The ISME Journal (2010) 4, 315–326

www.nature.com/ismej

ORIGINAL ARTICLE

Community shifts of soybean stem-associated bacteria responding to different nodulation phenotypes and N levels

Seishi Ikeda[1], Takashi Okubo[1], Takakazu Kaneko[2], Shoko Inaba[1], Tomiya Maekawa[3], Shima Eda[1], Shusei Sato[2], Satoshi Tabata[2], Hisayuki Mitsui[1] and Kiwamu Minamisawa[1]
[1]*Graduate School of Life Sciences, Tohoku University, Katahira, Aoba-ku, Sendai, Miyagi, Japan;* [2]*Kazusa DNA Research Institute, Kazusa-kamatari, Kisarazu, Chiba, Japan and* [3]*Graduate School of Agricultural Science, Tohoku University, Amamiya-machi, Sendai, Miyagi, Japan*

The diversity of stem-associated bacteria of non-nodulated (Nod^-), wild-type nodulated (Nod^+) and hypernodulated (Nod^{++}) soybeans were evaluated by clone library analyses of the 16S ribosomal RNA gene. Soybeans were dressed with standard nitrogen (SN) fertilization (15 kg N ha^{-1}) and heavy nitrogen (HN) fertilization (615 kg N ha^{-1}). The relative abundance of Alphaproteobacteria in Nod^+ soybeans (66%) was smaller than that in Nod^- and Nod^{++} soybeans (75–76%) under SN fertilization, whereas that of Gammaproteobacteria showed the opposite pattern (23% in Nod^+ and 12–16% in Nod^- and Nod^{++} soybeans). Principal coordinate analysis showed that the bacterial communities of Nod^- and Nod^{++} soybeans were more similar to each other than to that of Nod^+ soybeans under SN fertilization. HN fertilization increased the relative abundance of Gammaproteobacteria in all nodulation phenotypes (33–57%) and caused drastic shifts of the bacterial community. The clustering analyses identified a subset of operational taxonomic units (OTUs) at the species level in Alpha- and Gammaproteobacteria responding to both the nodulation phenotypes and nitrogen fertilization levels. Meanwhile, the abundance of Betaproteobacteria was relatively constant in all libraries constructed under these environmental conditions. The relative abundances of two OTUs in Alphaproteobacteria (*Aurantimonas* sp. and *Methylobacterium* sp.) were especially sensitive to nodulation phenotype and were drastically decreased under HN fertilization. These results suggested that a subpopulation of proteobacteria in soybeans is controlled in a similar manner through both the regulation systems of plant–rhizobia symbiosis and the nitrogen signaling pathway in plants.
The ISME Journal (2010) **4,** 315–326; doi:10.1038/ismej.2009.119; published online 19 November 2009
Subject Category: microbial population and community ecology
Keywords: autoregulation; community analysis; *Glycine max*; nitrogen; stem

Introduction

Diverse microorganisms live as endophytes and epiphytes of plants in nature. However, many questions regarding the driving forces and ecological rules underlying these relationships remain unanswered (Saito *et al.*, 2007; Hardoim *et al.*, 2008). During their evolution, legumes developed nodulation and mycorrhization systems to attain mutual symbiosis with rhizobia and mycorrhizae (Harrison, 2005; Stacey *et al.*, 2006). Although plants are known to control the degrees of nodulation and mycorrhization of roots by rhizobia and mycorrhizae, respectively, through autoregulation systems (Carroll *et al.*, 1985; Meixner *et al.*, 2007), the effects of the autoregulation systems on the other plant-associated microorganisms in the phytosphere remain unclear (Parniske, 2000). Recent studies have shown that the wild-type and symbiosis-defective mutants of legumes such as *Medicago truncatula* and *Glycine max* have different microbial community structures, and certain microbes associate preferentially with nodulated or mycorrhizal root systems (Offre *et al.*, 2007; Ikeda *et al.*, 2008a). Furthermore, transcriptional analyses revealed that both nodulation and mycorrhization systemically induced a series of defense genes in a host plant and consequently increased its disease resistance (Liu *et al.*, 2007; Kinkema and Gresshoff, 2008). These examples indicate that genetic alteration in the nodulation or mycorrhization signaling pathways can in turn alter the accompanying plant microflora, aside from rhizobia and mycorrhizae.

The autoregulatory mechanism in plants occurs through long-distance signaling between shoot and

Correspondence: K Minamisawa, Graduate School of Life Sciences, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan.
E-mail: kiwamu@ige.tohoku.ac.jp
Received 15 June 2009; revised 5 October 2009; accepted 9 October 2009; published online 19 November 2009

root (Oka-Kira and Kawaguchi, 2006), and a heavy supply of nitrate to the roots of leguminous plants inhibits nodulation through autoregulation (Carroll *et al.*, 1985; Okamoto *et al.*, 2009). Thus, there is a possibility that nodulation phenotype and host nitrogen status may affect the structure of the microbial community in shoots, as observed earlier in roots (Ikeda *et al.*, 2008a). However, the culture-independent analysis of soybean stem-associated microbial communities failed because of serious interference by plant DNA in the PCR amplification (Ikeda *et al.*, 2008a). To overcome this problem, a method was recently developed for enriching the bacterial cells collected from soybean stems (Ikeda *et al.*, 2009). The enrichment method allows the construction of a clone library of the bacterial 16S ribosomal RNA (rRNA) genes, which provides better resolution for examining the microbial community structures compared with the limited analyses in previous studies based mainly on fingerprinting methods (Offre *et al.*, 2007; Ikeda *et al.*, 2008a).

In this study, the effects of nodulation phenotypes and nitrogen application levels were examined on the stem-associated bacterial community including both epiphytes and endophytes in field-grown soybeans. Clone libraries of the 16S rRNA genes were constructed in conjunction with the method of enriching bacterial cells from soybean stems, and the bacterial diversities in different nodulation phenotypes were evaluated based on 16S rRNA gene sequences.

Materials and methods

Plant materials and field experimental design

The plants included cultivar Enrei (wild-type nodulating cultivar, Nod^+), lines En 1314 and En 1282 (non-nodulating mutants derived from Enrei, Nod^-; Francisco and Akao, 1993) and lines En 6500 and Sakukei 4 (hypernodulating mutants derived from Enrei, Nod^{++}; Akao and Kouchi, 1989; Matsunami *et al.*, 2004). Mutations in En 1314 and En 1282 soybeans were found in the *NFR1* gene, which is known to be responsible for a LysM-type receptor kinase for Nod factor recognition (Radutoiu *et al.*, 2003 Masaki Hayashi (National Institute of Agrobiological Sciences, Tsukuba, Japan), personal communication). A mutation in En 6500 was found in the *GmNARK* (*NTS1*) gene (Nishimura *et al.*, 2002), which mediates systemic autoregulation of nodulation (Searle *et al.*, 2003), and Sakukei 4 is a descendant with a hypernodulating phenotype that was derived from En 6500 (Matsunami *et al.*, 2004).

The seeds were planted on 10 June 2008 in the experimental field at Tohoku University (Kashimadai, Miyagi, Japan), which has been cultivated with a rotation of rice (paddy field condition) and soybeans (upland field condition) every year since 1997. The field soil was classified as gray lowland soil (pH(H_2O), 5.9; pH(KCl), 4.3; total carbon content, 1.21%; total nitrogen content, 0.11%; Truog phosphorus content, 69 mg P_2O_5 kg^{-1}). Soybeans of each nodulation genotype were grown in two neighboring fields (standard nitrogen (SN) and heavy nitrogen (HN) fields). The SN field was dressed with SN fertilization (15 kg N ha^{-1}) as a basal fertilization, and the HN field was also dressed four times with HN fertilization (N1 to N4 in Figure 1; each application 150 kg N ha^{-1} as urea, total 600 kg N ha^{-1}) in addition to the basal fertilization.

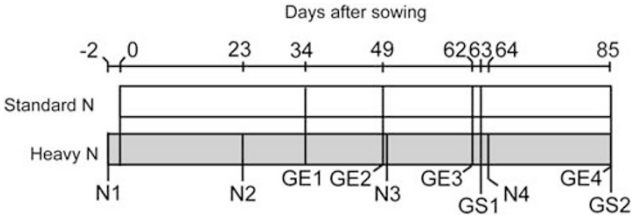


Figure 1 Time table and design of the field experiment. Basal fertilization was supplied for both standard and heavy nitrogen (HN) fields with 15 kg N ha^{-1} of a commercial fertilizer. The HN field was also supplied with nitrogen fertilization as urea four times (N1 to N4; 150 kg N ha^{-1} for each fertilization). GE1 to GE4 represent the days on which soybean growth indexes (Supplementary Figure 1) were measured. GS1 and GS2 indicate growth stages 1 and 2 for measuring dry weights and nitrogen contents of soybean tissues (Supplementary Figure 2). The sampling for the DNA extraction was carried out at GS2.

Growth evaluation and sampling

To define the environmental factors that are relevant to changes in bacterial community structures, the plant growth stage was monitored by measuring vegetative and reproductive indexes, plant dry weights, and nitrogen contents in plant tissues (GE1 to GE4 in Figure 1) and in field soils. Plant growth stages were evaluated according to the criteria of Fehr *et al.* (1971) four times throughout the growing period (GE1 to GE4 in Figure 1). Dry weights of shoots, roots, and nodules as well as the number of nodules were measured twice during the growing period (GS1 and GS2 in Figure 1). Total nitrogen of soybean shoots at GS1 was analyzed by the Kjeldahl method. Two plants were processed as a composite sample for measuring the N content of soybean shoots. The nitrogen content in soils under the SN and HN fertilization conditions was also measured three times before the additional fertilization of the HN field (N2, N3 and GS2 in Figure 1). Soil was air-dried and sieved through a 2-mm mesh. Dried soil was extracted by shaking with 2 M KCl solution for 30 min and then filtered through a filter paper (No. 2, Advantec, Tokyo, Japan). Contents of NH_4^+ and NO_3^- in the filtrates were determined by titrimetry after steam distillation (UDK127, Actac, Tokyo, Japan).

The aerial parts of soybean plants were harvested on 3 September 2008 (GS2 in Figure 1) and immediately transported on ice to the laboratory. Leaves were removed manually. The stems were serially washed well with tap and distilled water and stored at −80 °C until they were used. Four plants were processed as a composite sample for the DNA extraction. Three composite samples were analyzed for each genotype as a triplicate.

Clone library construction and sequencing

Stems (approximately 100 g) of a composite sample were homogenized without surface sterilization to prepare the stem-associated bacterial cells containing both epiphytes and endopytes, and the bacterial cells were extracted and purified using an enrichment method (Ikeda *et al.*, 2009). Total bacterial DNA was extracted from each bacterial cell fraction by the DNA extraction method described by Ikeda *et al.* (2008b). The final DNA samples were resuspended in 100 µl of sterilized water. The quality and quantity of DNA were assessed spectrophotometrically by calculating absorbance at a wavelength of 260 nm (A_{260}) and the A_{260}/A_{230} and A_{260}/A_{280} ratios. The triplicate DNA samples were prepared from composite samples for each soybean genotype under the SN and HN fertilization conditions.

PCR clone libraries for 16S rRNA genes were constructed as follows. Briefly, 50 ng of total bacterial DNA was used as a template in a final reaction volume of 60 µl, including 90 pmol of each primer and 9 U of Ex *Taq* DNA polymerase (Takara Bio, Otsu, Japan). The universal primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1525R (5′-AAGGAGGTGWTCCARCC-3′) were used (Lane, 1991). Cycling conditions were as follows: initial denaturation for 2 min at 94 °C; then 25 cycles consisting of 30 s at 94 °C, 30 s at 55 °C and 2 min at 72 °C; and a final extension for 10 min at 72 °C. Three PCR products derived from triplicate DNA samples were combined into a sample, and the PCR products were resolved by 1% agarose gel electrophoresis in 1 × TAE buffer. PCR products of the predicted size (approximately 1400 bp) were extracted from the gels using a QIAquick Gel Extraction Kit (Qiagen, Tokyo, Japan) and ligated into the plasmid vector pGEM-T Easy (Promega Japan, Tokyo, Japan) at 16 °C overnight using a DNA ligation kit, Mighty Mix (Takara Bio). ElectroTen-Blue Electroporation-Competent Cells (Stratagene, La Jolla, CA, USA) were then electroporated with the ligated DNA using a Gene Pulser (Bio-Rad Laboratories, Tokyo, Japan). After the transformants were cultured overnight at 37 °C on Luria-Bertani agar plates containing ampicillin (50 µg ml^{-1}), 384 colonies were randomly selected from each library for sequencing. Sequencing analysis was conducted using a Type 3730 × l DNA Analyzer (Applied Biosystems, Foster City, CA, USA) using a BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems). Template DNAs were prepared using an Illustra TempliPhi DNA Amplification Kit (GE Healthcare, Uppsala, Sweden). A partial sequence of the 16S rRNA gene was obtained using the 27F primer. Sequences were processed with Phred (Ewing and Green, 1998; Ewing *et al.*, 1998) at a cutoff value of 0.001 for eliminating low-quality regions. Then the 5′ region of the 16S rRNA gene (corresponding to bp 109–665 of the *Escherichia coli* 16S rRNA gene) was used for the sequence analyses.

Additional cloning and sequencing was carried out for three samples (Enrei of SN, Enrei of HN and En 1282 of HN) because of heavy contamination of chloroplast DNA. Libraries of these samples were screened by PCR with a soybean chloroplast-specific primer set, SCS-F (5′-AGGCGGTGGAAACTACCA AG-3′), and SCS-R (5′-CAGCGCCTAGTATCCATC GT-3′), designed in this study to eliminate clones derived from the chloroplast 16S rRNA gene. After the negative screening by PCR with the soybean chloroplast-specific primer set, 192, 96, and 96 clones were additionally subjected to sequencing for Enrei of SN, Enrei of HN and En 1282 of HN, respectively, as potentially nonbacterial clones.

Sequence analysis

The sequences were analyzed for orientation and detection of non-16S rDNA sequences using OrientationChecker (Ashelford *et al.*, 2006). The presence of chimeras was assessed using MALLARD (Ashelford *et al.*, 2006). A sequence identified at the 99.9% and 99% cutoffs or suspected at the 99.9% cutoff as a chimeric sequence was discarded. The remaining sequences were aligned using CLUSTAL W (Thompson *et al.*, 1994). On the basis of the alignment, a distance matrix was constructed using the DNADIST program from PHYLIP ver. 3.66 (http://evolution.genetics.washington.edu/phylip.html) with the default parameters. The resulting matrices were run in DOTUR (Schloss and Handelsman, 2005) to generate diversity indexes. The default DOTUR settings were used with threshold values of 97% sequence identity for the definition of operational taxonomic units (OTUs). Library coverage was calculated with the non-parametric estimator *C* (Good, 1953), as described by Kemp and Aller (2004). The reciprocal of Simpson's index (1/*D*) was used as a measure of diversity to evaluate the level of dominance in a community (Zhou *et al.*, 2002). The number of shared OTUs between libraries was calculated using SONS (Schloss and Handelsman, 2006).

UniFrac (Lozupone and Knight, 2005) was applied to examine the differences between clone libraries. A tree file generated by CLUSTAL W (Thompson *et al.*, 1994) and an environment file, which links a file to a library, were uploaded to UniFrac (Lozupone and Knight, 2005). Principal coordinate analysis (PCoA) was performed by using UniFrac

with the abundance-weighted option (Lozupone and Knight, 2005).

Phylogenetic analysis
The phylogenetic composition of the sequences in each library was evaluated using the Classifier program of RDP-II release 10 (Wang *et al.*, 2007), with confidence levels of 80%. BLASTN (Altschul *et al.*, 1990) was also used to classify the clones and to identify the closest relatives in the GenBank database.

For the phylogenetic analysis, sequences were aligned using the CLUSTAL W program (Thompson *et al.*, 1994). The neighbor-joining method was used for building the trees (Saitou and Nei, 1987). The PHYLIP format tree output was obtained using the bootstrapping procedure (Felsenstein, 1985); 1000 bootstrap trials were used. The trees were constructed using TreeView software (Page, 1996).

Accession numbers of nucleotide sequences
Nucleotide sequences of 16S rRNA genes for the clone libraries have been deposited in the DDBJ database under accession numbers shown in Table 1.

Results

Soybean growth
The soybean stem-associated bacteria including endophytes and epiphytes at the reproductive stage (GS2 in Figure 1) were subjected to clone library analyses. The growth indexes revealed no difference in the stages of the reproductive growth phase among the soybean genotypes or nitrogen treatments throughout the growth period, although slight variations were observed for the vegetative growth stages among the soybean genotypes in the early phase of growth (Supplementary Figure 1). At the sampling time, the reproductive growth indexes were R6 (pod-maturing stage; Fehr *et al.*, 1971) for all samples in the SN and HN fields, and the vegetative growth indexes ranged approximately from 12 to 15 for most samples. Thus, we concluded that the effects of soybean growth stage on the community analysis would not be significant among the genotypes examined under the SN or HN field conditions.

Statistics of clone libraries
The statistics of the clone libraries are summarized in Table 2. The library coverage ranged from 87.8 to 94.9%, showing a relatively high coverage for all libraries. Under the SN fertilization condition, Nod^+ soybeans (Enrei) showed the highest number of OTUs and slightly higher diversity for all indexes compared with those of mutants, whereas no clear effect of nodulation phenotype was observed for the number of OTUs and diversity indexes under the HN fertilization condition.

Table 1 Accession numbers of sequences deposited in the DDBJ

Genotypes	*N fertilization levels*	
	SN level	*HN level*
Enrei	AB482224–AB482598	AB483853–AB484222
En 1282	AB482599–AB482867	AB484223–AB484583
En 1314	AB482868–AB483207	AB484584–AB484930
En 6500	AB483208–AB483518	AB484931–AB485234
Sakukei 4	AB483519–AB483852	AB485235–AB485589

Abbreviations: HN, heavy nitrogen; SN, standard nitrogen.

Phylogenetic diversities of stem-associated bacteria with different nodulation phenotypes under standard nitrogen fertilization condition
The assessment of phylogenetic compositions using RDP Classifier revealed that proteobacteria was the most dominant taxon (88.7–96.5%) in all libraries at the phylum level under the SN fertilization condition (Figures 2a). In proteobacteria, the relative abundances of Alphaproteobacteria and Gammaproteobacteria varied according to either nodulation phenotype, whereas the abundance of Betaproteobacteria was relatively constant in all libraries constructed under these environmental conditions (Figure 2a). The ratio of Alpha/Gammaproteobacteria in Nod^- and Nod^{++} soybeans (4.6–6.1) was higher than that in Nod^+ soybeans (2.9; Figure 2a).

Further analyses of phylogenetic compositions at lower taxonomic levels revealed that three genera (*Aurantimonas*, *Methylobacterium* and *Rhizobium*) in the order Rhizobiales for Alphaproteobacteria (Figure 2c) and three orders (Pseudomonadales, Xanthomonadales and Enterobacteriales) for Gammaproteobacteria (Figure 2e) were mainly responsible for the community shifts of Alpha/Gammaproteobacteria under the SN fertilization condition.

The results of PCoA showed two tight clusters for Nod^{++} and Nod^- soybeans distinctly separated from Nod^+ soybeans, suggesting that the community structures of stem-associated bacteria were mainly dependent on nodulation phenotype under SN fertilization (Figure 3). In addition, the overall structures of the microbial communities in Nod^{++} and Nod^- soybeans were more similar to each other than to that of Nod^+ soybeans under the SN fertilization condition (Figure 3).

Clustering analysis of 16S rRNA sequences allowed the identification of the OTUs responsible for the population shifts of Alpha- and Gammaproteobacteria at the species level. Two OTUs (AP6 and AP30) noticeably increased the ratio of the Alphaproteobacterial population in mutants of both Nod^- and Nod^{++} soybeans compared with Nod^+ soybeans under the SN fertilization condition (Figures 4 and 5). The representative clones for these two OTUs were identical to *Aurantimonas ureilytica* (DQ883810) and *Methylobacterium chloromethanicum* (CP001298), respectively.

Table 2 Statistical analysis of 16S rRNA gene clone libraries derived from soybean stems

Libraries	*SN condition*					*HN condition*				
Nodulation phenotypes	*Enrei Nod*$^{+}$	*En 1282 Nod*$^{-}$	*En 1314 Nod*$^{-}$	*En 6500 Nod*$^{++}$	*Sakukei 4 Nod*$^{++}$	*Enrei Nod*$^{+}$	*En 1282 Nod*$^{-}$	*En 1314 Nod*$^{-}$	*En 6500 Nod*$^{++}$	*Sakukei 4 Nod*$^{++}$
Statistics										
No. of sequences	375	269	340	311	334	370	361	347	304	355
Library coverage (%)[a]	89.3	91.1	91.2	91.1	94.9	90.5	87.8	92.5	92.1	90.1
No. of OTUs[b]	75	52	60	60	57	71	91	56	55	71
Diversity indexes										
Chao1	153.0	86.5	114.4	97.8	88.7	130.5	143.6	110.2	85.7	130.5
ACE	158.7	83.5	110.2	101.1	75.3	132.3	164.0	95.6	87.9	132.3
Shannon index	3.5	3.1	3.0	3.3	3.3	3.4	3.8	3.2	3.2	3.4
Simpson index (1/*D*)	19.7	12.5	11.0	15.6	14.5	16.9	27.4	14.8	15.2	16.9

Abbreviations: HN, heavy nitrogen; OTUs, operational taxonomic units; rRNA, ribosomal RNA; SN, standard nitrogen.
[a]$C_X = 1-(n/N)$, where n_x is the number of singletons that are encountered only once in a library and N is the total number of clones.
[b]OTUs were defined at 97% sequence identity.

In contrast, the relative abundance of OTU GP39 in Enterobacteriales was lower in mutants of both Nod^{-} and Nod^{++} soybeans compared with that in Nod^{+} soybeans (Figures 4 and 6). The representative clone of OTU 39 was classified as *Enterobacter* sp. and was identical to an isolate (AY822505) that was derived from a plant-associated bacterium, and its closest relative in known species was *Klebsiella oxytoca* (EU982878) with 99% identity (Figure 6).

Comparisons of phylogenetic diversity of stem-associated bacteria between standard and heavy nitrogen fertilization conditions

The HN fertilization condition was designed to compare the community structures of shoot-associated bacteria among different nodulation phenotypes by eliminating the bias of nitrogen status of host plants and to suppress the nodulation. When HN fertilization was applied to the soybeans, the effect of nodulation phenotype on the bacterial communities was obscured (Figure 2b). However, the ratio of Alpha/Gammaproteobacteria was drastically decreased for all nodulation phenotypes under HN fertilization (ranging from 0.8 to 2.1) compared with those under SN fertilization (2.9–6.1; Figures 2a and b). Further analyses of phylogenetic compositions at lower taxonomic levels revealed that the same taxonomic groups that were responding to the community shifts between the Nod^{+} soybeans and the mutants (Nod^{-} and Nod^{++} soybeans) under SN fertilization were also responding to the effects of HN fertilization. Thus, three genera (*Aurantimonas*, *Methylobacterium* and *Rhizobium*) in the order of Rhizobiales for Alphaproteobacteria (Figure 2d) and three orders (Pseudomonadales, Xanthomonadales and Enterobacteriales) for Gammaproteobacteria (Figure 2f) were responding to the population shifts of Alpha- and Gammaproteobacteria in all libraries under the HN fertilization condition.

The results of PCoA clearly showed that HN fertilization caused drastic shifts of community structures of shoot-associated bacteria for all nodulation phenotypes from those found under SN fertilization (Figure 3). PC1 (49.4%) explained the effects of HN fertilization on the community structures. In addition, the community structures among the soybean genotypes under HN fertilization were far more dissimilar to each other than those under SN fertilization.

Clustering analysis of 16S rRNA sequences identified the same three OTUs (AP6, AP30 and GP39) that were responding to the population shifts of Alpha- and Gammaproteobacteria between the Nod^{+} soybeans and the mutants (Nod^{-} and Nod^{++} soybeans) under SN fertilization were also responding markedly to the effects of HN fertilization for all soybean genotypes (Figures 4–6). In addition, OTUs GP7, GP37 and GP38 were identified as responsible for the increase of the relative abundance of Gammaproteobacteria in all soybean genotypes under HN fertilization (Figures 4 and 6). The representative clone of OTU GP7 was classified as *Lysobacter* sp. and its closest relative among known species was *Lysobacter koreensis* (AB166878) with 96% identity (Figure 6). The representative clone of OTU GP37 was identical to several *Enterobacter* spp., including *Enterobacter aerogenes* (FJ796202) and *Enterobacter asburiae* (AB004744). The representative clone of OTU GP38 was identical to several species of Enterobacteriaceae, including *Pantoea agglomerans* (FJ603033), *Enterobacter ludwigii* (FJ462703), *Enterobacter cloacae* (FJ194527), *Enterobacter asburiae* (EU301774) and *Citrobacter freundii* (DQ133536). OTUs GP37, GP38 and GP39 were shown to be closely related to each other, as indicated by clustering in a branch of the Enterobacteriales (Figures 4 and 6).

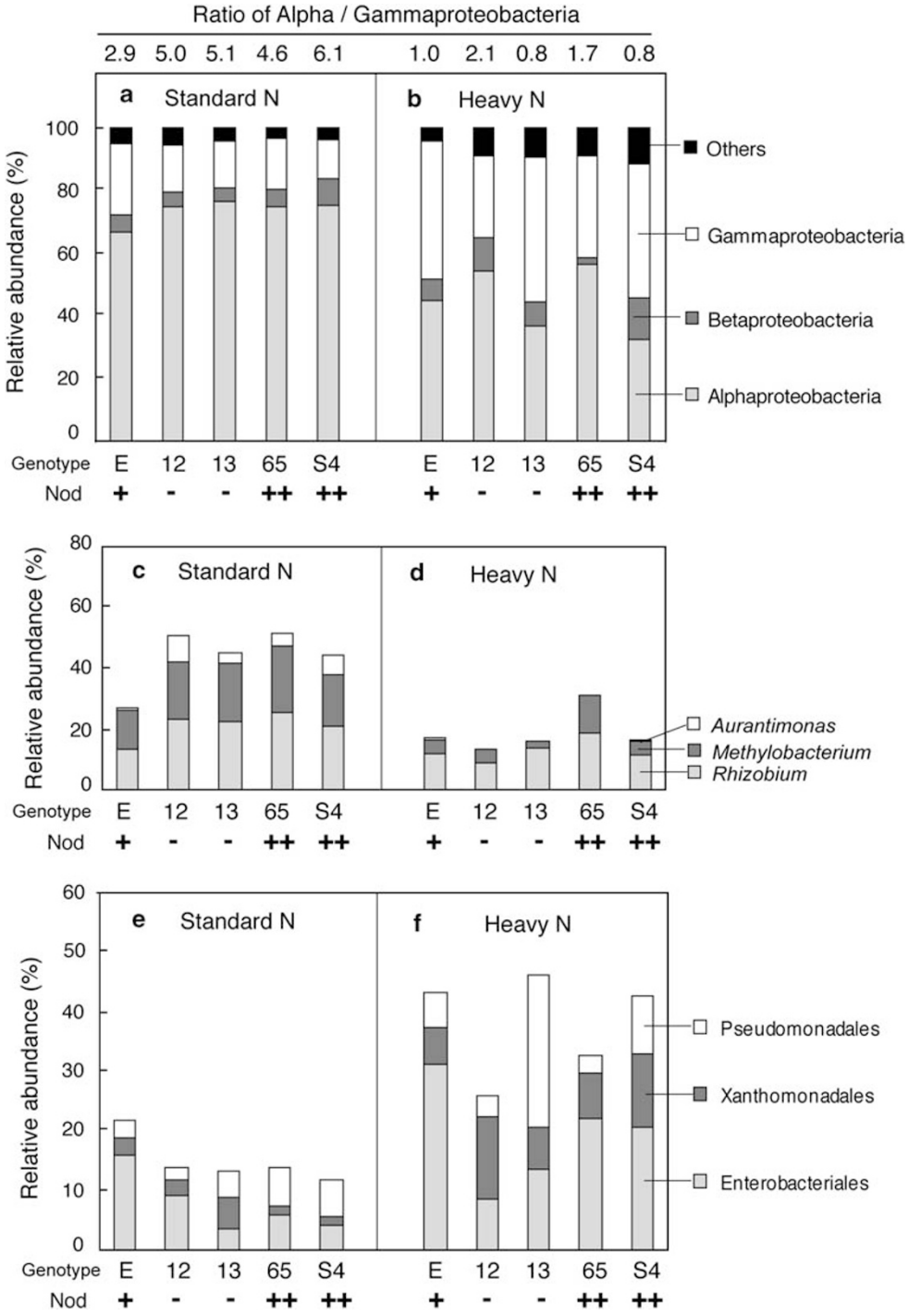


Figure 2 The phylogenetic compositions of 16S ribosomal RNA (rRNA) gene clone libraries of soybean shoot-associated bacteria in different nodulation phenotypes under standard and heavy nitrogen (HN) fertilization conditions. (**a**, **b**) Phylogenetic compositions at the phylum level; (**c**, **d**) relative abundance of three genera of Rhizobiales; (**e**, **f**) relative abundance of three orders of Gammaproteobacteria. Genotype stand for the name of plant materials: Enrei (**e**), En 1282 (12), En 1314 (13), En 6500 (65), Sakukei 4 (S4). Nod stands for the nodulation phenotype: wild-type (+), non-nodulated (−) and hypernodulated (++).

Discussion

The phylogenetic analyses based on 16S rRNA gene sequences revealed the striking effects of both nodulation phenotype and HN fertilization on bacterial communities in soybean stems (Figures 2 and 3). The population shifts of Alpha- and Gammaproteobacteria under SN fertilization indicated that the nodulation phenotypes affected the bacterial community structures not only in roots (Ikeda *et al.*, 2008a), in which nodule formation occurs, but also systemically in shoots.

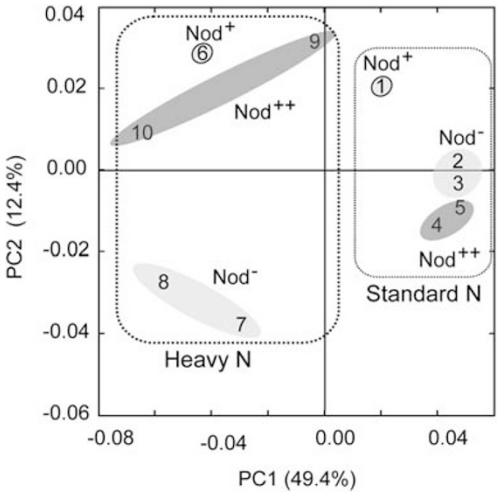


Figure 3 Principal-coordinate analysis for the 16S ribosomal RNA (rRNA) gene clone libraries of soybean stem-associated bacteria in different nodulation phenotypes. The ordination was constructed using Unifrac distances weighted by the relative abundances. Each number represents the phylogenetic composition of soybean stem-associated bacteria in different soybean genotypes under standard and heavy nitrogen (HN) conditions. Numbers 1 to 5 and 6 to 10 were sampled under standard and HN conditions, respectively. Soybean genotypes: 1 and 6: wild-type (Enrei, Nod^{+}); 2 and 7: non-nodulated (En 1282, Nod^{-}); 3 and 8: non-nodulated (En 1314, Nod^{-}); 4 and 9: hypernodulated (En 6500, Nod^{++}); 5 and 10: hypernodulated (Sakukei 4, Nod^{++}).

Surprisingly, further phylogenetic analyses at lower taxonomic levels and the results of PCoA revealed that soybeans of two very different nodulation phenotypes, Nod^{-} and Nod^{++}, hosted highly similar microbial community structures under the SN fertilization condition, even at the species level (Figures 3 and 4). A possible explanation for this may be the absence of mycorrhization in Nod^{-} soybeans under the present experimental conditions (data not shown). The rotation of the experimental field between upland and paddy conditions could reduce the population of mycorrhizal fungi. In the absence of mycorrhization, Nod^{-} soybeans do not need to activate autoregulation systems, whereas the autoregulation systems of Nod^{++} soybeans were genetically impaired. Thus, both Nod^{-} and Nod^{++} soybeans may be considered as hosts with no activated autoregulation in this study. Therefore, a bacterial population related to rhizobia may be passively affected by a part of the autoregulation systems for regulating rhizobia. In fact, the three genera (*Aurantimonas*, *Methylobacterium* and *Rhizobium*) in Rhizobiales found in the present study are all closely related to *Bradyrhizobium* spp., endosymbionts of soybeans (Gupta and Mok, 2007). Another surprising finding was the high similarity observed for the population shifts of Alpha- and Gammaproteobacteria between two different environmental factors: the effects of genotypes under SN fertilization and the effects of the nitrogen fertilization under the HN fertilization condition. Thus, the same taxonomic groups that were found in the population shifts of Alpha- and Gammaproteobacteria among different nodulation phenotypes under the SN fertilization condition were also responsible for the drastic population shifts of Alpha- and Gammaproteobacteria between the SN and HN fertilization conditions, regardless of nodulation phenotype. These results suggest that exogenous or endogenous nitrogen levels could be an important signal for regulating a subpopulation of proteobacteria by activating a plant signaling pathway, such as autoregulation systems in legume–rhizobia symbiosis.

The clustering analysis of 16S rRNA sequences (Figure 4) identified OTUs responsible for the population shifts of Alpha- and Gammaproteobacteria at the species level. The type strain of *A. ureilytica* was originally reported as an isolate from air dust (Weon *et al.*, 2007), but the closely related isolates (AB291857 and AB291900) have been identified on the surface and inside of rice leaves (Mano *et al.*, 2007). The genus *Aurantimonas* has been recently established (Denner *et al.*, 2003), and the biology of this genus in relation to the plant symbiosis is largely unknown, and needs to conduct further study. *Methylobacterium chloromethanicum* is currently considered as a synonym of *Methylobacterium extorquens* (Kato *et al.*, 2005). Several *Methylobacterium* spp., including *M. extorquens*, are well-known plant-associated bacteria (Corpe, 1985) as both an epiphyte and an endophyte, and the scavenging monocarbon wastes such as methanol in plant tissues is speculated for the putative role of this bacterial group (Corpe and Basil, 1982). Therefore, the population shifts of *Methylobacterium* sp. could be considered as an indication for the change of physiological status in plants. In fact, it has been reported that both the nodulation and the nitrate sensing in plant activate the autoregulation system that causes the drastic shift of internal hormone balance in plants (Caba *et al.*, 2000; Okamoto *et al.*, 2009). The representative clones for these OTUs showed high similarity to known species. Therefore, it would be feasible to isolate bacteria corresponding to these OTUs and to examine the interactions of these bacteria with soybeans by an inoculation test to clarify the speculations described above. In particular, it would be interesting to analyze the interaction between the *Aurantimonas* sp. representing OTU AP6 and soybeans because of its high sensitivity to nodulation genotype and nitrogen fertilization level.

On the basis of a culture-dependent method, Gammaproteobacteria were reported as the most dominant group of stem-associated bacteria in soybeans (up to 94% and 100% for epiphyte and endophyte isolates, respectively), whereas the Alphaproteobacteria isolated from shoots were shown to be far less abundant (5.9% for epiphytes; Kuklinsky-Sobral *et al.*, 2004). Okubo *et al.* (2009) also reported similar results of the strong dominance

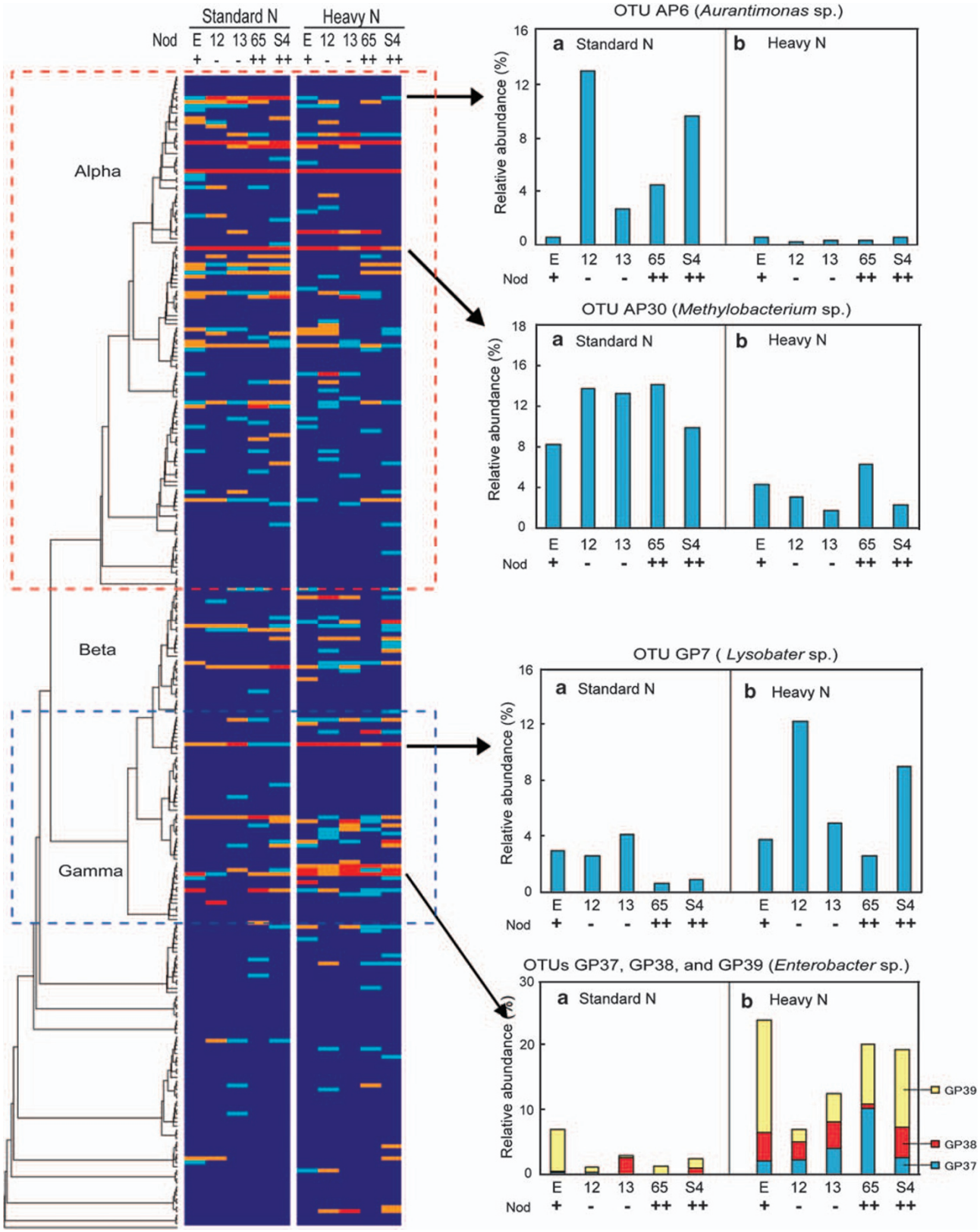
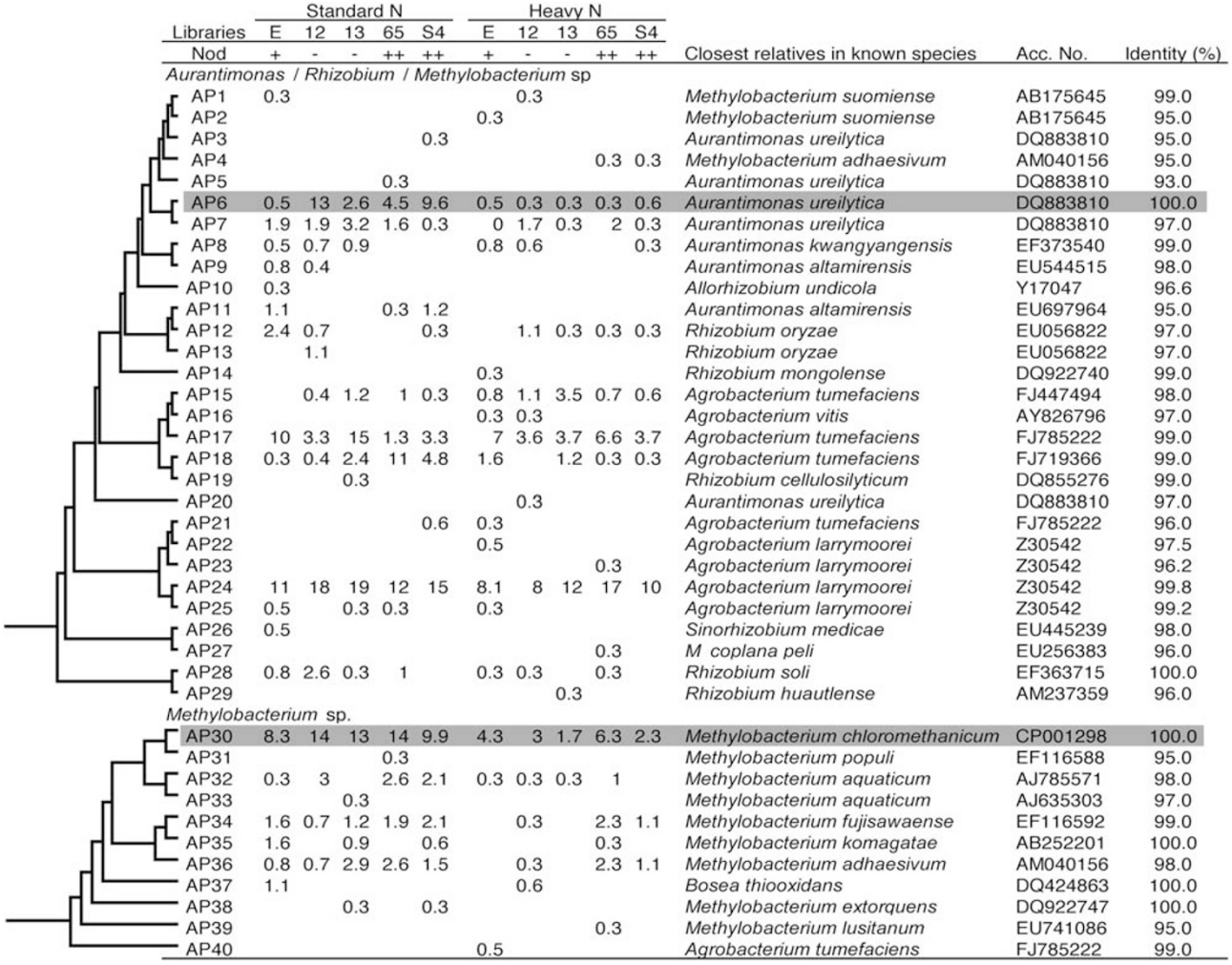


Figure 4 Heat map for the relative abundance of operational taxonomic units (OTUs) for the 16S ribosomal RNA (rRNA) gene clone libraries of soybean stem-associated bacteria under standard and heavy nitrogen (HN) fertilization conditions. Soybean genotypes are as follows: E: Enrei; 12: En 1282; 13: En 1314; 65: En 6500; and S4: Sakukei 4. Nod stands for the nodulation phenotype: wild-type (+), non-nodulated (−) and hypernodulated (++). The relative abundance of each OTU in each library is indicated by different colors: sky blue (<1%), orange (1–3%) and red (>3%), with a background of dark blue (0%). The relative abundance for the OTUs described in the main text is shown at the right side.



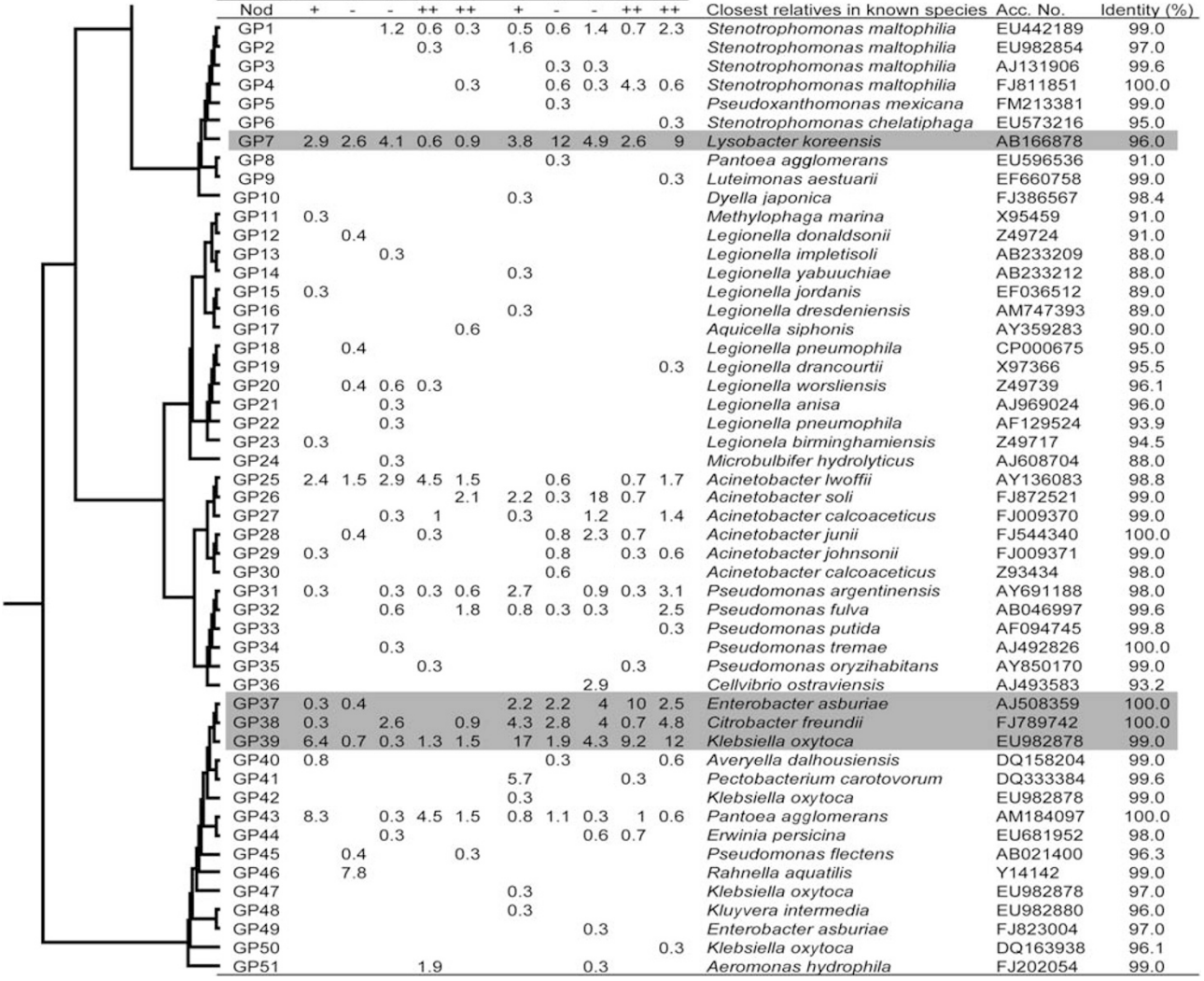
Libraries	Standard N					Heavy N					Closest relatives in known species	Acc. No.	Identity (%)
Nod	E	12	13	65	S4	E	12	13	65	S4			
	+	-	-	++	++	+	-	-	++	++			
Aurantimonas / *Rhizobium* / *Methylobacterium* sp													
AP1	0.3						0.3				*Methylobacterium suomiense*	AB175645	99.0
AP2						0.3					*Methylobacterium suomiense*	AB175645	95.0
AP3					0.3						*Aurantimonas ureilytica*	DQ883810	95.0
AP4									0.3	0.3	*Methylobacterium adhaesivum*	AM040156	95.0
AP5				0.3							*Aurantimonas ureilytica*	DQ883810	93.0
AP6	0.5	13	2.6	4.5	9.6	0.5	0.3	0.3	0.3	0.6	*Aurantimonas ureilytica*	DQ883810	100.0
AP7	1.9	1.9	3.2	1.6	0.3	0	1.7	0.3	2	0.3	*Aurantimonas ureilytica*	DQ883810	97.0
AP8	0.5	0.7	0.9			0.8	0.6			0.3	*Aurantimonas kwangyangensis*	EF373540	99.0
AP9	0.8	0.4									*Aurantimonas altamirensis*	EU544515	98.0
AP10	0.3										*Allorhizobium undicola*	Y17047	96.6
AP11	1.1			0.3	1.2						*Aurantimonas altamirensis*	EU697964	95.0
AP12	2.4	0.7			0.3		1.1	0.3	0.3	0.3	*Rhizobium oryzae*	EU056822	97.0
AP13		1.1									*Rhizobium oryzae*	EU056822	97.0
AP14						0.3					*Rhizobium mongolense*	DQ922740	99.0
AP15		0.4	1.2	1	0.3	0.8	1.1	3.5	0.7	0.6	*Agrobacterium tumefaciens*	FJ447494	98.0
AP16						0.3	0.3				*Agrobacterium vitis*	AY826796	97.0
AP17	10	3.3	15	1.3	3.3	7	3.6	3.7	6.6	3.7	*Agrobacterium tumefaciens*	FJ785222	99.0
AP18	0.3	0.4	2.4	11	4.8	1.6		1.2	0.3	0.3	*Agrobacterium tumefaciens*	FJ719366	99.0
AP19			0.3								*Rhizobium cellulosilyticum*	DQ855276	99.0
AP20							0.3				*Aurantimonas ureilytica*	DQ883810	97.0
AP21					0.6	0.3					*Agrobacterium tumefaciens*	FJ785222	96.0
AP22						0.5					*Agrobacterium larrymoorei*	Z30542	97.5
AP23									0.3		*Agrobacterium larrymoorei*	Z30542	96.2
AP24	11	18	19	12	15	8.1	8	12	17	10	*Agrobacterium larrymoorei*	Z30542	99.8
AP25	0.5		0.3	0.3		0.3					*Agrobacterium larrymoorei*	Z30542	99.2
AP26	0.5										*Sinorhizobium medicae*	EU445239	98.0
AP27									0.3		*M coplana peli*	EU256383	96.0
AP28	0.8	2.6	0.3	1		0.3	0.3		0.3		*Rhizobium soli*	EF363715	100.0
AP29								0.3			*Rhizobium huautlense*	AM237359	96.0
Methylobacterium sp.													
AP30	8.3	14	13	14	9.9	4.3	3	1.7	6.3	2.3	*Methylobacterium chloromethanicum*	CP001298	100.0
AP31				0.3							*Methylobacterium populi*	EF116588	95.0
AP32	0.3	3		2.6	2.1	0.3	0.3	0.3	1		*Methylobacterium aquaticum*	AJ785571	98.0
AP33			0.3								*Methylobacterium aquaticum*	AJ635303	97.0
AP34	1.6	0.7	1.2	1.9	2.1		0.3		2.3	1.1	*Methylobacterium fujisawaense*	EF116592	99.0
AP35	1.6		0.9		0.6				0.3		*Methylobacterium komagatae*	AB252201	100.0
AP36	0.8	0.7	2.9	2.6	1.5		0.3		2.3	1.1	*Methylobacterium adhaesivum*	AM040156	98.0
AP37	1.1						0.6				*Bosea thiooxidans*	DQ424863	100.0
AP38			0.3		0.3						*Methylobacterium extorquens*	DQ922747	100.0
AP39									0.3		*Methylobacterium lusitanum*	EU741086	95.0
AP40						0.5					*Agrobacterium tumefaciens*	FJ785222	99.0

Figure 5 Phylogenetic distribution of operational taxonomic units (OTUs) for Alphaproteobacteria of the clone libraries of 16S ribosomal RNA (rRNA) genes for soybean stem-associated bacteria. The dendrogram (left) indicates the phylogenetic relationships among the representative sequences of OTUs (defined by 97% identity). The table indicates the relative abundance of clones belonging to each OTU in each library and the results of a BLAST search using the representative sequences. The soybean genotypes are as follows: E: Enrei; 12: En 1282; 13: En 1314; 65: En 6500; and S4: Sakukei 4. Nod stands for the nodulation phenotype: wild-type (+), non-nodulated (−) and hypernodulated (++).

of Gammaproteobacteria were obtained when culturing endophytic bacteria in soybean stems. In contrast, based on the culture-independent method used in this study, Alphaproteobacteria were the most dominant group in soybean shoots under SN fertilization. These differences may simply reflect experimental biases between the culture-dependent and culture-independent methods used for analyzing bacterial communities. By considering the known diversity of plant-associated bacteria, however, the phylogenetic compositions strongly and unusually biased toward to the Gammaproteobacteria are not likely to simply reflect the nature of soybean-associated bacterial communities. As the relative abundance of Alphaproteobacteria was increased in nodulation mutants under the SN fertilization condition in the present study, a plant signaling pathway such as autoregulation systems in the legume–rhizobia symbiosis may be responsible for the dominance of Gammaproteobacteria in the culture of soybean-associated bacteria.

Exogenous nitrate suppresses the infection of rhizobia and inhibits nodule formation (Carroll *et al*., 1985; Barbulova *et al*., 2007; Okamoto *et al*., 2009). Recently, Okamoto *et al*. (2009) suggested that a nitrate-induced *CLE* gene in *Lotus japonica* is involved in autoregulation systems through the *HAR1* gene, which is a homolog of *GmNARK* in soybeans.

In this study, the HN treatment was initially performed to eliminate the effects of nitrogen deficiency in Nod^- soybeans on the microbial community and to inhibit nodulation in wild-type soybeans (Nod^+). The HN fertilization increased nitrate contents in the soil (7–10 mmol NO_3^- per kg dry soil), which were four- to sixfold higher than those in the SN field soil (1.2–2.6 mmol NO_3^- per kg dry soil) throughout the soybean cultivation period



	Standard N					Heavy N							
Libraries	E	12	13	65	S4	E	12	13	65	S4			
Nod	+	-	-	++	++	+	-	-	++	++	Closest relatives in known species	Acc. No.	Identity (%)
GP1			1.2	0.6	0.3	0.5	0.6	1.4	0.7	2.3	*Stenotrophomonas maltophilia*	EU442189	99.0
GP2				0.3		1.6					*Stenotrophomonas maltophilia*	EU982854	97.0
GP3							0.3	0.3			*Stenotrophomonas maltophilia*	AJ131906	99.6
GP4					0.3		0.6	0.3	4.3	0.6	*Stenotrophomonas maltophilia*	FJ811851	100.0
GP5							0.3				*Pseudoxanthomonas mexicana*	FM213381	99.0
GP6										0.3	*Stenotrophomonas chelatiphaga*	EU573216	95.0
GP7	2.9	2.6	4.1	0.6	0.9	3.8	12	4.9	2.6	9	*Lysobacter koreensis*	AB166878	96.0
GP8							0.3				*Pantoea agglomerans*	EU596536	91.0
GP9										0.3	*Luteimonas aestuarii*	EF660758	99.0
GP10						0.3					*Dyella japonica*	FJ386567	98.4
GP11	0.3										*Methylophaga marina*	X95459	91.0
GP12		0.4									*Legionella donaldsonii*	Z49724	91.0
GP13			0.3								*Legionella impletisoli*	AB233209	88.0
GP14						0.3					*Legionella yabuuchiae*	AB233212	88.0
GP15	0.3										*Legionella jordanis*	EF036512	89.0
GP16						0.3					*Legionella dresdeniensis*	AM747393	89.0
GP17					0.6						*Aquicella siphonis*	AY359283	90.0
GP18		0.4									*Legionella pneumophila*	CP000675	95.0
GP19										0.3	*Legionella drancourtii*	X97366	95.5
GP20		0.4	0.6	0.3							*Legionella worsliensis*	Z49739	96.1
GP21			0.3								*Legionella anisa*	AJ969024	96.0
GP22			0.3								*Legionella pneumophila*	AF129524	93.9
GP23	0.3										*Legionela birminghamiensis*	Z49717	94.5
GP24			0.3								*Microbulbifer hydrolyticus*	AJ608704	88.0
GP25	2.4	1.5	2.9	4.5	1.5		0.6		0.7	1.7	*Acinetobacter lwoffii*	AY136083	98.8
GP26					2.1	2.2	0.3	18	0.7		*Acinetobacter soli*	FJ872521	99.0
GP27			0.3	1		0.3		1.2		1.4	*Acinetobacter calcoaceticus*	FJ009370	99.0
GP28		0.4		0.3			0.8	2.3	0.7		*Acinetobacter junii*	FJ544340	100.0
GP29	0.3						0.8		0.3	0.6	*Acinetobacter johnsonii*	FJ009371	99.0
GP30							0.6				*Acinetobacter calcoaceticus*	Z93434	98.0
GP31	0.3		0.3	0.3	0.6	2.7		0.9	0.3	3.1	*Pseudomonas argentinensis*	AY691188	98.0
GP32			0.6		1.8	0.8	0.3	0.3		2.5	*Pseudomonas fulva*	AB046997	99.6
GP33										0.3	*Pseudomonas putida*	AF094745	99.8
GP34			0.3								*Pseudomonas tremae*	AJ492826	100.0
GP35				0.3					0.3		*Pseudomonas oryzihabitans*	AY850170	99.0
GP36								2.9			*Cellvibrio ostraviensis*	AJ493583	93.2
GP37	0.3	0.4				2.2	2.2	4	10	2.5	*Enterobacter asburiae*	AJ508359	100.0
GP38	0.3		2.6		0.9	4.3	2.8	4	0.7	4.8	*Citrobacter freundii*	FJ789742	100.0
GP39	6.4	0.7	0.3	1.3	1.5	17	1.9	4.3	9.2	12	*Klebsiella oxytoca*	EU982878	99.0
GP40	0.8						0.3			0.6	*Averyella dalhousiensis*	DQ158204	99.0
GP41						5.7			0.3		*Pectobacterium carotovorum*	DQ333384	99.6
GP42						0.3					*Klebsiella oxytoca*	EU982878	99.0
GP43	8.3		0.3	4.5	1.5	0.8	1.1	0.3	1	0.6	*Pantoea agglomerans*	AM184097	100.0
GP44			0.3					0.6	0.7		*Erwinia persicina*	EU681952	98.0
GP45		0.4			0.3						*Pseudomonas flectens*	AB021400	96.3
GP46		7.8									*Rahnella aquatilis*	Y14142	99.0
GP47						0.3					*Klebsiella oxytoca*	EU982878	97.0
GP48						0.3					*Kluyvera intermedia*	EU982880	96.0
GP49								0.3			*Enterobacter asburiae*	FJ823004	97.0
GP50										0.3	*Klebsiella oxytoca*	DQ163938	96.1
GP51				1.9				0.3			*Aeromonas hydrophila*	FJ202054	99.0

Figure 6 Phylogenetic distribution of operational taxonomic units (OTUs) for Gammaproteobacteria of the clone libraries of 16S ribosomal RNA (rRNA) genes for soybean stem-associated bacteria. The dendrogram (left) indicates the phylogenetic relationships among the representative sequences of OTUs (defined by 97% identity). The table indicates the relative abundance of clones belonging to each OTU in each library and the results of a BLAST search using the representative sequences. The soybean genotypes are as follows: E: Enrei; 12: En 1282; 13: En 1314; 65: En 6500; and S4: Sakukei 4. Nod stands for the nodulation phenotype: wild-type (+), non-nodulated (−) and hypernodulated (++).

(Supplementary Table 1). Plant dry weights and nitrogen contents showed no significant difference between nodulated Enrei and the Nod$^-$ mutant under the HN field conditions (Supplementary Figures 2a–d, Figure 2i). Under the HN fertilization conditions, nodule dry weight and nodule number of wild-type soybean Enrei were also markedly inhibited.

These results suggest that HN fertilization gave rise to nitrate-induced nodulation inhibition through the soybean autoregulation system. Thus, a GmNARK-mediated signaling pathway may regulate a subpopulation of Alphaproteobacteria, which may be closely related to rhizobia, beyond the regulation of nodule formation by *Bradyrhizobium japonicum* in soybeans. As a consequence, this may have resulted in the low ratio of Alpha/Gammaproteobacteria in wild-type soybeans.

It is also interesting that population shifts were observed for Nod^{++} soybeans under both SN and the HN fertilization conditions (Figures 2 and 4). As Nod^{++} soybeans do not harbor the functional *GmNARK* gene (Nishimura *et al.*, 2002), the *GmNARK*-mediated nitrate sensing systems could not regulate nodule formation (Okamoto *et al.*, 2009). Nevertheless, the effects of HN fertilization on Nod^{++} soybeans occurred systemically, and the mode of the community shifts was very similar to those for other nodulation phenotypes (Nod$^+$ and Nod$^-$ soybeans). Therefore, the *GmNARK* gene is not essential for the population shifts of Alpha- and Gammaproteobacteria responding to

HN fertilization. These results also may imply that presence of *GmNARK*-independent nitrogen-sensing system(s) affects the shoot-associated bacterial community in soybeans. However, there is a possibility that the Gammaproteobacteria population may be more adaptive than Alphaproteobacteria in soybeans under the influence of autoregulation and the effects of HN fertilization.

In conclusion, this study showed that the community structure of stem-associated bacteria was affected by nodulation phenotype and nitrogen fertilization level. Both nodulation phenotypes and nitrogen levels caused a similar directional change in the stem-associated bacterial communities, with a population shift of Alpha- and Gammaproteobacteria. Further analyses identified several OTUs in Alpha- and Gammaproteobacteria responding to the population shifts. The relative abundances of *Aurantimonas* and *Methylobacterium* sp. were especially sensitive to the nodulation phenotype and nitrogen level, being decreased in the wild-type nodulation phenotype and under HN fertilization. These results imply that the nitrogen signaling pathway in plants may be related to the symbiotic regulation systems between legumes and rhizobia and could have a significant influence on the diversity of plant-associated microbes.

In this study, the clone library analyses were conducted on the whole community of stem-associated bacteria including both the epiphytic and endophytic bacteria rather than on the each of epiphytic and endophytic bacterial communities. Therefore, the important issues for the tissue localization of bacteria responding to nodulation phenotypes and soil N levels remain unanswered, and should be addressed in the following study. Quantitative analyses of the population shifts of Alpha- and Gammaproteobacteria and the isolation of bacteria responsible for this population shift should also be conducted, which would facilitate our understanding of how plants regulate and shape the structures of microbial communities.

Acknowledgements

We thank M. Kokubun (Tohoku University) for providing the soybean seeds. This work was supported in part by Special Coordination Funds for Promoting Science and Technology and in part by a Grant-in-Aid for Scientific Research on Priority Areas 'Comparative Genomics' from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and in part by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics for Agricultural Innovation, PMI-0002).

References

Akao S, Kouchi H. (1989). A supernodulating mutant isolated from soybean cultivar Enrei. *Soil Sci Plant Nutr* **38**: 183–187.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ. (2006). New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* **72**: 5734–5741.

Barbulova A, Rogato A, D'Apuzzo E, Omrane S, Chiurazzi M. (2007). Differential effects of combined N sources on early steps of the Nod factor-dependent transduction pathway in *Lotus japonicus*. *Mol Plant Microbe Interact* **20**: 994–1003.

Caba JM, Centeno ML, Fernández B, Gresshoff PM, Ligero F. (2000). Inoculation and nitrate alter phytohormone levels in soybean roots: differences between a supernodulating mutant and the wild type. *Planta* **211**: 98–104.

Carroll BJ, McNeil DL, Gresshoff PM. (1985). Isolation and properties of soybean [*Glycine max* (L.) Merr.] mutants that nodulate in the presence of high nitrate concentrations. *Proc Natl Acad Sci USA* **82**: 4162–4166.

Corpe WA. (1985). A method for detecting methylotrophic bacteria on solid surfaces. *J Microbiol Methods* **3**: 215–221.

Corpe WA, Basil DV. (1982). Methanol-utilizing bacteria, associated wit green plants. *Dev Indust Microbiol* **23**: 483–493.

Denner EBM, Smith GW, Busse H-J, Schumann P, Narzt T, Polson SW *et al.* (2003). *Aurantimonas coralicida* gen, nov., sp. nov., the causative agent of white plague type II on Caribbean scleractinian corals. *Int J Syst Evol Microbiol* **53**: 1115–1122.

Ewing B, Green P. (1998). Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* **8**: 186–194.

Ewing B, Hillier L, Wendl MC, Green P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* **8**: 175–185.

Fehr RW, Caviness CE, Burmood DT, Pennington JS. (1971). Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci* **11**: 929–931.

Felsenstein J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Francisco PB, Akao S. (1993). Autoregulation and nitrate inhibition of nodule formation in soybean cv. Enrei and its nodulation mutants. *J Exp Bot* **44**: 547–553.

Good IJ. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40**: 237–264.

Gupta RS, Mok A. (2007). Phylogenomics and signature proteins for alpha proteobacteria and its main groups. *BMC Microbiol* **7**: 106.

Hardoim PR, Van Overbeek LS, Van Elsas JD. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* **16**: 463–471.

Harrison MJ. (2005). Signaling in the arbuscular mycorrhizal symbiosis. *Annu Rev Microbiol* **59**: 19–42.

Ikeda S, Kaneko T, Okubo T, Rallos EEL, Eda S, Mitsui H *et al.* (2009). Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. *Microb Ecol* **58**: 703–714.

Ikeda S, Rallos LEE, Okubo T, Eda S, Inaba S, Mitsui H *et al.* (2008a). Microbial community analysis of field-grown soybeans with different nodulation phenotypes. *Appl Environ Microbiol* **74**: 5704–5709.

Ikeda S, Tsurumaru H, Wakai S, Noritake C, Fujishiro K, Akasaka M *et al.* (2008b). Evaluation of the effects of different additives in improving the DNA extraction yield and quality from Andosol. *Microbes Environ* **23**: 159–166.

Kato Y, Asahara M, Arai D, Goto K, Yokota A. (2005). Reclassification of *Methylobacterium chloromethanicum* and *Methylobacterium dichloromethanicum* as later subjective synonyms of *Methylobacterium extorquens* and of *Methylobacterium lusitanum* as a later subjective synonym of *Methylobacterium rhodesianum*. *J Gen Appl Microbiol* **51**: 287–299.

Kemp PF, Aller JY. (2004). Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol Ecol* **47**: 161–177.

Kinkema M, Gresshoff PM. (2008). Investigation of downstream signals of the soybean autoregulation of nodulation receptor kinase *GmNARK*. *Mol Plant Microbe Interact* **21**: 1337–1348.

Kuklinsky-Sobral J, Araujo WL, Mendes R, Geraldi IO, Pizzirani-Kleiner AA, Azevedo JL. (2004). Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Environ Microbiol* **6**: 1244–1251.

Lane DJ. (1991). 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds). *Nucleic Acid Techniques in Bacterial Systematics*. John Wiley & Sons Ltd. pp 115–175.

Liu J, Maldonado-Mendoza I, Lopez-Meyer M, Cheung F, Town CD, Harrison MJ. (2007). Arbuscular mycorrhizal symbiosis is accompanied by local and systemic alterations in gene expression and an increase in disease resistance in the shoots. *Plant J* **50**: 529–544.

Lozupone C, Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**: 8228–8235.

Mano H, Tanaka F, Nakamura C, Kaga H, Morisaki H. (2007). Culturable endophytic bacterial flora of the maturing leaves and roots of rice plants (*Oryza sativa*) cultivated in a paddy field. *Microbes Environ* **22**: 175–185.

Matsunami T, Kaihatsu A, Maekawa T, Takahashi M, Kokubun M. (2004). Characterization of vegetative growth of a supernodulating soybean genotype, Sakukei 4. *Plant Prod Sci* **7**: 165–171.

Meixner C, Vegvari G, Ludwig-Müller J, Gagnon H, Steinkellner S, Staehelin C *et al.* (2007). Two defined alleles of the LRR receptor kinase *GmNARK* in supernodulating soybean govern differing autoregulation of mycorrhization. *Physiologia Plantarum* **130**: 261–270.

Nishimura R, Hayashi M, Wu G, Kouchi H, Imaizumi-Anraku H, Murakami Y *et al.* (2002). HAR1 mediates systemic regulation of symbiotic organ development. *Nature* **420**: 426–429.

Offre P, Pivato B, Siblot S, Gamalero E, Corberand T, Lemanceau P *et al.* (2007). Identification of bacterial groups preferentially associated with mycorrhizal roots of *Medicago truncatula*. *Appl Environ Microbiol* **73**: 913–921.

Oka-Kira E, Kawaguchi M. (2006). Long-distance signaling to control root nodule number. *Curr Opin Plant Biol* **9**: 496–502.

Okamoto S, Ohnishi E, Sato S, Takahashi H, Nakazono M, Tabata S *et al.* (2009). Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. *Plant Cell Physiol* **50**: 67–77.

Okubo T, Ikeda S, Kaneko T, Eda S, Mitsui H, Sato S *et al.* (2009). Nodulation-dependent communities of culturable bacterial endophytes from stems of field-grown soybeans. *Microbes Environ* **24**: 253–258.

Page RDM. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Bioinformatics* **12**: 357–358.

Parniske M. (2000). Intracellular accommodation of microbes by plants: a common developmental program for symbiosis and disease? *Curr Opin Plant Biol* **3**: 320–328.

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M *et al.* (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**: 585–592.

Saito A, Ikeda S, Ezura H, Minamisawa K. (2007). Microbial community analysis of the phytosphere using culture-independent methodologies. *Microbes Environ*. **22**: 93–105.

Saitou N, Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Schloss PD, Handelsman J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.

Schloss PD, Handelsman J. (2006). Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl Environ Microbiol* **72**: 6773–6779.

Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ *et al.* (2003). Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science* **299**: 109–112.

Stacey G, Libault M, Brechenmacher L, Wan J, May GD. (2006). Genetics and functional genomics of legume nodulation. *Curr Opin Plant Biol* **9**: 110–121.

Thompson JD, Higgins DG, Gibson TJ. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Weon H-Y, Kim B-Y, Yoo S-H, Joa J-H, Lee KH, Zhang Y-S Z *et al.* (2007). *Aurantimonas ureilytica* sp. nov., isolated from an air sample. *Int J Syst Evol Microbiol* **57**: 1717–1720.

Zhou J, Xia B, Treves DS, Wu LY, Marsh TL, O'Neill RV *et al.* (2002). Spatial and resource factors influencing high microbial diversity in soil. *Appl Environ Microbiol* **68**: 326–334.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)