Comparative study of endophytic and endophytic diazotrophic bacterial communities across rice landraces grown in the highlands of northern Thailand

Chakrapong Rangjaroen[1], Benjavan Rerkasem[2], Neung Teaumroong[3], Rungroch Sungthong[4] and Saisamorn Lumyong[1*]

[1]*Microbiology Division, Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*

[2]*Lanna Rice Research and Cultural Centre, Faculty of Agriculture, Chiang Mai University, Chiang Mai 50200, Thailand*

[3]*School of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhonratchasima 30000, Thailand*

[4]*Departamento de Agroquímica y Conservación de Suelos, Instituto de Recursos Naturales y Agrobiología de Sevilla, Consejo Superior de Investigaciones Científicas, Seville 41012, Spain*

***Corresponding author**:

Saisamorn Lumyong

Microbiology Division, Department of Biology, Faculty of Science,

Chiang Mai University, Chiang Mai 50200, Thailand

Tel.: +66 53 941946, Fax: +66 53 941949

E-mail: scboi009@gmail.com

Running headline: Bacterial communities living in rice landraces

Abstract

Communities of bacterial endophytes within the rice landraces cultivated in the highlands of northern Thailand were studied using fingerprinting data of 16S rRNA and *nifH* genes profiling by PCR-DGGE. The bacterial communities' richness, diversity index, evenness, and stability varied depending on the plant tissues, stages of growth and rice cultivars examined. The diversity indices of the endophytic diazotrophic bacteria within the landrace rice Bue Wah Bo were the lowest. The endophytic non-diazotrophic bacteria revealed greater diversity by cluster analysis with 7 clusters compared to the endophytic diazotrophic bacteria (3 clusters). Principal component analysis suggested that the endophytic non-diazotrophic bacteria showed higher stability of the community structures across the rice landraces than those of the endophytic diazotrophic bacteria. Uncultured bacteria were found dominantly in both bacterial communities, while higher generic varieties were observed in the endophytic diazotrophic bacterial community. These differences in the bacterial communities might be influenced either by genetic variation of the rice landraces or the rice cultivation system, where the nitrogen input could strongly affect the endophytic diazotrophic bacterial community.

Keywords: Bacterial community; Bacterial endophytes; 16S rRNA gene; *nifH* gene; Cluster analysis; Principal component analysis

Introduction

Rice landraces or local rice cultivars are often named locally and distinguished individually on the basis of their morphologies, geographies or ecologies (Bandara et al. 2006; Bonman et al. 1992; Pusadee et al. 2009). Among a total 120 rice varieties, 90% are the local cultivars grown annually in deep water rice areas (Oupkaew et al. 2011; Sommut 2003). In supplementary farming conditions, the rice landraces usually produce lower yields compared to the modern high-yield varieties but their products are still sufficient for household consumption. Although the majority of locally cultivated rice landraces are cultivated by traditional or natural rice farming without any chemical inputs and pest controls, they have been able to produce reasonable yields year by year (Oupkaew et al. 2011; Parzies et al. 2004; Purnomo et al. 2005). This traditional cultivation of the rice landraces requires both specific farming practices and unique local conditions (Chakhonkaen et al. 2012; Oupkaew et al. 2011; Saito et al. 2006). Investigations have revealed that genetic diversity of the rice landraces is one of the reasons which enable their evolutionary adaptation and selection to those traditional practices and farming systems (McCouch 2004). Pusadee et al. (2009) reported that the genetic diversity of a landrace rice cultivar Bue Chomee grown in the villages of the highland Karen people in northern Thailand, is explained with the isolation by distance model. The significant differences of genetic structure either in a single field or across fields of different villages of this rice cultivar support its genetic variation. Moreover, the authors also found that the gene flow of this rice cultivar is due to the farmers' seed sharing network. The rice landraces are not only a valuable resource for farmers who, for economic or ecological reasons, do not have access to modern, improved varieties, but also for future improvement by plant breeding. However, in addition to the sustainability of traditional rice production systems, these genetically diverse rice landraces may also have benefited from microbes associated with them.

In traditional farming systems, where fertilizers and pesticides are rarely used, the plant-associated microorganisms can play important roles in nutrients cycling and preventing attacks by phytopathogens. In general, the plant-microbe interactions may take place at the rhizosphere and/or within the plant tissues. The microorganisms that have lived for one period of their life cycle within the plant tissues without causing apparent symptoms in the host plant, may be considered as microbial endophytes (Azevedo et al. 2000; Bandara et al. 2006; Kaga et al. 2009; Mano and Morisaki 2008; Loaces et al. 2011). The different *in planta* roles of the microbial endophytes so far described such as nitrogen fixation (Wartiainen et al. 2008; Prakamhang et al. 2009), production of phytohormones (Arun et al. 2012; Feng et al. 2006), production of ACC-deaminase to reduce the level of ethylene (Tittabutr et al. 2008; Jha et al. 2012), siderophore production (Loaces et al. 2011), etc.

There are many reports demonstrated that several endophytic bacteria are living in plant tissues of rice such as *Bukholderia*, *Herbaspirillum*, *Rhizobium*, *Methylobacterium,* and *Bacillus* in root (Mano et al. 2007; Verma et al. 2004), *Azospirillum* and *Herbaspirillum* in stem (Koomnok et al. 2007; Mano and Morisaki 2008), and *Pantoea*, *Bacillus,* and *Sphingomonas* in seed (Kaga et al. 2009; Mano and Morisaki 2008; Mano et al. 2007). Some of these are diazotrophic bacteria, which are able to fix atmospheric nitrogen and to transform it to ammonium (NH4+) that are further to amino acid that required for growth and reproduction of rice. Recently, an evidence suggested that rice's genotypes and stages of growth together with nitrogen level and soil processing influence the plant-microbe association of *Azospirillum* sp. strain B510 (Sasaki et al. 2010). A similar study of nitrogen-fixing bacterium *A. caulinodans* was conducted by Van Nieuwenhove et al. (2000). The authors found that this bacterium requires an inoculum carrier plant *Sesbania rostrata* for its long term survival and greater persistence in the paddy rice rhizosphere. Moreover, the difference of rice varieties has no influence on both survival and nitrogenase activity of this bacterium. However, little is known about the complex community of endophytic and endophytic diazotrophic bacteria that live in association with high-genetic-variation rice landraces cultivated in highland areas.

Molecular techniques based on finger printing data have been applied worldwide for analysis of diverse microbial communities. Among these tools, polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE) is an acceptable and consistent approach to evaluate the community structures of various microorganisms living in association with rice plants (García de Salamone et al. 2010; Hardoim et al. 2011; Jia et al. 2007). In this study, we aim to compare the community structures of endophytic non-diazotrophic and endophytic diazotrophic bacteria living in different plant tissues of 5 highland rice landraces at different stages of growth. The 16S rRNA gene was selected for the evaluation of the total bacterial endophytes, while the *nifH* gene was used for the endophytic diazotrophic bacteria. The finger printing data of both genes were profiled by PCR-DGGE. The influencing factors that affect the bacterial communities within the rice landraces are addressed and discussed in this article.

Materials and Methods

Collection and preparation of rice samples

Highland rice landrace; Bue Polo (BP), Bue Pra Taw (BPT), Bue Pra Do (BPD), and Bue Saw Mi (BSM) were collected from fields farmed without chemical input, except for a cultivar Bue Wah Bo (BWB) that

was collected from nitrogen fertilizer-supplied fields (**Table 1**). All rice cultivars were cultivated in a farm at Chom Thong District, Chiang Mai, Thailand. This area was established since 2005, aiming to grow several crops including the rice landraces for household consumption of nearby villagers. The location of each collected site was recorded by GPS (GPS-12XL, Garmin, Kansas, USA). The rice samples were collected in 2010, when the mean of rainfall of Chiang Mai was 1,156 mm with 112 rainy days (data reported by the Thai Meteorology Department in 2010). Four growth stages including pre-planting in April (only mature rice seeds), nursery seedlings in June, vegetative plants in August, and reproductive cultivars in September were sampled. Mature rice seeds were obtained directly from the farmers, while 10 rice plants were randomly taken from the field of each rice cultivar and kept (not exceeding 12 h) in an ice-box and transferred to the laboratory, where they were gently washed with tap water and rinsed several times with sterilized water. The rice plant tissues including leaf, stem, and root were excised into 2-3 cm length. The rice grains of the reproductive stage were removed directly from the ear of rice. Each sample derived from 10 plants at the same stage of growth was pooled together and surface sterilized following the method described by Koomnok et al. (2007).

Extraction of total DNA from plant materials

Total DNA was extracted from the surface sterilized tissue of the rice using a modified potassium acetate method (Dellaporta et al. 1983). Briefly, the rice tissue were ground in liquid nitrogen and transferred into sterilized 1.5-mL microcentrifuge tube. Pre-heated extraction buffer (720 µL) containing 100 mM Tris-HCl, 50 mM EDTA, 500 mM NaCl and 1.25% (w/v) SDS was added and mixed before incubation at 65 °C for 20 min. Proteins were precipitated by adding 225 µL of 5 M potassium acetate and cooling on ice for 20 min before the decanting supernatant into a new tube. Remaining protein residue was removed by adding an equal volume of buffer-equilibrated phenol-chloroform-isoamyl alcohol (25:24:1 (v/v/v)). The DNA was precipitated by adding an equal volume of cold isopropanol, and was washed twice with 70% (v/v) ethanol. The DNA extraction protocol was repeated 3 times from the same pooled rice tissue. All extracted DNA from the same rice tissue was pooled together and resuspended with 30 µL of TE buffer prior to storage at -20 °C for further DNA manipulations.

PCR-DGGE analyses of 16S rRNA and nifH genes

Amplification of a variable region 3 of the 16S rRNA gene was performed using a pair of the universal primers [forward primer 341f (5'-CCTACGGGAGGCAGCAG-3') with a GC-clamp, and reverse primer 534r

(5'-ATTACCGCGGCTGCTGG-3')] (Muyzer et al. 1993). The *nifH* gene was amplified using a nested PCR (Coelho et al. 2009). At the first step, the amplification was performed with the forward primer FGPH19 (5'-TACGGCAARGGTGGNATHG-3') (Simonet et al. 1991) and the reverse primer PolR (5'-ATSGCCATCATYTCRCCGGA-3') (Poly et al. 2001a), generating the PCR products of 429 bp. The second step was performed with the forward primer PolF (5'-TGCGAYCCSAARGCBGACTC-3') (Poly et al. 2001a) with a GC-clamp and the reverse primer AQER (5'-GACGATGTAGATYTCCTG-3') (Poly et al. 2001b). The PCR products from the same sample were pooled and purified by GeneJET^{TM} PCR Purification Kit (Fermentas, Lithuania) following the manufacturer's instructions and stored at -20 °C before performing DGGE.

The DGGE was carried out following the manufacturer's protocol of Bio-Rad DCode® Universal Mutation Detection System (Bio-Rad, Hercules, CA, USA). Briefly, 25 µL of an individual PCR product was loaded on 8% (w/v) polyacrylamide gel with a linear gradient of 30-70% denaturant for the 16S rRNA gene and 30-60% denaturant for the *nifH* gene (100% denaturant corresponds to 40% (v/v) of formamide plus 7 M urea). Electrophoresis was performed at 130 V for 6 h at a constant temperature of 60 °C. Gels were then stained with 0.5 µg mL^{-1} ethidium bromide for 30 min. The DGGE bands profile was visualized and photographed by GeneFlash Syngene Bio Imaging (Syngene, Cambridge, UK).

Numerical analyses of PCR-DGGE bands profiling data

The DGGE bands profiling data were analyzed in different aspects for investigation of the living bacterial communities associated with the rice landraces. Mathematic parameters that describe the bacterial community in a defined habitat including richness, Shannon-Wiener diversity index, evenness, and stability were computed on the basis of number and relative intensity of the DGGE bands. The mathematical formulae of the indices were explained as following. First, a band was determined as a phylotype of any endophytic bacteria, while the richness of a phylotype was calculated by equation (**i**) using the total difference of DGGE bands present in an individual (Fromin et al. 2002).

$$d = \frac{S-1}{\log N} \qquad \textbf{(i)}$$

Where, d is a richness of the bacterial community in a respective habitat; S is the number of different bands; N is the total number of individuals.

The Shannon-Wiener diversity index (H') was used to determine the diversity impacts of the bacterial community in any defined habitat. The H' index was calculated by equation (**ii**), where P_i is the relative intensity of band i ($P_i = N_i / N$; N_i, an intensity of band i; N, a sum of all band intensities) (Shannon and Weaver 1963).

$$H' = -\sum P_i \ln P_i \qquad \textbf{(ii)}$$

The evenness (E) of the bacterial community is a proportion of the H' index divided by the natural logarithm of the richness (d) of that respective community (Asakawa and Kimura 2008), which was calculated by equation (**iii**).

$$E = \frac{H'}{\ln d} \qquad \textbf{(iii)}$$

The stability (S) of the bacterial community was calculated with the formula (**iv**), where $P_i^{\max}$ is the maximum intensity of band i and n is the number of samples in the profile (Asakawa and Kimura 2008).

$$S = \frac{\sum \frac{P_i}{P_i^{\max}}}{n} \qquad \textbf{(iv)}$$

A percent proportion of the endophytic diazotrophic bacteria (N) per total number of the endophytic bacteria (T) in any respective habitat was calculated using a sum of all DGGE bands' intensities obtained from the endophytic diazotrophic bacteria (PN) divided by those obtained from the endophytic bacteria (PT), where n is the total number of all bands present in the profile (**v**).

$$N/T(\%) = \frac{\sum_{i=1}^{n} PN_i}{\sum_{i=1}^{n} PT_i} \times 100 \qquad \textbf{(v)}$$

Cluster and principal component analyses

The numerical scores of the DGGE bands profiling the research data were also used to compare structural similarity between the bacterial communities that live with the rice landraces using cluster analysis. The percent similarities among the band profiles were calculated using Pearson's correlation coefficient generated by a similarity matrix. An agglomerative hierarchical clustering was performed using the unweight pair group method with an arithmetic average (UPGMA) and displayed as a dendrogram. The multiple parameters of each band profile comprised of rice plant tissue types, stages of growth and rice cultivars, were presented as a 2-dimentional plot by principal component analysis (PCA), using the band intensity data. The 2-dimentional plot represented the community pattern of either endophytic or endophytic diazotrophic bacteria within the rice landraces. The new generating axes of the 2-dimentional plot were selected by comparing the correlation matrix based on eigenvalues and eigenvectors calculated. The percent variations of the new axes were indicated in the plot diagram. Both cluster and principal component analyses were carried out using an Excel-XLSTAT 2011 program.

DNA and protein sequencings and construction of phylogenetic trees

Representative DGGE bands were selected, marked and excised from the DGGE gels and transferred to the sterilized 1.5-mL microcentrifuge tube. Thereafter, 20 μL of sterilized distilled water was added into each tube. The tubes were kept at 4 ºC for overnight to allow the DNA to passively diffuse out of the gel strips (Prakamhang et al. 2009). Eluted DNA was used as a DNA template for the PCR amplification following the same conditions as previously mentioned, but the GC-clamp was removed from those respective primers. The amplicons were ligated into the pGEM®-T Easy Vector System (Promega, USA). The resultant ligated products were transformed into *Escherichia coli* JM109 competent cells. The purified plasmid DNAs of selected transformant colonies were sequenced by Macrogen Inc. (Seoul, South Korea). The *nifH* gene sequences were reverse translated to the protein sequences before alignment. The alignments of both 16s rRNA gene sequences and *nifH*-derived protein sequences were performed by BLASTN program (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The phylogenetic trees were constructed by the neighbor-joining method using Mega 4 (Tamura et al. 2007), and confidence levels were estimated for 1,000 replicates.

234 ***Comparative study by statistical analysis***

235 Comparisons of multiple means with standard deviations (SDs) obtained from individual mathematic
236 parameters of both endophytic and endophytic diazotrophic bacteria living in association with different rice
237 landraces were performed by non-parametric analysis using the SPSS 18.0 computer program (SPSS, Chicago
238 IL, USA), with one-way analysis of variance (ANOVA) and Tukey's *post hoc* tests at appropriate significant
239 level of $P = 0.05$ or lower.

240

241 **Results**

242 ***Numerical analysis of bacterial communities living in rice landraces***

243 The bacterial communities within the rice landraces were evaluated based on the PCR-DGGE profiling
244 data. Existence and intensity of each DGGE band were recorded and used for numerical analysis of all
245 community parameters (richness, Shannon-Wiener diversity index, evenness, and stability). These parameters
246 varied depending on the rice cultivar, rice plant tissue, stages of growth and husbandry conditions of the rice
247 landraces **(Table 2)**. The highest in both the richness and diversity index of both endophytic and endophytic
248 diazotrophic bacterial communities was found in the root tissues of the BPT rice landrace. Higher richness and
249 diversity indices of the endophytic bacteria were found mostly in the root tissues rather than the other rice plant
250 tissues; while these two parameters of the endophytic diazotrophic bacteria were not constant across the rice
251 plant tissues of any rice landrace. Upon examination of the non-parametric statistical comparison was carried
252 out, the means of each mathematic parameter of the endophytic bacteria were not significantly different across
253 the rice landraces (P=0.0001). However, only the community's stability of the endophytic bacteria derived from
254 any stages of growth of the BPD rice landrace was significantly different (P=0.05) with the highest mean
255 compared to those other rice cultivars. In case of the endophytic diazotrophic bacteria, all the community
256 parameters obtained from the BWB rice landrace were significantly different (P=0.001) to those other rice
257 cultivars. The means of most parameters derived from this rice landrace were lower than those of the others
258 except for the community's stability was uniformly high.

259 Ratios of the endophytic diazotrophic bacteria per total number of the endophytic bacteria in any
260 conditions tested were quantified on the basis of the DGGE band's intensity (**Table 2**). The means of the percent
261 proportion obtained during the cultivation periods across the rice landraces was relatively high (40.99±16.20 -
262 91.68±7.18%) except for the BWB rice landrace which showed the lowest percentage proportion of 14.33±4.97
263 to 29.89±23.38%. These mean percentages found across the different stages of growth in the individual rice

cultivar were not considered significantly different (P=0.0001). However, these means percentages of the BWB rice landrace were significantly lower (P=0.01) when compared to the other rice landraces in any stages of growth.

Community structures of bacterial endophytes in rice landraces

The DGGE bands profiling data of both bacterial communities which live in association with the rice landraces examined in this study were clustered using UPGMA (**Fig. 1**). There were 3 dominant DGGE bands of the endophytic bacteria found in all rice plant tissues obtained from every stages of growth (**Fig. 1A**). The % similarities of the endophytic bacterial communities were in the range of 34.88% to 95.55%, which were categorized into 7 clusters. The community structures derived from root tissues of the BPT rice landrace (cluster V) and BSM (cluster VII) grown in reproductive stage revealed unique molecular fingerprints which separated them from the others. The endophytic bacteria within the BP rice landrace showed similar community structures, belonging to cluster II (**Fig. 1A)**. Most of the band profiles across the different rice plant tissues or stages of growth within the same rice cultivar exhibited higher % similarities (similarity indices ≥ 80%). Contrary to those observations, across the same rice plant tissues or stages of growth but looking at different rice cultivars, % similarities were much lower. For example, in cluster III (**Fig. 1A)** where the community structures of the endophytic bacteria found in root and stem of the BPT rice landrace grown in nursery stage, exhibited a relatively high similarity of 82.60%. Similar observations were also found in cluster II (**Fig. 1A)** where the community structure found in stem of BPD rice landrace at nursery stage showed high similarity to those found in leaf at reproductive stage (95.55%) and grain (88.24%) of the same rice landrace.

In case of the endophytic diazotrophic bacteria, no dominant DGGE band was observed in all 3 clusters analyzed (**Fig. 1B**). Their % similarities were in the range of -1.77% to 100%, while the high similarities (≥ 80%) within the same rice cultivars could be found frequently in all clusters. However, it was clearly demonstrated that most of the community structures derived from the BWB rice landrace were separated from the others and clustered in clusters II and III (**Fig. 1B)**. These 2 clusters were obviously different to cluster I, supported by low % similarity of 18.95%.

The community structures of both bacteria were also viewed 2-dimensionally by PCA (**Fig. 2**). The overall variances of data plots based on the DGGE band's intensity of these bacterial community structures influenced by the environmental factors were depicted in percent variances of axes F1 and F2. The percent variances were ranged varied depending on the rice cultivars, where F1 (49.24-59.39%) and F2 (8.40-14.33%)

were found in the endophytic bacterial community, whereas F1 (32.29-65.26%) and F2 (14.76-23.44%) were found in the endophytic diazotrophic bacterial community. The community structures of the endophytic bacteria were similar across the rice landraces as the data plots were located more tightly in the graphs compared to those of the endophytic diazotrophic bacteria that exhibited a straggly distribution pattern **(Fig. 2)**.

Phylogenetic evaluation of bacterial endophytes in rice landraces

The representative DGGE bands of both bacterial endophytes derived from the rice landraces (**Fig. 1A** and **B**) were sequenced to determine their phylogenetic relationships and identifications. Most selected band sequences of the 16S rRNA gene were matched to the uncultured bacteria, while some of them revealed phylogenetic relationship to *Pseudomonas* and *Klebsiella* (**Table 3** and **Fig. 3**). The band sequence A11-4 showed 100% identity to the sequence of *Pantoea* sp. A1128, which is only a 16S rDNA sequence that has the origin from rice paddy soil (**Table 3**). The same endophytic bacterium was found in root (band sequences A10-8; HE860550) and leaf (band sequences A2-7; HE860547) of the BPD rice landrace grown in nursery stage (**Fig. 1A**). These two band sequences were closely related to the same uncultured bacterium supported by the sequence similarity of 95 and 100% (**Table 3**). The band sequences A1-4, A11-3, P2, and P7 derived from different rice plant tissues of the rice landraces grown in various stages of growth (**Fig. 1A**) were clustered in the same phylogenetic cluster with the uncultured bacteria. These bacteria have their origins from pond water and the herbivorous-freshwater fish, *Ctenopharyngodon idellus* (**Table 3**). The band sequences P1 (HE860534; from mature seed) and P6 (HE860538; from stem of vegetative stage) of the BP rice landrace (**Fig. 1A**) were clustered in the same phylogenetic clade (**Fig. 3**), which were closely related to the uncultured bacteria living in root and rhizosphere of grasses.

For the endophytic diazotrophic bacteria, most of the band sequences were also closely related to the uncultured bacteria. However, the other sequences revealed higher varieties of matched genera (*Novosphingobium*, *Pelomonas*, *Herbaspirillum*, *Enterobacter*, *Klebsiella,* and *Spirochaeta*) than those of the endophytic bacteria (**Table 3** and **Fig. 4**). The band sequences 3-1 and 3-2 derived from pre-planting mature seed and reproductive roots of the BP rice landrace were clustered in the same phylogenetic clade with endophytic *Herbaspirillum* sp. B501 (**Table 3** and **Fig. 4)**. This was evidence to ensure that the band sequences were derived from the bacterial endophytes. The same bacterium, *Enterobacter* sp. was also found in root of rice at reproductive stage (band sequences 17-5) and grain (band sequences 17-3) of the BPT rice landrace (**Fig. 1B** and **Table 3**). This was in contrast to the same bacterium but found in different rice landraces such as the band

sequences 24-1 (JX042298; from the BPD rice landrace) and 24-4 (JX042299; from the BSM rice landrace) that were matched to *Novosphingobium nitrogenifgens* (**Fig. 1B** and **Table 3**). This band sequence 37-1 (JX042307) was closely related to the sequence of an uncultured bacterium (98% identity) which was the only a sequence found with its origin from roots of *O. officinalis* (**Table 3**). The other 2 band sequences 31-4 (JX042302) and 37-8 (JX042308) were closely related to the sequences of the uncultured bacteria from the rhizosphere samples (**Table 3**).

In addition, using this phylogenetic tree (**Fig. 4**), it was able to visualize 5 phylogenetic clades including α-, β-, γ-, and δ-Proteobacteria and Spirochaetes. Most of the band sequences were belonged to clade II of β-Proteobacteria followed by clade III of γ-Proteobacteria.

Discussion

Within this study, the community structures of bacterial endophytes within every plant tissue examined from the highland rice landraces grown in different stages of growth exhibited similarities but also dissimilarities depending on either environmental factors relating to the cultivating practices or the varieties of the rice landraces.

These factors can affect the interactions, colonization traits, survival, and distribution of the bacterial endophytes living in the rice plant tissues (Hardoim et al. 2008; Knief et al. 2012; Mano and Morisaki 2008). In general, high diversity of bacteria has been found in soil, especially in rhizosphere habitats, where plants release root exudates and mutually beneficial nutrient exchange between host plant and microorganisms take place (Bais et al. 2006; Hardoim et al. 2011; Li et al. 2008; Roesch et al. 2008; Sessitsch et al. 2012). It was in accordance to our observations that the community's richness of endophytic bacteria examined was higher in root tissues than those other tissues of most rice cultivars. However, it was a different picture when endophytic diazotrophic bacteria were examined, as their richness values were relatively low and varied across the rice plant tissues.

The diversity indices of both bacterial groups found in the rice plant tissues also had a similar trend in terms of their richness values. Some authors assumed that lower diversity of the endophytic diazotrophic bacteria within wild rice cultivars might due to the genetic variation of the rice (Elbeltagy and Ando 2008; Engelhard et al. 2000). However, high density of the endophytic diazotrophic bacteria was found in nitrogen-poor environment where more than 50% of the endophytes were diazotrophic bacteria (Sessitsch et al. 2012). In our study, the proportions of the endophytic diazotrophic bacteria per the total endophytic bacteria were

354 relatively high (> 40%) with no significant different in most of the rice landraces, except for the BWB rice
355 landrace that revealed significantly lower values. This might strongly relate to the difference of agricultural
356 practices, where the BWB rice landrace was supplemented with nitrogen fertilizer at the nursery stage of growth
357 (**Table 1**). There are some evidences suggested that the use of chemical input in rice paddy soil, particularly for
358 the nitrogen fertilizer, can influence the population, diversity, function, and host-interaction of nitrogen fixing
359 bacteria with their mutual rice plants (Prakamhang et al. 2009; Wartiainen et al. 2008).

360 The results derived from cluster analysis and PCA also ensured that the community structures of the
361 endophytic bacteria in any rice landraces tested revealed higher similarity and stability than those of the
362 endophytic diazotrophic bacteria. According to the multiple parameters involved, these two cultivation
363 techniques could have a significant impact upon the whole picture of the bacterial community structures.
364 However, a better understanding of the bacterial community in generic and/or species levels was needed to
365 provide greater resolution of this complex bacterial consortium and its agricultural significance. This has been
366 partly addressed, by the molecular sequencing tools applied in this study. Here, we selected the representative
367 DGGE bands for bacterial identification. The 16S rDNA fragments were closely related to uncultured bacteria,
368 some of them were related to *Pseudomonas* sp., *Pantoea* sp. and *Klebsiella* sp. which are members of the γ-
369 Proteobacteria. This bacterial phylum has been found currently in association with rice plant and paddy soil
370 either in symbiotic (Reinhold-Hurek and Hurek 2011; Sun et al. 2008; Takahashi et al. 2011) or free living
371 patterns (Sooksa-Nguan et al. 2010; Tago et al. 2011). Previous studies concerning the endophytic bacteria of
372 rice reported that the Proteobacteria are the dominant bacteria found in association with rice plant (Hardoim et al.
373 2011; Knief et al. 2012; Reinhold-Hurek and Hurek 2011; Sessitsch et al. 2012). Some isolates of
374 Actinobacteria, Firmicutes (Reinhold-Hurek and Hurek 2011), Bacteroidetes and Deinococcus-Thermus phyla
375 (Knief et al. 2012) have previously been found in relation with rice plants. In the case of the endophytic
376 diazotrophic bacteria, most of our isolates belonged to the Proteobacteria; mainly in the β-Proteobacteria.
377 Interestingly, some members of the phylum Spirochaetae living in association with the rice landraces were
378 firstly reported in this study. Nonetheless, some members of this bacterial phylum play an important role in
379 nitrogen fixation, as evidence revealed that some Spirochetes, either living in termite hindguts or free-living in
380 freshwater, can fix nitrogen through the *nifH* homolog genes (Lilburn et al. 2001).

381 It was clear from this study, that the endophytic bacterial communities across the highland rice landraces
382 revealed flexibility of phenotype and genotype depending on various factors including different rice cultivars
383 that possess high genetic variation, cultivation processes with natural farming and/or supplied with chemical

inputs. An apparent result of different treatments by supplying nitrogen fertilizer ensured the dissimilarity of the bacterial community structures, which decreased the community's richness and diversity of nitrogen fixing bacteria. However, the results in this study indicated that both endophytic and endophytic diazotrophic bacteria could exist in the rice various tissues examined and were distributed according to the growth practices of the rice plants. Moreover, some of them might exist in the mature seeds and persist even in the post-harvest stage. These endophytic bacteria of the rice landraces might be an important key to maintain yield and health of the rice plants in poor or deficient farming condition. Application of this plant-microbe interaction data, particularly in case of endophytic diazotrophic bacteria might be an alternative way to reduce chemical usage, and at the same time improve soil quality and sustain an organic farming system for future agricultural husbandry of rice.

Acknowledgements

The authors would like to acknowledge the supports from Office of the Higher Education Commission, Thailand under the National Research University Project, the Graduated School of Chiang Mai University, Chiang Mai, Thailand and Thailand Research Fund RTA5580007. We also thank the Model Farm Project, the Royal Initiatives of Her Majesty the Queen, Khun Tae Village, Thailand for assistance in rice samples collection and Mr. Chainarong Kesamoon, a Ph.D. candidate at the Department of Mathematics, Autonomous University of Barcelona Spain, for his kindly suggest in numerical analysis of this work. We are grateful to Mr. Keegan H. Kennedy, Department of Biology, Chiang Mai University and Dr. Peter Green, Aberdeen, UK for helping improve English text.

References

Arun B, Gopinath B, Sharma S (2012) Plant growth promoting potential of bacteria isolated on N free media from rhizosphere of *Cassia occidentalis*. World J Microbiol Biotechnol 28:2849-2857.

Asakawa S, Kimura M (2008) Comparison of bacterial community structures at main habitats in paddy field ecosystem based on DGGE analysis. Soil Biol Biochem 40:1322-1329.

Azevedo JL, Maccheroni JW, Pereira JO, de Araújo WL (2000) Endophytic microorganisms: a review on insect control and recent advances on tropical plants. Electron J Biotechnol 3:40-65.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. Annu Rev Plant Biol 57:233-266.

413 Bandara W, Seneviratne G, Kulasooriya S (2006) Interactions among endophytic bacteria and fungi: Effects and
414 potentials. J Biosciences 31:645-650.
415 Bonman JM, Khush GS, Nelson RJ (1992) Breeding rice for resistance to pests. Annu Rev Phytopathol 30:507-
416 528.
417 Chakhonkaen S, Pitnjam K, Saisuk W, Ukoskit K, Muangprom A (2012) Genetic structure of Thai rice and rice
418 accessions obtained from the International Rice Research Institute. Rice 5:1-13.
419 Coelho MRR, Marriel IE, Jenkins SN, Lanyon CV, Seldin L, O'Donnell AG (2009) Molecular detection and
420 quantification of *nifH* gene sequences in the rhizosphere of sorghum (*Sorghum bicolor*) sown with two
421 levels of nitrogen fertilizer. Appl Soil Ecol 42:48-53.
422 Dellaporta S, Wood J, Hicks J (1983) A plant DNA minipreparation: Version II. Plant Mol Biol Rep 1:19-21.
423 Elbeltagy A, Ando Y (2008) Expression of nitrogenase gene (*NIFH*) in roots and stems of rice, *Oryza sativa*, by
424 endophytic nitrogenfixing communities. Afr J Biotechnol 7:1950-1957.
425 Engelhard M, Hurek T, Reinhold-Hurek B (2000) Preferential occurrence of diazotrophic endophytes, *Azoarcus*
426 spp., in wild rice species and land races of *Oryza sativa* in comparison with modern races. Environ
427 Microbiol 2:131-141.
428 Feng Y, Shen D, Song W (2006) Rice endophyte *Pantoea agglomerans* YS19 promotes host plant growth and
429 affects allocations of host photosynthates. J Appl Microbiol 100:938-945.
430 Fromin N, Hamelin J, Tarnawski S, Roesti D, Jourdain-Miserez K, Forestier N, Teyssier-Cuvelle S, Gillet F,
431 Aragno M, Rossi P (2002) Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting
432 patterns. Environ Microbiol 4:634-643.
433 García de Salamone I, Di Salvo L, Escobar Ortega J, Boa Sorte P, Urquiaga S, Teixeira K (2010) Field response
434 of rice paddy crop to *Azospirillum* inoculation: physiology of rhizosphere bacterial communities and
435 the genetic diversity of endophytic bacteria in different parts of the plants. Plant Soil 336:351-362.
436 Hardoim PR, Andreote FD, Reinhold-Hurek B, Sessitsch A, van Overbeek LS, van Elsas JD (2011) Rice root-
437 associated bacteria: insights into community structures across 10 cultivars. FEMS Microbiol Ecol
438 77:154-164.
439 Hardoim PR, van Overbeek LS, van Elsas JD (2008) Properties of bacterial endophytes and their proposed role
440 in plant growth. Trends Microbiol 16:463-471.
441 Jha B, Gontia I, Hartmann A (2012) The roots of the halophyte *Salicornia brachiata* are a source of new
442 halotolerant diazotrophic bacteria with plant growth-promoting potential. Plant Soil 356:265-277.

443 Jia Z, Kikuchi H, Watanabe T, Asakawa S, Kimura M (2007) Molecular identification of methane oxidizing
444 bacteria in a Japanese rice field soil. Biol Fert Soils 44:121-130.
445 Kaga H, Mano H, Tanaka F, Watanabe A, Kaneko S, Morisaki H (2009) Rice seeds as sources of endophytic
446 bacteria. Microbes Environ 24:154-162.
447 Knief C, Delmotte N, Chaffron S, Stark M, Innerebner G, Wassmann R, von Mering C, Vorholt JA (2012)
448 Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice.
449 ISME J 6:1378-1390.
450 Koomnok C, Teaumroong N, Rerkasem B, Lumyong S (2007) Diazotroph endophytic bacteria in cultivated and
451 wild rice in Thailand. ScienceAsia 33:429-435.
452 Li YL, Fan XR, Shen QR (2008) The relationship between rhizosphere nitrification and nitrogen-use efficiency
453 in rice plants. Plant Cell Environ 31:73-85.
454 Lilburn TG, Kim KS, Ostrom NE, Byzek KR, Leadbetter JR, Breznak JA (2001) Nitrogen fixation by symbiotic
455 and free-living Spirochetes. Science 292:2495-2498.
456 Loaces I, Ferrando L, Fernández Scavino A (2011) Dynamics, diversity and function of endophytic siderophore-
457 producing bacteria in rice. Microbial Ecol 61:606-618.
458 Mano H, Morisaki H (2008) Endophytic bacteria in the rice plant. Microbes Environ 23:109-117.
459 Mano H, Tanaka F, Nakamura C, Kaga H, Morisaki H (2007) Culturable endophytic bacterial flora of the
460 maturing leaves and roots of rice plants (*Oryza sativa*) cultivated in a paddy field. Microbes Environ
461 22:175-185.
462 McCouch S (2004) Diversifying selection in plant breeding. PLoS Biol 2:1507-1512.
463 Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing
464 gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S
465 rRNA. Appl Environ Microbiol 59:695-700.
466 Oupkaew P, Pusadee T, Sirabanchongkran A, Rerkasem K, Jamjod S, Rerkasem B (2011) Complexity and
467 adaptability of a traditional agricultural system: case study of a gall midge resistant rice landrace from
468 northern Thailand. Genet Resour Crop Evol 58:361-372.
469 Parzies HK, Spoor W, Ennos RA (2004) Inferring seed exchange between farmers from population genetic
470 structure of barley landrace Arabi Aswad from northern Syria. Genet Resour Crop Evolution 51:471-
471 478.

472 Poly F, Monrozier LJ, Bally R (2001a) Improvement in the RFLP procedure for studying the diversity of *nifH*
473 genes in communities of nitrogen fixers in soil. Res Microbiol 152:95-103.
474 Poly F, Ranjard L, Nazaret S, Gourbiere F, Monrozier LJ (2001b) Comparison of *nifH* gene pools in soils and
475 soil microenvironments with contrasting properties. Appl Environ Microbiol 67:2255-2262.
476 Prakamhang J, Minamisawa K, Teamtaisong K, Boonkerd N, Teaumroong N (2009) The communities of
477 endophytic diazotrophic bacteria in cultivated rice (*Oryza sativa* L.). Appl Soil Ecol 42:141-149.
478 Purnomo E, Mursyid A, Syarwani M, Jumberi A, Hashidoko Y, Hasegawa T, Honma S, Osaki M (2005)
479 Phosphorus solubilizing microorganisms in the rhizosphere of local rice varieties grown without
480 fertilizer on acid sulfate soils. Soil Sci Plant Nutr 51:679-681.
481 Pusadee T, Jamjod S, Chiang Y-C, Rerkasem B, Schaal BA (2009) Genetic structure and isolation by distance in
482 a landrace of Thai rice. Proc Nat Acad Sci 106:13880-13885.
483 Reinhold-Hurek B, Hurek T (2011) Living inside plants: bacterial endophytes. Curr Opin Plant Biol 14:435-443.
484 Roesch L, Camargo F, Bento F, Triplett E (2008) Biodiversity of diazotrophic bacteria within the soil, root and
485 stem of field-grown maize. Plant Soil 302:91-104.
486 Saito K, Linquist B, Keobualapha B, Shiraiwa T, Horie T (2006) Farmers' knowledge of soils in relation to
487 cropping practices: A case study of farmers in upland rice based slash-and-burn systems of northern
488 Laos. Geoderma 136:64-74.
489 Sasaki K, Ikeda S, Eda S, Mitsui H, Hanzawa E, Kisara C, Kazama Y, Kushida A, Shinano T, Minamisawa K,
490 Sato T (2010) Impact of plant genotype and nitrogen level on rice growth response to inoculation with
491 *Azospirillum* sp. strain B510 under paddy field conditions. Soil Sci Plant Nutr 56:636-644.
492 Sessitsch A, Hardoim P, Döring J, Weilharter A, Krause A, Woyke T, Mitter B, Hauberg-Lotte L, Friedrich F,
493 Rahalkar M, Hurek T, Sarkar A, Bodrossy L, van Overbeek L, Brar D, van Elsas JD, Reinhold-Hurek
494 B (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by
495 metagenomic analysis. Mol Plant-Microbe Interact 25:28-36.
496 Shannon CE, Weaver W (1963) The mathematical theory of communication. University of Illinois Press,
497 Urbana,
498 Simonet P, Grosjean MC, Misra AK, Nazaret S, Cournoyer B, Normand P (1991) *Frankia* genus-specific
499 characterization by polymerase chain reaction. Appl Environ Microbiol 57:3278-3286.
500 Sommut W (2003) Changes in flood-prone rice ecosystems in Thailand, crop year 2000-2001. Prachinburi

501 Sooksa-Nguan T, Gypmantasiri P, Boonkerd N, Thies JE, Teaumroong N (2010) Changes in bacterial
502 community composition in the system of rice intensification (SRI) in Chiang Mai, Thailand. Microbes
503 Environ 25:224-227.
504 Sun L, Qiu F, Zhang X, Dai X, Dong X, Song W (2008) Endophytic bacterial diversity in rice (*Oryza sativa* L.)
505 roots estimated by 16S rDNA sequence analysis. Microbial Ecol 55:415-424.
506 Tago K, Ishii S, Nishizawa T, Otsuka S, Senoo K (2011) Phylogenetic and functional diversity of denitrifying
507 bacteria isolated from various rice paddy and rice-soybean rotation fields. Microbes Environ 26:30-35.
508 Takahashi H, Sekiguchi H, Ito T, Sasahara M, Hatanaka N, Ohba A, Hase S, Ando S, Hasegawa H, Takenaka S
509 (2011) Microbial community profiles in intercellular fluid of rice. J Gen Plant Pathol 77:121-131.
510 Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA)
511 software version 4.0. Mol Biol Evol 24:1596-1599.
512 Tittabutr P, Awaya JD, Li QX, Borthakur D (2008) The cloned 1-aminocyclopropane-1-carboxylate (ACC)
513 deaminase gene from *Sinorhizobium* sp. strain BL3 in *Rhizobium* sp. strain TAL1145 promotes
514 nodulation and growth of *Leucaena leucocephala*. Syst Appl Microbiol 31:141-150.
515 Van Nieuwenhove C, Van Holm L, Kulasooriya SA, Vlassak K (2000) Establishment of *Azorhizobium*
516 *caulinodans* in the rhizosphere of wetland rice (*Oryza sativa* L.). Biol Fert Soils 31:143-149.
517 Verma SC, Singh A, Chowdhury SP, Tripathi AK (2004) Endophytic colonization ability of two deep-water rice
518 endophytes, *Pantoea* sp. and *Ochrobactrum* sp. using green fluorescent protein reporter. Biotechnol
519 Lett 26:425-429.
520 Wartiainen I, Eriksson T, Zheng W, Rasmussen U (2008) Variation in the active diazotrophic community in rice
521 paddy--*nifH* PCR-DGGE analysis of rhizosphere and bulk soil. Appl Soil Ecol 39:65-75.
522
523

Table 1 Sampling locations and conditions of the rice landraces grown in highland of northern Thailand

Rice cultivars	Collection site	Area above sea level (m)	Chemical fertilized
Bue Polo (BP)	N 18.38958 E 098.50671	1191	No
Bue Pra Taw (BPT)	N 18.39079 E 098.51444	1194	No
Bue Pra Do (BPD)	N 18.38804 E 098.51085	1201	No
Bue Saw Mi (BSM)	N 18.39682 E 098.51631	1137	No
Bue Wah Bo (BWB)	N 18.38798 E 098.51092	1199	Applied with N-P-K (46-0-0) of 6.25 g/m^2 in nursery stage

Table 2 Numerical comparison of the 16S rRNA and *nifH* genes obtained from endophytic bacteria of the rice landraces profiling by PCR-DGGE

Different growth stages and plant tissues of rice cultivars			16S rRNA gene				*nifH* gene				% Proportion of endophytic diazotrophic bacteria
			d	*H′*	*E*	*S*	*d*	*H′*	*E*	*S*	
Pre-planting mature seed	BP		8.73	2.61	1.21	0.70	7.75	2.04	1.00	0.66	
	BPT		7.41	2.47	1.23	0.74	6.43	1.76	0.95	0.81	
	BPD		6.72	2.37	1.25	0.80	5.72	1.59	0.91	0.81	59.10±19.22
	BSM		6.00	2.29	1.28	0.75	4.19	1.09	0.76	0.92	
	BWB		5.24	2.18	1.31	0.72	3.32	0.68	0.57	0.87	
Nursery	BP	Root	8.73	2.61	1.21	0.64	8.38	2.18	1.03	0.80	
		Stem	6.72	2.37	1.24	0.66	7.10	1.92	0.98	0.83	80.18±3.84
		Leaf	8.08	2.54	1.21	0.63	7.10	1.93	0.98	0.80	
	BPT	Root	9.35	2.68	1.20	0.56	7.10	1.92	0.98	0.75	
		Stem	6.72	2.38	1.25	0.68	7.10	1.92	0.98	0.77	76.67±4.56
		Leaf	7.41	2.45	1.22	0.61	7.10	1.91	0.97	0.76	
	BPD	Root	8.08	2.54	1.22	0.69	5.72	1.60	0.92	0.83	
		Stem	4.43	2.07	1.39	0.80	5.72	1.60	0.92	0.84	68.91±7.47
		Leaf	6.72	2.38	1.25	0.80	5.72	1.59	0.91	0.81	
	BSM	Root	8.08	2.54	1.22	0.64	4.98	1.38	0.86	0.88	
		Stem	8.73	2.61	1.21	0.63	6.43	1.76	0.94	0.78	65.82±10.75
		Leaf	8.08	2.54	1.22	0.66	7.10	1.93	0.98	0.81	
	BWB	Root	3.55	1.93	1.52	0.76	-	0.00	-	1.00	
		Stem	9.97	2.73	1.19	0.61	3.32	0.69	0.57	0.89	19.48±17.37
		Leaf	4.43	2.06	1.39	0.71	3.32	0.69	0.57	0.91	
Vegetative	BP	Root	6.72	2.38	1.25	0.67	7.75	2.04	1.00	0.63	
		Stem	4.43	2.06	1.38	0.74	6.43	1.77	0.95	0.76	82.10±6.32
		Leaf	6.72	2.36	1.24	0.64	6.43	1.77	0.95	0.76	
	BPT	Root	9.35	2.69	1.20	0.64	10.19	2.45	1.06	0.70	
		Stem	6.72	2.37	1.24	0.65	7.10	1.89	0.96	0.66	89.65±9.40
		Leaf	3.55	1.92	1.52	0.69	7.10	1.89	0.96	0.75	
	BPD	Root	6.72	2.37	1.25	0.64	7.10	1.93	0.98	0.81	
		Stem	4.43	2.06	1.39	0.76	4.98	1.34	0.84	0.69	81.92±17.22
		Leaf	2.57	1.78	1.89	0.85	6.43	1.77	0.95	0.81	
	BSM	Root	6.72	2.37	1.24	0.58	7.10	1.93	0.98	0.84	
		Stem	4.43	2.06	1.38	0.71	7.75	2.04	0.99	0.74	90.98±8.95
		Leaf	3.55	1.92	1.51	0.70	6.43	1.77	0.95	0.79	
	BWB	Root	2.57	1.79	1.89	0.80	3.32	0.69	0.57	0.91	
		Stem	9.35	2.68	1.20	0.64	-	0.00	-	1.00	21.38±19.62
		Leaf	9.35	2.67	1.19	0.65	3.32	0.68	0.57	0.87	
Reproductive	BP	Root	9.97	2.75	1.20	0.63	6.43	1.78	0.95	0.76	
		Stem	6.72	2.36	1.24	0.70	5.72	1.56	0.89	0.62	71.03±8.13
		Leaf	5.24	2.16	1.30	0.62	5.72	1.53	0.88	0.56	
		Grain	3.55	1.90	1.50	0.66	5.72	1.57	0.90	0.64	
	BPT	Root	12.30	2.97	1.18	0.59	7.10	1.94	0.99	0.94	
		Stem	8.08	2.53	1.21	0.57	7.75	2.05	1.00	0.77	67.66±10.01
		Leaf	6.72	2.35	1.23	0.54	5.72	1.58	0.90	0.74	
		Grain	6.72	2.38	1.25	0.66	4.98	1.35	0.84	0.79	
	BPD	Root	9.35	2.68	1.20	0.69	6.43	1.78	0.96	0.79	
		Stem	3.55	1.94	1.53	0.89	4.98	1.37	0.85	0.85	66.99±6.65
		Leaf	5.24	2.18	1.32	0.80	5.72	1.60	0.91	0.83	
		Grain	6.72	2.37	1.24	0.63	4.98	1.37	0.85	0.82	
	BSM	Root	10.57	2.81	1.19	0.66	6.43	1.79	0.96	0.86	
		Stem	9.97	2.73	1.19	0.61	8.38	2.19	1.03	0.85	82.34±14.21
		Leaf	6.00	2.26	1.26	0.61	8.38	2.18	1.02	0.74	
		Grain	5.24	2.15	1.30	0.64	7.10	1.93	0.98	0.77	
	BWB	Root	9.97	2.76	1.20	0.72	3.32	0.69	0.57	0.92	
		Stem	10.57	2.78	1.18	0.58	3.32	0.69	0.58	0.98	22.34±16.38
		Leaf	10.57	2.78	1.18	0.59	4.19	1.10	0.77	0.93	
		Grain	8.73	2.61	1.20	0.63	-	0.00	-	1.00	

The 16S rRNA and *nifH* genes were taken from any rice plant tissues of landrace rice cultivars Bue Polo (BP), Bue Pra Taw (BPT), Bue Pra Do (BPD), Bue Saw Mi (BSM) and Bue Wah Bo (BWB) grown in any stages, which were further profiled by PCR-DGGE where the DGGE bands profiling data were used to analyze numerical parameters comprised of richness (*d*), diversity index (Shannon-Wiener index: *H′*), evenness (*E*) and stability (*S*). The results indicated with (-) are incomputable mathematics.

Table 3 Data summary of the 16S rRNA gene and *nifH*-derived protein sequences obtained from the respective bands present in DGGE profiles and their closest match sequences from GenBank database

	Band	Accession no.	Database match with accession no. in parentheses	Origin	% Identity
16S rRNA gene sequence	P1	HE860534	Uncultured bacterium clone M3a (AJ851120.1)	Root of *Molinia caerulea*	98
	P2	HE860535	Uncultured bacterium DGGE band DS30-2 (EU585922.1)	*Ctenopharyngodon idellus*	96
	P3	HE860536	Uncultured bacterium DGGE band 7 (HQ876068)	*Danio rerio*	94
	P4	HE860537	Uncultured bacterium clone M3a (AJ851120.1)	Root of *Molinia caerulea*	99
	P5	HE860538	Uncultured bacterium clone M3a (AJ851120.1)	Root of *Molinia caerulea*	97
	P6	HE860539	Uncultured bacterium clone M6a (AJ851126.1)	Rhizosphere of grasses	95
	P7	HE860540	Uncultured bacterium DGGE band DS31-5 (EU585924.1)	*Ctenopharyngodon idellus*	97
	P8	HE860541	Uncultured bacterium DGGE band C-3 (EF669488)	*Ctenopharyngodon idellus*	96
	P9	HE860542	Uncultured bacterium clone M15a (AJ851144.1)	Rhizosphere of grasses	97
	P10	HE860543	Uncultured bacterium DGGE band DS31-5 (EU585924.1)	*Ctenopharyngodon idellus*	94
	P11	HE860544	Uncultured bacterium DGGE band DS34-3 (EU585925.1)	*Ctenopharyngodon idellus*	92
	A1-4	HE860545	Uncultured bacterium DGGE band DS31-5 (EU585924.1)	*Ctenopharyngodon idellus*	100
	A2-7	HE860547	Uncultured bacterium DGGE band 7 (HQ876068.1)	*Danio rerio*	100
	A7-9	HE860548	Uncultured *Pseudomonas* sp. clone 19318_ZC5M6_G12 (JQ897449)	Argillite geological	100
	A10-2	HE860549	Uncultured bacterium clone nby488c06c1 (HM831057.1)	Skin of *Mus musculus*	95
	A10-8	HE860550	Uncultured bacterium DGGE band 7 (HQ876068.1)	*Danio rerio*	95
	A11-3	HE860551	Uncultured bacterium DGGE band he-49 (FJ235679.1)	Pond water	100
	11-4	HE860552	*Pantoea* sp. A1128 (JX266309.1)	Taihu paddy soil	100
	12-1	HE860553	Uncultured bacterium DGGE band 7 (HQ876068.1)	*Danio rerio*	100
	12-2	HE860554	Uncultured bacterium DGGE band 7 (HQ876068.1)	*Danio rerio*	100
	14-2	HE860555	Uncultured bacterium clone ncd2380h03c1 (JF207567.1)	Volar of *Homo sapiens*	95
nifH-derived protein sequences	3-1	JX042294	*Herbaspirillum* sp. B501 *nifH* genes (AB196476.1)	Endophytic *Herbaspirillum* sp. B501	99
	3-2	JX042295	*Herbaspirillum* sp. B501 *nifH* genes (AB196476.1)	Endophytic *Herbaspirillum* sp. B501	89
	17-3	JX042296	*Enterobacter* sp. SP1 *nifH* genes (JQ001785.1)	*Saccharum officinarum* L.	98
	17-5	JX042297	*Enterobacter* sp. SP1 *nifH* genes (JQ001785.1)	*Saccharum officinarum* L.	97
	24-1	JX042298	*Novosphingobium nitrogenifigens* Y88 NifH (DQ660368.1)	A New Zealand pulp and paper wastewater	97
	24-4	JX042299	*Novosphingobium nitrogenifigens* Y88 NifH (DQ660368.1)	A New Zealand pulp and paper wastewater	99
	26-1	JX042300	*Klebsiella* sp. AL060225_04 *nifH* genes (FJ593866.1)	*Trachymyrmex* sp.	99
	26-5	JX042301	*Spirochaeta aurantia* clone 2 *nifH* genes (AF325792.1)	*Spirochaeta aurantia*	89
	31-4	JX042302	Uncultured bacterium DGGE band 13r NifH (JN648871.1)	Horticultural soil	96
	31-6	JX042303	Uncultured bacterium clone 16Z65C *nifH* genes (AY787543.1)	N-ViroTech pulp and paper wastewater	87
	33-3	JX042304	*Pelomonas saccharophila nifH* gene (AB188120.1)	Genomic DNA of *Pelomonassaccharophila*	94
	33-5	JX042305	*Klebsiella variicola* strain 6A2 *nifH* genes (AY367394.1)	Genomic DNA of *Klebsiella variicola*	99
	33-6	JX042306	Uncultured bacterium clone P67 *nifH* genes (GU196859.1)	Genomic DNA *Ornithocercus quadratus*	96
	37-1	JX042307	Uncultured bacterium clone g1(102) *nifH* genes (AF331982.1)	Roots of *Oryza officinalis*	98
	37-8	JX042308	Uncultured bacterium clone 4-15 (FJ807383.1)	Mangrove rhizosphere soil	85

Figure legends

Fig. 1 Cluster analysis of the 16S rRNA (**A**) and *nifH* (**B**) genes obtained from endophytic bacteria of the rice landraces profiling by PCR-DGGE. Rice cultivars including Bue Polo (BP), Bue Pra Taw (BPT), Bue Pra Do (BPD), Bue Saw Mi (BSM), and Bue Wah Bo (BWB), growing in different stages comprised of pre-planting mature seed (MS), nursery stage (N), vegetative stage (V) and reproductive stage (R), were used. The parts (root, stem, leaf, and grain) of rice tissues are indicated behind the codes of rice cultivars grown in different stages. The marked codes upper the respective bands in the band profile were determined as selected bands for further sequencing and phylogenetic analysis (see also **Figs. 3** and **4**).

Fig. 2 Principal component analysis of the 16S rRNA and *nifH* genes obtained from endophytic bacteria of the rice landraces profiling by PCR-DGGE

Plant tissues	Growth stages			
	Pre-planting	Nursery	Vegetative	Reproductive
Mature seed	△			
Root		◇	□	○
Stem		◇	□	○
Leaf		◆	■	●
Grain				○

Fig. 3 Unrooted phylogenetic tree of 16S rRNA gene sequences respective to the marked bands present in the DGGE bands profile (**Fig. 1A**). Bootstrap analysis was based on 1,000 replicates. Neighbor Joining tree and the scale bar represent 5% dissimilarity, where the accession numbers of the sequences are indicated in the parentheses.

Fig. 4 Unrooted phylogenetic tree of *nifH*-derived protein sequences respective to the marked bands present in the DGGE bands profile (**Fig. 1B**). Bootstrap analysis was based on 1,000 replicates. Neighbor Joining tree and the scale bar represent 5% dissimilarity, where the accession numbers of the sequences are indicated in the parentheses. The phylogenetic tree was divided into 5 different clades including α-Proteobacteria (**I**), β-Proteobacteria (**II**), γ-Proteobacteria (**III**), δ-Proteobacteria (**IV**) and Spirochaetes (**V**).

Fig. 1

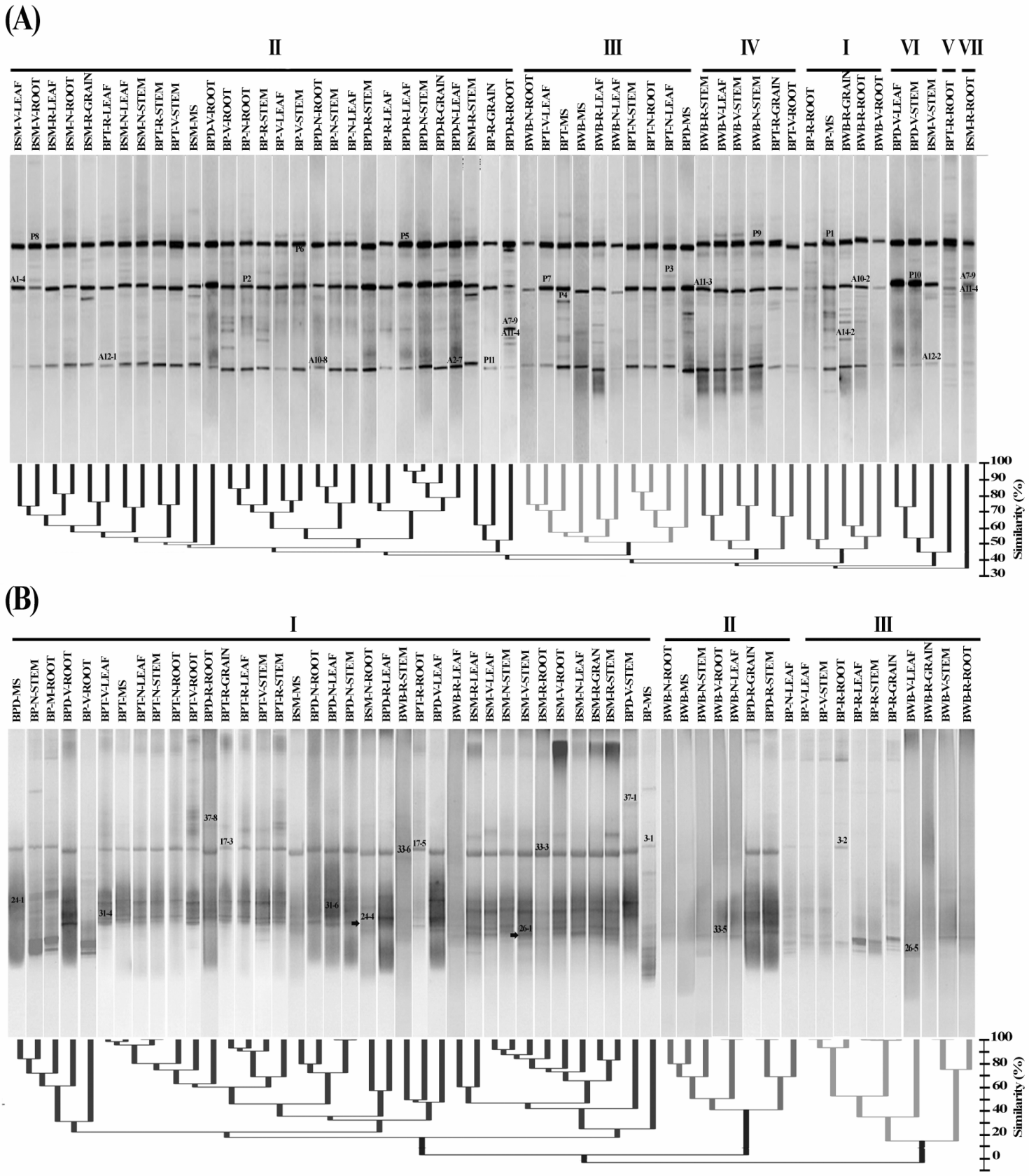
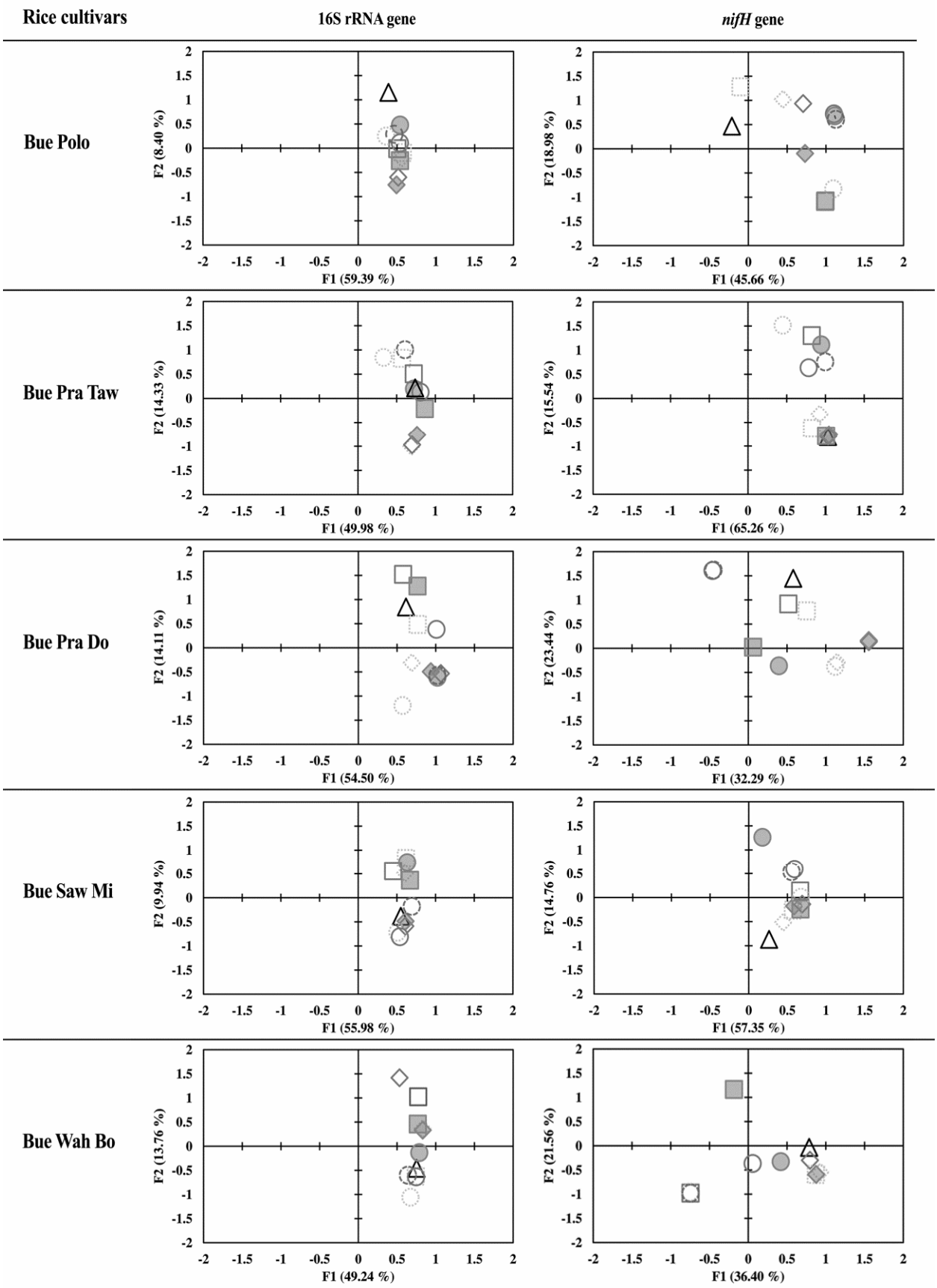
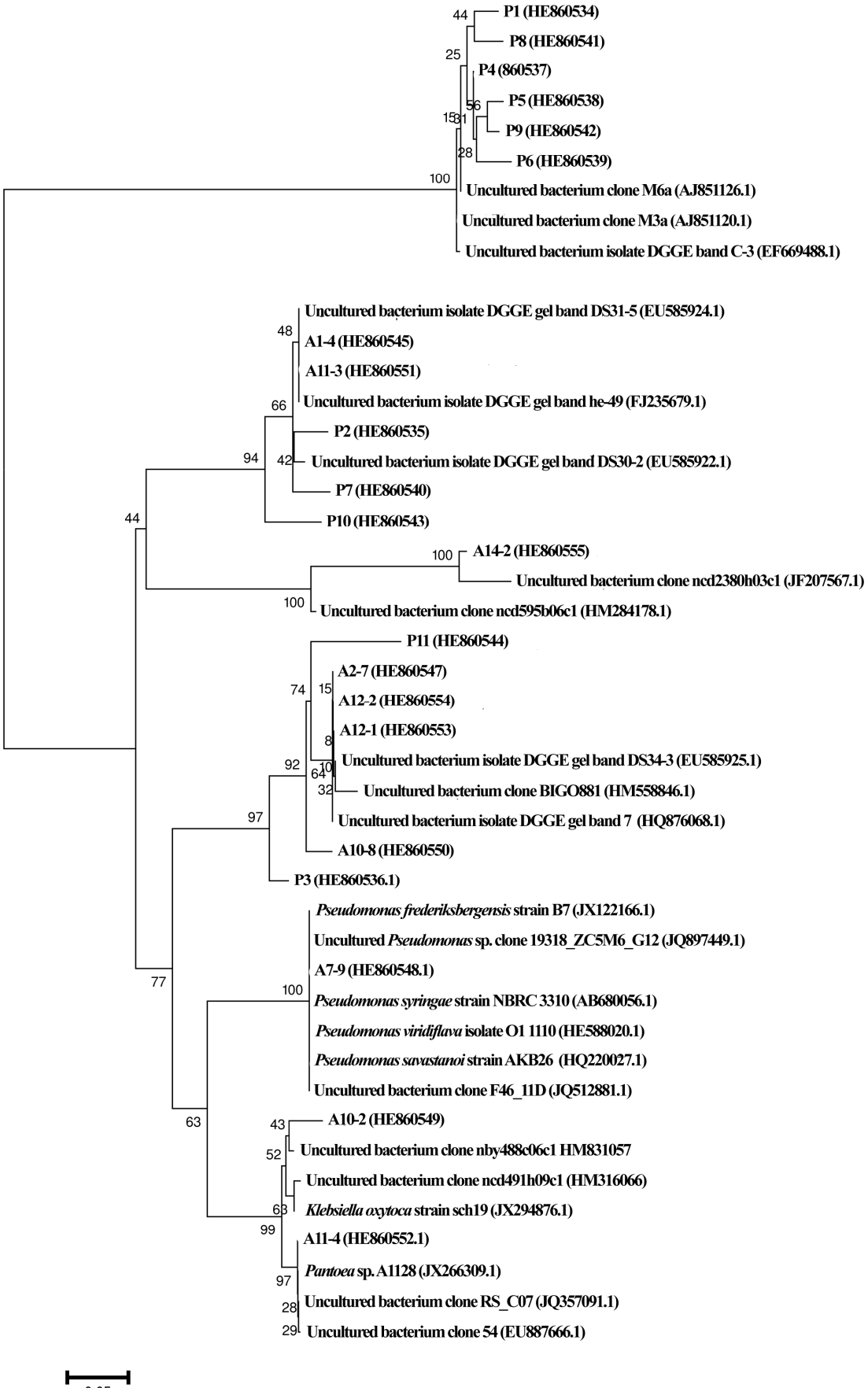

(A)
II
III
IV
I
VI
V
VII
BSM-V-LEAF
BSM-V-ROOT
BSM-R-LEAF
BSM-N-ROOT
BSM-R-GRAIN
BPT-R-LEAF
BSM-N-LEAF
BSM-N-STEM
BPT-R-STEM
BPT-V-STEM
BSM-MS
BPD-V-ROOT
BP-V-ROOT
BP-N-ROOT
BP-R-STEM
BP-V-LEAF
BP-V-STEM
BPD-N-ROOT
BP-N-STEM
BP-N-LEAF
BPD-R-STEM
BP-R-LEAF
BPD-R-LEAF
BPD-N-STEM
BPD-R-GRAIN
BPD-N-LEAF
BSM-R-STEM
BP-R-GRAIN
BPD-R-ROOT
BWB-N-ROOT
BPT-V-LEAF
BPT-MS
BWB-MS
BWB-R-LEAF
BWB-N-LEAF
BPT-N-STEM
BPT-N-ROOT
BPT-N-LEAF
BPD-MS
BWB-R-STEM
BWB-V-LEAF
BWB-V-STEM
BWB-N-STEM
BPT-R-GRAIN
BPT-V-ROOT
BP-R-ROOT
BP-MS
BWB-R-GRAIN
BWB-R-ROOT
BWB-V-ROOT
BPD-V-LEAF
BPD-V-STEM
BSM-V-STEM
BPT-R-ROOT
BSM-R-ROOT
A1-4
P8
A12-1
P2
P6
A10-8
P5
A2-7
P11
A7-9
A11-4
P7
P4
P3
A11-3
P9
P1
A14-2
A10-2
P10
A12-2
A7-9
A11-4
100
90
80
70
60
50
40
30
Similarity (%)
(B)
I
II
III
BPD-MS
BP-N-STEM
BP-M-ROOT
BPD-V-ROOT
BP-V-ROOT
BPT-V-LEAF
BPT-MS
BPT-N-LEAF
BPT-N-STEM
BPT-N-ROOT
BPT-V-ROOT
BPD-R-ROOT
BPT-R-GRAIN
BPT-R-LEAF
BPT-V-STEM
BPT-R-STEM
BSM-MS
BPD-N-ROOT
BPD-N-LEAF
BPD-N-STEM
BSM-N-ROOT
BPD-R-LEAF
BWB-R-STEM
BPT-R-ROOT
BPD-V-LEAF
BWB-R-LEAF
BSM-R-LEAF
BSM-V-LEAF
BSM-N-STEM
BSM-V-STEM
BSM-R-ROOT
BSM-V-ROOT
BSM-N-LEAF
BSM-R-GRAIN
BSM-R-STEM
BPD-V-STEM
BP-MS
BWB-N-ROOT
BWB-MS
BWB-N-STEM
BWB-V-ROOT
BWB-N-LEAF
BPD-R-GRAIN
BPD-R-STEM
BP-N-LEAF
BP-V-LEAF
BP-V-STEM
BP-R-ROOT
BP-R-LEAF
BP-R-STEM
BP-R-GRAIN
BWB-V-LEAF
BWB-R-GRAIN
BWB-V-STEM
BWB-R-ROOT
24-1
37-8
17-3
31-4
31-6
24-4
33-6
17-5
26-1
33-3
37-1
3-1
33-5
3-2
26-5
100
80
60
40
20
0
Similarity (%)

Fig. 2

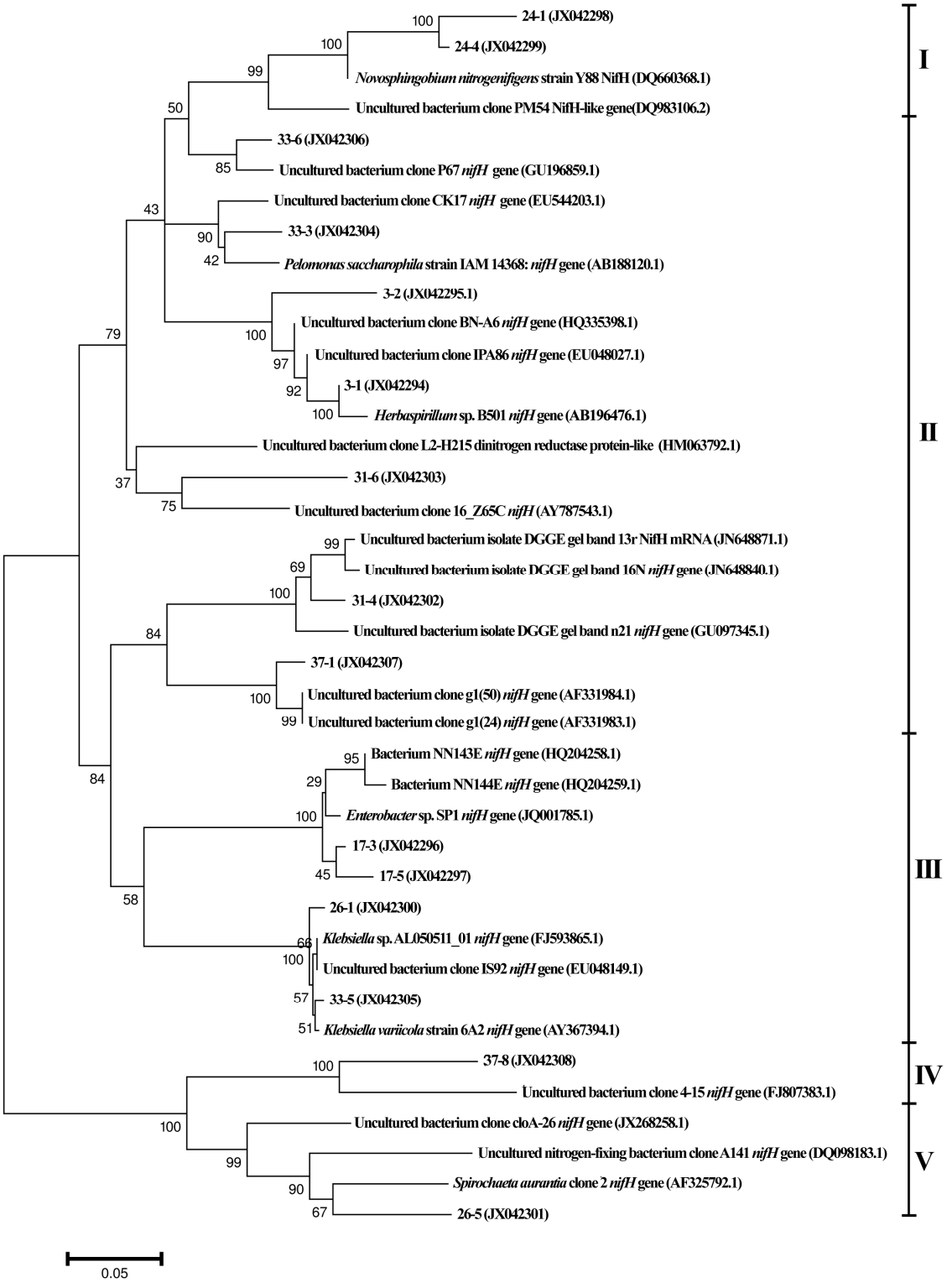

Rice cultivars
16S rRNA gene
nifH gene
Bue Polo
Bue Pra Taw
Bue Pra Do
Bue Saw Mi
Bue Wah Bo
F1 (59.39 %)
F2 (8.40 %)
F1 (45.66 %)
F2 (18.98 %)
F1 (49.98 %)
F2 (14.33 %)
F1 (65.26 %)
F2 (15.54 %)
F1 (54.50 %)
F2 (14.11 %)
F1 (32.29 %)
F2 (23.44 %)
F1 (55.98 %)
F2 (9.94 %)
F1 (57.35 %)
F2 (14.76 %)
F1 (49.24 %)
F2 (13.76 %)
F1 (36.40 %)
F2 (21.56 %)

62 **Fig. 3**


P1 (HE860534)
P8 (HE860541)
P4 (860537)
P5 (HE860538)
P9 (HE860542)
P6 (HE860539)
Uncultured bacterium clone M6a (AJ851126.1)
Uncultured bacterium clone M3a (AJ851120.1)
Uncultured bacterium isolate DGGE band C-3 (EF669488.1)
Uncultured bacterium isolate DGGE gel band DS31-5 (EU585924.1)
A1-4 (HE860545)
A11-3 (HE860551)
Uncultured bacterium isolate DGGE gel band he-49 (FJ235679.1)
P2 (HE860535)
Uncultured bacterium isolate DGGE gel band DS30-2 (EU585922.1)
P7 (HE860540)
P10 (HE860543)
A14-2 (HE860555)
Uncultured bacterium clone ncd2380h03c1 (JF207567.1)
Uncultured bacterium clone ncd595b06c1 (HM284178.1)
P11 (HE860544)
A2-7 (HE860547)
A12-2 (HE860554)
A12-1 (HE860553)
Uncultured bacterium isolate DGGE gel band DS34-3 (EU585925.1)
Uncultured bacterium clone BIGO881 (HM558846.1)
Uncultured bacterium isolate DGGE gel band 7 (HQ876068.1)
A10-8 (HE860550)
P3 (HE860536.1)
Pseudomonas frederiksbergensis strain B7 (JX122166.1)
Uncultured *Pseudomonas* sp. clone 19318_ZC5M6_G12 (JQ897449.1)
A7-9 (HE860548.1)
Pseudomonas syringae strain NBRC 3310 (AB680056.1)
Pseudomonas viridiflava isolate O1 1110 (HE588020.1)
Pseudomonas savastanoi strain AKB26 (HQ220027.1)
Uncultured bacterium clone F46_11D (JQ512881.1)
A10-2 (HE860549)
Uncultured bacterium clone nby488c06c1 HM831057
Uncultured bacterium clone ncd491h09c1 (HM316066)
Klebsiella oxytoca strain sch19 (JX294876.1)
A11-4 (HE860552.1)
Pantoea sp. A1128 (JX266309.1)
Uncultured bacterium clone RS_C07 (JQ357091.1)
Uncultured bacterium clone 54 (EU887666.1)
0.05

63

64 **Fig. 4**


24-1 (JX042298)
24-4 (JX042299)
Novosphingobium nitrogenifigens strain Y88 NifH (DQ660368.1)
Uncultured bacterium clone PM54 NifH-like gene(DQ983106.2)
33-6 (JX042306)
Uncultured bacterium clone P67 nifH gene (GU196859.1)
Uncultured bacterium clone CK17 nifH gene (EU544203.1)
33-3 (JX042304)
Pelomonas saccharophila strain IAM 14368: nifH gene (AB188120.1)
3-2 (JX042295.1)
Uncultured bacterium clone BN-A6 nifH gene (HQ335398.1)
Uncultured bacterium clone IPA86 nifH gene (EU048027.1)
3-1 (JX042294)
Herbaspirillum sp. B501 nifH gene (AB196476.1)
Uncultured bacterium clone L2-H215 dinitrogen reductase protein-like (HM063792.1)
31-6 (JX042303)
Uncultured bacterium clone 16_Z65C nifH (AY787543.1)
Uncultured bacterium isolate DGGE gel band 13r NifH mRNA (JN648871.1)
Uncultured bacterium isolate DGGE gel band 16N nifH gene (JN648840.1)
31-4 (JX042302)
Uncultured bacterium isolate DGGE gel band n21 nifH gene (GU097345.1)
37-1 (JX042307)
Uncultured bacterium clone g1(50) nifH gene (AF331984.1)
Uncultured bacterium clone g1(24) nifH gene (AF331983.1)
Bacterium NN143E nifH gene (HQ204258.1)
Bacterium NN144E nifH gene (HQ204259.1)
Enterobacter sp. SP1 nifH gene (JQ001785.1)
17-3 (JX042296)
17-5 (JX042297)
26-1 (JX042300)
Klebsiella sp. AL050511_01 nifH gene (FJ593865.1)
Uncultured bacterium clone IS92 nifH gene (EU048149.1)
33-5 (JX042305)
Klebsiella variicola strain 6A2 nifH gene (AY367394.1)
37-8 (JX042308)
Uncultured bacterium clone 4-15 nifH gene (FJ807383.1)
Uncultured bacterium clone cloA-26 nifH gene (JX268258.1)
Uncultured nitrogen-fixing bacterium clone A141 nifH gene (DQ098183.1)
Spirochaeta aurantia clone 2 nifH gene (AF325792.1)
26-5 (JX042301)
I
II
III
IV
V
0.05

65