Diverse Genomic Backgrounds Vs. Highly Conserved Symbiotic Genes in *Sesbania*-Nodulating Bacteria: Shaping of the Rhizobial Community by Host and Soil Properties

Zhenpeng Zhang¹,² · Wei Liu¹,³ · Shuai Shao¹,⁴ · En-Tao Wang⁵ · Yan Li¹,³

Received: 5 November 2019 / Accepted: 19 January 2020 / Published online: 29 January 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Aiming at investigating the overall diversity, biogeography, and symbiosis gene evolutionary history of the *Sesbania cannabina*-nodulating rhizobia in China, a total of 874 rhizobial isolates originating from the root nodules of this plant grown at different sites were characterized and compared with those of some reference strains. All of the *S. cannabina*-nodulating rhizobia were classified into 16 (geno) species, including seven novel genospecies in the genera *Ensifer*, *Rhizobium*, *Neorhizobium*, and *Agrobacterium*, with *Ensifer sesbaniae* and *Neorhizobium huautlense* as the dominant and universal species. Ten of these species were found to nodulate other leguminous hosts or to lack nodulating abilities and were defined as symbiovar sesbania. Biogeographic patterns were observed, for which pH, TN, AK, and AP were the main determinants. The effects of pH were opposite to those of TN and AK, while AP presented effects independently of TN, AK, and pH. Symbiotic genes of these rhizobia showed a common origin, but *nodA* evolved faster than *nifH*. Point mutation is the main driving force in the evolution of both *nodA* and *nifH*, and lateral transfer of symbiotic genes might play an important role in the formation of diverse *S. cannabina*-nodulating rhizobial species. *S. cannabina* only nodulates with *Sesbania* rhizobia, demonstrating its severe selection on rhizobial symbiosis genes. Soil pH and physiochemical characteristics could affect rhizobial survival and competitive nodulation. This study provides insight into the community shifts and evolution of rhizobia in relation to their host and soil environments.

Keywords *Sesbania cannabina* · Rhizobia · Symbiotic genes · Horizontal transfer · Soil properties · Adaptation

Introduction
Rhizobia or root nodule bacteria are soil bacteria that induce the formation of root and/or stem nodules on their leguminous hosts [1]. In nodules, rhizobial cells differentiate into bacteroids and reduce the atmospheric nitrogen into ammonia available to the host as a nutrient [2]. This symbiotic nitrogen fixation (SNF) system contributes approximately 40 million tons of fixed nitrogen to agricultural systems per year [3]. Relating to their complicated life cycle, the genome of rhizobia is composed of two gene categories: (1) chromosomal genes for survival and adaptation to environments and (2) symbiotic genes relating to nodulation (*nod/noe/nol*) and nitrogen fixation (*nif/fix/fdx*). The symbiotic genes are usually located on mobile plasmids or genomic islands that are transferable between different bacteria [4].

Belonging to the tribe *Sesbanieae*, the leguminous genus *Sesbania* includes more than 70 species distributed throughout tropical and subtropical areas [5], in which nodulation has been reported for approximately 40 species [6]. *Sesbania*
cannabina is an annual semishrub native to the South Pacific Islands and has been introduced to other areas, including China (Malcolm, 2012; http://www.iucnredlist.org/details/summary/168726/0), as green manure to improve saline-alkaline soils and to remove heavy metals from contaminated soils [7, 8]. Previous studies indicated that S. cannabina nodulated with rhizobia belonging to Ensifer sesbaniae, Ensifer terangae bv. sesbaniae, Ensifer saheli, Agrobacterium pusense, and Neorhizobium huautlense in areas of Africa and Asia [9–12]. This plant is widely spread in acidic soils in the southern part of China, and it has been introduced to northern China for the remediation of saline-alkaline soils. Recently, we revealed diverse species with conserved symbiotic genes in the rhizobia nodulating S. cannabina in alkali soils in China [13]. However, the rhizobial community of S. cannabina in Chinese acidic soils has not been studied and may differ from that in saline-alkaline soils, as is the case for soybean rhizobia [14, 15].

Considering that different types of soil might contain distinct microbial populations [14, 15], the characterization of rhizobial populations in soils with different properties might help us to learn how the interaction among rhizobia, their host legumes, and the environment drives the distribution and evolution of rhizobia species. Therefore, we performed the present study to isolate and characterize the rhizobia associated with S. cannabina in soils with different properties (1) to investigate the genetic diversity and symbiotic genotypes; (2) to analyze the relationship between rhizobial richness, rhizobial distribution, and soil characteristics; and (3) to assess the evolutionary history of symbiotic genes. Our results identified additional S. cannabina-nodulating rhizobia, revealed clear biogeographic patterns, and evidenced the common origin and evolutionary history of symbiotic genes.

**Experimental Procedures**

**Soil and Nodule Sampling and Rhizobial Isolation**

S. cannabina root nodules were collected from 20 sites in China in September 2013 and from May to July 2018 (Table 1). At each site, 5–10 randomly selected S. cannabina plants were uprooted, and root nodules were carefully cut off and placed into a sterilized tube filled with silica gel particles for transportation to the laboratory for subsequent bacterial isolation. Soil samples simultaneously collected from the root zone (0–20-cm depth) of each plant were mixed and air-dried to form a composite sample for the site, which was used for physicochemical characterization in the laboratory. For rhizobial isolation, the rehydrated nodules (in sterilized 0.85% NaCl solution at 4 °C for 6 h) were surface sterilized, washed, crushed into juice, and inoculated by streaking on yeast mannitol agar (YMA) plates as described previously [16]. The inoculated plates were incubated at 28 °C for 3–7 days, and a single colony was randomly selected from each plate and purified by repeated streaking on YMA plates. The purified isolates were preserved at −80 °C in 20% (v/v) glycerol. For soil characterization, 2-mm mesh-sieved soil samples were used to determine the chemical properties, including soil pH [17], organic carbon (OC) [18], total nitrogen (TN) [19], available nitrogen (AN) [20], available phosphorus (AP) [18], and available potassium (AK) [20], using the corresponding standard protocols.

**PCR Amplification and Sequencing**

The genomic DNA of each isolate extracted with the TIANGEN genomic DNA extraction kit for bacteria (TIANGEN, China) was used as template to amplify the recA (DNA recombination and repair protein) and nifH (nitrogenase iron protein) genes by using the primer pairs recA41F/640R [21] and SnifHF/R (3′-CGATCATCTCGCCGCRAAGGA-5′ and 3′-GATAATCCGGGCTGCTTG-5′), respectively, in which SnifHF/R were specifically designed for Sesbania rhizobia in this study. All the obtained sequences were aligned using Clustal W integrated in MEGA 7.0 [22]. The recA and nifH haplotypes were then defined by using DnaSP v5 [23], and a representative strain was randomly selected from each haplotype. For the representative strains, the housekeeping genes atpD (ATP synthase subunit beta) and glnII (glutamine synthetase II protein) were amplified using the corresponding primers and protocols [24], respectively, while a symbiosis gene nodA (acyltransferase) was amplified by using the specific primers designed for sesbania rhizobia in this study: SnodAF (3′-CGATCACCGGAAACTTCTGCG-5′) and SnodAR (3′-GATGCCGGGAAAGCAYTCGTGC-5′). The PCR product was commercially sequenced at Beijing AuGCT DNA-SYN Biotechnology Co., Ltd. by using the method of Sanger [25], and their quality was checked by using BioEdit software [26]. All of the obtained sequences were deposited in the GenBank database (Table S7).

**Phylogenetic and Statistical Analyses**

All the obtained sequences were aligned by BLAST in GenBank, and the corresponding reference sequences were downloaded. Multilocus sequence analysis (MLSA) was performed by combining the recA-atpD and recA-atpD-glnII sequences separately to determine the phylogenetic relationships for all the representative strains. Phylogenetic trees were reconstructed with Kimura’s two-parameter model and the neighbor-joining method in MEGA 7.0 with 1000 bootstrap replicates. Genospecies were defined with a threshold of 97.3% [27]. The rhizobial diversity, richness, and evenness in the sampling sites were evaluated by the Shannon-Wiener
index ($H$), Simpson index ($D$), and Pielou index ($J$) by using the Vegan package (version 1.17-4) in the R statistical language platform (version 3.4.0) [28]. The correlation between rhizobial distribution and soil properties was evaluated by using CANOCO software version 4.5 [29]. Due to the length (4.248) of the gradient (first axis) in the DCA of the rhizobial community analysis, CCA was selected to analyze the relationship between the soil characteristics and rhizobial genospecies.

### Table 1
Properties of soil samples and the distribution of different rhizobial genospecies at 11 sites

| Soil property | Sampling site and date (month/year, 18 = 2018; 13 = 2013) | JJ | FZ | ZZ | GZ | HZ | ZJ | BH | CQ | BN | YRD | RD |
|---------------|----------------------------------------------------------|----|----|----|----|----|----|----|----|----|-----|----|
| pH            |                                                          | 6.55 | 6.79 | 6.17 | 7.00 | 5.92 | 7.34 | 7.33 | 7.14 | 7.32 | 7.88 | 8.22 |
| ANa (mg kg$^{-1}$) |                                                      | 119.00 | 65.84 | 200.00 | 28.28 | 88.90 | 73.50 | 125.12 | 105.53 | 141.38 | 55.25 | 72.8 |
| APb (mg kg$^{-1}$) |                                                      | 0.67 | 2.40 | 20.26 | 1.03 | 3.47 | 0.63 | 6.12 | 17.01 | 21.95 | 7.35 | 15.65 |
| AKc (mg kg$^{-1}$) |                                                      | 246.60 | 196.61 | 390.80 | 195.20 | 189.07 | 251.20 | 321.40 | 274.01 | 227.02 | 150.74 | 75.31 |
| TNd (mg kg$^{-1}$) |                                                      | 1797.40 | 1483.08 | 4097.77 | 1227.57 | 1797.40 | 1201.64 | 1018.98 | 2859.06 | 735.18 | 401.43 |
| OCe (g kg$^{-1}$) |                                                      | 10.06 | 14.95 | 45.56 | 31.80 | 11.01 | 19.98 | 56.83 | 11.22 | 28.62 | 14.81 | 10.12 |

- **Number of genospecies**
- **Number of isolates in sampling site**
- **Total**

| Number of genospecies | Number of isolates in sampling site | Total |
|-----------------------|-------------------------------------|-------|
| 1                     | 119                                 | 121   |
| 2                     | 10                                 | 14    |
| 3                     | 1                                  | 1     |
| 4                     | 21                                 | 57    |
| 5                     | 1                                  | 291   |
| 6                     | 9                                  | 13    |
| 7                     | 1                                  | 1     |
| 8                     | 105                                | 105   |
| 9                     | 23                                 | 23    |
| 10                    | 8                                  | 8     |
| 11                    | 5                                  | 5     |
| 12                    | 3                                  | 3     |
| 13                    | 161                                | 161   |
| Diversity index       | Shannon-Wiener                      | 0.52  |
|                       | Simpson                            | 0.34  |
|                       | Inverse Simpson                    | 1.52  |
|                       | J                                  | 0.76  |

- **AN**, available nitrogen
- **AP**, available phosphorus
- **AK**, available potassium
- **TN**, total nitrogen
- **OC**, organic carbon

**Nucleotide Polymorphisms and Evolutionary Analyses**

DnaSP v5 [23] was selected to investigate the nucleotide polymorphisms of *nodA* and *nifH*. The number of haplotypes ($h$), haplotype diversity ($Hd$), nucleotide diversity ($\pi$), nucleotide diversity for synonymous substitutions ($\pi_s$), and nucleotide diversity for nonsynonymous substitutions ($\pi_N$) was calculated and statistically compared. The “ancestral” for *nodA* and *nifH* was estimated with the admixture LOCPRIOR model in
STRUCTURE [30]. According to this model, individual i acquired the gene from K ancestral subpopulations, and suitable K values were selected according to the program CLUMPP and STRUCTURE HARVESTER analyses [31, 32]. The split phylogenetic networks (neighbor-net tree) (1000 bootstraps) for nodA and nifH were inferred by using the SplitsTree4 program [33].

Nodulation Test

The nodulation ability of \textit{S. cannabina} was examined for all the representative strains with standard procedures [34]. For the genospecies reported previously for symbionts of other leguminous hosts, the representative strains of \textit{Sesbania} rhizobia, as well as the corresponding reference strains, were also employed for cross-nodulation tests with the corresponding hosts. Surface-sterilized and pregerminated seeds (on 0.6% agar, 48 h) were transferred (in five replicates) to a Leonard jar (1 seed per jar) filled with sterilized vermiculite moistened with low-nitrogen nutrient solution. An aliquot of 1-mI rhizobial culture (approximately \(10^8\)) cells was inoculated into each seedling, and the plants were cultured in an automated greenhouse with day/night cycles of 12 h at 24 °C. Plants were harvested at 35 days postinoculation (dpi), and plants with green leaves and the presence of pink nodules were deemed effective nitrogen fixation hosts.

Survival and Competitive Nodulation in Soils

To compare the effects of soil pH on rhizobial distribution and nodule occupancy, alkaline soil collected from the Yellow River Delta (pH 8.0) and acid soil (pH 6.0) collected from Huizhou were separately autoclaved twice at 121 °C for 2 h. The effectiveness of sterilization was confirmed by the absence of bacterial growth on YMA plates inoculated with the samples. Acidic soil collected from Huizhou, and alkaline soil collected from the Yellow River Delta [13]. According to the recA and nifH sequence analyses, the isolates were classified into 27 recA haplotypes (Fig. 1 and Table S8) and 10 nifH haplotypes (Supplementary Fig. S6 and Table S8). Through MLSA (Figs. S2 and S3, Table S1), the representatives of the 27 recA haplotypes were divided into 13 genospecies within the genera \textit{Ensifer} (5 genospecies), \textit{Agrobacterium} (1 genospecies), \textit{Rhizobium} (5 genospecies), and \textit{Neorhizobium} (2 genospecies), in which \textit{E. sesbaniae} (291 isolates), \textit{N. huautlense} (235 isolates), \textit{E. alkalisoli} (121 isolates), and \textit{Rhizobium} sp. II (105 isolates) were the dominant groups (> 10%) accounting for 33.30%, 26.89%, 13.84%, and 12.01% of the total isolates, respectively (Fig. 1 and Table 1). \textit{N. huautlense} and \textit{E. sesbaniae} strains have wide distributions, recorded in 9 and 8 of the 11 sampling sites, respectively. Among the 11 sites, the highest \(H'\) was found in Beihai (\(H' = 1.28\), 5 genospecies), and the lowest was found in Huizhou (\(H' = 0.08\), 2 genospecies) (Table 1).

Correlation Between Distribution of Rhizobia and Soil Characteristics

As shown in Fig. 2, TN, AP, AK, and pH were the main factors affecting the distribution of \textit{Sesbania} rhizobia in this study. According to the lengths and angles among the fleches (factors), TN and AK have similar effects on the rhizobial genospecies, e.g., they are positively correlated with \textit{Ensifer} sp. I, \textit{Neorhizobium} sp. I, \textit{Rhizobium} sp. IV, \textit{E. sesbaniae}, and...
Rhizobium sp. I but negatively correlated with *R. binae*, *Rhizobium* sp. III, *Rhizobium* sp. II, *E. melioti*, *E. alkalisoli*, *A. pusense*, and *Ensifer* sp. II. The effects of pH were opposite to those of TN and AK. AP presented effects independently of TN, AK, and pH. AN and OC showed minor correlations with the distribution of rhizobia genospecies.

**Symbiotic Genes of Sesbania Rhizobia**

All the isolates obtained in this study were selected to amplify *nifH*. Most of the isolates amplified successfully, except for 59 isolates in 5 genospecies (30 in *E. sesbaniae*, 4 in *A. pusense*, 4 in *R. binae*, 10 in *Rhizobium* sp. II, and 11 in *E. alkalisoli*) (Table S6) that were failed in amplification. The 59 isolates failed to amplify *nifH*, and 38 representative strains for different *recA* and *nifH* haplotypes were used in the amplification of *nodA*. As a result, *nodA* amplification was successful for all the representative strains and for the *nifH*-absent isolates *E. sesbaniae* BH005 and BH012 (Fig. S6). In the phylogenetic analysis, 15 *nodA* haplotypes and 12 *nifH* haplotypes were detected among representative strains of *S. cannabina*-nodulating rhizobia across the world (Fig. S7). The symbiotic
genes nodA and nifH of the Sesbania rhizobia formed a unique clade in the corresponding phylogenetic trees (Fig. 3 and Fig. S6). Except for five strains (YIC4313, HZ044, YIC4240, YIC4261, and BH010), most of the strains belonging to the same nodA haplotype were also classified in the same nifH haplotype (Fig. S7).

At the genus level, 12 nodA and 7 nifH haplotypes in the Ensifer strains were defined, while three nodA and two nifH haplotypes were recorded in the Neorhizobium strains (Fig. S7). Most of the representative strains for Rhizobium (geno) species were classified into single nodA and nifH haplotypes, except Rhizobium sp. II HZ044. The strains in A. pusense presented a unique nodA and nifH haplotype. At the (geno) species level, Ensifer sp. II presented 4 nodA and 4 nifH haplotypes, respectively (Fig. S7), followed by E. sesbaniae with 4 nodA and 3 nifH haplotypes, respectively (Table S2 and Fig. S7). The nifH haplotypes 5, 2, 1, 1, 3, and 6 were dominant in E. alkalisoli (accounting for 89.92%), E. meliloti (87.5%), Ensifer sp. II (52.63%), E. sesbaniae (63.57%), N. huautlense (94.04%), and Rhizobium sp. II (55.24%), respectively (Tables S2 and S5).

**Molecular Analyses of Symbiotic Genes**

In STRUCTURE analyses, K values of 4 and 3 were chosen for nodA and nifH, respectively (Fig. S8), which were congruent with the phylogenetic trees (Fig. S7) and the neighbor-net network trees (Fig. S9). The Ensifer species were grouped into 3 nodA and 3 nifH clusters; the Neorhizobium species were found in two nodA clusters and one nifH cluster, while both the Rhizobium and Agrobacterium strains were grouped into single nodA and nifH clusters, respectively (Figs. S7, S8, and S9). Thus, the symbiotic genes of Ensifer species were more diversified, which was consistent with the haplotype analysis results.

In STRUCTURE and neighbor-net analyses (Figs. S8 and S9), nodA and nifH of N. huautlense FZ054 and CQ173 experienced the most apparent recombination events because they had two clear original ancestors and showed gene exchange with E. sesbaniae GZ015 and Ensifer sp. II GZ008 and YIC4261 (Figs. S8 and S9). Furthermore, nodA showed a higher nucleotide polymorphism than nifH, since the segregating sites h/Hd and π values of nodA were higher than those of nifH (Table S4). The minimum recombination (Rn) of nodA was also higher than that of nifH. The Rn values for nodA lie outside the 95% interval (upper limit) of values obtained with coalescence simulations with an intermediate level of recombination, indicating that many recombination events have occurred. However, point mutation contributes more than recombination to the evolutionary history of both genes because both the r/m and ρ/θ values were less than 1.0. A synonymous mutation was the main mutation source because πS > πN and πN/πS < 1 (Table S4).

**Cross-Nodulation Test**

All 38 representative strains with both nodA and nifH formed effective nodules on S. cannabina. Strains E. sesbaniae BH005 and BH012 with nodA but without nifH formed white nodules with a normal shape on S. cannabina. The reference strains originating from hosts other than S. cannabina failed to nodulate S. cannabina, and the Sesbania rhizobia did not nodulate with other legumes in the cross-nodulation tests (Table S3).

**Survival of Rhizobia in Acidic and Alkaline Soils**

In acidic soil, strains isolated from alkaline soil, such as E. alkalisoli YIC4027T and Ensifer sp. II YIC4108, disappeared after incubation for 6 and 12 days, respectively, while the strains E. meliloti YIC4071 from alkaline soil and E. psoraleae 65,632 from neutral soil disappeared after incubation for 21 days. N. huautlense CQ179 and Neorhizobium sp. I BN813 could not be isolated after incubation for 27 and 45 days, respectively. Strains Rhizobium sp. IV FZ027 and E. sesbaniae FZ001 from acidic soil and strains Rhizobium sp. II BH001 and R. binae BH003 from neutral soil, as well as the strain A. pusense YIC4072 from alkaline soil, were maintained at a relatively high level of 3.85–5.07 (log CFU g⁻¹) in acidic soil after incubation for 45 days.

In alkaline soil (Fig. 4b), all the selected strains showed a relatively high survival level after incubation for 45 days. A. pusense YIC4072 (from alkaline soil) was the most abundant, and Neorhizobium sp. I BN813 (from neutral soil) presented the lowest CFUs. Strains N. huautlense CQ179 and E. psoraleae 65,632 from neutral soil also showed lower...
showed excellent competitive abilities (54–85%) over *Ensifer* sp. II 4108, *E. meliloti* YIC4071, *Rhizobium* sp. IV FZ027, *R. binae* BH007, and *A. psusense* YIC4072 under all three pH conditions (Fig. 5a, b, and c), except that *E. meliloti* YIC4071 was more competitive (56%) than *N. huautlense* CQ179 (44%) in alkaline soil (Fig. 5b).

The intraspecies competitive ability ranked *E. alkalisoli* YIC4006 > YIC4027 > YIC4240; *E. meliloti* YIC4071 > YIC5069; *N. huautlense* CQ179 > CQ173; and *Rhizobium* sp. II HZ044 > BH001 (Fig. 6). These patterns were consistent with their isolation frequencies (Table S5). However, strains *Ensifer* sp. II GZ008 and *Ensifer* sp. II YIC4313 showed better competitive abilities than *Ensifer* sp. II YIC4108; *E. sesbaniae* GZ015 and GZ030 were more competitive than FZ001 (Fig. 6).

**(Phylogenetic tree of nodA sequences)**

The intraspecies competitive ability ranked *E. alkalisoli* YIC4006 > YIC4027 > YIC4240; *E. meliloti* YIC4071 > YIC5069; *N. huautlense* CQ179 > CQ173; and *Rhizobium* sp. II HZ044 > BH001 (Fig. 6). These patterns were consistent with their isolation frequencies (Table S5). However, strains *Ensifer* sp. II GZ008 and *Ensifer* sp. II YIC4313 showed better competitive abilities than *Ensifer* sp. II YIC4108; *E. sesbaniae* GZ015 and GZ030 were more competitive than FZ001 (Fig. 6).

**(Phylogenetic tree of nodA sequences)**

The neighbor-joining method. The tree was reconstructed by the neighbor-joining method. The tree was reconstructed by the neighbor-joining method.

**Competitive Inter- and Intraspecies Nodulation**

According to the nodule occupancy, the competitive abilities were as follows: *N. huautlense* CQ179 > *E. sesbaniae* FZ001 > *Rhizobium* sp. II BH001 > *E. alkalisoli* 4027 in acidic soil (Fig. 5a), *E. alkalisoli* YIC4027 > *E. sesbaniae* FZ001 > *N. huautlense* CQ179 > *Rhizobium* sp. II BH001 > *E. alkalisoli* YIC4027 in vermiculite (neutral conditions) (Fig. 5c). Against the representative strains of less abundant genospecies, *N. huautlense* CQ179 showed excellent competitive abilities (54–85%) over *Ensifer* sp. II 4108, *E. meliloti* YIC4071, *Rhizobium* sp. IV FZ027, *R. binae* BH007, and *A. psusense* YIC4072 under all three pH conditions (Fig. 5a, b, and c), except that *E. meliloti* YIC4071 was more competitive (56%) than *N. huautlense* CQ179 (44%) in alkaline soil (Fig. 5b).

The intraspecies competitive ability ranked *E. alkalisoli* YIC4006 > YIC4027 > YIC4240; *E. meliloti* YIC4071 > YIC5069; *N. huautlense* CQ179 > CQ173; and *Rhizobium* sp. II HZ044 > BH001 (Fig. 6). These patterns were consistent with their isolation frequencies (Table S5). However, strains *Ensifer* sp. II GZ008 and *Ensifer* sp. II YIC4313 showed better competitive abilities than *Ensifer* sp. II YIC4108; *E. sesbaniae* GZ015 and GZ030 were more competitive than FZ001 (Fig. 6).
These intraspecies competition results were inconsistent with their isolation frequencies in fields (Table S5).

Discussion

Diversity and Biogeographic Distributions of *S. cannabina* Rhizobia

Together with the previously described *E. saheli*, *E. terangae*, and *E. psoraleae* [9, 11] and the 13 genospecies identified in our present study (Fig. 1), a total of 16 genospecies in the genera *Agrobacterium*, *Ensifer*, *Neorhizobium*, and *Rhizobium* were defined for the *S. cannabina* rhizobia across the world, demonstrating that *S. cannabina* is a promiscuous host for rhizobia. This promiscuous symbiotic property might have helped *S. cannabina* adapt to different types of soils. Among the 16 genospecies, seven were putative novel species, and 10 were described previously for rhizobia nodulating with other legumes (*E. alkalisoli*, *E. meliloti*, *E. psoraleae*, *E. saheli*, *E. sesbaniae*, *E. terangae*, *Rhizobium* sp. I, and *R. binae*) or for nonsymbiotic bacteria (*A. pusense* and *Ensifer* sp. II) (Fig. 1).

![Fig. 4](image-url) Survival rate of representative strains incubated in (a) acidic soil (pH = 5.92) and (b) alkaline soil (pH = 8.22). The strains in dashed lines indicate that they disappeared during the period of survival. Abscissa for the cultivation days and ordinate for the logarithm of CFU in 1-g soil. Data are mean ± s.e.m. based on three replicates.

![Fig. 5](image-url) Pairwise comparisons of interspecies pairs in the competition experiment using sterilized soils (a, acidic soil, pH = 5.92; b, alkaline soil, pH = 8.22; c, vermiculite, pH = 7). Each pair of strains for the competitive experiment was mixed 1:1; data are mean ± s.e.m. based on three replicates. The genospecies of interspecies pairs are shown on the bottom right.
and Table S3). Therefore, the results of our study greatly enlarged the diversity of rhizobia and the host range for several defined rhizobial species.

Despite the diversity information, the dominance of *E. sesbaniae*, *N. huautlense*, *E. alkalisoli*, and *Rhizobium* sp. II at different sites, as well as the absence of *E. saheli*, *E. terangae*, and *E. psoraleae*, clearly evidenced the biogeographic patterns of *S. cannabina* rhizobia. Similar to results in previous studies [13, 36], environmental factors including the soil nutrient factors TN, AP, and AK and the soil physico-chemical factor pH were revealed as the main determinants of the distribution of the rhizobial genospecies associated with *S. cannabina*. In general, *E. alkalisoli* and *E. meliloti* were only spread and dominant in alkaline soils; *Rhizobium* sp. II was superdominant in acidic soil (pH 5.9 at Huizhou) but not distributed in alkaline soils at the Yellow River Delta and Rudong; *E. sesbaniae* was found in weakly acidic to weakly alkaline soils (pH 6.17–7.34) but rarely in alkaline soils at Rudong; and *N. huautlense* could be dominant in weakly acidic to alkaline soils but not in acidic soil at Huizhou (Table 1). *N. huautlense* presented the widest distribution, being spread at 9 sites and dominant (>10% total isolates) at 7 sampling sites (Table 1).

**Symbiotic Gene Conservation and Description of Symbiovar Sesbaniae**

Although great diversity at the genus and species levels was detected in *S. cannabina* rhizobia, the formation of a unique clade in the phylogenetic trees of *nodA* and *nifH* (Fig. 3 and Fig. S6) evidenced a common origin of symbiotic genes. This situation further confirmed that *S. cannabina* is a host that selects strongly the symbiosis gene background but less for the chromosomal background (species) of its symbionts [13], similar to common bean [37]. The discrepancies between the phylogenies of housekeeping genes (Fig. 1) and symbiotic genes (Fig. 3) might evidence the horizontal transfer of symbiotic genes among the *S. cannabina* rhizobia, which was also supported by the results of STRUCTURE and neighbor-net analyses (Figs. S8 and S9). Based upon these results, all strains originating from *S. cannabina* should be classified as symbiovar sesbaniae, regardless of their genus/species affiliation. The strains defined as bv. sesbaniae in *E. saheli* and *E. terangae* [38] should be renamed sv. sesbaniae.

The greater h and greater divergence in *nodA* than in *nifH* (Fig. S6, Tables S2 and S5) among the *S. cannabina* rhizobia indicated a faster evolution of *nodA* than *nifH*. Neighbor-net and recombination event analyses by DnaSP indicate that *nodA* experienced more recombination events than *nifH* (Figs. S8 and S9, Table S4). Mutation contributes more than recombination to the evolutionary history of both *nodA* and *nifH*, which is consistent with the *Glycine max*-nodulating *Ensifer* (*Sinorhizobium*) species [39]. Compared with *nif* genes, *nod* genes may suffer more selection pressure from the hosts, which has resulted in more haplotypes and higher π values in *nod* than in *nifH* (Table S4). However, the similar topologies of the *nodA* and *nifH* phylogenetic trees (Fig. 3 and Fig. S6) demonstrated that these two genes (even the *nod* and *nif* gene clusters) have cotransferred or coevolved. In addition, lateral transfer might occur mainly intragenus, as found in a previous study [4], since the strains of *Agrobacterium*, *Rhizobium*, and *Neorhizobium* formed subclades according to their general affiliations that were intermingled with the *Ensifer* strains in the phylogeny of the symbiotic genes (Table S2 and Fig. S7). Thus, the *nodA* and *nifH* genes (Fig. S7) might have evolved to adapt to the genome of the recipient rhizobia, such as codon usage preference [40]. In this study, failures in *nodA* or *nifH* gene amplification and nodule formation on *S. cannabina* were detected in strains of *Agrobacterium* *delta* (YIC4121, ZJ121) and *Agrobacterium salinitolerans* (YIC4104, YIC4260, YIC5082) (results not shown), which were previously defined as symbionts of

---

**Fig. 6** Pairwise comparisons of intraspecies pairs in the competition experiment using sterilized vermiculite. Each pair of strains for the competitive experiment was mixed 1:1; data are mean ± s.e.m. based on three replicates. The genospecies of intraspecies pairs are shown on the right of the histogram.
S. cannabina [13]. These strains might be recent recipients of symbiotic genes, whose cooperation with chromosomes has not been well established, and symbiotic genes might be lost easily, as predicted in *Ensifer morelense* [41].

**Survival and Competition Ability**

In general, our results (Fig. 4) showed that the survival and competition ability of the rhizobia affected the rhizobial distribution, as suggested previously [13, 42–44]. It is interesting that all the representative strains maintained a very low survival rate in acidic soil after incubation for 21 days (Fig. 4a), while most of the strains showed a relatively high survival rate in alkaline soil after incubation for 45 days (Fig. 4b). This is consistent with the usual distribution of *Ensifer* strains in alkaline soil, which have incorporated many genes in the genome to adapt to alkaline conditions [36] and endured a pH of 10 but not pH 6 [45]. Our results indicate that soil characteristics could regulate the rhizobial community structure by affecting bacterial vitality and enhancing or weakening the competitive abilities of the bacteria. However, inconsistent results between the distribution/richness results and interspecies/inspecies competition were also observed, indicating that competition may be affected not only by environmental conditions but also by biological factors [46].

**Conclusion**

Diverse symbionts, including seven novel genospecies within *Agrobacterium*, *Ensifer*, *Neorhizobium*, and *Rhizobium*, were identified among the *S. cannabina* rhizobia, and all of them were designated as symbiovar sesbaniae based on their highly conserved symbiosis genes and nodulation test results. Clear biogeographic patterns of these rhizobia were observed, with *E. sesbaniae* and *N. huautlense* as the most dominant and widely distributed *S. cannabina* rhizobia in China. The soil characteristics pH, TN, AK, and AP were the main factors affecting the rhizobial distribution. The survival rate and competitive nodulating abilities of rhizobia demonstrated greater tolerance to alkaline conditions than to acidic conditions, which are correlated with their soil origins. The lateral transfer of *nod* and *nif* genes accelerated the intraspecies diversification of rhizobia, but point mutation was the main driving force in the evolution of symbiotic genes, with faster evolution in *nodA* than in *nifH*.

**Funding Information** This work was financed by the National Natural Science Foundation of China (31600009, 31800099), Yantai Science and Technology Project (2018ZHGY074), Key Research and Development Program of Hebei province (19273802D), Joint fund of Jilin province and Chinese Academy of Sciences for high-tech industrialization (2019SYHZ0036), and Key Deployment Project of Chinese Academy of Sciences (KFZD-SW-112). En-Tao Wang was supported by the projects SIP 2017259 and 20180734 authorized by IPN, Mexico.

**References**

1. Long SR (1996) *Rhizobium* symbiosis: nod factors in perspective. Plant Cell 8(10):1885–1898
2. Jones KM, Hajime K, Davies BW, Taga ME, Walker GC (2007) How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol 5(8):619–633
3. Herridge DF, Peoples MB, Boddey RM (2008) Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil 311(1–2):1–18
4. Remigi P, Zhu J, Young J PW, Masson-Boivin C (2016) Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. Trends Microbiol. 24(1):63–75
5. Allen ON, Allen EK (1981) The *Leguminosae*. A source book of characteristics, uses and nodulation, University of Wisconsin Press/Macmillan Publishing, Madison, WI/London
6. Webb JR, Sprent JI (2002) Nodulation in Legumes. Kew Bull 57(3):634
7. Rao D LN, Gill HS (1995) Biomass and biofertilizer production by *Sesbania cannabina* in alkaline soil. Bioresour Technol. 53(2):169–172
8. Ye ZH, Yang ZY, Chan GYS, Wong MH (2001) Growth response of Sesbania rostrata and *S. cannabina* to sludge-amended lead/zinc mine tailings: a greenhouse study. Environ Int. 26(5):449–455
9. De Lajudie P (1994) Polyphasic taxonomy of rhizobia: emendation of the genus *Sinorhizobium* and description of *Sinorhizobium melliloti* comb. nov., *Sinorhizobium saheli* sp. nov., and *Sinorhizobium teranga* sp. nov. Int J Syst Evol Microbiol. 44(4):715–733
10. Chen WM, Lee T (2001) Genetic and phenotypic diversity of rhizobial isolates from sugarcane- *Sesbania cannabina*-rotation fields. Biol Fert Soils 34(1):14–20
11. Wang YC, Wang F, Hou BC, Wang ET, Chen WF, Xi FH, Chen WX, Li Y, Zhang YB (2013) Proposal of *Ensifer psoraleae* sp. nov., *Ensifer sesbaniae* sp. nov., *Ensifer morelense* comb. nov. and *Ensifer americanum* comb. nov. Syst Appl Microbiol 36(7):467–473
12. Cummings SP, Gyaneshwar P, Vinuesa P, Farruggia FT, Andrews M, Humphry D, Elliott GN, Nelson A, Orr C, Pettitt D (2009) Nodulation of *Sesbania* species by Rhizobium (*Agrobacterium*) strain IRBG74 and other rhizobia. Environ Microbiol. 11(10):2510–2525
13. Li Y, Li X, Liu Y, Wang ET, Ren C, Liu W, Xu H, Wu H, Jiang N, Li Y (2016) Genetic diversity and community structure of rhizobia nodulating *Sesbania cannabina* in saline-alkaline soils. Syst Appl Microbiol. 39(3):195–202
14. Man CX, Wang H, Chen WF, Sui XH, Wang ET, Chen WX (2008) Diverse rhizobia associated with soybean grown in the subtropical and tropical regions of China. Plant Soil 310(1–2):77–87
15. Zhang YM, Li Y, Chen WF, Wang ET, Tian CF, Li QQ, Zhang YZ, Sui XH, Chen WX (2011) Biodiversity and biogeography of rhizobia associated with soybean plants grown in the North China plain. Appl Environ Microbiol 77(18):6331–6342
16. Vincent JM (1970) A manual for the practical study of the root-nodule bacteria. International Biological Program, Blackwell Scientific, Oxford
17. Donohue SJ (1992) Reference soil and media diagnostic procedures for the southern region of the United States. J Am Ceram Soc 71(4):219–221
18. Westerman RL (1990) Soil testing and plant analysis. Soil Science Society of America, Inc

19. Page AL, Miller RH, Keeney Dennis R (1982) Methods of soil analysis. Part 2. Chem-ical and microbiological properties, American Soc. of Agronomy, Inc, Madison, Wisconsin, USA

20. Shen J, Li R, Zhang F, Fan J, Tang C, Rengel Z (2004) Crop yields, soil fertility and phosphorus fractions in response to long-term fertilization under the rice monoculture system on a calcareous soil. Field Crop Res 86(2):225–238

21. Vinuesa P, Silva C, Werner D, Martinez-Romeroa (2005) Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in Bradyrhizobium species cohesion and delineation. Mol Phylogenet Evol, 34(1):29–54

22. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874

23. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452

24. Tan Z, Hurek T, Vinuesa P, Müller P, Ladha JK, Reinholdhurek B (2001) Specific detection of Bradyrhizobium and Rhizobium strains colonizing rice (Oryza sativa) roots by 16S-23S ribosomal DNA intergenic spacer-targeted PCR. Appl Environ Microbiol 67(8):3655–3664

25. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 74(12):5463–5467

26. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp Series 41(41):95–98

27. Martens M, Dawyndt P, Coopman R, Gillis M, De VP, Willems A (2003) Versalovic J, Schneider M, De Bruijn F, Lupski JR (1994) Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. Methods Mol Cell Biol 5(1):25–40

28. Tian CF, Zhou YJ, Zhang YM, Li QQ, Zhang YZ, Li DF, Wang S, Wang J, Gilbert LB, Li YR (2012) Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. Proc Natl Acad Sci U S A 109(22):8629–8634

29. Leps J, Smilauer P (2003) Multivariate analysis of ecological data. Cambridge University Press, Cambridge

30. Rogel MA, Ormeno-Orrillo E, Romero EM (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst Appl Microbiol 34(2):96–104

31. Guo HJ, Wang ET, Zhang XX, Li QQ, Zhang YM, Tian CF, Chen WX (2014) Replicon-dependent differentiation of symbiosis-related genes in Sinorhizobium strains nodulating Glycine max. Appl Environ Microbiol 80(4):1245–1255

32. Bedhomme S, Amorós-Moya D, Valero LM, Bonifaci N, Pujana M, Bravo IG (2019) Evolutionary changes after translational challenges imposed by horizontal gene transfer. Genome Biol Evol 11(3):814–831

33. Wang ET, Tan ZY, Willems A, Fernández-López M, Reinhold-Hurek B, Martinez-Romero E (2002) Sinorhizobium morelense sp. nov., a Leucaena leucocephala-associated bacterium that is highly resistant to multiple antibiotics. Int J Syst Evol Microbiol. 52(5):1687–1693

34. Alexandre A, Brígido C, Laranjo M, Rodrigues S, Oliveira S (2009) Survey of chickpea rhizobia diversity in Portugal reveals the predominance of species distinct from Mesorhizobium ciceri and Mesorhizobium mediterraneum. Microb Ecol 58(4):930–941

35. Han LL, Wang ET, Han TX, Liu J, Sui XH, Chen WF, Chen WX (2009) Unique community structure and biogeography of soybean rhizobia in the saline-alkaline soils of Xinjiang, China. Plant Soil 324(1/2):291–305

36. Xiong HY, Zhang XX, Guo HJ, Ji YY, Li Y, Wang XL, Zhao W, Mo FY, Chen JC, Yang T (2017) The epidemicity of facultative microsymbionts in faba bean rhizosphere soils. Soil Biol Biochem 115:243–252

37. Li Y, Yan J, Yu B, Wang ET, Li X, Yan H, Liu W, Xie Z (2016) Ensifer alkalisoli sp. nov., isolated from root nodules of Sesbania cannabina grown in saline-alkaline soils. Int J Syst Evol Microbiol 66(12):5294–5300

38. Zhang J, Yu T, Lou K, Mao PH, Wang ET, Chen WF, Chen WX (2014) Genotypic alteration and competitive nodulation of Mesorhizobium muleiense against exotic chickpea rhizobia in alkaline soils. Int J Syst Evol Microbiol 37(7):520–524