

Dominant associations of *Ensifer medicae*-*Medicago polymorpha* and *Ensifer meliloti*-*Medicago lupulina* in farmland and natural ecosystem

Mingxing Tang

Hebei University

Hao Yu Wang

Hebei University

Xin Qi

Hebei University

Bao Juan Yuan

Hebei University

Zhang Bin

Hebei University

En Tao Wang

Instituto Politecnico Nacional, 11340 Mexico City

Bei Nan Wang

Hebei University

Fang Wang

Southwest Forestry University

Zhong Kuan Liu

Institute of Agricultural Resources and Regional Planning of CAAS

Xiaoyun Liu (✉ liuxiaoyunly@126.com)

CAS: Chinese Academy of Sciences <https://orcid.org/0000-0002-0472-6461>

Research Article

Keywords: Medicago, Ensifer, genetic diversity, gene phylogeny, plant host, soil nutrient

Posted Date: June 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-615146/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Aims

The nitrogen-fixing rhizobia associated with *Medicago polymorpha* and *M. lupulina* in Yunnan, China have been poorly documented. This study aims to analyze the diversity of rhizobia isolated from these two *Medicago* species and investigate the impact of abiotic (soil properties) and biotic (plant hosts) factors on *Medicago*-associated rhizobia in this region.

Methods

91 rhizobial isolates were characterized by RFLP of 16S rDNA and 16S–23S IGS, BOX-PCR fingerprinting, nodulation assays and phylogeny analyses based on housekeeping and symbiosis genes. The genetic diversity of the rhizobial isolates was assessed by the BOX AIR pattern and Shannon index. Additionally, the correlation of soil properties and rhizobial distribution was determined by the constrained analysis of principle coordinates (CAP) based on Bray-Curtis distance of presence/absence (PA) transformed species data.

Results

All the tested strains fell in the genus *Ensifer* and divided into two species, *E. medicae* and *E. meliloti*. Both host plants and soil properties contributed to the rhizobial diversity. For either *E. meliloti* or *E. medicae*, isolates from native host plants tended to be more genetically diverse than those of the same species from non-native hosts. The soil edaphic factor analysis elucidated that nitrogen, organic matter as well as Ca^{2+} and Na^+ are the key factors to shape the biogeographical distribution of rhizobia.

Conclusions

This study evidenced the microsymbiont preference of *M. polymorpha* to *E. medicae* and *M. lupulina* to *E. meliloti*, but also revealed the considerable impacts of both plant hosts and soil factors on the rhizobial diversity and biodistribution.

Introduction

Symbiotic nitrogen-fixing system, root and/or stem nodules, established between legume plants and rhizobia is the most significant and efficient biological nitrogen fixation system, which accounts for one quarter of the global nitrogen fixation. Inside the nodules, rhizobia reduce the atmospheric nitrogen to ammonium and provide it to plants as nitrogen nutrient, while the plant partners supply carbohydrates to rhizobia as carbon and energy source. Therefore, both the plant and bacterial symbionts obtain benefits from symbiosis. As a kind of extensively important forages and pasture plants, the *Medicago* species can form root nodules with their specific rhizobia, and the nitrogen fixation by the *Medicago*-associated rhizobia plays an important role in the world-wild pasture production.

Efforts of decades have been focused on the diversity analysis of *Medicago*-associated rhizobia aiming at screening the rhizobial strains with high nitrogen-fixing efficiency. Up to date, rhizobia isolated from the nodules of *Medicago* species (*M. sativa*, *M. truncatula*, *M. lupulina*, and *M. orbicularis*) were mainly classified into the genera *Ensifer* and *Rhizobium* (Hou et al., 2009; Rome et al., 1996; Roumiantseva et al., 2002; Sebbane et al., 2006; van Berkum et al., 1998; Wang et al., 2019; Zakhia et al., 2004). Among them, *Ensifer meliloti* was found predominantly to nodulate *Medicago* spp. (Alías-Villegas et al., 2015; Brunel et al., 1996; Djedidi et al., 2011; Elboutahiri et al., 2010; Howieson et al., 2000; Rome et al., 1996; Silva et al., 2007; Villegas et al., 2006). *E. medicae* was reported as the microsymbiont for multiple *Medicago* species in Tunisia and France (Zribi et al., 2004). Although this species shared a common host *M. sativa* with its sister species *E. meliloti*, it could be distinguished from *E. meliloti* by its capacity to nodulate and fix nitrogen with *M. polymorpha* L. (Rome et al., 1996) and *E. medicae* was reported as the unique symbiont of this plant (Brunel et al., 1996). Although another report showed that *E. meliloti* could nodulate with *M. polymorpha* as well (Howieson and Ewing, 1986), only inefficient nodules or root swellings are induced (Alías-Villegas et al., 2015). In addition, some strains isolated from root nodules of *Medicago* species were identified as members of *Rhizobium* species, including *R. mongolense*, *R. galegae*, *R. gallium*, *R. etli* and *R. leguminosarum* (Hou et al., 2009; Sebbane et al., 2006; Van Berkum et al., 1998; Zakhia et al., 2004).

There are 6 annual and 12 perennial *Medicago* species in China (Fang and Li, 2019). Among these plants, species *M. polymorpha* and *M. lupulina* were naturalized in Yunnan Province located in the tropical region in China, which are annual broadleaf herb and annual or short-lived perennial herb, respectively, cultivated as green manure in rice field or grown spontaneously. These two plants are important forage serving as green feed for livestock, due to their ability to form nitrogen-fixing nodules. Both *Medicago* species play an integral role in sustainable agriculture, enriching soil with bioavailable nitrogen source in rice field. In particular, *M. polymorpha* has been cultivated as green manure in some regions of Yunnan, where the soils are mainly acidic and contain relatively more aluminum than those in the northern regions of China. *M. polymorpha* is found adaptive to the humid and warm climate of Yunnan, turning out to be the only high-nitrogen forage in this region. However, efficient rhizobia associated with *M. polymorpha* might be lacking, this plant species has not spread widely in rice field in Yunnan, and was only restricted to regions of Chuxiong Prefecture and Lufeng County. Therefore, selection and application of highly effective rhizobial strains is essential for *M. polymorpha* cultivation in these regions, where this

plant is needed as green manure or forage for improving the extensive agriculture. However, up to date there has been no document about microsymbionts associated with *M. polymorpha* in China.

To better understand the specific interactions and synchronal evolution between rhizobia and *M. polymorpha*, but also collect rhizobial strains for screening highly effective isolates as inoculant applied in rice green manure cultivating, it would be helpful to investigate the rhizobial diversity. In the present study, rhizobia associated with *M. polymorpha* and *M. lupulina* were collected from Chuxiong Prefecture, Dehong Prefecture, Yuxi County and Kunming District of Yunnan. A total of 91 strains were isolated and characterized by 16S rDNA sequencing, 16S-23S rRNA IGS PCR-RFLP, BOX-PCR fingerprinting assay and nodulation tests. In order to analyze the correlation between rhizobial distribution and the soil properties, soil samples were also collected for physicochemical characterization. The results revealed that *M. polymorpha* preferred *Ensifer medicae* and *M. lupulina* preferred *Ensifer meliloti* in the studied region. Additionally, both host plants and soil nutrients had strong impacts on the genetic diversity of *Medicago*-associated rhizobia.

Materials And Methods

Nodule collection and soil physicochemical characterization

Root nodules of *M. polymorpha* and *M. lupulina*, as well as soil samples in the root zone of these plants were collected from farmlands and natural ecosystem on 7 sampling locations (Dabanqiao, Jinshan, Yanzhan, Cangling, Jiangchuan, Yinjiang and Yiliang towns) with 39 sample sites in Yunnan (Fig. S1). In each sampling site, 5-10 plants were sampled for collecting the efficient root nodules. Soils were sampled compositely from the root zone of nodule sampled plants (5-20 cm in depth) and were mixed as a single sample for each sampling sites. Complete nodules were dissected immediately from rinsed roots and stored over dehydrated silica gel in closed drying tube until their use for rhizobial isolation in the laboratory. Soil samples were ground and passed through 2-mm mesh screens for determining the physicochemical properties. Soil available nitrogen (N), available phosphate (P) (using Bray's hydrochloric acid fluoride ammonium by extraction method), and available potassium (K^+) (by ammonium acetate extraction plus flame photometry) were determined with the standard procedures (Du and Gao, 2006). Soil pH was measured using a pH meter (Mettler Toledo) by suspending 5 g soil in 5 mL of distilled water, and organic matter (OM) was measured using the potassium dichromate volumetric method (Du and Gao, 2006). Soil contents of Cl^- and HCO_3^- were determined by silver nitrate titration and potentiometer titration, respectively. Concentrations of Ca^{2+} , Mg^{2+} , Na^+ , Fe^{2+} , K^+ and SO_4^{2-} were measured using inductively coupled plasma atomic emission spectrometry.

Rhizobial isolation and ITS/BOX fingerprinting

Nodules stored in closed drying tubes were immersed in sterile water for 1-2 hours. The rehydrated nodules were surface sterilized by immersing in 95% (v/v) ethanol for 30 s and in 0.2% mercuric chloride for 2 to 3 min (depending on the nodule diameter), following by rinsing six times in sterile water. Then, the nodules were crushed separately and the liquid from each nodule was spread on plates of yeast-mannitol agar (YMA) and incubated at 28°C for 48h (Vincent, 1970). The obtained bacterial colonies were isolated and purified by repeatedly streaking on the same medium plates. Cultures of pure isolates were stored in YM broth supplied with 30% of glycerol at -70 °C.

For each isolate, as well as the type strains *E. meliloti* USDA 1002^T and *E. medicae* A321^T, genomic DNA was extracted by guanidine isothiocyanate method using FastPure Bacteria DNA isolation Mini kit (Nanjing Vazyme Biotech CO., Ltd) from 5 mL of YM culture agitated (120 rpm) overnight at 28°C. Using the DNA extract as template, the BOXAIR primer 5'-CTA CGG CAA GGC GAC GCT GAC G-3' (Versalovic et al., 1991) was used for BOX-PCR in a total volume of 25 μ L reaction mixture with the PCR procedure of Nick et al. (1999). The PCR products were separated by electrophoresis in 1.5% (w/v) agarose gels containing ethidium bromide (0.5 mg ml^{-1}) and were photographed under UV light. The BOX profiles were distinguished by their different band patterns, e.g. the isolates sharing the same pattern were designed as the same BOX pattern and were treated as colones of the same strain. For analysis of restriction fragment length polymorphism (RFLP), 16S–23S rRNA intergenic spacer (IGS) was amplified in 25 ml volume with primers FGPS1490 (5'-TGC GGC TGG ATC ACC TCC TT-3') and FGPS132 (5'-CCG GGT TTC CCC ATT CGG-3'), and the corresponding PCR protocol (Laguerre et al., 1996). Aliquot of 5 ml of the PCR products were used to verified the IGS amplification (about 900 bp) by electrophoresis in 1% (w/v) agarose gel. Aliquot of 5-10 μ L, depending on the concentration, was digested separately with the restriction endonucleases *Hae*II (GG|CC), *Rsa*I (GT|AC), *Hif*I (G|ANTC) and *Msp*I (C|CGG) (Laguerre et al., 1994) at 37°C for 6h, as specified by the manufacturer with an excess of enzyme (5U per reaction). The restriction fragments were separated by horizontal electrophoresis in agarose (2%, w/v) gels (14 cm in length) at 80V for 3h and were visualized by staining with ethidium bromide. Strains or isolates with different RFLP patterns were designated into distinct IGS types.

Phylogenetic analyses of housekeeping genes and symbiotic genes

Based on the results of IGS-RFLP analyses and BOX-PCR patterns, isolates representing different clusters and sampling sites were chosen for the analysis of multiple gene sequencing. The 16S rRNA gene amplified by PCR as described previously with the primers fd1 (5'-AGA GTT TGA TCC TGG CTC AGA-3') and rd1 (5'-AAG GAG GTG ATC CAG CC-3') (Weisburg et al., 1991). Multilocus sequence analysis (MLSA) based on the five housekeeping genes *atpD* (encoding for the ATP synthase beta-chain), *recA* (recombinase A), *dnaK* (DnaK chaperone), *gyrB* (DNA gyrase, beta-subunit) and *glnA* (glutamine synthetase I) was also performed in the present study, which has been widely used to differentiate rhizobial species (Martens et al., 2007; 2008; Vinuesa et al., 2005). The primer pairs recA41F/recA640R described by Vinuesa et al. (2005), glnA144F/glnA1142R and

dnaK1466F/dnaK1777R described by Martens et al. (2007), atpD352F/atpD871R and gyrB343F/gyrB1043R described by Martens et al. (2008) were used to amplify the corresponding genes in 25 ml volume by PCR. The PCR products were checked by electrophoresis in 1% (w/v) agarose gel. After purified with the Solarbio DNA purification kit (Beijing Solarbio Science & Technology Co., Ltd.), the amplicons were sequenced commercially using the same primers in Beijing Genomics Institute (BGI). The sequences acquired in this study were aligned with those from type strains of the defined bacterial species (obtained from the NCBI database by blasting), using Clustal W (Thompson et al., 1997). Maximum likelihood phylogenetic trees were constructed and were bootstrapped with 1000 pseudo-replicates using Mega 6.1 (Tamura et al., 2013). Phylogenies were also constructed using the concatenated sequences of 16S rRNA and the five housekeeping genes by Maximum likelihood method.

Fragments of the *nifH* gene (about 800 bp) and *nodC* gene (about 700 bp) were amplified with primer pairs nifHF/nifHR and nodCF540/nodCR1160, respectively, using the protocols of Laguerre et al. (2001). The visualization, purification and sequencing of the *nifH* and *nodC* amplicons were performed same as that mentioned for the housekeeping genes. All of the acquired nucleotide sequences were used for alignment with related genes extracted from GenBank database by Blast, and construction of the phylogenies was performed using the same methods described above for the keeping housekeeping genes.

All the obtained nucleic acid sequences were submitted in GenBank database under the accession numbers MT863814-863838 for *nodC* gene, MT863789-863813 for *nifH*, and others listed in Table S1.

Nodulation assays

A total of 35 representative strains were used in the nodulation tests that were selected according to their different 16S-23S IGS genotypes and BOX patterns, as well as phylogenies of 16S rRNA and five keep housing gene sequencing. *M. polymorpha* and *M. lupulina* seeds were scarified using concentrated sulfuric acid for 10 min, rinsed several times with sterile water, and then surface-sterilized in 3.2% (w/v) sodium hypochlorite followed by several rinses with sterile water. They were then placed on 0.8% water-agar at 4°C for 3 days, and then germinated at 28°C until the seedlings developed roots of 0.5-1 cm in length. Two seedlings were transplanted into a sterile glass tubes (30 × 200 cm) with nitrogen-free plant nutrient solution (Vincent, 1970) in 0.8% agar. The seedlings were then inoculated separately with 0.1 mL YM broth culture of each test strain (about 10⁸ cells mL⁻¹). Five replicates were used and controls without inoculation were included. The plants were placed in a growth cabinet under conditions described previously (Zhang et al., 2012). Plants were checked for nodule formation at 35 d after inoculation and BOX-PCR was performed for nodule crushes to verify the nodule occupation by inoculant strains.

Data analysis

The genetic differentiation and rhizobial species abundance determined by MLSA of the representative strains were calculated by DnaSP 5 software (Librado and Rozas, 2009). And the gene diversity and Shannon diversity index (H') (Shannon and Weaver, 1949) of total test strains were estimated by using the PopGene software 3.1 resourced from BOX AIR profiles data (BOX profiles were converted into a binary dataset). The genetic differentiation levels (Kst^*) (Hudson et al., 1992a, 1992b) between rhizobial communities and populations of the same species in farmlands and natural ecosystem were calculated. Correlations between soil characteristics and rhizobial species were analysed with constrained analysis of principle coordinates (CAP). Based forward selection of each environment (Anderson and Willis, 2003). Only variables that significantly increase the explained variance ($p < 0.05$) were used for CAP analysis.

Results

Genetic diversity of rhizobial strains isolated from *M. polymorpha* and *M. lupulina* nodules

Totally 91 rhizobial isolates were obtained from 39 sampling sites in 7 locations (Fig. S1), in which 70 were from *M. polymorpha* and 21 were from *M. lupulina* (Table 1). All of these strains were divided into five rDNA types by IGS-RFLP analysis and displayed 39 BOX-AIR fingerprinting patterns (Table 1). The IGS type A containing 19 BOX-AIR patterns was mainly composed of 63 strains from *M. polymorpha* (Table 1) collected in five locations, as well as 4 strains from *M. lupulina* sampled in Yiliang, Kunming. The IGS type C represented 24 strains from both *M. polymorpha* and *M. lupulina* with 17 BOX-AIR patterns. In contrast, the IGS types B, D and E each contained only one or two strains, and each strain has its own specific BOX-AIR pattern.

Phylogenies based on housekeeping and symbiotic genes

Based on the BOX-PCR patterns as well as the IGS-RFLP profiles of all the bacterial isolates, 25 representatives were selected for sequencing the 16S rDNA and five housekeeping genes (*recA*, *glnII*, *atpD*, *dnaK* and *glnA*) to determine their species affiliation (Fig. S2-S7).

Based on the 16S rDNA phylogeny, the representatives were divided into two clades, with 10 isolates closely related to *E. meliloti* LMG6133^T (>99% similarity) and the remaining 15 clustered with *S. medicae* WSM A321^T (>99% similarity) (Fig. S2). The phylogenetic tree based on MLSA of the 5 concatenated housekeeping genes (Fig. 1) showed a similar topology to that of the 16S rDNA tree, and similarities greater than 97% were observed among the strains within each of the *E. meliloti* and *S. medicae* clades (including the type strain in each). Thus, these two clades were identified as *E. meliloti* and *S. medicae*, respectively. *E. medicae* covered 63 test strains belonging to the IGS types A and B, while *E. meliloti* clade included 28 strains

in IGS types C, D and E. *E. medicae* was predominated in *M. polymorpha* isolates (occupied 85.7%, 60/70), while *E. meliloti* was dominant in *M. lupulina* (occupied 80.9%, 17/21).

Intriguingly, some *E. meliloti* strains exhibited traces of genetic incongruence, which were evidenced by comparison of the housekeeping gene phylogenies of MLSA. In the phylogenetic tree of *aptD* (Fig. S3), the 15 *E. medicae* representative strains together with the type strain clustered in one clade that was consistent to the result of MLSA. Nevertheless, the other 10 isolates formed 3 lineages within the clade intermingling with *E. meliloti* and *E. kummerowiae* type strains. Similar relationships were also observed in the *recA* phylogeny (Fig. S5), and the genetic incongruence between *aptD/recA* phylogeny and MLSA phylogeny suggested that the representative strains SWF67487 and SWF67523 might acquire genetic materials via horizontal gene transfer from *S. kummerowiae* strains during the natural evolution process. The gene transfer event of *E. meliloti* isolates from *E. kummerowiae* can be further supported by the phylogeny of symbiotic genes, since the 25 representatives were divided into 3 groups in the *nodC* phylogeny (Fig. 2), which were different from the results of MLSA. Particularly, 8 *E. meliloti* strains formed a monoclade together with *S. kummerowiae* CCBAU 71714^T at 100% identity. Likewise, these 8 strains were also clustered together with *S. kummerowiae* CCBAU 71714^T in the *nifH* phylogeny (Fig. S8).

Taken together, the phylogenetic analyses based on multiple concatenated housekeeping gene sequences (MLSA) divided the 25 representatives into two groups, corresponding to *E. meliloti* and *E. medicae*. Traces of HGT events in some *E. meliloti* strains were evidenced by the differences between phylogenies of single housekeeping genes, and between the phylogenies of MLSA and symbiotic genes.

Symbiotic performance of representative strains on *Medicago polymorpha*

Given the colonial morphology, hosts, and the genomic/phylogenetic analyses, 32 representative strains belonged to *E. meliloti* and *E. medicae* from both *Medicago* species were chosen for nodulation assays (Table 1). All the selected strains formed nodules on roots of *M. polymorpha*. Based on the patterns of BOX-PCR of nodule crushes, most the recovered strains (31/32) were identical to the corresponding inoculants, despite their host of origin. Thus, they were undoubtedly defined as the microsymbionts of *M. polymorpha*. Only strain SWF67450 was not recovered from nodules on plants inoculated with this strain, which might be explained that the plants in this test were partially or totally contaminated by other strains. Notably, two *E. meliloti* strains, SWF65114 and SWF65115, originally isolated from *M. lupulina* plants were capable of inducing efficient nodules on *M. polymorpha* as well.

Biogeographical patterns and genetic differentiation of *Medicago* rhizobia

The aforementioned phylogenetic analysis indicated genetic incongruence in *E. meliloti* strains but not in *E. medicae* representatives. Then, we assessed the gene diversity of the rhizobial isolates by assessing the BOX AIR pattern and Shannon index. *E. meliloti* populations showed significantly higher gene diversity index than the *E. medicae* counterparts. Moreover, a high level of genetic differentiation ($F_{ST}=0.92054$) was observed between *E. meliloti* and *E. medicae*, based on the sequences of 5 housekeeping genes (Table 2). To know whether host specificity would affect genetic diversity of rhizobial isolates, we firstly compared the genetic diversity of symbiotic strains within the *Ensifer* species. The total *E. medicae* strains from *M. polymorpha* in both farmland and nature ecosystem showed higher genetic diversity to those from *M. lupulina* nodules according to the BOX-PCR pattern and Shannon index. Similarly, *E. meliloti* isolates from *M. lupulina* showed significantly higher genetic diversity than those from *M. polymorpha* nodules (Fig. 3). Thus, these data suggested that rhizobial strains from their native host tend to be more genetically diverse.

Then, we evaluated whether soil type would affect rhizobial diversity by performing the PCoA based on Bray-Curtis distance (Table S2, Fig. 4). The amount of multiple nutrient factors was extensively various in natural ecosystem while it tended to be more similar in the farmland sites, as the observation that the component 1 represented 53.4% of relative eigenvalues, and component 2 represented 20.04% of relative eigenvalues, after Cailliez correction. Furthermore, multivariate analysis of variance (MANOVA) revealed that the *E. medicae* distribution significantly differed between the two habitats ($P=0.047$), while no overt difference was observed for *E. meliloti* species between farmland and nature ecosystem. In addition, permutational multivariate analysis of variance (PERMANOVA) corroborated that soil type (farmland/natural ecosystem) accounted for 1.61% ($P = 0.15$) of variation in the observed beta-diversity of rhizobia (Bray-Curtis distance metric).

On the one hand, all the *Ensifer* strains originated from natural ecosystem had a higher level of nucleotide diversity than those from farmland (Table 2), suggesting that the soil type might contribute to rhizobial genetic diversity. The greater variation in the content of several nutrients in natural ecosystem might explain the higher genetic diversity of rhizobia. Consistently, moderate level of genetic differentiation was observed within *E. medicae* ($F_{ST}=0.05251$) or *E. meliloti* ($F_{ST}=0.0501$) between natural ecosystem and farmland. Furthermore, gene flow was detected for both *E. meliloti* and *E. medicae* strains between farmland and natural ecosystem. Altogether, these data suggested that *Ensifer* strains from natural ecosystem were more genetically diverse compared to those from farmland.

Deterministic factors for diversity of rhizobia from two *Medicago* plants

To identify which specific soil factor could determine the genetic diversity of rhizobial populations, analysis of variance (ANOVA) based on a set of environmental factors was performed. Compared to natural ecosystem, the alfalfa farmland had a significantly higher content of multiple nutrient factors, including organic matter (OM), sodium (Na^+), magnesium (Mg^{2+}) and bicarbonate (HCO_3^-) (Table S3, S4). As shown in the CAP analysis based on Bray-Curtis distance of PA (presence/absence) transformed species data (Fig. 5A), the soil physiochemical data could explain 53.4% of variation in

the diversity of five IGS type species. Soil Ca^{2+} , Na^+ , HCO_3^- , Cl^- , SO_4^{2-} and pH contributed to a significant partition in the first component (63.51% of total variance). Meanwhile, nitrogen (N), OM, potassium (K), and phosphorus (P) played a minor role in the second component (14.31 % of total variance), showing negative effects on rhizobial distribution. Variation partition analysis revealed that 11.7%, 5.8%, 2.2% and 0.5% could be significantly explained by 6 ions, NPK, pH and OM variables respectively. These results indicated that Ca^{2+} and Na^+ were major factors in shaping the genetic diversity of five IGS types (Fig. 5B).

Rhizobial species correlated with different sets of explanatory variables were further identified. IGS type A was more likely found in the farmland sites with lower HCO_3^- . IGS type C was more likely found in the nature ecosystem with high soil Ca^{2+} , Na^+ , HCO_3^- , Cl^- , lower nitrogen, K and OM. The IGS types B, D and E more likely appeared in the nature ecosystem sites with high P and low HCO_3^- . In summary, IGS types A and C have similar nutrition utilization, and both preferred soils with low potassium, OM, phosphate and nitrogen contents, while IGS type C preferred soil with high contents of Ca^{2+} and Na^+ (Fig. 5A).

To evaluate the correlation of rhizobial abundance with environment factors, CAP analysis based on Hellinger transformed species data was conducted (Fig. 5C). IGS type A had a significantly higher abundance in farmlands that were positively associated with pH. While more stringent than type A, type C was also negatively correlated with soil N, k, P and OM. IGS types B, D and E were not considered due to their extremely low abundance (Fig. 5C). Variation partition analysis of Hellinger transformed species data (Fig. 5D) exhibited almost the same pattern as the PA transformed species data (Fig. 5B). The range of soil physiochemical content for two rhizobial species identified in this study was given in Table S4.

Taken together, distribution of *E. medicae* strains was positively correlated with Na^+ in soil, and their abundance was also associated with low level of Nit and OM. *E. meliloti* isolates appeared to prefer good-quality soil conditions with high Ca^{2+} and Na^+ contents, and negatively correlated with soil N and OM. To conclude, the soil contents of N, OM as well as Ca^{2+} and Na^+ were the major soil factors to shape the distribution of rhizobial strains belonging to the two *Ensifer* species detected in the present study.

Discussion

In the present study, we systematically investigated, for the first time, the diversity of rhizobia associated with *M. polymorpha* in China (Yunnan Province) and compared them with those of *M. lupulina* grown in the same region. Based upon the results of BOX-PCR patterns, IGS-RFLP and MLSA-based phylogenies, the rhizobia of both the *Medicago* species were identified as diverse populations belonging to *E. medicae* and *E. meliloti*. The dominance of *E. medicae* associated with *M. polymorpha* (86.7 % nodule occupation), and of *E. meliloti* in the root nodules of *M. lupulina* (80.95% of nodule occupation) was in accordance with previous reports about the rhizobia from *M. polymorpha* and *M. lupulina* (Silva et al., 2007). Furthermore, in nodulation tests the *E. meliloti* strains originated from *M. lupulina* induced effective nodules on *M. polymorpha*. This situation might be explained by the adaptation of rhizobia to the local conditions or by the interactions between rhizobia and the microbiota in soil, similar to the case of chickpea plant and its rhizobia: *Mesorhizobium ciceri* presented great compatibility to form effective nodules with chickpea (*Cicer arietinum L.*) in sterilized soil, but it was absent in fields of China (Zhang et al. 2014).

Our results demonstrated that the specificity for the two *Medicago* species and their microsymbionts appeared not so stringent, although *M. polymorpha* was recognized as the sole host of *E. medicae* (Biondi et al., 2003; Brunel et al., 1996; Rome et al., 1996). While, it has been found that *E. medicae* could also nodulate the perennial *M. sativa L.* (Ardley et al., 2015) and *M. lupulina* (Silva et al., 2007). On the other hand, *E. meliloti* was once identified ineffective symbionts for *M. polymorpha*, but commonly nodulate *M. sativa*, *M. truncatula* and *M. laciniata* (Denton et al. 2007; Villegas et al. 2006). However, nodulation induced by *E. meliloti* strains on *M. polymorpha* has been recorded lately (Alías-Villegas et al., 2015), while our results in the present study further evidenced the effective nodulation of *M. polymorpha* by both *E. medicae* and *E. meliloti* strains, and an *E. meliloti* strain was even one of the most effective symbionts for *M. polymorpha* (Liu et al., 2011). Therefore, it could be estimated that the associations between the hosts *M. polymorpha*/ *M. lupulina* and the microsymbionts *E. medicae*/ *E. meliloti* were not only determined by the preference (or specificity) between the hosts and the rhizobia, but also might be regulated by the soil abiotic and biotic conditions.

Furthermore, soil edaphic could explain the distinct geographical distribution for these two rhizobial species. From CAP analysis, *E. meliloti* were displayed strong negative correlation to soil contents of nitrogen and organic matter in nature ecosystem, while the negative correlation of *E. medicae* with nitrogen and organic matter contents was even more stringent. Garau et al. (2005) reported that *E. medicae* was mostly associated with medics that well adapted to moderately acid soils, such as *M. polymorpha*, *M. arabica* and *M. murex*; whereas *E. meliloti* was predominantly isolated from *M. littoralis* and *M. tornata* that naturally grow in soils with alkaline or neutral pH, and also from *M. sativa* in acidic soils as well (Ramirez-Bahena et al., 2015). In our studies, both *E. medicae* and *E. meliloti* strains were mostly isolated from alkaline soils, but the nutrient contents and iron patterns were different among the sampling sites, which formed the determinants to regulate the nodule occupancy of *E. medicae* and *E. meliloti* on *M. polymorpha* and *M. lupulina*.

By analyzing the BOX-AIR fingerprinting profiles of all the 91 isolates, we found that *E. meliloti* populations showed significantly higher gene diversity index than the *E. medicae* counterparts in two plant hosts and two habitats in Yunnan, China. Analysis on the nucleotide diversity of 25 selected representatives also supported this observation. This difference may be due to the strain number of the species in BOX profile analysis. A high level of stringency in *Medicago*-rhizobia specificity might constrain the genetic diversity of strains with poor symbiosis efficiency. Since the more efficient rhizobial strains would outcompete in the nodulation process, thus being thrived and accumulated since they can favor plant growth by fixing nitrogen

more efficiently. Such selection of symbionts by plant host may not only lead to the formation of the dominant rhizobial populations in the sampling sites, but also could enhance the diversity of advantaged strains as contrast to the inefficient ones.

Furthermore, in accordance with the another report (Bailly et al. 2007), our results indicated that the soil type (or ecological environment) have effect on rhizobial diversity. Rhizobial strains tested in this study (both *Ensifer* species strains) from natural ecosystem have higher nucleotide diversity than those from farmland. This suggested that *Medicago* symbionts might undergo more gene exchange events with other bacterial habitants in the natural ecosystem when compared to that in the farmland. Indeed, a greater value of genetic differentiation was observed in *E. meliloti* stains than in *E. medicae* strains that might be a result of ecological niche barrier. *E. meliloti* was more frequently recovered from nature ecosystem; while *E. medicae* strains were more readily isolated from *M. polymorpha* in farmland. Selection force by environment and plant host may block rhizobial migration and prevent the subsequent gene exchange. This may also challenge the hypothesis that *E. medicae* is an evolutionary divergence from *E. meliloti* population (Biondi et al., 2003). Genome sequence analysis of more rhizobial species is needed to clarify the evolutionary relationship between *E. medicae* and *E. meliloti*.

Instead, these two *Ensifer* species retained their own genetic characteristics. We found high levels of gene flow (Nm) within *E. medicae* or *E. meliloti* populations in the two habitats. Various subpopulations of the same *Ensifer* species could intermingle their genes through the horizontal gene transfer processes such as conjugation and transformation, which conferred sufficient genetic exchanges among populations. This may be an important evolutionary force to contribute to the genetic diversity. Interestingly, phylogenies of some housekeeping and symbiotic genes suggested that gene transfer from *E. kummerowiae*, another *Ensifer* species that can nodulate *M. marina* and *Kummerowia stipulacea* (Wei et al., 2002; Alías-Villegas., 2015), but not *M. lupulina*.

PCoA showed that *E. meliloti* distributed geographically more scattered than *E. medicae*, the rhizobial distribution might be determined by their hosts. Indeed, *M. polymorpha* was mostly distributed in farmlands of Dabangqiao, Jinshan, Yaozhan, Cangling and Jiangchuan, while *M. lupulina* is mainly distributed around Yunnan (Fig.S1). So, during the long process of rhizobium-legume interaction, rhizobia populations and legume hosts can converge within distribution range. For example, *M. polymorpha* was mainly nodulated by strains of rRNA typeA in locations of Fufeng and Chuxiong, and group rRNA type D in Dehong. These data indicated that hosts and geographical environment deeply influence the distribution of rhizobia (Béna et al., 2005; Chen et al., 2004; Wang et al., 2019). In this study, only *E. meliloti* isolates were obtained from *M. polymorpha* in Yinjiang, the west of Yunnan, while the other *E. meliloti* strains were recovered from *M. lupulina*.

Taken together, the diverse ecological environments and leguminous plants accumulate a great variety of *E. medicae* and *E. meliloti* strains in association with *M. polymorpha* and *M. lupulina* in Yunnan. This study not only provided pioneering efforts for screening efficient nitrogen-fixating strains with *Medicago* spp. in the local regions, but also revealed tremendous impact of host specificity and soil factors on rhizobial diversity.

Declarations

This research was financially supported by the National Natural Science Foundation of China (grant No. 31370051 and 31360003), China Agricultural Research System-Green Manure (CARS-22), Key Bioengineering Discipline of Hebei Province (No.1050-5030023) and the Public Service Sectors (agriculture) Research and Special Funds (grant No. 201303061), the Key Discipline Project for Biotechnology of Hebei Province, and First-Class Discipline Construction Funding of Hebei Province.

There is no conflict of interest among authors. And all the authors listed have read the manuscript and approved the submission.

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgements:

This research was financially supported by the National Natural Science Foundation of China (grant No. 31370051 and 31360003), China Agricultural Research System-Green Manure (CARS-22), Key Discipline Project for Biotechnology of Hebei Province, and First-Class Discipline Construction Funding of Hebei Province. We thank Hua jie Liu (Hebei University, China) for reviewing the statics.

References

1. Alías-Villegas C, Cubo MT, Lara-Dampier V et al (2015) Rhizobial strains isolated from nodules of *Medicago marina* in southwest Spain are abiotic-stress tolerant and symbiotically diverse. *Syst Appl Microbiol* 38:506–514. <https://doi.org/10.1016/j.syapm.2015.07.003>
2. Ardley J, Tian R, O'Hara G et al (2015) High-quality permanent draft genome sequence of *Ensifer medicae* strain WSM244, a microsymbiont isolated from *Medicago polymorpha* growing in alkaline soil. *Stand Genomic Sci* 10:1–8. <https://doi.org/10.1186/s40793-015-0119-5>
3. Bailly X, Olivieri I, Brunel B et al (2007) Horizontal gene transfer and homologous recombination drive the evolution of the nitrogen-fixing symbionts of *Medicago* species. *J Bacteriol* 189:5223–5236. <https://doi.org/10.1128/JB.00105-07>
4. Béna G, Lyet A, Huguet T, Olivieri I (2005) *Medicago - Sinorhizobium* symbiotic specificity evolution and the geographic expansion of *Medicago*. *J Evol Biol* 18:1547–1558. <https://doi.org/10.1111/j.1420-9101.2005.00952.x>

5. Biondi EG, Pilli E, Giuntini E et al (2003) Genetic relationship of *Sinorhizobium meliloti* and *Sinorhizobium medicae* strains isolated from Caucasian region. FEMS Microbiol Lett 220:207–213. [https://doi.org/10.1016/S0378-1097\(03\)00098-3](https://doi.org/10.1016/S0378-1097(03)00098-3)
6. Fang QG, Li YB (2019) Classification and distribution of domestic species of genus *Medicago* in China. Grass Prataculture 31:1–7
7. Brunel B, Rome S, Ziani R, Cleyet-Marel JC (1996) Comparison of nucleotide diversity and symbiotic properties of *Rhizobium meliloti* populations from annual *Medicago* species. FEMS Microbiol Ecol 19:71–82. [https://doi.org/10.1016/0168-6496\(95\)00076-3](https://doi.org/10.1016/0168-6496(95)00076-3)
8. Denton MD, Hill CR, Bellotti WD, Coventry DR (2007) Nodulation of *Medicago truncatula* and *Medicago polymorpha* in two pastures of contrasting soil pH and rhizobial populations. Appl Soil Ecol 35:441–448. <https://doi.org/10.1016/j.apsoil.2006.08.001>
9. Du S, Gao X (2006) Technical Specification of Soil Analysis. China Agriculture Press, Beijing
10. Hou BC, Wang ET, Li Y et al (2009) Rhizobial resource associated with epidemic legumes in tibet. Microb Ecol 57:69–81. <https://doi.org/10.1007/s00248-008-9397-4>
11. Howieson JG, Ewing MA (1986) Acid Tolerance in the *Rhizobium meliloti*-*Medicago* Symbiosis. 55–64
12. Howieson JG, Nutt B, Evans P (2000) Estimation of host-strain compatibility for symbiotic N-fixation between *Rhizobium meliloti*, several annual species of *Medicago* and *Medicago sativa*. Plant Soil 219:49–55
13. Laguerre G (1994) Rapid Identification of Rhizobia by Restriction Fragment Length Polymorphism Analysis of PCR-Amplified 16S rRNA Genes. 56–63
14. Laguerre G, Mavingui P, Allard MR et al (1996) Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: Application to *Rhizobium leguminosarum* and its different biovars. Appl Environ Microbiol 62:2029–2036
15. Laguerre G, Nour SM, Macheret V et al (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. Microbiology 147:981–993. <https://doi.org/10.1099/00221287-147-4-981>
16. Liu XY, LI GUOZG, Liu QX, Xue GX, Wand SM YP (2011) Acta Ecologica Sinica. Acta Ecol Sin 31:4034–4041
17. Martens M, Dawyndt P, Coopman R, Gillis M, deVos PWA (2008) Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). Int J Syst Evol Microbiol 58:200–214
18. Martens M, Delaere M, Coopman R et al (2007) Multilocus sequence analysis of *Ensifer* and related taxa. Int J Syst Evol Microbiol 57:489–503
19. Nick G, de Lajudie P, Eardly BD et al (1999) *Sinorhizobium arboris* sp nov and *Sinorhizobium kostiense* sp nov., isolated from leguminous trees in Sudan and Kenya. Int J Syst Bacteriol 49:1359–1368
20. Rome S, Fernandez MP, Brunel B et al (1996) *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. Int J Syst Bacteriol 46:972–980. <https://doi.org/10.1099/00207713-46-4-972>
21. Roumiantseva ML, Andronov EE, Larissa A et al (2002) Diversity of *Sinorhizobium meliloti* from the Central Asian Alfalfa Gene Center Diversity of *Sinorhizobium meliloti* from the Central Asian Alfalfa Gene Center. <https://doi.org/10.1128/AEM.68.9.4694>
22. Sebbane N, Sahounne M, Zakhia F et al (2006) Phenotypical and genotypical characteristics of root-nodulating bacteria isolated from annual *Medicago* spp. in Soummam Valley (Algeria). Lett Appl Microbiol 42:235–241. <https://doi.org/10.1111/j.1472-765X.2005.01846.x>
23. Silva C, Kan FL, Martínez-Romero E (2007) Population genetic structure of *Sinorhizobium meliloti* and *S. medicae* isolated from nodules of *Medicago* spp. in Mexico. FEMS Microbiol Ecol 60:477–489
24. Tamura K, Stecher G, Peterson D et al (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
25. Thompson JD, Gibson TJ, Plewniak F et al (1997) The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882
26. van Berkum P, Beyene D, Bao GP et al (1998) *Rhizobium mongolense* sp. nov. is one of three rhizobial genotypes identified which nodulate and form nitrogen-fixing symbioses with *Medicago ruthenica* [(L.) Ledebour]. Int J Syst Bacteriol 48:13–22
27. Versalovic J, Koeuth T, Lupski JR (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res 19:6823–6831
28. Villegas MDC, Rome S, Mauré L et al (2006) Nitrogen-fixing sinorhizobia with *Medicago laciniata* constitute a novel biovar (bv. medicaginis) of *S. meliloti*. Syst Appl Microbiol 29:526–538. <https://doi.org/10.1016/j.syapm.2005.12.008>
29. Vincent JM (1970) The cultivation, isolation and maintenance of rhizobia. In A Manual for the Practical Study of the Root-Nodule Bacteria, pp. 1–13. Edited by J. M. Vincent. Oxford: Blackwell Scientific
30. Vinuesa P, Silva C, Lorite MJ et al (2005a) Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. Syst Appl Microbiol 28:702–716
31. Vinuesa P, Silva C, Werner D, Martínez-Romero E (2005b) Population genetics and phylogenetic inference in bacterial molecular systematics: The roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. Mol Phylogenet Evol 34:29–54
32. Wang ET, Chen WF, Tian CF, Young JPW, Chen WX (2019) Ecology and Evolution of Rhizobia: principles and applications. Springer Nature Singapore Pte Ltd, Singapore
33. Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703

34. Zakhia F, Jeder H, Domergue O et al (2004) Characterisation of Wild Legume Nodulating Bacteria (LNB) in the Infra-arid Zone of Tunisia. 395:380–395
35. Zhang JJ, Lou K, Jin X et al (2012) Distinctive *Mesorhizobium* populations associated with *Cicer arietinum* L. in alkaline soils of Xinjiang, China. Plant Soil 353:123–134
36. Zhang JJ, Yu T, Lou K, Mao PH, Wang ET, Chen WF, Chen WX (2014) Genotypic alteration and competitive nodulation of the *Mesorhizobium muleiense* against exotic chickpea rhizobia in alkaline soils from Xinjiang, China. Syst Appl Microbiol 37:520–524
37. Zribi K, Mhamdi R, Huguet T, Aouani ME (2004) Distribution and genetic diversity of rhizobia nodulating natural populations of *Medicago truncatula* in tunisian soils. Soil Biol Biochem 36:903–908. <https://doi.org/10.1016/j.soilbio.2004.02.003>

Tables

Table 1 Strains obtained in this study and information related to references strains and plant hosts.

Strain name	Rhizobial isolates	Host	IGS type	BOX AIR patterns	Sampling location	Habitat
<u><i>Ensifer medicae</i> genotype</u>	63 (total)					
SWF65100, SWF65103, SWF65105, SWF65107	4	<i>Medicago polymorpha</i>	A	a, b, c	Dabanqiao of Kunming City	Farmland
SWF67395, SWF67451, SWF67455, SWF67456, SWF67457, SWF67462, SWF67463, SWF67464, SWF67466, SWF67470, SWF67477, SWF67479, SWF67480, SWF67481, SWF67350, SWF67356, SWF67489, SWF67492, SWF67494	19	<i>Medicago polymorpha</i>	A	d, e, f, g	Jinshan of Lufeng Town	Farmland
SWF67394, SWF67397, SWF67398, SWF67400, SWF67403, SWF67407, SWF67409, SWF67410, SWF67412, SWF67413, SWF67414, SWF67416, SWF67417, SWF67418, SWF67421, SWF67422, SWF67423, SWF67424, SWF67425, SWF67431, SWF67432, SWF67436, SWF67438, SWF67441, SWF67443, SWF67446, SWF67447	27	<i>Medicago polymorpha</i>	A	h, i, j, k, l, m, n, o, p	Yaozhan of Lufeng Town	Farmland
SWF67303, SWF67448, SWF67498, SWF67499, SWF67500, SWF67501, SWF67503, SWF67347	8	<i>Medicago polymorpha</i>	A	i, q, i, l, r	Cangling of Chuxiong City	Farmland
SWF67405	1	<i>Medicago polymorpha</i>	A	o	Yaozhan of Lufeng Town	Natural ecosystem
SWF67343, SWF67370	2	<i>Medicago lupulina</i>	A	i, r,	Yiliang of Kunming City	Natural ecosystem
SWF65116, SWF66320	2	<i>Medicago lupulina</i>	A	s	Yiliang of Kunming City	Farmland
<u><i>E. medicae</i> genotype</u>	1					
SWF67497	1	<i>Medicago polymorpha</i>	B	t	Cangling of Chuxiong City	Natural ecosystem
<u><i>E. meliloti</i> genotype</u>	25 (total)					
SWF67521, SWF67524	2	<i>Medicago polymorpha</i>	C	u, v	Yinjiang of Dehong City	Farmland
SWF67526, SWF67527, SWF67528, SWF67529, SWF67534, SWF67537	6	<i>Medicago polymorpha</i>	C	w, x	Yinjiang of Dehong City	Farmland
SWF66326	1	<i>Medicago lupulina</i>	C	y, z, aa	Jiangchuan of Yuxi City	Natural ecosystem
SWF66437	1	<i>Medicago lupulina</i>	C	ab	Yiliang of Kunming City	Natural ecosystem
SWF67486, SWF67487, SWF67344	3	<i>Medicago lupulina</i>	C	ac, ad, ae	Cangling of Chuxiong	Natural ecosystem
SWF66329, SWF66436	2	<i>Medicago lupulina</i>	C	af, ag	Yiliang of Kunming City	Natural ecosystem
SWF67371, SWF67373, SWF67383, SWF67393	4	<i>Medicago lupulina</i>	C	ah, ai	Jinshan of Lufeng Town	Natural ecosystem
SWF65112, SWF65114, SWF65115, HBU65332, HBU65334	5	<i>Medicago lupulina</i>	C	aj, ak	Yiliang of Kunming City	Farmland
<u><i>E. meliloti</i> genotype</u>	2					
SWF67522, SWF67523	2	<i>Medicago polymorpha</i>	D	am	Yinjiang of Dehong City	Natural ecosystem
<u><i>E. meliloti</i> genotype</u>	1					

SWF66332	1	<i>Medicago lupulina</i>	E	al	Yiliang of Kunming City	Natural ecosystem
Total	91			39		
<i>Rhizobium legumiarum</i> USDA2370 ^T		<i>Pisum sativum, Trifolium pratense, vicia faba, Phaseolus vulgaris</i>				
<i>Rhizobium legumiarum</i> 162K68		<i>Pisum sativum, Trifolium pratense, vicia faba, Phaseolus vulgaris</i>				
<i>Ensifer meliloti</i> USDA1002 ^T		<i>Medicago sativa, Melilotus, Trigonella</i>				
<i>Ensifer medicae</i> USDA1037 ^T		<i>M. truncatula, M. orbicularis, M. polymorpha, and M. rugosa</i>				

Note: The bold-marked 32 stains are selected for nodulation test.

Table 2 Genetic diversity and differentiation analysis of concatenated keeping-house gene (16S rRNA-5 genes) sequences of rhizobial species from two *Medicago* species

Species	habitat	Host	Nucleotide Diversity Estimates			Genetics Differentiation and Gene Flow Estimates			
			n	h	π	Kst*	P-value	Fst	Nm
<i>E. medicae</i>	Farmland	<i>M. polymorpha</i>	10	10	0.00413	0.01644	0.2150	-0.0109	-2.53
Nature ecosystem	<i>M. polymorpha</i>	2		2	0.01366				
<i>E. meliloti</i>	Farmland	<i>M. lupulina</i>	3	3		0.00872	-0.02607	0.9390	-0.02547
Nature ecosystem	<i>M. lupulina</i>	5		5	0.01076				
<i>E. medicae</i>	Farmland	<i>M. polymorpha/M. lupulina</i>	11	11		0.00416	0.01245	0.1310	0.05251
Nature ecosystem	<i>M. polymorpha/M. lupulina</i>	4		4	0.01179				
<i>E. meliloti</i>	Farmland	<i>M. polymorpha/M. lupulina</i>	4	4		0.00848	0.00939	0.6630	0.0501
Nature ecosystem	<i>M. polymorpha/M. lupulina</i>	6		6	0.01042				
<i>E. medicae</i>	Farmland/ Nature	<i>M. polymorpha/M. lupulina</i>	15	15		0.00441	0.31532	<0.0001***	0.92054
<i>E. meliloti</i>	Farmland/ Nature	<i>M. polymorpha/M. lupulina</i>	10	10	0.00820				
<i>E. medicae</i>	Farmland	<i>M. polymorpha</i>		10		0.00431	0.29381	<0.0001***	0.90810
<i>E. meliloti</i>	Nature ecosystem	<i>M. lupulina</i>		5	0.01076				0.02

Note: n, Number of sequences; h, Number of haplotypes; π, Nucleotide diversity per site; Nm, the minimum number of migration events. Fst, F-statistics of the genetic differences among population. For Fst, the range 0.0 to 0.05 may be considered as indicating little genetic differentiation; 0.05 to 0.15, moderate genetic differentiation; 0.15 to 0.25, great genetic differentiation; above 0.25 indicate very great genetic differentiation. For Kst*, if the observed value of the statistic had P-value < 0.05, the null hypothesis of no genetic differentiation was rejected. Probability obtained by the Monte Carlo permutation test with 1000 replicates.

Figures

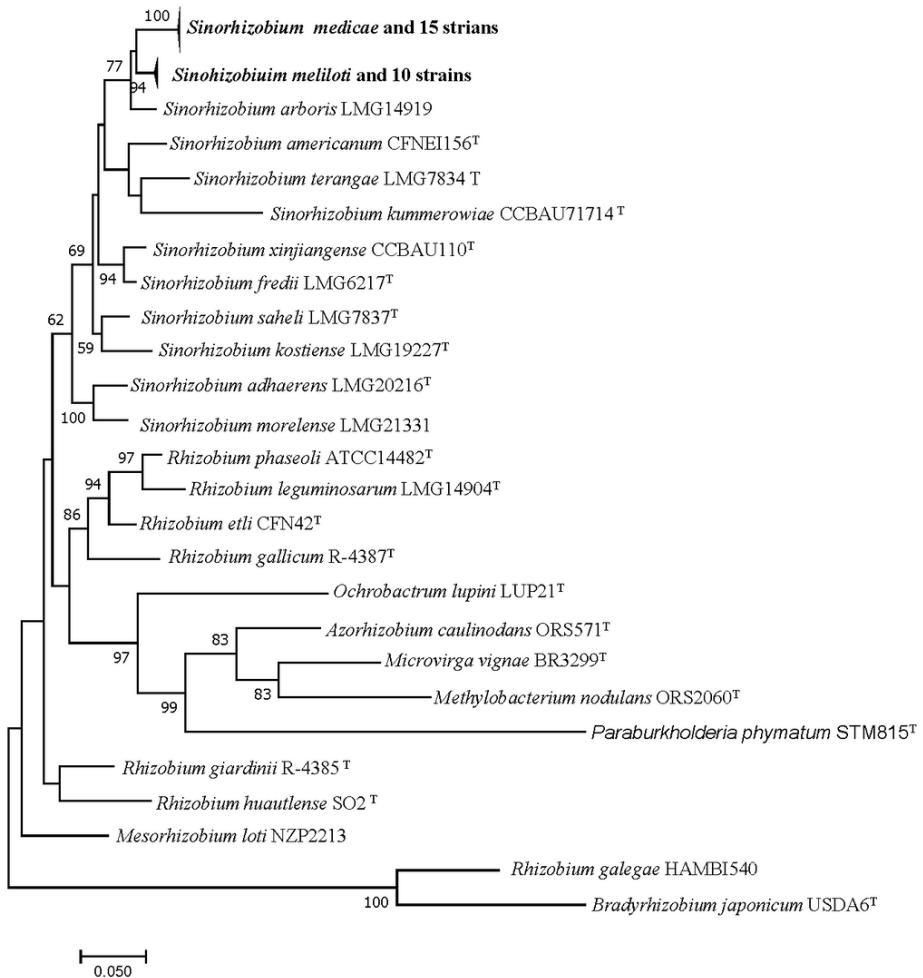


Figure 1

Phylogenetic tree based on concatenated 16S rRNA and 4 housekeeping gene (atpD, recA, dnaK, gyrB and glnA) sequences (4074bp). The tree was constructed by using the Maximum likelihood method. Bootstrap values are indicated (>50) in the main nodes in a bootstrap analysis of 1,000 replicates, 5% substitutions per site

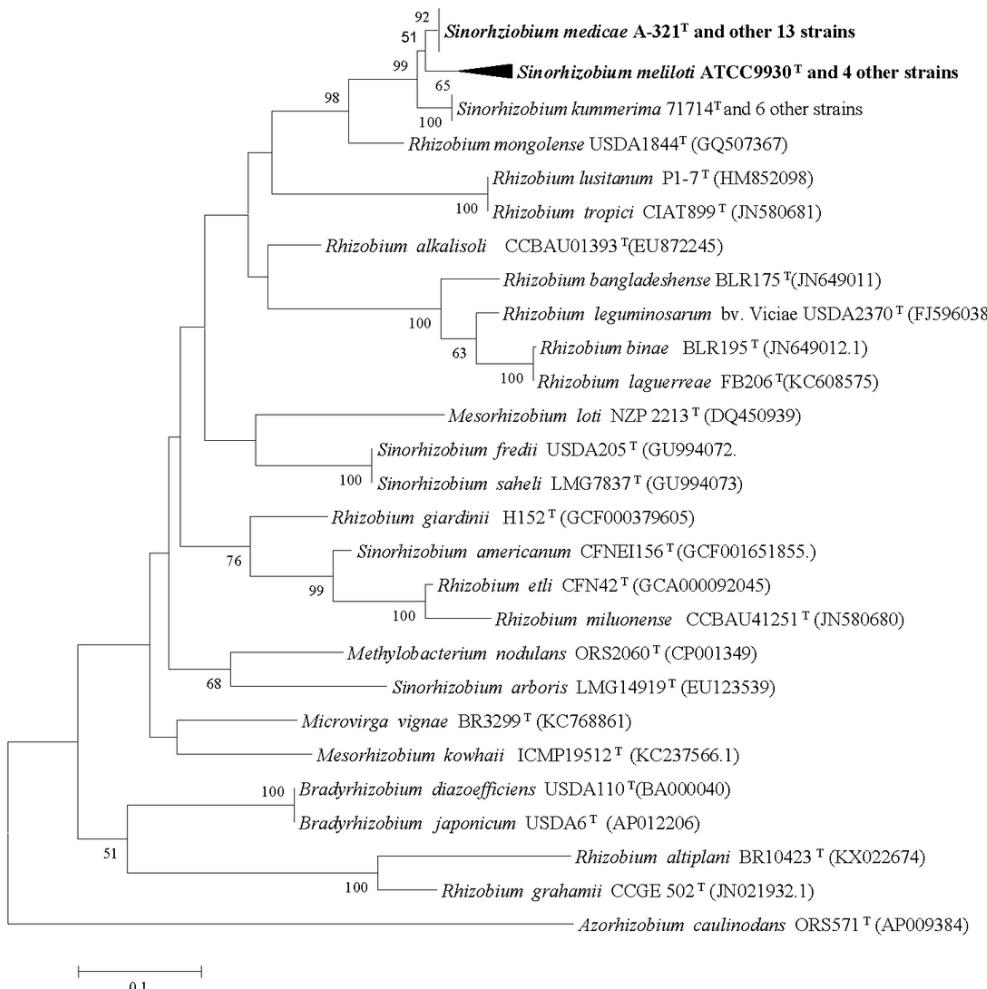


Figure 2

Phylogenetic tree of nodC gene showing the groups of the rhizobial strains isolated from *Medicago* spp. The tree was constructed by using the Maximum likelihood method. Bootstrap values are indicated (> 50) in the main nodes in a bootstrap analysis of 1,000 replicates, 10% substitutions per site

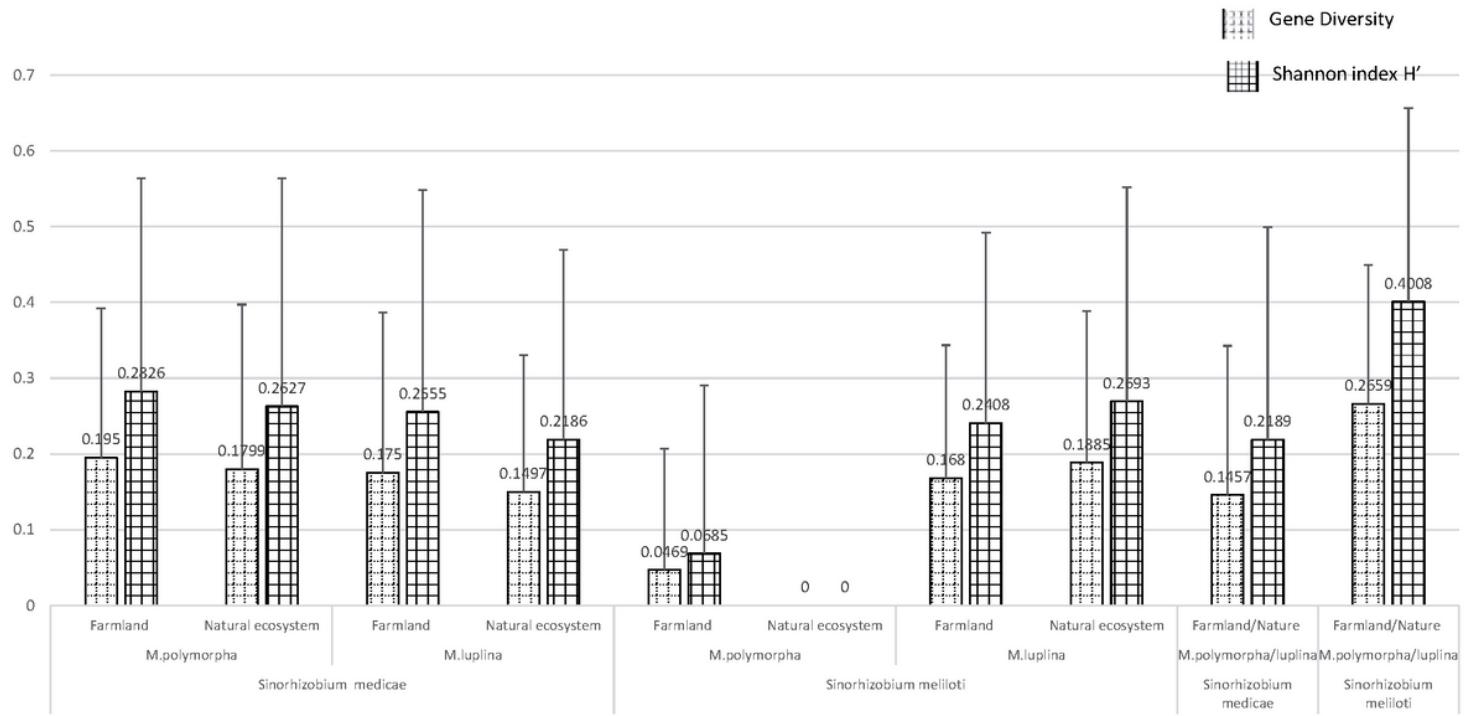


Figure 3

Diversity analysis between habitats (gene diversity and Shannon index (H') of rhizobial species from plants of *M. polymorpha* and *M. lupulina* in base on BOX AIR profiles

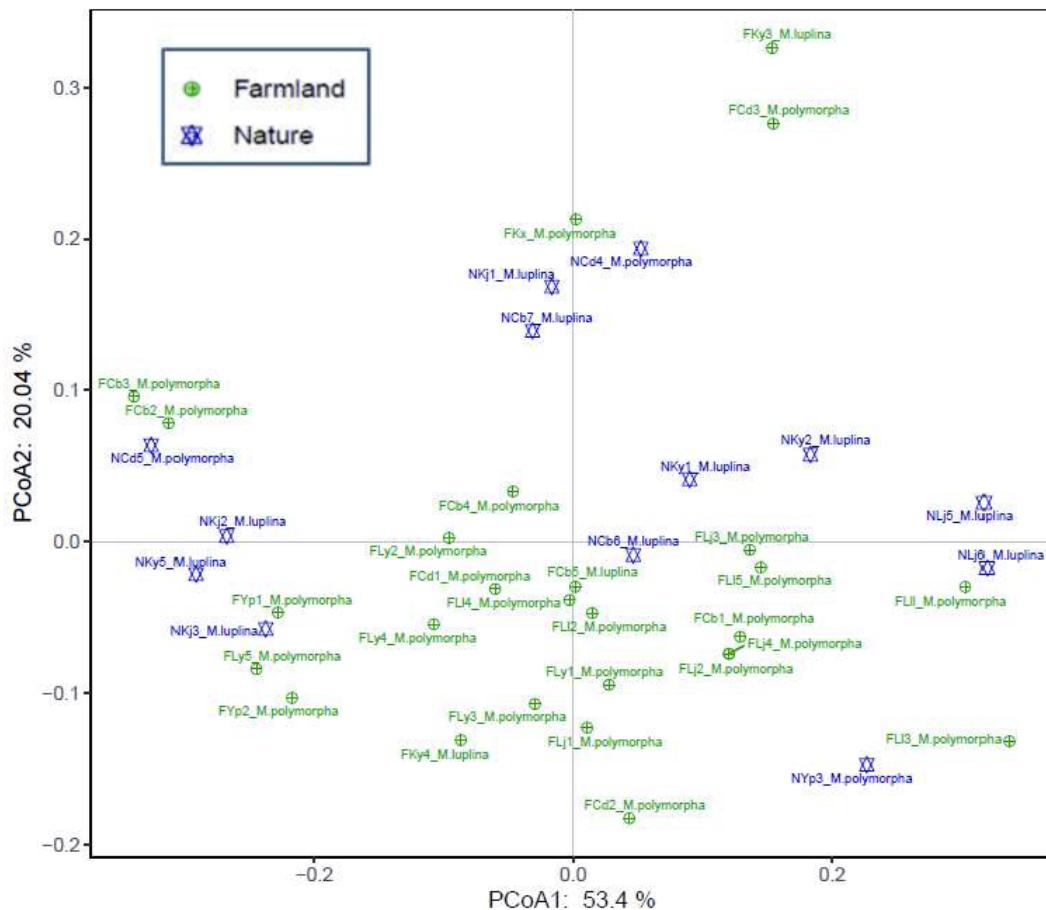


Figure 4

Principal coordinates analysis (PCoA) based on Bray-Curtis distance. The percentage of contribution to the variance(inertia) of the rhizobial diversity is given in parenthesis

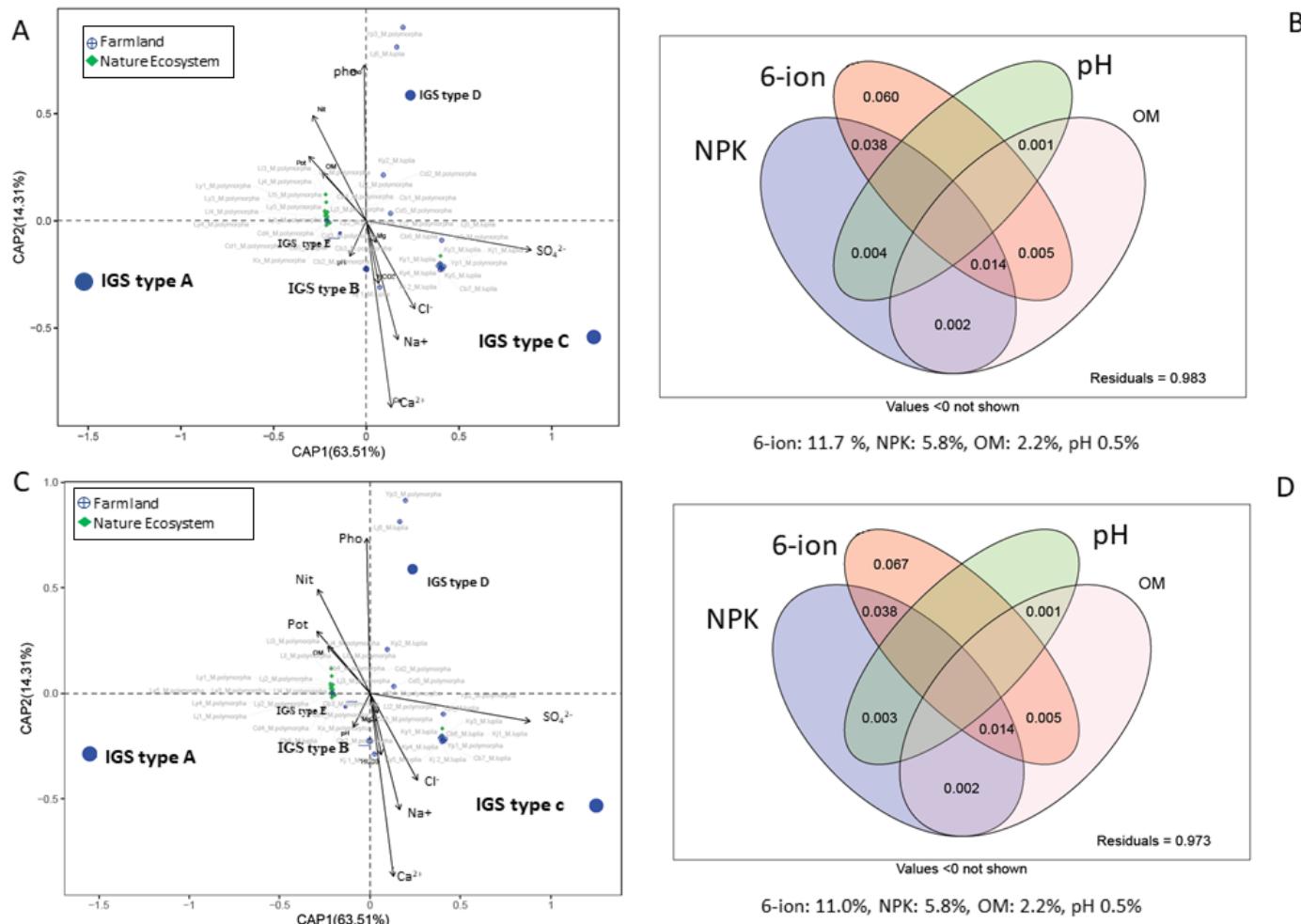


Figure 5

CAP (constrained analysis of principle coordinates) and variation partition analysis

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.accessionnumbersofsequencesforstrains.xlsx
- TableS2.Detailedinformationofsilsamples5.27.xlsx
- TableS3MeanANOVAofenvironmentconditions.xlsx
- TableS4.xlsx
- supplementalfigurecollectionPDF.pdf