Soil Biology & Biochemistry 78 (2014) 128–137

Contents lists available at ScienceDirect

Soil Biology & Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Diversity and distribution of rhizobia nodulated with *Phaseolus vulgaris* in two ecoregions of China

Ying Cao [a], En-Tao Wang [b], Liang Zhao [a], Wei-Min Chen [a], Ge-Hong Wei [a, *]

[a] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, College of Life Sciences, Northwest A&F University, 712100 Yangling, Shaanxi, China
[b] Departamento de Microbiología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, 11340 México, D.F., Mexico

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form
30 July 2014
Accepted 31 July 2014
Available online 12 August 2014

Keywords:
Rhizobia
Common bean
Phylogeny
Biogeography
Evolution
Alpha diversity
Systematic

ABSTRACT

To reveal the diversity and geographic distribution of bean (*Phaseolus vulgaris* L.) rhizobia in China, a total of 226 strains of nodule bacteria were isolated from bean plants grown in fields of nine sites in northern (ecoregion I) and southern (ecoregion II) regions of China, with soil pH ranged from 8.06 to 4.97. Based upon the phylogenies of housekeeping and symbiotic genes, 96.9% of the isolates were *Rhizobium* corresponding to four defined species and four unnamed genospecies, while the remaining were five *Bradyrhizobium* genospecies. *Rhizobium leguminosarum*, *Rhizobium etli* and *Rhizobium* sp. II or IV were dominant in different ecoregions, with varied relative abundances. Therefore, diverse and distinctive bean rhizobial communities exist in different ecoregions of China, demonstrating that the introduction of bean plants has driven the evolution of rhizobia to fit the necessity to nodulate the host plant under the local conditions, mainly the soil pH and the nutrient availability. The predominance of *R. etli* and the high similarities in the symbiotic genes (*nodC* and *nifH*) of all the *Rhizobium* genospecies with *R. etli* evidenced that symbiotic genes were transferred from *R. etli*, possibly introduced to China accompanying the seeds, to the related indigenous rhizobia. These findings enlarged the diversity of bean rhizobia, evidenced the biogeographic patterns of these rhizobia, and demonstrated the possible evolution and emerging of novel rhizobia under the combined selection of host plant and soil conditions.

1. Introduction

Bean or common bean (*Phaseolus vulgaris* L.), an annual leguminous herb, is originated in the Americas (van Schoonhoven and Voysest, 1991). For its high concentration of protein, fiber, and complex carbohydrates, it is cultivated worldwide (Kaplan, 1981) and plays an important role in feeding human being. Although bean is considered to have poor nitrogen fixing capacity (Mnasri et al., 2007), it can acquire their nitrogen through an association with rhizobia. Meanwhile, the bean is a relatively promiscuous host, nodulating with *Rhizobium etli*, *Rhizobium tropici*, *Rhizobium leguminosarum* bv. phaseoli, *Rhizobium gallicum*, *Rhizobium giardinii*, *Rhizobium lusitanum*, *Rhizobium phaseoli*, *Rhizobium azibense* (Mnasri et al., 2014), *Rhizobium freirei* (Dall'Agnol et al., 2013), *Rhizobium mesoamericanum* (López-López et al., 2012), *Sinorhizobium meliloti* (Zurdo-Piñeiro et al., 2009), *Sinorhizobium americanum* (Mnasri et al., 2012), and *Bradyrhizobium* sp. (Han et al., 2005) in different regions: predominated by *R. etli* in the South and Middle Americas (Amarger, 2001), Europe (García-Fraile et al., 2010) and Jordan (Tamimi and Young, 2004); by *R. leguminosarum* in the Andean region and Nepal (Bernal and Graham, 2001; Adhikari et al., 2013; Ribeiro et al., 2013); by *R. tropici* in regions with acid soils and high temperature (Martínez-Romero et al., 1991; Anyango et al., 1995; Grange and Hungria, 2004); and by *R. phaseoli*, *R. etli* and a novel *Rhizobium* group in Africa (Aserse et al., 2012b).

The great diversity of bean rhizobia might be a result of the worldwide cultivation of this plant, since microbial distribution is influenced by environmental variation with time and space (Martiny et al., 2006). In previous studies, soil pH has been revealed as a major factor to affect legume–rhizobia symbiotic process and nitrogen fixation (Wolff et al., 1991); while the temperature may have affected the evolution of *Bradyrhizobium* communities (Stepkowski et al., 2012). In China, biogeographic patterns have been observed in the soybean rhizobia which were correlated to the soil pH and salinity (Han et al., 2009; Li et al., 2011). Furthermore, several strains belonged to *R. leguminosarum* bv. phaseoli and *Bradyrhizobium* sp. have been reported as microsymbionts of bean

* Corresponding author. Fax: +86 2987091175.
E-mail address: weigehong@nwsuaf.edu.cn (G.-H. Wei).

http://dx.doi.org/10.1016/j.soilbio.2014.07.026
0038-0717/

in different regions of China, depending on the soil pH (Han et al., 2005). It has been known that the extreme ecological conditions could drive the rhizobia to develop active mechanisms to adapt the environments, such as accumulating intracellular compatible solutes (glutamate and trehalose) for resistance of high salinity (Faghire et al., 2012).

Bean was directly introduced from the Americas to China in the 15th century (van Schoonhoven and Voysest, 1991). To date, production of bean in China is the highest among the Asian countries, up to 15,702,000 tons over 2011 (http://faostat.fao.org/). However, the diversity of bean rhizobia in China has not been fully studied yet. Considering the long history and a vast area of bean cultivation, as well as its promiscuous feature for microsymbionts, diverse rhizobia and varied community compositions of bean rhizobia might be expected in China. The study on diversity and distribution of bean rhizobia collected from different geographical regions might improve our knowledge about the rhizobial diversity, evolution and biogeography. Thus, the goals of this study were (i) to identify the diversity of bean rhizobia isolated from different ecoregions; and (ii) to assess the geographic distribution of rhizobial species and its determinants.

2. Materials and methods

2.1. Bacterial strains and sampling sites

In this study, root nodules were collected from bean plants grown at 9 different sites located in northern and southern regions of China (Table S1). Nodules were sterilized by immerging in 95% ethanol and 0.2% $HgCl_2$ following by rinsing six times with distilled water as described previously (Kamicker and Brill, 1986). The surface-sterilized nodules were crushed and streaked on the yeast–mannitol agar (YMA) (Vincent, 1970) and incubated at 28 °C for three days to two weeks, until the single colonies occurred. Single colonies were picked up and repeatedly streaked on the same medium until the colony morphology was homogeneous. The purified isolates were suspended in the YM broth supplemented with 20% (w/v) of glycerol and stored at −70 °C.

Soil samples were collected at the same sites from the surface layer (0–10 cm in depth) together with the bean root system. Physicochemical properties of the soils, including pH and contents of nitrogen, phosphorus and potassium were measured with the routine methods (Rayment and Higginson, 1992).

2.2. DNA extraction and sequencing of multilocus genes

For extraction of DNA, the isolates were cultured in TY broth (Tryptone 5 g; Yeast extract 3 g; $CaCl_2$ 0.6 g; Distilled water 1 L; pH 7.0) with agitation at 28 °C up to the early stationary stage and cells were collected by centrifugation as reported elsewhere (Cava et al., 1989). Genomic DNA of each isolate was prepared using the CTAB method (Wilson, 2001) as template for amplifying ribosomal RNA gene and symbiotic genes. The 16S rRNA gene was amplified with primer P1 and P6 (Chen et al., 1997) and the related PCR protocol (Chun and Goodfellow, 1995). Gene *nodC* was amplified using the primer pairs *nodC*F/*nodC*I and *nodC*Fu/*nodC*I for *Rhizobium* and *Bradyrhizobium*, respectively (Laguerre et al., 2001). Gene *nifH* was amplified using the primer pair *nifH*F/*nifH*I (Laguerre et al., 2001). The PCR products digested with three restriction endonucleases *Msp*I, *Hha*I and *Hae*III separately at 37 °C for 4 h which were resolved by electrophoresis in 2% (w/v) agarose gel (Amreso) in 0.5 × TBE (Tris–borate–EDTA) buffer. Then the gels were stained with ethidium bromide (0.5 μg/ml) and the RFLP patterns were scanned with Bio-red Gel Imager. The RFLP patterns for these authenticated genes of rhizobia were defined.

According to the result of RFLP analyses of 16S rRNA, *nodC* and *nifH* genes, the isolates were designed into genotypes. Representative isolates of different genotypes were selected for the subsequent multilocus sequence analysis. The 16S rRNA, *nifH* and *nodC* genes were amplified again and three housekeeping genes *atpD* (ATP synthase β-subunit), *glnII* (glutamine synthase II) and *recA* (recombinase A protein) were amplified by using primer pairs *atpD*255F/*atpD*782R, *glnII*12F/*glnII*689R and *recA*41F/*recA*640R, and the corresponding protocols (Turner and Young, 2000; Vinuesa et al., 2005; Han et al., 2008). The amplified genes were sequenced directly with the method of Sanger (Sanger et al., 1977) in Beijing AuGCT DNA-SYN Biotechnology Co., Ltd.

2.3. Phylogenetic analysis

The acquired nucleotide sequences were applied for BLAST (Altschul et al., 1990) to obtain the closely related sequences in the GenBank database. The sequences obtained in this study together with the related sequences from the database were aligned by using the program ClustalW in the package of MEGA version 5 (Tamura et al., 2011). Sensitivity and quality analyses of data were firstly involved in assessing phylogeny (Grant and Kluge, 2003). For detecting the quality of data, the index of substitution saturation (I_{ss} and $I_{ss.c}$) was measured with DAMBE (Xia, 2009, 2013). Combined sequences of housekeeping genes were subject to define the MLST (Multilocus sequences type). Meanwhile, it was necessary to analyze the sensitivity of concatenated sequences. To evaluate whether the data sets containing different evolutionary histories are incongruent, incongruently length difference (ILD) test (Farris et al., 1994, 1995) was performed with the program of PAUP*4.0b10. ILD analyses of combined housekeeping genes sequences were conducted with 1000 replicates of random addition heuristic search with TBR branch swapping (Yoder et al., 2001).

Phylogenetic trees were constructed by using the MEGA version 5 (Tamura et al., 2011) with the neighbor joining (NJ) method. The robust of the NJ tree was evaluated by bootstrap analyses with 1000 replications (Sy et al., 2001). Genospecies were defined according to the MLSA relationships using 97% of sequence similarity as the threshold as suggested previously (López-Guerrero et al., 2012; Ribeiro et al., 2013).

2.4. Statistical analysis and community diversity estimation

Sequence-based markers are used to evaluate species abundance and biogeography by non-parametric multivariate methods (Li et al., 2012). CCA (Canonical Correspondence Analysis) (ter Braak, 1986), as a constrained ordination method, was utilized to understand the relationship between environmental variables of sampling sites and genospecies abundance defined based on the results of MLSA. Before the CCA, a linear or unimodal ordination model was determined by DCA (Detrended Canonical Analysis) (Lepš and Šmilauer, 2003). The maximal value of the lengths of the gradient in four ordination axes was below 3, suggesting that the linear gradient analysis model was more suitable, but the unimodal also could be used. The analysis of species abundance and environmental data was performed by CANOCO 5 (Microcomputer Power, Ithaca, NY).

The community diversity of rhizobia in each of the ecoregions were estimated by the Shannon–Weaver index (H) and species richness (d), calculated with the following formula: $H = (N \log N - \Sigma\, n_i \log n_i)C/N$ and $d = (S - 1)/\log N$. Here N is the number of total isolates in the ecoregion, n_i presents the number of isolates in genospecies I; C is a constant of 2.3; S is the number of genospecies detected in the ecoregion.