and its involvement in plant growth promotion and yield increase. Appl. Microbiol. Biotechnol. **65:**497–503. http://dx.doi.org/10.1007/s00253-004-1696-1.
5. **Fauvart M, Michiels J.** 2008. Rhizobial secreted proteins as determinants of host specificity in the rhizobium-legume symbiosis. FEMS Microbiol. Lett. **285:**1–9. http://dx.doi.org/10.1111/j.1574-6968.2008.01254.x.
6. **Yang S, Tang F, Gao M, Krishnan HB, Zhu H.** 2010. R gene-controlled host specificity in the legume-rhizobia symbiosis. Proc. Natl. Acad. Sci. U. S. A. **107:**18735–18740. http://dx.doi.org/10.1073/pnas.1011957107.
7. **Flores M, Morales L, Avila A, Gonzalez V, Bustos P, Garcia D, Mora Y, Guo X, Collado-Vides J, Pinero D, Davila G, Mora J, Palacios R.** 2005. Diversification of DNA sequences in the symbiotic genome of Rhizobium etli. J. Bacteriol. **187:**7185–7192. http://dx.doi.org/10.1128/JB.187.21.7185-7192.2005.
8. **Gonzalez V, Acosta JL, Santamaria RI, Bustos P, Fernandez JL, Hernandez Gonzalez IL, Diaz R, Flores M, Palacios R, Mora J, Davila G.** 2010. Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of Rhizobium etli. Appl. Environ. Microbiol. **76:**1604–1614. http://dx.doi.org/10.1128/AEM.02039-09.
9. **Johnston AWB, Beynon JL, Buchanan-Wollaston AV, Setchell SM, Hirsch PR, Beringer JE.** 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. Nature **276:**634–636. http://dx.doi.org/10.1038/276634a0.
10. **Gonzalez V, Santamaria RI, Bustos P, Hernandez-Gonzalez I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramirez MA, Jimenez-Jacinto V, Collado-Vides J, Davila G.** 2006. The partitioned Rhizobium etli genome: genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. U. S. A. **103:**3834–3839. http://dx.doi.org/10.1073/pnas.0508502103.
11. **Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiumi T, Sasamoto S, Watanabe A, Idesawa K, Iriguchi M, Kawashima K, Kohara M, Matsumoto M, Shimpo S, Tsuruoka H, Wada T, Yamada M, Tabata S.** 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res. **9:**189–197. http://dx.doi.org/10.1093/dnares/9.6.189.
12. **Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S.** 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. DNA Res. **7:**331–338. http://dx.doi.org/10.1093/dnares/7.6.331.
13. **Mavingui P, Flores M, Guo X, Davila G, Perret X, Broughton WJ, Palacios R.** 2002. Dynamics of genome architecture in Rhizobium sp. strain NGR234. J. Bacteriol. **184:**171–176. http://dx.doi.org/10.1128/JB.184.1.171-176.2002.
14. **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. U. S. A. **95:**5145–5149. http://dx.doi.org/10.1073/pnas.95.9.5145.
15. **Viprey V, Rosenthal A, Broughton WJ, Perret X.** 2000. Genetic snapshots of the Rhizobium species NGR234 genome. Genome Biol. **1:**RESEARCH0014. http://dx.doi.org/10.1186/gb-2000-1-6-research0014.
16. **Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ.** 2009. Gibberellin biosynthesis in bacteria: separate *ent*-copalyl diphosphate and *ent*-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett. **583:**475–480. http://dx.doi.org/10.1016/j.febslet.2008.12.052.
17. **Tully RE, van Berkum P, Lovins KW, Keister DL.** 1998. Identifcation and sequencing of a cytochrome P450 gene cluster from *Bradyrhizobium japonicum*. Biochim. Biophys. Acta **1398:**243–255. http://dx.doi.org/10.1016/S0167-4781(98)00069-4.
18. **Keister DL, Tully RE, Van Berkum P.** 1999. A cytochrome P450 gene cluster in the Rhizobiaceae. J. Gen. Appl. Microbiol. **45:**301–303. http://dx.doi.org/10.2323/jgam.45.301.
19. **Tully RE, Keister DL.** 1993. Cloning and mutagenesis of a cytochrome P-450 locus from *Bradyrhizobium japonicum* that is expressed anaerobically and symbiotically. Appl. Environ. Microbiol. **59:**4136–4142.
20. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
21. **Kawasaki T, Kuzuyama T, Kuwamori Y, Matsuura N, Itoh N, Furihata K, Seto H, Dairi T.** 2004. Presence of copalyl diphosphate synthase gene in an actinomycete possessing the mevalonate pathway. J. Antibiot. **57:**739–747. http://dx.doi.org/10.7164/antibiotics.57.739.
22. **Cyr A, Wilderman PR, Determan M, Peters RJ.** 2007. A modular approach for facile biosynthesis of labdane-related diterpenes. J. Am. Chem. Soc. **129:**6684–6685. http://dx.doi.org/10.1021/ja071158n.
23. **Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ.** 2010. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl. Microbiol. Biotechnol. **85:**1893–1906. http://dx.doi.org/10.1007/s00253-009-2219-x.
24. **Salazar E, Diaz-Mejia JJ, Moreno-Hagelsieb G, Martinez-Batallar G, Mora Y, Mora J, Encarnacion S.** 2010. Characterization of the NifA-RpoN regulon in Rhizobium etli in free life and in symbiosis with Phaseolus vulgaris. Appl. Environ. Microbiol. **76:**4510–4520. http://dx.doi.org/10.1128/AEM.02007-09.
25. **Jones KM.** 2012. Increased production of the exopolysaccharide succinoglycan enhances Sinorhizobium meliloti 1021 symbiosis with the host plant Medicago truncatula. J. Bacteriol. **194:**4322–4331. http://dx.doi.org/10.1128/JB.00751-12.
26. **Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS, Mauchline TH, East AK, Quail MA, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, Whitehead S, Parkhill J.** 2006. The genome of

Rhizobium leguminosarum has recognizable core and accessory components. Genome Biol. **7:**R34. http://dx.doi.org/10.1186/gb-2006-7-4-r34.

27. **Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, et al.** 2001. The composite genome of the legume symbiont Sinorhizobium meliloti. Science **293:**668–672. http://dx.doi.org/10.1126/science.1060966.
28. **Schneiker-Bekel S, Wibberg D, Bekel T, Blom J, Linke B, Neuweger H, Stiens M, Vorholter FJ, Weidner S, Goesmann A, Puhler A, Schluter A.** 2011. The complete genome sequence of the dominant Sinorhizobium meliloti field isolate SM11 extends the S. meliloti pan-genome. J. Biotechnol. **155:**20–33. http://dx.doi.org/10.1016/j.jbiotec.2010.12.018.
29. **Weidner S, Baumgarth B, Gottfert M, Jaenicke S, Puhler A, Schneiker-Bekel S, Serrania J, Szczepanowski R, Becker A.** 2013. Genome sequence of Sinorhizobium meliloti Rm41. Genome Announc. **1**(1)**:**e00013–12. http://dx.doi.org/10.1128/genomeA.00013-12.
30. **Schmeisser C, Liesegang H, Krysciak D, Bakkou N, Le Quere A, Wollherr A, Heinemeyer I, Morgenstern B, Pommerening-Roser A, Flores M, Palacios R, Brenner S, Gottschalk G, Schmitz RA, Broughton WJ, Perret X, Strittmatter AW, Streit WR.** 2009. Rhizobium sp. strain NGR234 possesses a remarkable number of secretion systems. Appl. Environ. Microbiol. **75:**4035–4045. http://dx.doi.org/10.1128/AEM.00515-09.
31. **Peters RJ.** 2010. Two rings in them all: the labdane-related diterpenoids. Nat. Prod. Rep. **27:**1521–1530. http://dx.doi.org/10.1039/c0np00019a.
32. **Hamano Y, Dairi T, Yamamoto M, Kawasaki T, Kaneda K, Kuzuyama T, Itoh N, Seto H.** 2001. Cloning of a gene cluster encoding enzymes responsible for the mevalonate pathway from a terpenoid-antibiotic-producing Streptomyces strain. Biosci. Biotechnol. Biochem. **65:**1627–1635. http://dx.doi.org/10.1271/bbb.65.1627.
33. **Durr C, Schnell H-J, Luzhetskyy A, Murillo R, Weber M, Welzel K, Vente A, Bechthold A.** 2006. Biosynthesis of the terpene phenalinolactone in Streptomyces sp. Tu6071: analysis of the gene cluster and generation of derivatives. Chem. Biol. **13:**365–377. http://dx.doi.org/10.1016/j.chembiol.2006.01.011.
34. **Hayashi Y, Matsuura N, Toshima H, Itoh N, Ishikawa J, Mikami Y, Dairi T.** 2008. Cloning of the gene cluster responsible for the biosynthesis of brasilicardin A, a unique diterpenoid. J. Antibiot. **61:**164–174. http://dx.doi.org/10.1038/ja.2008.126.
35. **Kim SY, Zhao P, Igarashi M, Sawa R, Tomita T, Nishiyama M, Kuzuyama T.** 2009. Cloning and heterologous expression of the cyclooctatin biosynthetic gene cluster afford a diterpene cyclase and two p450 hydroxylases. Chem. Biol. **16:**736–743. http://dx.doi.org/10.1016/j.chembiol.2009.06.007.
36. **Mann FM, Xu M, Davenport EK, Peters RJ.** 2012. Functional characterization and evolution of the isotuberculosinol operon in *Mycobacterium tuberculosis* and related mycobacteria. Front. Microbiol. **3:**368. http://dx.doi.org/10.3389/fmicb.2012.00368.
37. **Smanski MJ, Yu Z, Casper J, Lin S, Peterson RM, Chen Y, Wendt-Pienkowski E, Rajski SR, Shen B.** 2011. Dedicated ent-kaurene and ent-atiserene synthases for platensimycin and platencin biosynthesis. Proc. Natl. Acad. Sci. U. S. A. **108:**13498–13503. http://dx.doi.org/10.1073/pnas.1106919108.
38. **Pessi G, Ahrens CA, Rehrauer H, Lindemann A, Hauser F, Fischer H-M, Hennecke H.** 2007. Genome-wide transcript analysis of Bradyrhizobium japonicum bacteroids in soybean root nodules. Mol. Plant Microbe Interact. **20:**1353–1363. http://dx.doi.org/10.1094/MPMI-20-11-1353.
39. **Hauser F, Pessi G, Friberg M, Weber C, Rusca N, Lindemann A, Fischer H-M, Hennecke H.** 2007. Dissection of the Bradyrhizobium japonicum NifA+sigma54 regulon, and identification of a ferredoxin gene (fdxN) for symbiotic nitrogen fixation. Mol. Genet. Genomics **278:**255–271. http://dx.doi.org/10.1007/s00438-007-0246-9.
40. **Hedden P, Phillips AL, Rojas MC, Carrera E, Tudzynski B.** 2001. Gibberellin biosynthesis in plants and fungi: a case of convergent evolution? J. Plant Growth Regul. **20:**319–331. http://dx.doi.org/10.1007/s003440010037.
41. **Boiero L, Perrig D, Masciarelli O, Penna C, Cassán F, Luna V.** 2007. Phytohormone production by three strains of *Bradyrhizobium japonicum* and possible physiological and technological implications. Appl. Microbiol. Biotechnol. **74:**874–880. http://dx.doi.org/10.1007/s00253-006-0731-9.
42. **Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A, Perret X.** 1997. Molecular basis of symbiosis between Rhizobium and legumes. Nature **387:**394–401. http://dx.doi.org/10.1038/387394a0.
43. **Oke V, Long SR.** 1999. Bacteroid formation in the Rhizobium-legume symbiosis. Curr. Opin. Microbiol. **2:**641–646. http://dx.doi.org/10.1016/S1369-5274(99)00035-1.
44. **Oono R, Schmitt I, Sprent JI, Denison RF.** 2010. Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. New Phytol. **187:**508–520. http://dx.doi.org/10.1111/j.1469-8137.2010.03261.x.