

Functional Characteristics of an Endophyte Community Colonizing Rice Roots as Revealed by Metagenomic Analysis

A. Sessitsch,¹ P. Hardoim,^{2,3} J. Döring,⁴ A. Weilharter,¹ A. Krause,⁴ T. Woyke,⁵ B. Mitter,¹ L. Hauberg-Lotte,⁴ F. Friedrich,⁴ M. Rahalkar,⁴ T. Hurek,⁴ A. Sarkar,⁴ L. Bodrossy,¹ L. van Overbeek,³ D. Brar,⁶ J. D. van Elsas,² and B. Reinhold-Hurek⁴

¹AIT Austrian Institute of Technology GmbH, Bioresources Unit, Tulln, Austria; ²University of Groningen, Department of Microbial Ecology, Haren, The Netherlands; ³Plant Research International, 6708PB Wageningen, The Netherlands; ⁴University of Bremen, Department of Microbe-Plant Interactions, Bremen, Germany; ⁵Department of Energy (DOE) Joint Genome Institute, Walnut Creek, CA, U.S.A.; ⁶International Rice Research Center (IRRI), Los Baños, Philippines

Submitted 1 August 2011. Accepted 9 September 2011.

Roots are the primary site of interaction between plants and microorganisms. To meet food demands in changing climates, improved yields and stress resistance are increasingly important, stimulating efforts to identify factors that affect plant productivity. The role of bacterial endophytes that reside inside plants remains largely unexplored, because analysis of their specific functions is impeded by difficulties in cultivating most prokaryotes. Here, we present the first metagenomic approach to analyze an endophytic bacterial community resident inside roots of rice, one of the most important staple foods. Metagenome sequences were obtained from endophyte cells extracted from roots of field-grown plants. Putative functions were deduced from protein domains or similarity analyses of protein-encoding gene fragments, and allowed insights into the capacities of endophyte cells. This allowed us to predict traits and metabolic processes important for the endophytic lifestyle, suggesting that the endorhizosphere is an exclusive microhabitat requiring numerous adaptations. Prominent features included flagella, plant-polymer-degrading enzymes, protein secretion systems, iron acquisition and storage, quorum sensing, and detoxification of reactive oxygen species. Surprisingly, endophytes might be involved in the entire nitrogen cycle, as protein domains involved in N₂-fixation, denitrification, and nitrification were detected and selected genes expressed. Our data suggest a high potential of the endophyte community for plant-growth promotion, improvement of plant stress resistance, biocontrol against pathogens, and bioremediation, regardless of their culturability.

Rice is the staple food for the largest number of people on Earth, and rice cultivation is resource intensive. Annually, irrigated rice production consumes approximately 10% of global fertilizer N production and approximately 40% of developed water supplies in Asia; however, it also contributes to greenhouse gas emissions by microbial conversions. The competition for arable land and water and the rise in rice consumption

demand increased yields and improved management practices (FAO 2008). In particular, endophytes that reside inside tissues of healthy plants are likely to positively affect the host (Weyens et al. 2009); however, the potential of applying endophytic bacteria is still underexplored (Mano and Morisaki 2008). Rice roots harbor endophytes (Hurek et al. 1994), which can reach up to 10⁸ cultivable N₂-fixing bacteria per gram of root dry weight (Barraquio et al. 1997) and even larger numbers of bacteria that have defied cultivation so far (Engelhard et al. 2000; Knauth et al. 2005; Mano and Morisaki 2008). Thus, endophytes are likely to considerably affect their host plant.

To understand and manipulate their contribution, it is important to decode metabolic processes, adaptations, and beneficial characteristics. However, assessing microbial functions is impeded by difficulties in cultivating most prokaryotes, and endophytes inside host tissues are not easily amenable to biochemical or genetic analyses. Endophytic bacterial cells inside plant tissues, tightly attached to host cells, are difficult to extract and separate from plant cells and contaminating surface bacteria. Mechanical removal of rhizoplane populations had been demonstrated previously, using vigorous shaking with glass beads (Reinhold et al. 1986). Because the culturable endorhizosphere population was found to be entirely different from the rhizoplane population (Reinhold et al. 1986), this treatment appeared to be efficient. In roots of flooded plants with large aerenchymatic tissue such as Kallar grass (Reinhold et al. 1986) or rice (Barraquio et al. 1997), the population of culturable root endophytes was reported to be high compared with the rhizoplane population, making rice roots a promising microhabitat in which to study endophytes.

Our knowledge of functions, niche adaptations, host-microbe interactions, and putative beneficial traits depends, thus far, on cultivated endophytes from rice (Krause et al. 2006; Yan et al. 2008) and other plants (Compant et al. 2010). Cellular capacities of uncultured microbial communities can be deciphered using metagenomic approaches (Dinsdale et al. 2008), in which DNA extracted from the entire population is analyzed for its gene content. We applied this strategy to investigate, for the first time in a culture-independent manner, key biological processes contributed by an endophytic bacterial community in roots of *Oryza sativa*. Analysis of protein-coding genes allowed identification of metabolic and interactive potential of the endophyte community. In addition, we performed mRNA expression analysis for selected bacterial genes.

Corresponding author: B. Reinhold-Hurek; Telephone: +49 421-21862860; Fax: +49 421-21862873; E-mail: breinhold@uni-bremen.de

*The e-Xtra logo stands for “electronic extra” and indicates that three supplementary tables and one supplementary figure are published online.

RESULTS AND DISCUSSION

Metagenome characteristics.

We first established successfully a protocol for endophyte cell extraction and enrichment from roots, using flowering *Oryza sativa* 'APO' plants grown on experimental fields at the International Rice Research Institute, Philippines (IRRI). For our experiment, the stepwise removal of non-endophytic bacteria from roots and the stability of the endophyte consortium was confirmed by *nifH*-DNA-based community analysis (restriction fragment length polymorphism [T-RFLP]) of N₂-fixing bacteria in samples from different stages of the procedure (Supplementary Fig. S1A). The *nifH* gene profiles obtained from total roots were altered by surface treatment: the relative abundance of several OTU ((operational taxonomic units)) decreased. Surface-treated roots harbored OTU profiles similar to those of endophyte cells extracted from the same root pool (plot 1), indicating that endophyte extraction was successful and endophytes had been enriched. Moreover, the phylogenetic distribution of *nifH* fragments retrieved from DNA of untreated roots was entirely different from that encountered in the endophyte metagenome (discussed below), suggesting that a major contamination with rhizoplane bacteria was unlikely.

Small-insert shotgun libraries were generated from DNA extracted from the endophyte community, and Sanger sequencing yielded approximately 47 Mb of bacterial DNA sequences. The average endophyte genome size based on typical single-copy genes (Raes et al. 2007) was calculated as 5.7 Mbp. Of more than 64,000 protein-encoding genes, 36% of the annotated genes could be categorized into a COG category and 37% encoded defined pfam domains.

Community composition.

Phylogenetic distribution of endophytes, based on rRNA genes in the metagenome library and, thus, not biased by poly-

merase chain reaction (PCR) amplification in addition to the cloning bias, corresponded very well with the taxonomic distribution of protein-coding genes. Members of phylum *Proteobacteria* dominated the endophyte community, *Gammaproteobacteria*, comprising mostly *Enterobacter*-related endophytes, and *Alphaproteobacteria*, including many rhizobia, being highly abundant (Fig. 1). Thus, they are likely to contribute the predominant genomes covered by our functional analysis below. Within the family *Enterobacteriaceae*, predominant sequences were related to an endophyte strain previously isolated from rice (*Enterobacter* sp. strain CBMB30) (Lee et al. 2006). In addition to *Proteobacteria*, *Firmicutes* and a few members of other phylogenetic lineages were detected that typically occur in soil but were not expected inside plants (phyla *Verrucomicrobiae*, *Planctomycetes*, and *Fusobacteria*). Thus, the taxonomic range of putative endophytes was extended by our study. Also, other 16S rDNA-based studies suggested a quite high diversity of rice root-associated bacteria (Sun et al. 2008). Other common soil prokarya, such as *Acidobacteria* or *Archaea*, were absent, despite methanogens being generally important members of the rice rhizosphere community (Lu and Conrad 2005; Shrestha et al. 2009). Rice roots that develop aerenchyma provide a source of oxygen in an otherwise anoxic flooded soil. Therefore, a redox gradient is formed, allowing diverse metabolic activity to take place inside root tissues (discussed below). Oxygen may deter methanogens from colonizing the rice root interior; however, the detection of various genes required for methane oxidation in the metagenome, including *pmo*, *mmo*, and *mx*a genes, provided evidence for the presence of endophytic methane-oxidizing bacteria. In agreement with that, a 16S rRNA gene-based cultivation-independent analysis confirmed the presence of methanotrophs (P. Haridoim, unpublished data). Thus, the predominant endophyte community was quite different from the consortia that are typically encountered in rice paddy soil (Lu and Conrad 2005; Shrestha et al. 2009).

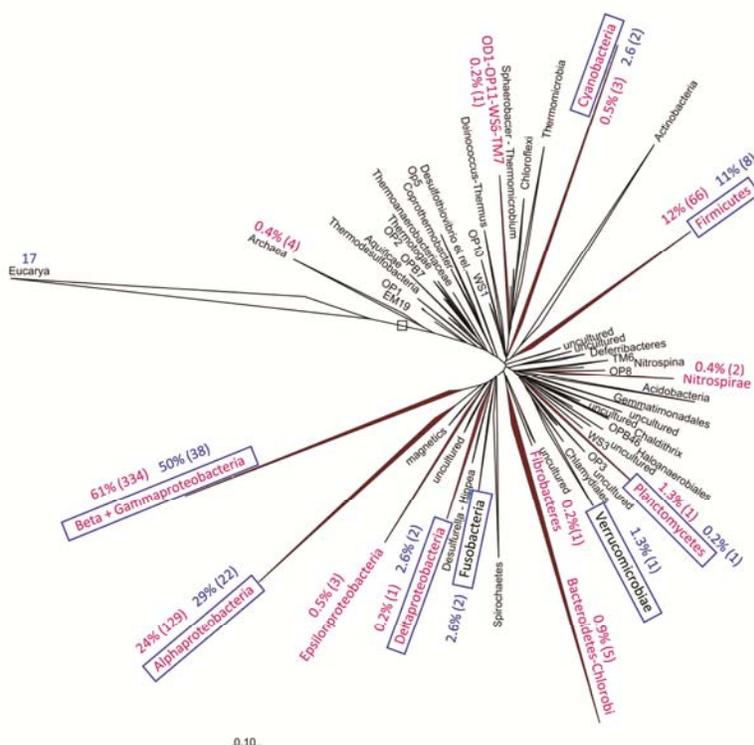


Fig. 1. Neighbor-joining phylogenetic tree of prokaryotic and eukaryotic 16S/18S rRNA sequences showing the phylogenetic distribution of microbial endophytes detected in the metagenome sequences. Sequences were aligned and the tree was constructed using the ARB phylogenetic software package. The bar indicates 10% evolutionary distance. Approximately one-third of the 16S rRNA genes found in the metagenomic library showed less than 97% identity with sequences deposited in public databases. In parenthesis: numbers of clones, percentage given for prokaryotic sequences.

Establishment of endophytes in the host.

Due to Sanger sequencing, sufficiently large sequence tags were obtained to permit us to predict more than 64,000 protein-coding genes. Annotation based on protein sequence identities and protein domain analyses allowed us to identify metabolic capacities, potentials related to plant-growth promotion, and features presumably important for interaction with the plant, and provided novel targets for future research (e.g., on interaction mechanisms).

Several sets of protein-coding genes were abundant in the endophyte metagenome, which allowed us to ascribe numerous capacities to this community. Some of these features might reflect adaptations of the bacterial community to the endorhizosphere as an exclusive microhabitat.

The outcome of plant–bacterium interactions is often determined by bacterial protein secretion systems (Downie 2010). In the endophyte metagenome, all known protein secretion systems for translocation across the cytoplasmic and outer membranes were present (Supplementary Table S1). An exception was the type III secretion system (T3SS) commonly used by symbiotic (Downie 2010) and pathogenic bacteria to inject effector proteins directly into the host cytoplasm and, thereby, to modulate the host response. For the T3SS, not all essential elements could be found in the metagenome, suggesting that it was not highly conserved among the rice endophytes. This correlates with the suggestion that endophytes, though plant associated, considerably differ from pathogens (Krause et al. 2006); genomes of cultivated endophytes do often not contain a T3SS (Reinhold-Hurek and Hurek 2011), although the *Herbaspirillum seropedicae* genome, for example, contains one (Pedrosa et al. 2011).

New targets for functional studies are proteins whose role in interactions has not yet been studied in cultivated endophytes. A high representation of genes encoding components of type VI secretion systems, suspected to deliver effector proteins into cells of eukaryotic hosts (Pukatzki et al. 2009), suggests their importance in beneficial plant–microbe interactions. In rhizobia, the type VI secretion system may contribute to host specificity (Bladergroen et al. 2003). With *hcp* genes at approximately 0.15/Mbp and a calculated genome size of 5.7 Mbp, almost every endophyte genome would contain an *hcp* gene, which encodes a key protein forming the secretion channel. Interestingly, this distribution is also reflected in genomes of cultivated endophytic bacteria: among 12 published genomes, seven harbor at least two putative type VI secretion gene clusters (Reinhold-Hurek and Hurek 2011; Weilharter et al. 2011). Also striking was the high relative abundance of genes encoding domains typical for autotransporters and the chaperone/usher pathway. An abundance of 0.2 or 0.3/Mbp was detected for domains typical for autotransporters (pfam03797) or the chaperone/usher pathway (pfam00577), respectively. The autotransporter secretion pathway exports proteins with diverse pathogenicity-related functions through the outer membrane, including proteases, toxins, adhesins, and invasions, while the chaperone/usher pathway is used to translocate fimbrial compounds of the P or type I pili to the cell surface.

Hydrolytic, plant-polymer-degrading enzymes were identified as another putatively important feature. High gene abundance (domains represented in cellulases, xylanases, cellobiohydrolases, and cellulose-binding proteins, 2.2/Mbp; domains typical in pectinases, 0.25/Mbp) suggested more than one set per endophyte genome. Pectinases were especially prominent in the metagenome and may contribute to endophyte entry into and spreading inside roots by degrading middle lamella. The role of hydrolytic enzymes has been shown (e.g., for the endoglucanase EglA in the endophyte *Azoarcus* sp. strain BH72 involved in rice root invasion) (Reinhold-Hurek et al. 2006).

Although flagellins are known to elicit an innate immune response in *Arabidopsis* (Zipfel et al. 2004), all components of the flagellar apparatus were encoded in the rice endophyte metagenome. Because they were abundant among the endophytes, with pfam00669 (Flagellin_N)-containing genes at 0.4/Mbp, which approximates 2/genome, motility or flagella-mediated adhesion may be required for establishment in the rice endosphere. A putative role in root-associated lifestyle was also supported by detection of transcription of flagellin-encoding endophytic *fliC* in rice roots (Fig. 2A).

The rice endophyte metagenome contained a high number and diversity of genes encoding enzymes potentially involved in the detoxification of reactive oxygen species (ROS), as well

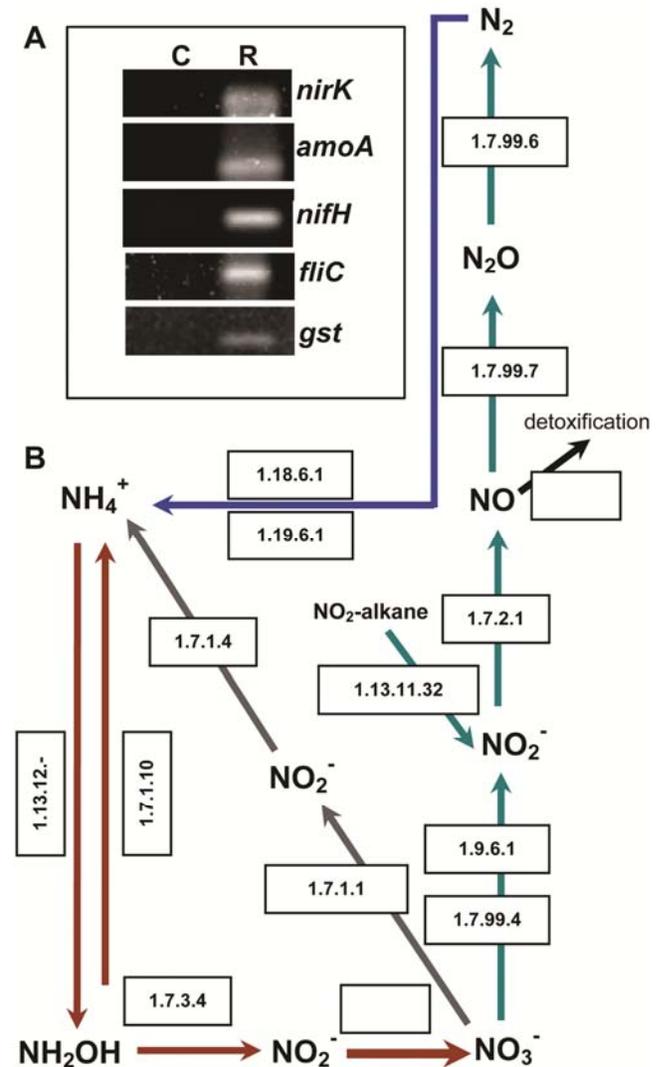


Fig. 2. A, Transcripts of selected gene fragments in RNA extracts of roots of *Oryza sativa* cv. Apo related to the N cycle (left) and other features (right), as detected by reverse-transcription polymerase chain reaction. C = control for DNA contamination (heat-inactivated reverse transcriptase) and R = reaction with reverse transcription. *nirK*, encoding nitrite reductase related to *Bradyrhizobium*, 416 bp; *amoA*, encoding gammaproteobacterial ammonium monooxygenase, 270 bp; *nifH*, encoding iron protein of nitrogenases, 362 bp; *fliC*, encoding gammaproteobacterial flagellin, 141 bp; *gst*, encoding *Bradyrhizobium*-related glutathione-S-transferase, 327 bp. **B**, Pathways of nitrogen metabolism represented in the endophytic metagenome of rice, adapted with modification from Kyoto Encyclopedia of Genes and Genomes. Arrows indicate nitrogen fixation in blue, denitrification in green, nitrification in brown, and nitrate assimilation in gray. Numbers inside boxes depict the EC number of the enzymatic reaction for which protein or proteins were detected in the endophytic metagenome.

as glutathione synthases and also glutathione-S-transferases (GST). Genes encoding glutathione synthetases (pfam02955 prokaryotic glutathione synthetase, ATP-grasp domain; pfam02951 prokaryotic glutathione synthetase, N-terminal domain; pfam04262 glutamate-cysteine ligase) occurred at 0.3/Mbp, while domains generally potentially involved in detoxification of ROS and products of oxidative stress were harbored at 3.3 genes/Mbp. In accordance with these findings, a glutathione synthetase gene was also found to be transcribed in association with rice roots (Fig. 2A). Because plants produce a range of ROS in response to abiotic stress or to colonizing microorganisms which elicit an oxidative burst, the abundance in the metagenome and in cultivated plant-associated bacteria (Fouts et al. 2008; Taghavi et al. 2010) suggests that endophytes require these enzymes to be able to successfully colonize plants.

Although, in an anoxic paddy soil, iron should not be limiting, the rice root itself obtains some oxygen through its aerenchyma and, therefore, the bioavailability of iron should be reduced. Indeed, our gene analysis indicates that the root interior appears to be a microenvironment extremely depleted of bioavailable iron: the endophyte metagenome showed a high number of genes encoding proteins that are potentially involved in siderophore biosynthesis (compounds of low molecular mass with high specific affinity toward ferric iron or other metal ions, typically secreted to the extracellular milieu, at 0.3 genes/Mbp). Accordingly, siderophore receptors, the so-called TonB-dependent receptors required for the uptake of ferric-siderophore complexes, or other elements for the uptake of iron as well as

iron storage proteins, were also found in abundance at 2.2 or 1.6 genes/Mbp, respectively. Also, in all genomes of cultivated and sequenced endophytes, these genes were abundantly present (Reinhold-Hurek and Hurek 2011), suggesting a high potential of the rice endophyte community to compete for iron with other members of the microbial community and, perhaps, with the plant host.

The endophyte microbiome comprised high diversity (more than 80% of the known pfam domains were detected) and abundance of transcriptional regulators. A subset was well represented, such as the LysR family (pfam00126 HTH_1 and pfam03466 LysR_substrate, at 4.2 genes/Mbp), AraC family (at 3.6 genes/Mbp), Crp family (at 0.5 genes/Mbp), and the IclR families (at 0.5 genes/Mbp) of transcriptional regulators, which are relevant for bacterial virulence, quorum sensing, and metabolic adaptation (Maddocks and Oyston 2008; Molina-Henares et al. 2006; Sun and Shi 2001). Analysis of the major protein domains responsible for production (GGDEF), binding (PilZ), and degradation (EAL and HD-GYP) of the second-messenger cyclic-di-GMP revealed that the amount of genes encoding phosphodiesterase A (PDEA) is high in the rice endophyte microbiome (at six genes/Mbp). PDEA reduce the level of c-di-GMP and, therefore, enable bacteria to stay motile and, in the case of pathogens, stay virulent, whereas high levels of c-di-GMP enhance biofilm formation and lower infectivity. Because the intensive investigation of c-di-GMP-mediated signaling started only recently, little is known about the regulation of this pathway. Cyclic-di-GMP-mediated signaling might be of great importance for endophytic communi-

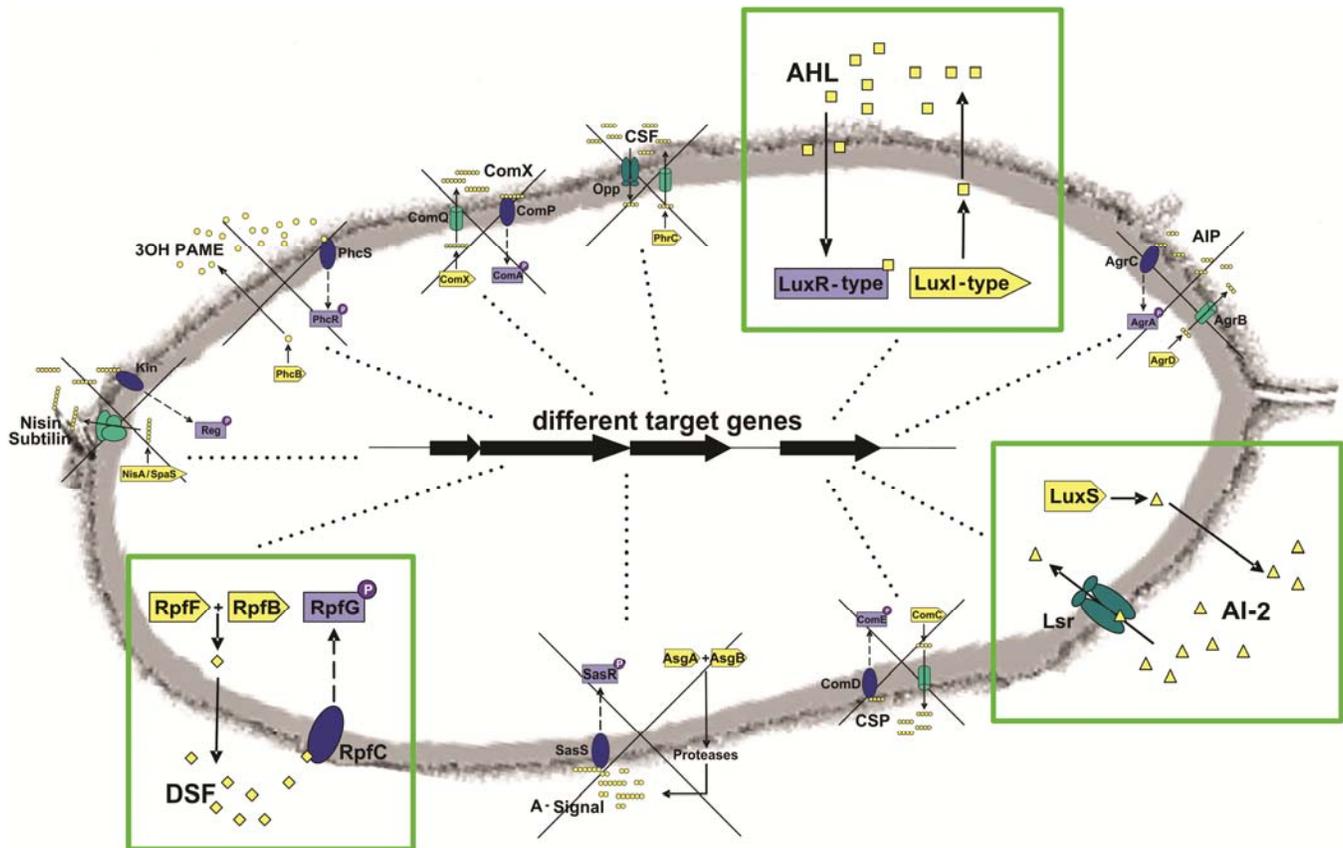


Fig. 3. Quorum sensing (QS) systems detected in metagenome by domain and sequence similarity analyses. Proteins for signal molecule synthesis and detection that were encoded in the endophyte metagenome are boxed in green. Known QS systems for which no evidence was found are crossed out. QS systems are labeled as AHL (acylhomoserine lactone system), AI-2 (autoinducer-2 system), DSF (diffusible signal factor system), A-signal system, CSF (sporulation factor system), ComX system, CSP (competence stimulating peptide system), AIP (autoinducing peptide system), 3-OH PAME system, and Nisin/Subtilisin system. Proteins that are involved in signal molecule synthesis (yellow arrows), detection (blue ellipses and blue boxes), and transport (green ellipses and barrels) are indicated with the particular protein name.

ties, as was already shown for prokaryotic biofilm lifestyle and pathogenesis (Cotter and Stibitz 2007).

Taken together, this suggests a very high degree of plasticity of responses to varying environmental stimuli such as plant signals, although the endosphere microenvironment was previ-

ously thought to be a rather stable, protected environment (Krause et al. 2006).

Microbial communication by autoinducer molecules might be highly important for the rice endophytic community. Three systems—autoinducer-2 system (AI-2), the diffusible

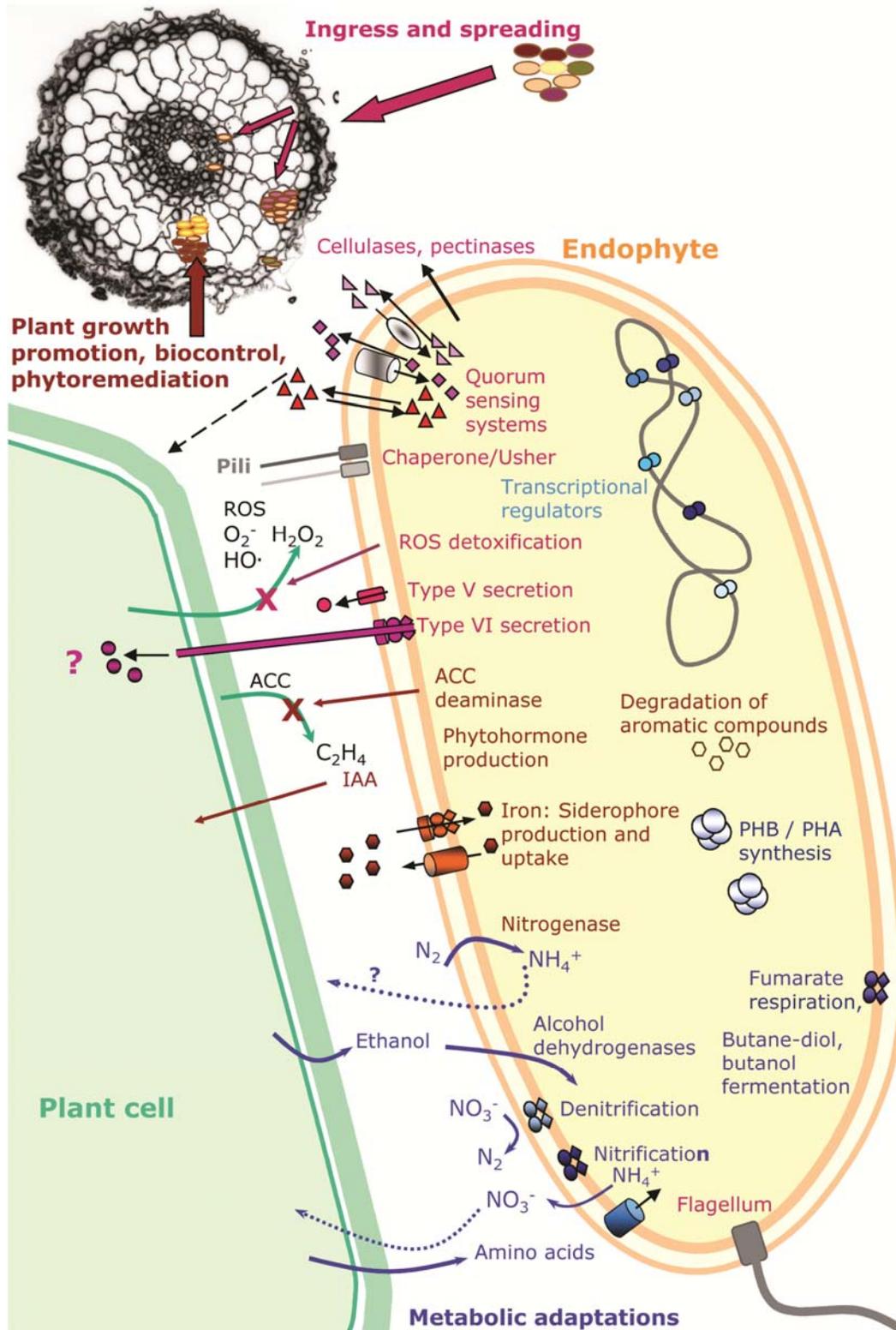


Fig. 4. Reconstruction of rice–endophyte interactions inferred from gene content analyses of the rice endophyte metagenome. The endophyte community is depicted as one cell for reasons of simplicity; however, it should not be inferred that all processes are active in one cell simultaneously. In red, processes putatively related to ingress into roots, spreading, and establishment within roots; in blue, processes putatively related to plant growth promotion, biocontrol, and phytoremediation; in brown, processes putatively related to plant growth promotion, biocontrol, and phytoremediation; in blue, putative metabolic adaptations. Question mark (?) indicates unknown effector proteins of protein secretion system and unknown transfer of fixed nitrogen to the plant.

signal factor system, and the acylhomoserine lactone system—were identified in the endophytic rice microbiome (Fig. 3; Supplementary Table S2). The AI-2 system (which has been commonly found in metagenomes) is known for both gram-negative and gram-positive bacteria, where it acts as a global signal molecule for interspecies communication (Williams et al. 2007). All LuxS sequences found in the rice endophytes were gammaproteobacterial. Interestingly, the DNA-binding transcriptional activator SdiA was detected in the rice endophyte metagenome. In *Escherichia coli* and *Salmonella enterica*, SdiA appears to encode a receptor that exclusively detects signal molecules of other species (Michael et al. 2001), indicating the importance of quorum-sensing molecules and their receptors in a diverse bacterial community. This diversity of systems is relatively high compared with other Sanger-based, similar-sized metagenomes of eukaryote-associated communities; in human gut (National Center for Biotechnology Information [NCBI] 16729) and termite gut (NCBI 19107), only the AI-2 system was detected whereas, in the *Olavius algarvensis* endosymbionts (NCBI 17779), none could be identified. Thus, a high diversity among endophytes probably reflects a need for concerted gene regulation for virulence and colonization, even across species borders.

Metabolic adaptations.

Our gene analysis suggested numerous metabolic adaptations of endophytes to their microenvironment. Partially, they might relate to a gradient of oxygen concentrations between anoxic soil, microaerobic aerenchymatic tissues, and O₂ depletion in bacterial microcolonies. Depending on the distance to the root base and on the root diameter, in aerenchymatic roots of typical flooded wetland plants in an anoxic soil, a concentration of 40 to 60% of atmospheric O₂ concentration can be expected, with a sharp decline at the outer wall layer and a decrease to 0 to 20% in the inner stele (Bodelier et al. 2006). Similar values were obtained for rice roots (Revsbech et al. 1999). Thus, different redox niches can be expected in rice roots and different metabolic capacities of the endophytes. In addition to fermentative abilities, which might be required to cope with anoxic niches, the abundance of alcohol dehydrogenases (at 2.3 genes/Mbp) indicated that ethanol accumulated by rice roots at low O₂ concentrations might be an important carbon source, as suspected for the grass endophyte *Azoarcus* sp. strain BH72 (Krause et al. 2006). Transport systems belonged to the most abundant COG category (11.3% of genes, which could be assigned to COG categories) (Supplementary Table S3), comprising mainly ATP-binding cassette family transporters for several amino acids or polyamines (11.1 genes/Mbp), indicating that they may be important N or C sources derived from the plant. The high abundance of genes involved in the degradation of aliphatic and aromatic compounds (4.5 genes/Mb) was striking, which suggests that the rice endophyte community is well equipped to degrade a wide range of aromatic compounds derived from the plants' secondary metabolism. The presence of genes encoding enzymes involved in the degradation of halogenated aromatic compounds (0.9 genes/Mb) may be required for degradation of lignin precursors, and might additionally play a role in the break down of pesticides and herbicides applied in agriculture.

The abundance of genes encoding proteins for the synthesis and degradation of carbon storage compounds such as polyhydroxyalkanoates or polyhydroxybutyrate (PHB) (1.6 genes/Mbp) indicates that the ability to produce storage compounds may enhance survival during starvation and tolerance to stress (Kadouri et al. 2003). PHB may also serve as a redox regulator for the removal of growth inhibitory metabolites (Aneja et al. 2004) such as plant secondary metabolites.

Plant growth promotion and biocontrol.

Our analysis also revealed features of endophytes that are likely to affect plant yield. We detected numerous traits in the endophyte metabiome that are likely to shape the fitness and performance of their host and, therefore, might be exploited for improvement of agricultural management practices. Plant growth may be promoted, and stress tolerance and nutrient uptake improved by providing phytohormones such as indole acetic acid (IAA) (Spaepen et al. 2007) or cytokinins (Timmusk et al. 1999) and by producing the enzyme ACC deaminase (Glick 2004). All corresponding genes were detected in the endophyte metagenome for IAA production even for the three pathways: the indole-3-acetamide, indole-3-pyruvate, and tryptamine pathways. Also, numerous antioxidant activities (discussed above) might foster stress tolerance of the host plant.

With respect to plant health, the presence of genes for siderophore production indicates strong biocontrol capacities allowing competition with pathogens for iron, as suspected for some cultured endophytes (Krause et al. 2006; van der Lelie et al. 2009). Also, bacterial quorum-sensing autoinducers may participate in the interaction of plants with pathogens and symbionts by induction of plant gene expression (Mathesius et al. 2003) (e.g., conferring systemic resistance to pathogens) (Schuhegger et al. 2006). The high diversity of autoinducer molecules represented in the endophyte microbiome (discussed above) and the possibility of novel autoinducers (Krause et al. 2006) advocate a strong, underexplored impact of endophytes.

Although degradative capacities for aromatic compounds are commonly found in soil bacteria, the endophyte metagenome surprisingly harbored a high number of genes for ring-cleaving enzymes (3.4 genes/Mbp). This emphasizes the potential of endophytes for bioremediation. Because iron and metal transport and storage mechanisms were well represented (discussed above), endophyte-plant systems might also be developed to affect heavy metal transport and phytoremediation.

Nitrogen cycle.

Nitrogen is the nutrient most limiting to plant productivity in terrestrial ecosystems and in rice production. Losses of N fertilizer depend on nitrogen use efficiency, which can be affected by microbial N cycling that is mainly thought to occur in soil (Li et al. 2008). Surprisingly, microbial nitrogen cycling was almost completely represented in the endophyte metagenome, although different steps require different oxygen concentrations: we found gene-based evidence for aerobic (nitrification), microaerobic (nitrogen fixation), and anaerobic (denitrification) parts of the N cycle (Fig. 2B). Among several domains of enzymes typical for denitrification, the NapE domain (pfam06796, periplasmic nitrate reductase) was detected. Furthermore, domains for formate and nitrate transporters and nitrous oxide reductase NosZ were also found. Additionally, components of the assimilatory nitrate reductase pathway to assimilate ammonium from nitrate were present. AmoA (pfam05145), which can be considered to specifically represent putative bacterial ammonia monooxygenases for nitrification, was represented in the endophyte metabiome. Transcript detection by reverse-transcription (RT)-PCR even suggested concurrent root-associated microbial activity in the three major branches of the N cycle, including nitrogen fixation (Fig. 2A). Nitrification and bacterial ammonia oxidation in the rhizosphere were thought to be major contributors to rice nutrition (Li et al. 2008); however, our data suggest that rice root endophytes also may impact growth and nitrogen use efficiency in rice. Their activities should particularly be taken into account when consequences of shifting flood regimes to aerobic rice cultivation are assessed.

The high apparent density of diazotrophic endophytes was remarkable, suggesting an N-poor microenvironment. Among an average of eight sequenced genomes, five *nifH* genes were detected, indicating that more than 50% of the endophytes were diazotrophic. Three were distantly related (89%) to *Bradyrhizobium* sp. strain BTAi1, one to *Xanthobacter autotrophicus* Py2 (88%), and one to *Dickeya dadantii* (84%). Despite moderate N-fertilizer application, root-associated *nifH* transcription was active (Fig. 2A), which underscores the role of biological nitrogen fixation (BNF) in this habitat. Phylogenetic analysis of *nifH* clone libraries prepared from DNA and mRNA of total roots showed that the entire community was rather diverse, also comprising *Geobacter* spp. typically occurring in soil. However, only a minority of root-associated diazotrophs was active in BNF as indicated by *nif* gene expression. Activity was mostly located in the alphaproteobacterial clade related to *Rhizobium* or *Bradyrhizobium* spp., clustering with a sequence found to be expressed in sugar cane roots (Burbano et al. 2011), which extends the typically root-nodule-associated rhizobial nitrogen fixation activity to cereals. Rhizobia have been reported as plant-growth-promoting endophytes of rice; however, considerable nitrogen fixation has not been reported (Saxena et al. 2000; Yanni et al. 1997). An exception is the recently discovered predominant expression of *nif*-genes of *Rhizobium rosettiformans* in roots of sugarcane from several continents (Burbano et al. 2011).

Conclusions.

Most functional information on bacterial endophytes is derived from the (genomic or mutational) analysis of model strains, whereas our metagenome analysis revealed insights beyond the genome information of individual bacterial strains. Without cultivation bias, we analyzed the genomic characteristics of the most abundant bacterial endophytes colonizing rice roots under field conditions. Our analysis revealed that bacterial endophyte communities seem to be highly adapted to proliferate and spread within plants (Fig. 4). Rice and other plants may be considered to be rich reservoirs of bacterial activities affecting plant growth and health. Bacterial gene expression analysis at different stages of rice growth will help us to understand the dynamics of the plant–endophyte system. A deeper understanding of endophyte functions and mechanisms for their establishment in the endosphere could be exploited to improve agricultural management practices with respect to biocontrol, bioremediation, and plant nutrition.

MATERIALS AND METHODS

Extraction of endophyte cells from rice roots.

Rice, *O. sativa* APO (IR55423-01), was grown on three plots at IRRI, Los Baños, Philippines, with N-P-K fertilizer at 90-30-30 kg/ha. Root samples from flowering plants were taken within 2 weeks (September to October 2006) and processed daily for endophyte extraction, and cells were pooled for metagenome analysis. Roots were thoroughly washed with tap water and small lateral roots were removed. The rhizoplane was chemically sterilized (approximately 5% sodium hypochlorite for 2 min), and rhizoplane bacteria were physically removed by established methods (Reinhold et al. 1986) using vigorous shaking with sterile glass beads in sterile water. Roots depleted of surface bacteria were then disrupted by scalpel in order to release the endophytes, and then shaken again with glass beads in sterile saline (0.9% NaCl) for 4 h at 30°C to detach microorganisms. Large plant and fungal cells and iron particles were removed from the supernatant by filtration (5- μ m pore size), and residual cells were pelleted at 4°C at 15,000 \times g and were stored in liquid nitrogen.

Extraction and purification of genomic DNA.

Endophyte cells were suspended in 200 mM Tris-HCl (pH 8.5), 25 mM EDTA, and 1% sodium dodecyl sulfate, and the lysis of the cells was performed by bead-beating (FastPrep instrument; BIO 101, La Jolla, CA, U.S.A.). The suspension was incubated with Proteinase K for 1 h at 60°C and extracted with phenol/chloroform/isoamylalcohol, and nucleic acids were precipitated with ethanol and dissolved in Tris-EDTA (TE) (pH 8.0). After purification by gel filtration and RNase One treatment for 2 h at 25°C, the DNA was again extracted with phenol/chloroform/isoamylalcohol, precipitated with ethanol, and resuspended in TE buffer (pH 8.0).

Analysis of *nifH* pools and bacterial transcripts.

For mRNA analysis and DNA-based clone library construction, roots were briefly washed and shock frozen in liquid nitrogen. Root total DNA for profiling (Engelhard et al. 2000) or RNA and DNA for clone libraries (Hurek et al. 2002; Knauth et al. 2005) were essentially extracted as previously described after homogenization of roots in liquid nitrogen and cell disruption by bead beating. Amplification of *nifH* fragments (Hurek et al. 2002) was slightly modified by using cDNA from RT (1 of 50 μ l) for amplification in a separate reaction (MolTaq; Molzyme, Bremen, Germany). Either T-RFLP analysis (Knauth et al. 2005) was performed after electroelution of fragments from agarose gels or clone libraries were constructed. *nifH* fragments were cloned into the pJET 1.2 vector using the CloneJet PCR cloning kit (Fermentas, St. Leon-Rot, Germany). Sequence analysis of 70 clones included rarefaction analysis (PAST program), construction of a phylogenetic tree based on *nifH* protein sequences from related cultured and uncultured diazotrophs, and in silico restriction analysis of the clones.

The following primers were designed based on metagenomic sequences and used for RT or amplification by PCR; underscored positions represent locked nucleic acid modifications to improve the sensitivity of RT (Burbano et al. 2010): *NirK*, nitrite reductase gene related to *nirK* of bradyrhizobia, NirK_RT_Brady_96 TTCACCTGGGTCATCAGAT; NirK_f_Brady_96 GACGAGAAGGGCAATTTC; NirK_r_Brady_96 ACTTGCC TTCGACCTTGAA. *fliC*, gene for gammaproteobacterial flagellin, FliCmet_RT GGMASCTGGTTGGCCTG, FliCmet_f TGGGTGCMTSCAGARCCG, FliCmet_r GCCGGCTATGCGMCGG. *gst*, gene for glutathione-S-transferase, GST_f CTGGAAGGCCAAGACCAAC; GST_r ACCAGATCTTGAC CTGAG. *amoA*, gene for gammaproteobacterial ammonium monooxygenase, AmoA_gamma_RT GCCGAMGCRGTCAC CATCAA; AmoA_f_gamma GTTCCKGCKGCRCTGTTGC; AmoA_r_gamma CCAGGTKCCGTCGTTCC. After RT for 30 min at 47°C, conditions for amplification were as follows. For *fliC*, an initial denaturation at 94°C for 1 min; then, 40 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 58°C, and extension for 30 s at 72°C; and a final extension for 5 min at 72°C were used. For *gst*, annealing was modified to 56°C. For *nirK*, initial denaturation was done at 95°C for 5 min; then, 40 cycles of denaturation for 1 min at 95°C, annealing for 2 min at 58°C, and extension for 2 min at 72°C; and a final extension for 10 min at 72°C were used. For *amoA*, annealing was modified to 51°C.

Shotgun library construction and Sanger sequencing.

A small insert library of the metagenome was constructed using approximately 1 μ g of metagenomic rice endophyte DNA. Briefly, the DNA was randomly sheared to 2- to 4-kbp fragments using a HydroShear (GeneMachines, San Carlos, CA, U.S.A.) and the sheared DNA purified using AMPure SRPI beads (Agencourt, Beverly, MA, U.S.A.) according to the

manufacturer's instructions. The metagenomic DNA was then end repaired using the End-it DNA End-Repair kit (Epicentre, Madison, WI, U.S.A.), followed by phenol/chloroform/isoamylalcohol extraction and ethanol precipitation. Approximately 20 ng of DNA was then blunt-end ligated into 100 ng of pUC19 vector overnight at 16°C using T4 DNA ligase (Roche Applied Science, Indianapolis, IN, U.S.A.) and 10% (vol/vol) polyethylene glycol (Sigma-Aldrich). After phenol-chloroform extraction and ethanol precipitation, 1 µl of ligation product was electroporated into ElectroMAX DH10B Cells (Invitrogen, Carlsbad, CA, U.S.A.), and clones were prepared and sequenced on an ABI PRISM 3730 capillary DNA sequencer (Applied Biosystems, Foster City, CA, U.S.A.) according to the JGI standard protocols. End sequencing yielded 175,872 reads totaling 108.29 Mbp of Sanger sequence.

For phylogenetic inference of *Bacteria* and *Archaea*, 16S rRNA gene sequences were analyzed using the ARB phylogenetic software package (Ludwig et al. 2004) against a representative subset of an internally curated 16S rRNA gene database.

Metagenome assembly.

The sequence reads were vector and quality trimmed with Lucy version 1.19p (Chou and Holmes 2001), resulting in 138,989 reads, which were screened for contamination with rice plant DNA using megablast against the *O. sativa* genome ($\geq 98\%$ i.d., min e-30). After removal of 59,331 rice sequences, the resulting 79,658 host plant-free metagenomic endophyte reads were assembled using the Paracel Genome Assembler 2.6.2. The assembly size comprised 14,739,389 bp in 10,583 contigs, 236 of which were major contigs defined as being at least 10 reads deep and 2 kbp long. The N50 contig length was 1,420 bp and the longest contig was 13.4 kbp.

To estimate the effective genome size, the method of Raes and associates (2007) was applied. In brief, it assesses the number of 35 selected marker genes that are generally single-copy genes; thus, their density is inversely linear to a representative microbial genome size prevalent in the metagenome analyzed.

Gene prediction and annotation.

The metagenomic assembly was loaded into IMG/M-ER for gene predictions and annotation as described in the "Standard Operating Procedure for the Annotations of Genomes and Metagenomes submitted to the Integrated Microbial Genomes Expert Review (IMG-ER) System" document. Additionally, proteins were routinely identified by domains for specific functions (pfam or TIGRFam); by COG category; or, in some cases (quorum-sensing systems, nitrogenase, ACC deaminase, IAA production, and so on), by sequence similarity searches (BLASTP, *e* value < 1e-5) compared with well-studied proteins.

The sequence data were deposited in GenBank under project accession number ADIE01000000.

ACKNOWLEDGMENTS

We thank C. M. Vera Cruz (IRRI) for providing her laboratory equipment during endophyte cell enrichment; J. Brons (University Groningen), B. Nikolic, and M. Fürnkranz (both AIT Seibersdorf) for help with endophyte extraction; and E. Prügger (University Bremen) for help with RNA extraction. This work was supported by a grant provided by the FWF (National Science Foundation, grant number P 21261-B03) to A. Sessitsch and partially by a grant awarded by the BMBF (Bundesministerium für Bildung und Forschung) in the framework of GABI-FUTURE (number 315068) to B. Reinhold-Hurek and T. Hurek. The sequencing for the project was provided through the United States Department of Energy (DOE) Community Sequencing Program. This work was performed under the auspices of the DOE's Office of Science, Biological, and Environmental Research Program and by the University of California, Lawrence Berkeley National Laboratory under contract number DE-AC02-05CH11231; Law-

rence Livermore National Laboratory under contract number DE-AC52-07NA27344; and Los Alamos National Laboratory under contract number DE-AC02-06NA25396.

LITERATURE CITED

- Aneja, P., Dai, M., Lacorre, D. A., Pillon, B., and Charles, T. C. 2004. Heterologous complementation of the exopolysaccharide synthesis and carbon utilization phenotypes of *Sinorhizobium meliloti* Rm1021 polyhydroxyalkanoate synthesis mutants. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett. 239:277-283.
- Barraquio, W. L., Revilla, L., and Ladha, J. K. 1997. Isolation of endophytic diazotrophic bacteria from wetland rice. Plant Soil 194:15-24.
- Bladergroen, M. R., Badelt, K., and Spaink, H. P. 2003. Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. Mol. Plant-Microbe Interact. 16:53-64.
- Bodelier, P. L. E., Frenzel, P., Drake, H. L., Hurek, T., Küsel, K., Lovell, C., Megonigal, P., Reinhold-Hurek, B., and Sorrell, B. 2006. Ecological aspects of microbes and microbial communities inhabiting the rhizosphere of wetland plants. Pages 205-238 in: Wetlands and Natural Resource Management. Ecological Studies Vol. 190. J. T. A. Verhoeven, B. Beltman, R. Bobbink, and D. F. Whigham, eds. Springer, Berlin.
- Burbano, C. S., Reinhold-Hurek, B., and Hurek, T. 2010. LNA-substituted degenerate primers improve detection of nitrogenase gene transcription in environmental samples. Environ. Microbiol. Rep. 2:251-257.
- Burbano, C. S., Liu, Y., Rösner, K. L., Reis, V. M., Caballero-Mellado, J., Reinhold-Hurek, B., and Hurek, T. 2011. Predominant *nifH* transcript phylotypes related to *Rhizobium rosetiformans* in field-grown sugarcane plants and in Norway spruce. Environ. Microbiol. Rep. 3:383-389.
- Chou, H. H., and Holmes, M. H. 2001. DNA sequence quality trimming and vector removal. Bioinformatics 17:1093-1104.
- Compant, S., Clément, C., and Sessitsch, A. 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. Soil Biol. Biochem. 42:669-678.
- Cotter, P. A., and Stibitz, S. 2007. c-di-GMP-mediated regulation of virulence and biofilm formation. Curr. Opin. Microbiol. 10:17-23.
- Dinsdale, E. A., Edwards, R. A., Hall, D., Angly, F., Breitbart, M., Brulc, J. M., Furlan, M., Desnues, C., Haynes, M., Li, L., McDaniel, L., Moran, M. A., Nelson, K. E., Nilsson, C., Olson, R., Paul, J., Brito, B. R., Ruan, Y., Swan, B. K., Stevens, R., Valentine, D. L., Thurber, R. V., Wegley, L., White, B. A., and Rohwer, F. 2008. Functional metagenomic profiling of nine biomes. Nature 452:629-632.
- Downie, J. A. 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Rev. 34:150-170.
- Engelhard, M., Hurek, T., and Reinhold-Hurek, B. 2000. Preferential occurrence of diazotrophic endophytes, *Azoarcus* spp., in wild rice species and land races of *Oryza sativa* in comparison with modern races. Environ. Microbiol. 2:131-141.
- FAO 2008. The State of Food and Agriculture 2008. Biofuels: Prospects, Risks and Opportunities. Food and Agriculture Organization of the United Nations, Rome.
- Fouts, D. E., Tyler, H. L., DeBoy, R. T., Daugherty, S., Ren, Q., Badger, J. H., Durkin, A. S., Huot, H., Shrivastava, S., Kothari, S., Dodson, R. J., Mohamoud, Y., Khouri, H., Roesch, L. F., Krogfelt, K. A., Struve, C., Triplett, E. W., and Methe, B. A. 2008. Complete genome sequence of the N₂-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. PLoS Biol. 4:e1000141. Published online.
- Glick, B. R. 2004. Bacterial ACC deaminase and the alleviation of plant stress. Adv. Appl. Microbiol. 56:291-312.
- Hurek, T., Reinhold-Hurek, B., Van Montagu, M., and Kellenberger, E. 1994. Root colonization and systemic spreading of *Azoarcus* sp. strain BH72 in grasses. J. Bacteriol. 176:1913-1923.
- Hurek, T., Handley, L., Reinhold-Hurek, B., and Piché, Y. 2002. *Azoarcus* grass endophytes contribute fixed nitrogen to the plant in an unculturable state. Mol. Plant-Microbe Interact. 15:233-242.
- Kadouri, D., Jurkevitch, E., and Okon, Y. 2003. Involvement of the reserve material poly-beta-hydroxybutyrate in *Azospirillum brasilense* stress endurance and root colonization. Appl. Environ. Microbiol. 69:3244-3250.
- Knauth, S., Hurek, T., Brar, D., and Reinhold-Hurek, B. 2005. Influence of different *Oryza* cultivars on repression of *nifH* gene pools in roots of rice. Environ. Microbiol. 7:1725-1733.
- Krause, A., Ramakumar, A., Bartels, D., Battistoni, F., Bekel, T., Boch, J., Böhm, M., Friedrich, F., Hurek, T., Krause, L., Linke, B., McHardy, A.,

- Sarkar, A., Schneiker, S., Syed, A. A., Thauer, R., Vorhölder, F., Weidner, S., Pühler, A., Reinhold-Hurek, B., Kaiser, A., and Goesmann, A. 2006. Genomic insights into the lifestyle of the mutualistic, N₂-fixing grass endophyte *Azoarcus* sp. strain BH72. *Nat. Biotechnol.* 24:1385-1391.
- Lee, H. S., Madhaiyan, M., Kim, C. W., J., C. S., Chung, K. Y., and Sa, T. M. 2006. Physiological enhancement of early growth of rice seedlings (*Oryza sativa* L.) by production of phytohormone of N₂-fixing methylophilic isolates. *Biol. Fertil. Soils* 42:402-408.
- Li, Y. L., Fan, X. R., and Shen, Q. R. 2008. The relationship between rhizosphere nitrification and nitrogen-use efficiency in rice plants. *Plant Cell Environ.* 31:73-85.
- Lu, Y., and Conrad, R. 2005. In situ stable isotope probing of methanogenic archaea in the rice rhizosphere. *Science* 309:1088-1090.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and Schleifer, K. H. 2004. ARB: A software environment for sequence data. *Nucleic Acids Res.* 32:1363-1371.
- Maddocks, S. E., and Oyston, P. C. 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* 154:3609-3623.
- Mano, H., and Morisaki, H. 2008. Endophytic bacteria in the rice plant. *Microbes Environment* 23:109-117.
- Mathesius, U., Mulders, S., Gao, M., Teplitski, M., Caetano-Anolles, G., Rolfe, B. G., and Bauer, W. D. 2003. Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. *Proc. Natl. Acad. Sci. U.S.A.* 100:1444-1449.
- Michael, B., Smith, J. N., Swift, S., Heffron, F., and Ahmer, B. M. 2001. SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. *J. Bacteriol.* 183:5733-5742.
- Molina-Henares, A. J., Krell, T., Eugenia Guazzaroni, M., Segura, A., and Ramos, J. L. 2006. Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Rev.* 30:157-186.
- Pedrosa, F. O., Monteiro, R. A., Wassem, R., Cruz, L. M., Ayub, R. A., Colauto, N. B., Fernandez, M. A., Fungaro, M. H., Grisard, E. C., Hungria, M., Madeira, H. M., Nodari, R. O., Osaku, C. A., Petzl-Erler, M. L., Terenzi, H., Vieira, L. G., Steffens, M. B., Weiss, V. A., Pereira, L. F., Almeida, M. I., Alves, L. R., Marin, A., Araujo, L. M., Balsanelli, E., Baura, V. A., Chubatsu, L. S., Faoro, H., Favetti, A., Friedermann, G., Glienke, C., Karp, S., Kava-Cordeiro, V., Raitz, R. T., Ramos, H. J., Ribeiro, E. M., Rigo, L. U., Rocha, S. N., Schwab, S., Silva, A. G., Souza, E. M., Tadra-Sfeir, M. Z., Torres, R. A., Dabul, A. N., Soares, M. A., Gasques, L. S., Gimenes, C. C., Valle, J. S., Ciferri, R. R., Correa, L. C., Murace, N. K., Pamphile, J. A., Patussi, E. V., Prioli, A. J., Prioli, S. M., Rocha, C. L., Arantes, O. M., Furlaneto, M. C., Godoy, L. P., Oliveira, C. E., Satori, D., Vilas-Boas, L. A., Watanabe, M. A., Dambros, B. P., Guerra, M. P., Mathioni, S. M., Santos, K. L., Steindel, M., Vernal, J., Barcellos, F. G., Campo, R. J., Chueire, L. M., Nicolas, M. F., Pereira-Ferrari, L., Silva, J. L., Gioppo, N. M., Margarido, V. P., Menck-Soares, M. A., Pinto, F. G., Simao Rde, C., Takahashi, E. K., and Yates, M. G. 2011. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. *PLoS Genet.* 7:e1002064. Published online.
- Pukatzki, S., McAuley, S. B., and Miyata, S. T. 2009. The type VI secretion system: Translocation of effectors and effector-domains. *Curr. Opin. Microbiol.* 12:11-17.
- Raes, J., Korbelt, J. O., Lercher, M. J., von Mering, C., and Bork, P. 2007. Prediction of effective genome size in metagenomic samples. *Genome Biol.* 8:R10.
- Reinhold, B., Hurek, T., Niemann, E.-G., and Fendrik, I. 1986. Close association of *Azospirillum* and diazotrophic rods with different root zones of Kallar grass. *Appl. Environ. Microbiol.* 52:520-526.
- Reinhold-Hurek, B., and Hurek, T. 2011. Living inside plants: Bacterial endophytes. *Curr. Opin. Plant Biol.* 14:435-43.
- Reinhold-Hurek, B., Maes, T., Gemmer, S., Van Montagu, M., and Hurek, T. 2006. An endoglucanase is involved in infection of rice roots by the not cellulose-metabolizing endophyte *Azoarcus* sp. BH72. *Mol. Plant-Microbe Interact.* 19:181-188.
- Revsbech, N. P., Pedersen, O., Reichardt, W., and Briones, A. 1999. Microsensor analysis of oxygen and pH in the rice rhizosphere under field and laboratory conditions. *Biol. Fertil. Soils* 29:379-385.
- Saxena, S., Ladha, J. K., Gyaneshwar, P., Reinhold-Hurek, B., Hernandez, R. J., and Biswas, J. C. 2000. Evaluation of *lacZ* and *gusA* markers to study rhizobial colonization in rice roots. *Ind. J. Microbiol.* 40:15-20.
- Schuhegger, R., Ihring, A., Gantner, S., Bahnweg, G., Knappe, C., Vogt, G., Hutzler, P., Schmid, M., Van Breusegem, F., Eberl, L., Hartmann, A., and Langebartels, C. 2006. Induction of systemic resistance in tomato by N-acyl-L-homoserine lactone-producing rhizosphere bacteria. *Plant Cell Environ.* 29:909-918.
- Shrestha, P. M., Kube, M., Reinhardt, R., and Liesack, W. 2009. Transcriptional activity of paddy soil bacterial communities. *Environ. Microbiol.* 11:960-970.
- Spaepen, S., Vanderleyden, J., and Remans, R. 2007. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Rev.* 31:425-448.
- Sun, H., and Shi, W. 2001. Analyses of *mnp* genes during *Myxococcus xanthus* development. *J. Bacteriol.* 183:6733-6739.
- Sun, L., Qiu, F., Zhang, X., Dai, X., Dong, X., and Song, W. 2008. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microb. Ecol.* 55:415-424.
- Taghavi, S., van der Lelie, D., Hoffman, A., Zhang, Y. B., Walla, M. D., Vangronsveld, J., Newman, L., and Monchy, S. 2010. Genome sequence of the plant growth promoting endophytic bacterium *Enterobacter* sp. 638. *PLoS Genet.* 6:e1000943. Published online.
- Timmus, S., Nicander, B., Granhall, U., and Tillberg, E. 1999. Cytokinin production by *Paenibacillus polymyxa*. *Soil Biol. Biochem.* 31:1847-1852.
- van der Lelie, D., Taghavi, S., Monchy, S., Schwender, J., Miller, L., Ferrieri, R., Rogers, A., Wu, X., Zhu, W., Weyens, N., Vangronsveld, J., and Newman, L. 2009. Poplar and its bacterial endophytes: Coexistence and harmony. *Crit. Rev. Plant Sci.* 28:346-358.
- Weilharter, A., Mitter, B., Shin, M. V., Chain, P. S., Nowak, J., and Sessitsch, A. 2011. Complete genome sequence of the plant growth-promoting endophyte *Burkholderia phytofirmans* strain PsJN. *J. Bacteriol.* 193:3383-3384.
- Weyens, N., van der Lelie, D., Taghavi, S., Newman, L., and Vangronsveld, J. 2009. Exploiting plant-microbe partnerships to improve biomass production and remediation. *Trends Biotechnol.* 27:591-598.
- Williams, P., Winzer, K., Chan, W. C., and Cámara, M. 2007. Look who's talking: Communication and quorum sensing in the bacterial world. *Phil. Trans. R. Soc. B* 362:1119-1134.
- Yan, Y., Yang, J., Dou, Y., Chen, M., Ping, S., Peng, J., Lu, W., Zhang, W., Yao, Z., Li, H., Liu, W., He, S., Geng, L., Zhang, X., Yang, F., Yu, H., Zhan, Y., Li, D., Lin, Z., Wang, Y., Elmerich, C., Lin, M., and Jin, Q. 2008. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. *Proc. Natl. Acad. Sci. U.S.A.* 105:7564-7569.
- Yanni, Y. G., Rizk, R. Y., Corich, V., Squartini, A., Ninke, K., Philip-Hollingsworth, S., Orgambide, G., De Bruijn, F., Stoltzfus, J., Buckley, D., Schmidt, T. M., Mateos, P. F., Ladha, J. K., and Dazzo, F. B. 1997. Natural endophytic association between *Rhizobium leguminosarum* bv. *trifolii* and rice roots and assessment of its potential to promote rice growth. *Plant Soil* 194:99-114.
- Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E. J., Jones, J. D., Felix, G., and Bolter, T. 2004. Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* 428:764-767.

AUTHOR-RECOMMENDED INTERNET RESOURCES

- National Center for Biotechnology Information's GenBank database: www.ncbi.nlm.nih.gov/Genbank
 PAlaontological Statistics (PAST) website: folk.uio.no/ohammer/past
 United States Department of Energy Joint Genome Institute (JGI) website: www.jgi.doe.gov
 JGI's Community Sequencing Program: www.jgi.doe.gov/CSP/index.html