Systematic and Applied Microbiology 42 (2019) 223–231

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: http://www.elsevier.com/locate/syapm

Genetic divergence among *Bradyrhizobium* strains nodulating wild and cultivated *Kummerowia* spp. in China

Zhaojun Ji[a,b,c], Tianyan Liu[b,c], Jixing Zhang[a], Hui Yan[b,c], Entao Wang[d], Qingguo Cui[e], Wenxin Chen[b,c], Wenfeng Chen[b,c,*]

[a] College of Life Science and Horqin Plant Stress Biology Research Institute, Inner Mongolia University for the Nationalities, Tongliao 028042, Inner Mongolia, China
[b] State Key Laboratory of Agrobiotechnology, Beijing 100193, China
[c] College of Biological Sciences and Rhizobium Research Center, China Agricultural University, Beijing 100193, China
[d] Departamento de Microbiología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, México D. F. 11340, México
[e] State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:
Received 9 August 2018
Received in revised form 7 October 2018
Accepted 9 October 2018

Keywords:
Bradyrhizobium
Kummerowia
Microevolution
Phylogeny
Recombination
Urban segregation

ABSTRACT

Distribution of rhizobial species is affected by geographical isolation and selected by leguminous hosts, however, little is known about the molecular evolution of rhizobia nodulating the same legume in different eco-environments. In present study, the microevolution of *Bradyrhizobium* associated with the leguminous grass *Kummerowia* grown in exurban areas and cultivated in urban areas in China was investigated. Total 14 genospecies, including seven new groups, were identified based on a concatenated sequence analysis of taxonomic markers (*SMc00019*, *truA* and *thrA*) for 94 representative strains. Results demonstrated that lower levels of nucleotide diversity were found in the strains isolated from urban areas compared with those isolated from exurban areas, based on the evolutional analyses of three housekeeping genes (*atpD*, *glnII* and *recA*), two symbiosis-related genes (*nodC* and *nifH*), and the taxonomic markers. Moreover, compared with urban areas, gene exchange and recombination occurred more frequently among the genospecies isolated from exurban areas, regardless of the geographical distribution. Finally, the evolutionary lineage of *Bradyrhizobium* strains isolated from urban areas was independent of that of the strains isolated from exurban areas. In summary, the evolutionary history of *Kummerowia* bradyrhizobia may have been gradually segregated to different evolutionary lineages, irrespective of distinct biogeography.

Introduction

Kummerowia stipulacea (Korean clover) and *Kummerowia striata* (Japanese clover) are legume grasses widely distributed along the roadside, hillside and lakeside as well as in other exurban areas habitats in China and other countries [22]. Because of its high oestrogenic activity and production of flavonoids, which are used for menopausal dysfunction treatment, anti-HIV and antitumor medicine, herbal *K. striata* has been deliberately cultivated [34,39]. Furthermore, *Kummerowia* has been grown as a flowering grass in university campuses in urban areas. Symbiotic nitrogen fixation of *Kummerowia* with rhizobia plays an important role in the healthy growth of the plant without the requirement of fertilisers. Rhizobial strains of *Kummerowia* have been recorded previously [19], and slow-growing strains, such as *Bradyrhizobium* spp., are the predominant microsymbionts in China. Some fast-growing strains grown in the Loess Plateau have been identified as *Sinorhizobium kummerowiae* [35,38], whereas some of those grown in campus lawns have been identified as *Rhizobium cauense* [20]. Furthermore, 14 reference strains (*Bradyrhizobium diazoefficiens*, *Bradyrhizobium canariense*, *Bradyrhizobium japonicum*, *Bradyrhizobium betae*, *Bradyrhizobium huanghuaihaiense*, *Bradyrhizobium daqingense*, *Bradyrhizobium liaoningense*, *Bradyrhizobium yuanmingense*, *Bradyrhizobium iriomotense*, *Bradyrhizobium elkanii*, *Bradyrhizobium pachyrhizi*, *Bradyrhizobium jicamae* and *Bradyrhizobium lablabi* and an outgroup genospecies *S. kummerowiae*) have been employed here to identify the phylogenetic relationship of these rhizobial strains isolated from nodules on the roots of *Kummerowia* distributed in exurban and urban areas [19].

---

[*] Corresponding author.
E-mail address: chenwf@cau.edu.cn (W. Chen).

https://doi.org/10.1016/j.syapm.2018.10.003

The biodiversity of the rhizobial strains of *Kummerowia* associated with biogeographic patterns has also been recently studied [19]; however, the genetic communication underlying their long-term evolutionary histories has not yet been elucidated. An important requirement is to systematically distinguish the rhizobial species and also the subdivisions of each species based on taxonomic markers, particularly when multi-genes are concatenated. Natural selection and genetic drift based on mutation, gene flow, rearrangement and recombination of the genome of these rhizobia in the microevolution process could then contribute to the understanding of the mechanisms of biodiversity generation and preservation. Rhizobial strains associated with *Kummerowia* isolated both from exurban and urban areas in the present study might be the best material to explore this microevolutionary mechanism. Recent studies have reported that the three taxonomic markers *SMc00019*, *truA* and *thrA* are useful for constructing phylogenetic trees with a gap of 2% for the similarity percentage of intraspecies/interspecies average nucleotide identity (ANI). This will support the predetermined species, thereby corresponding to a taxonomy system based on the rhizobial genomes [40]. However, the three housekeeping genes *atpD*, *glnII* and *recA*, which were often employed to differentiate species of rhizobia collected from different ecoregions [14,7], have not been previously compared with the taxonomic markers *SMc00019*, *truA* and *thrA* [30].

Here, evolutionary analyses of *Bradyrhizobium* strains nodulating *Kummerowia* collected from exurban and urban areas in China were performed by studying *SMc00019*, *truA* and *thrA* as taxonomic markers, *atpD*, *glnII* and *recA* (three housekeeping genes) as environmental condition adapters and *nodC* and *nifH* (two critical symbiosis-related genes) as host legume selectors. Biodiversity, genetic divergence, gene flow, recombination and evolutionary lineages were compared among the *Kummerowia* bradyrhizobia isolated from exurban and urban areas.

Materials and methods

Rhizobial strains

A total of 283 rhizobial isolates obtained from the root nodules of *Kummerowia* were collected from 28 sites (Supplementary Fig. S1) in south (covering southern, southwest, central and east China, 99 isolates, 6 provinces) and north (covering northern, northeast and northwest China, 184 isolates, 2 provinces) [19,8]. Geographically, these sites can be divided into two segregated ecoregions corresponding to exurban and urban areas, with 130 isolates and 153 isolates, respectively. These strains were previously preserved in the Culture Collection of Beijing Agricultural University (CCBAU) in tryptone-yeast extract medium [27] supplemented with 20% glycerol. Thirteen reference strains were used to determine the phylogenetic positions of the following *Kummerowia* bradyrhizobia: *B. betae* LMG 21987^T, *B. canariense* BTA-1^T, *B. daqingense* CCBAU 15774^T, *B. elkanii* USDA 76^T, *B. huanghuaihaiense* CCBAU 23303^T, *B. iriomotense* LMG 24129^T, *B. japonicum* USDA 6^T, *B. diazoefficiens* USDA 110^T, *B. jicamae* PAC68^T, *B. lablabi* CCBAU 23086^T, *B. liaoningense* USDA 3622^T, *B. pachyrhizi* PAC48^T and *B. yuanmingense* CCBAU 10071^T. And *S. kummerowiae* CCBAU 71714^T was employed as an outgroup genospecies to construct the phylogenetic trees. All these strains were grown in YMA medium at 28 °C.

Gene amplification and sequencing

The total template DNA of these strains was extracted using the method described by Terefework et al. [32]. The primer sequences and PCR protocols used for the amplification of *SMc00019*, *truA*, *thrA*, *atpD*, *glnII*, *recA*, *nodC* and *nifH* have been described previously [40,14,41]. The PCR products were purified and commercially sequenced using an ABI 3730xl sequenator (Beijing, China). All obtained sequences were manually checked using Chromas Pro software (V.1.7.6; Technelysium).

Phylogenetic analyses

The nucleotide sequences for all these genes were aligned and neighbour-joining (NJ) phylogenetic trees were constructed using CLUSTAL W integrated in MEGA5 [31] with a Kimura 2-parameter model. Genospecies were identified according to the similarity percentage of intraspecies ANI value (>96%) and interspecies ANI value (<94%) boundaries based on taxonomic markers (*SMc00019*, *truA* and *thrA*) for different species [40]. Next, the degree of tree-like structures for the alleles of each locus and concatenated sequences were assessed to reveal potentially incompatible signals in the evolutionary history with split phylogenetic networks (1000 bootstraps) using the SPLITSTREE 4.0 programme [12]. PHYML [5] was used in combination with MODELTEST 3.7 [24] for maximum likelihood (ML) tree construction, and the Shimodaira–Hasegawa test [26] was then performed to assess the topological consistency of trees inferred from different sets of sequence partitions as implemented in the PAUP* 4.0b1 programme [29]. Recombination effects observed among these populations during the phylogenetic history were judged using CLONALFRAME for five independent runs (500,000 burn-in iterations plus 1,000,000 sampling iterations for each run) based on three separate concatenated sequences [3].

Nucleotide polymorphism and population genetics analyses

Nucleotide polymorphisms, including the number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), number of synonymous substitutions per synonymous site (π_S), number of non-synonymous substitutions per non-synonymous site (π_N) and π_N/π_S ratio, were calculated using DNASP v5 software [18,21]. The average nucleotide divergence (Dxy) and number of migrants (Nm) were also calculated using DNASP v5 [18,21] to infer the genetic divergence and gene flow among these populations. Three different NeighborNet trees were also constructed to analyse the gene flow and genetic exchange among these rhizobial strains isolated from different ecoregions based on the taxonomic markers, housekeeping and symbiosis-related genes. CLONALFRAME was employed to calculate the two following recombination rate statistics: r/m (the relative impact of recombination compared with that of the point mutation in the genetic diversification of the lineage) and ρ/θ (the relative frequency of the occurrence of recombination compared with that of the point mutation in the history of the lineage) [3]. The DNASP software was used to calculate the minimal recombination events (Rm) for persuasively estimating the recombination within the populations and between the exurban and urban ecological environments [11]. Admixture levels of the bradyrhizobial genospecies inherited from the K 'ancestral' subpopulations were estimated using STRUCTURE with the admixture LOCPRIOR model [10,13].

Nucleotide sequence accession numbers

The 752 nucleotide sequences obtained in this study were deposited in the GenBank database (accession numbers presented in Supplementary Table S4). The other remaining 78 nucleotide sequences of the 13 type strains used in the phylogenetic trees were directly obtained directly from the GenBank database (Supplementary Table S5).

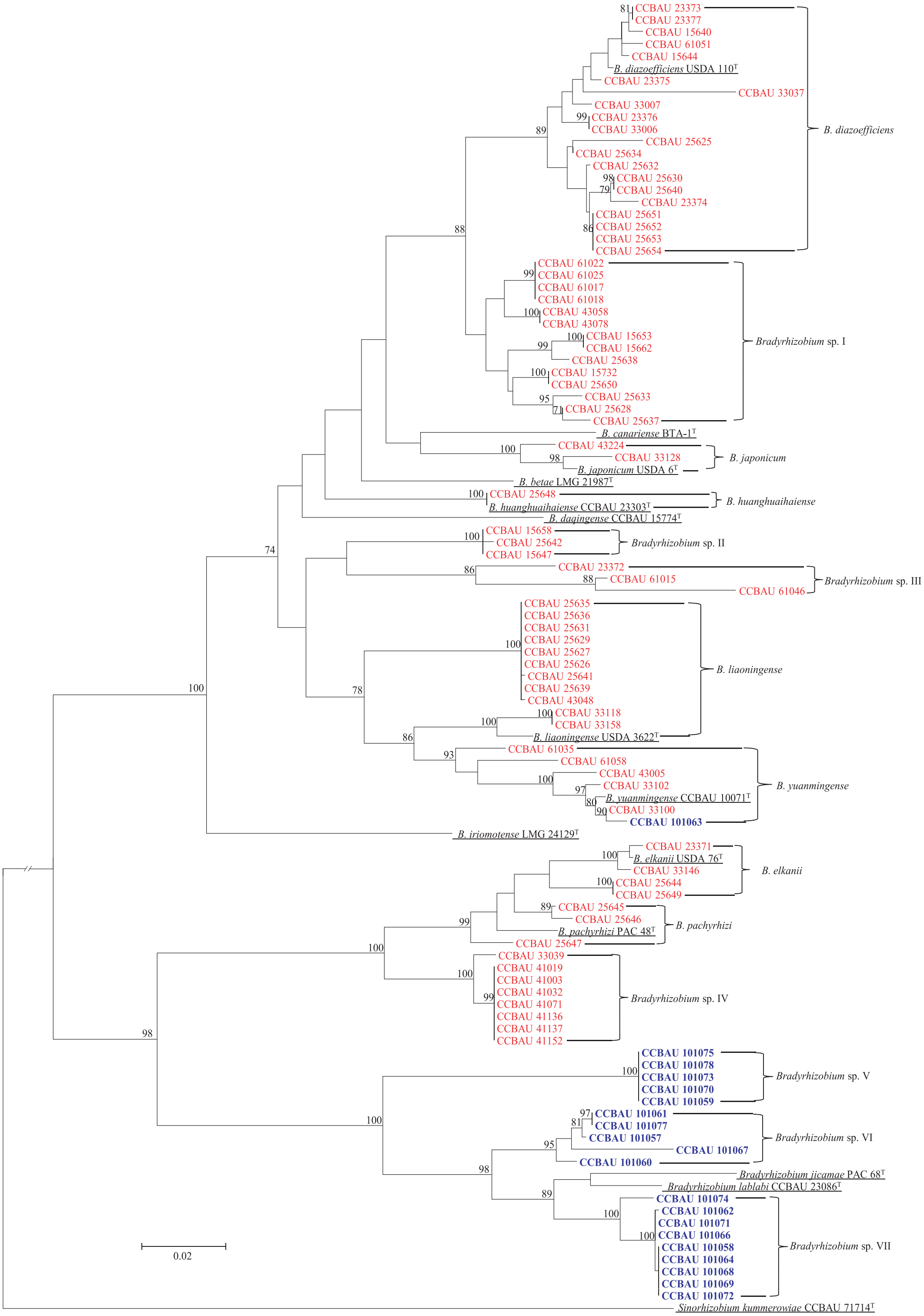

Fig. 1. Neighbor-Joining (NJ) tree constructed based upon the concatenated sequences of three taxonomic markers. Three genes *SMc00019*, *truA* and *thrA* were used as taxonomic markers. Bootstrap values greater than 70% are indicated at the branch points. Type (superscript T) strains underlined are used for references. Strains isolated from exurban and urban areas were highlighted with red and blue colours, respectively. The scale bar represents 1% nucleotide substitutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Results

Phylogenies of Bradyrhizobium *nodulate* Kummerowia *in China*

A total of 94 representative strains were selected from earlier studies of *Kummerowia* rhizobia (strain information shown in Table S1) based on the sequence analyses of the gene *thrA*, which is one of the useful phylogenetic and taxonomic markers (not shown). The topological structures of the NJ and ML phylogenetic trees showed no significant difference. In view of the NJ phylogenetic tree, a total of seven genospecies were identified as *B. diazoefficiens* (20 strains), *B. japonicum* (two strains), *B. huanghuaihaiense* (one strains), *B.*

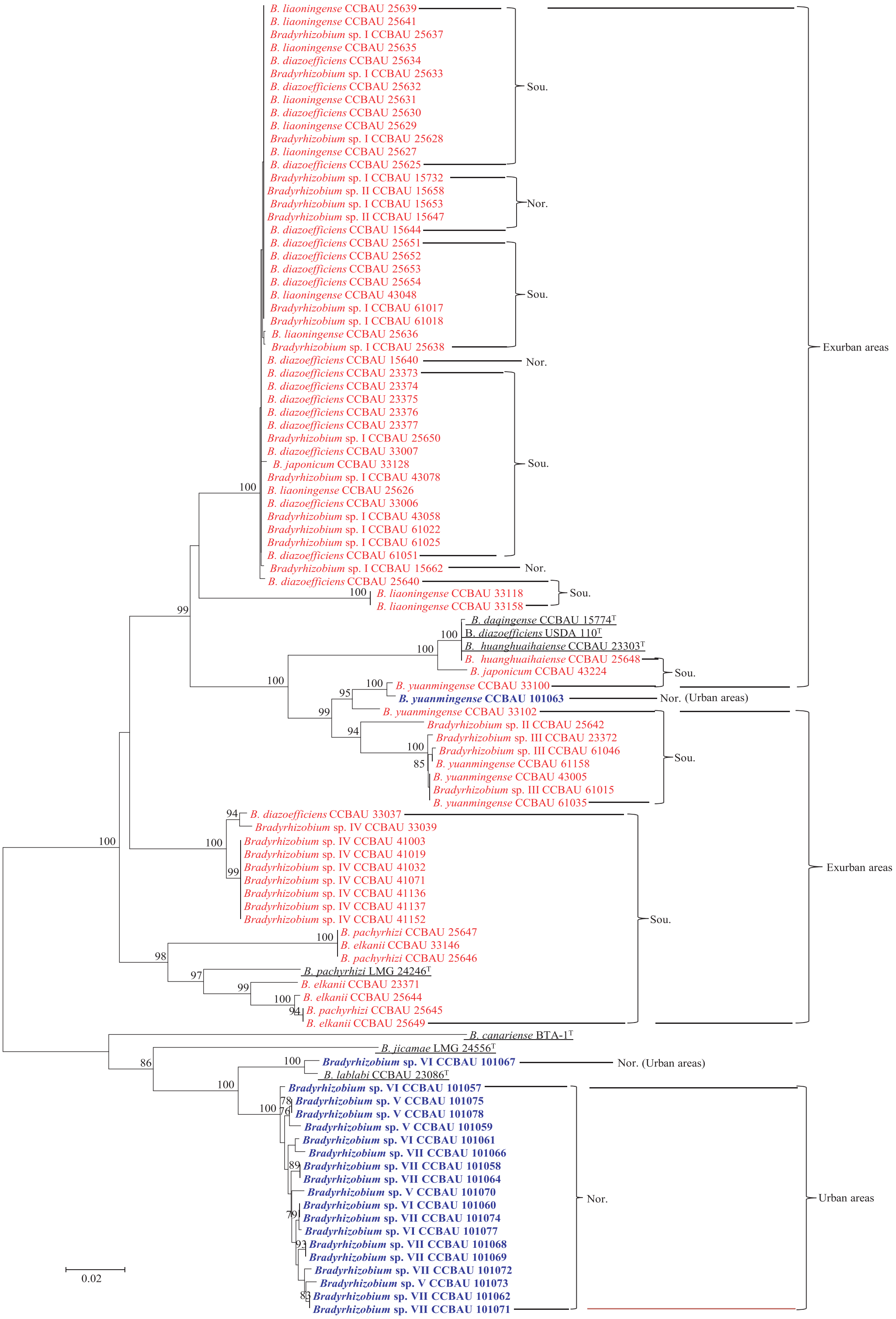

Fig. 2. Neighbor-Joining (NJ) tree constructed based upon the concatenated sequences of symbiosis-related genes. Two symbiosis-related genes (*nodC* and *nifH*) were used. Bootstrap values greater than 70% are indicated at the branch points. Sou. or Nor. indicates that the strains were isolated from southern or northern ecoregions of China, respectively. Type (superscript T) strains underlined are used for references. Strains isolated from exurban and urban areas were highlighted with red and blue colours, respectively. Sou. or Nor. indicates that the strains were isolated from southern or northern ecoregions of China, respectively. The scale bar represents 1% nucleotide substitutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

liaoningense (11 strains), *B. yuanmingense* (six strains), *B. elkanii* (four strains) and *B. pachyrhizi* (three strains), according to the similarity percentage of interspecies (<94%) and intraspecies (>96%) ANI value boundaries. And this values of another seven genospecies were identified as new genospecies (designated as *Bradyrhizobium* sp. I–VII). The topological structures of the taxonomic markers and the housekeeping genes were consistent for most strains, although very few variations could not be well grouped in their own genospecies (Fig. 1 and Supplementary Fig. S2). However, the phylogenetic tree was reconstructed for the concatenated sequences

Table 1
Nucleotide polymorphism for the three kinds of genes in *Kummerowia* bradyrhizobia.

Area (no. of strains)	Length (bp)	S	*Eta*	h/Hd	π	π_S	π_N	π_N/π_S
Concatenated taxonomic makers								
Exurban (74)	1212	396	512	47/0.977	0.09254	0.34082	0.03668	0.10762
Urban (20)	1212	284	331	10/0.884	0.07103	0.28939	0.02100	0.07257
Nor. (27)	1212	129	138	5/0.905	0.05198	0.20628	0.01025	0.04969
Sou. (67)	1212	395	507	43/0.973	0.09496	0.34800	0.03857	0.11083
Concatenated housekeeping genes								
Exurban (74)	1371	300	366	50/0.982	0.05494	0.20289	0.01580	0.07787
Urban (20)	1371	183	199	11/0.916	0.03816	0.15589	0.00568	0.03644
Nor. (27)	1371	83	88	5/0.905	0.03022	0.10819	0.00752	0.06951
Sou. (67)	1371	300	366	46/0.979	0.05637	0.20834	0.01632	0.07833
Concatenated symbiosis-related genes								
Exurban (74)	1083	320	403	24/0.839	0.05702	0.26290	0.00936	0.03560
Urban (20)	1083	217	236	15/0.974	0.02827	0.14134	0.00698	0.04938
Nor. (27)	1083	2	2	3/0.524	0.00070	0.00278	0	0
Sou. (67)	1083	320	402	23/0.860	0.06155	0.28488	0.01013	0.03556

Note: S, segregating sites or number of polymorphic (segregating) sites; *Eta*, total number of mutations. h, haplotype number; Hd, haplotype diversity; π, average number of nucleotide differences per site between two sequences; π_S, nucleotide diversity for synonymous substitutions; π_N, nucleotide diversity for nonsynonymous substitutions.

of the symbiotic genes of these strains, which were intermingled regardless of whether they were isolated from southern or northern China (Fig. 2). Interestingly, the strains with low levels of biodiversity isolated from urban areas were grouped together and distributed into three new genospecies, with the exception of strain *B. yuanmingense* CCBAU 101063. In contrast, high biodiversity was observed in strains isolated from exurban areas, which were distributed into seven previously identified and four new genospecies.

Table 2
Genetic divergence (presented as Dxy) and gene flow (presented as Nm) in the *Kummerowia* bradyrhizobia from exurban and urban areas.

Genes	Dxy	Nm
Concatenated taxonomic markers	0.16549^c	0.49
Concatenated housekeeping genes	0.08524^c	0.6
Concatenated symbiosis-related genes	0.15288^d	0.19

Note: Number of gene flow (Nm) and average nucleotide divergence between groups (Dxy) are shown in the upper and lower triangular of the table.
Statistical difference letter is marked as superscript after each number in the lower triangular of the table: b, $0.01 < P < 0.05$; c, $0.001 < P < 0.01$; d, $0.0001 < P < 0.01$; e, non-significant.

Nucleotide diversity inferred from different genes and ecological environments

Nucleotide polymorphisms of the 94 representative strains for the three kinds of genes (taxonomic markers, housekeeping and symbiosis-related genes) were calculated and the various parameters are listed in Supplementary Table S3. The sequence type and haplotype diversity (Hd) of symbiosis-related genes were obviously lower than those of the other genes. The nucleotide diversity (π) values showed little difference among the three taxonomic markers or housekeeping genes. However, the largest deviation of the π value (0.11886 for *nodC* and 0.06646 for *nifH*) was found for symbiosis-related genes (Supplementary Table S3); this was due to the large difference between their corresponding π_N (0.03299 for *nodC* and 0.00656 for *nifH*), and not their π_S values (0.35672 for *nodC* and 0.34485 for *nifH*), as shown in Supplementary Table S3. In addition, the strains isolated from urban areas had lower π, π_N and π_S values compared with those isolated from exurban areas for all the test genes (Table 1). And the highest π values were presented in exurban areas (0.09254) and southern China (0.09496) for the three taxonomic markers and the lowest π values were presented in urban areas (0.02827) and northern China (0.00070) for the symbiosis-related genes (Table 1).

The π_N/π_S values showed large differences (0.019 for *nifH* to 0.1819 for *SMc00019*) for the test genes (Table S3). The π_N/π_S values were higher in exurban than urban areas for the taxonomic markers and housekeeping genes; however, the π_N/π_S values were lower for the symbiosis-related genes. The π_N/π_S values were also higher for strains in southern China than for those in northern China for all the genes (Table 1).

Gene flow and genetic divergence among Bradyrhizobium *of* Kummerowia *in two segregated ecoregions*

Genetic exchange and gene flow occurred more frequently among 94 representative strains in the NeighborNet trees of the taxonomic markers (Fig. 3a) and symbiosis-related genes (Fig. 3b) than those of the housekeeping genes (Supplementary Fig. S3). Furthermore, the gene flow among strains isolated both from southern and northern China was frequent but seldom occurred between the strains from exurban and urban areas for the three kinds of test genes (Fig. 3 and S3).

The K value of 4 was chosen for each of the three kinds of concatenated genes with 100,000 burn-in and 1,000,000 sampling iterations in STRUCTURE analyses and the strains were passively arranged in order on the basis of genetic exchange (Fig. 4). Four lineages (I–IV) showing clear and similar structure patterns were observed in lineage I and II for the taxonomic markers and housekeeping genes, and in lineage III and IV for taxonomic markers and symbiosis-related genes, respectively. In addition, the gene flow between lineages I and II occurred frequently for the taxonomic markers and housekeeping genes. In contrast, all the isolates obtained from urban areas were located in one of the lineages that seldom intermingled with the isolates from exurban areas for all kinds of genes.

As shown in Tables 2 and 3, the highest genetic distances (Dxy) were detected and the p-values for the Dxy comparison pairs were statistically significant ($p < 0.05$) between exurban and urban areas for the taxonomic markers, housekeeping and symbiosis-related genes (0.16549, 0.08524 and 0.15288, respectively) and between southern and northern China (0.04931 and 0.03823 for the taxonomic markers and housekeeping genes, respectively). However, no significant difference was found for the symbiosis-related genes (0.03823) between the pairs of isolates obtained from southern and northern China ($p < 0.05$). Moreover, the Dxy values between exurban and urban areas (0.16549, 0.08524 and 0.15288 for the taxonomic markers, housekeeping and symbiosis-related genes, respectively) were much higher than the mean π value (0.08179, 0.04655 and 0.04265 for the taxonomic markers, housekeeping

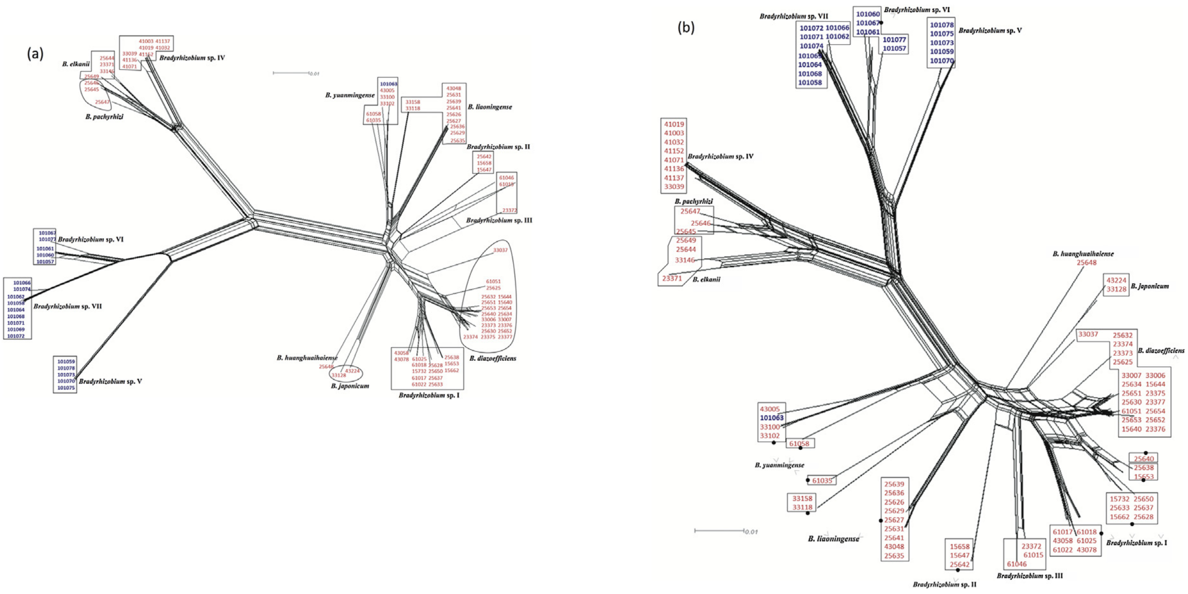

Fig. 3. The Neighbor-Nets generated using SplitsTree4 with the Hamming distance option. The trees are constructed based on taxonomic markers (a) and symbiosis-related gene sequences. (b) Strains isolated from exurban and urban areas were highlighted with red and blue colours, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

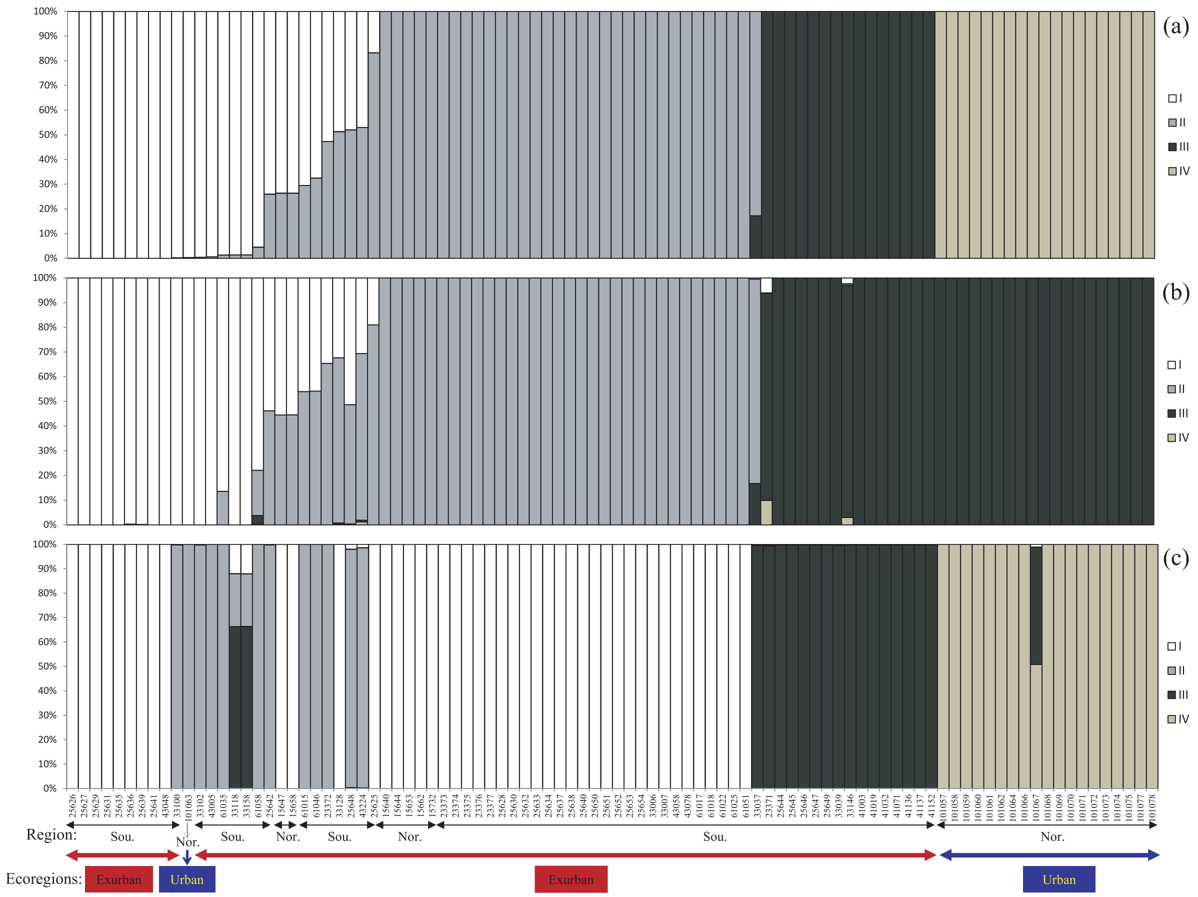

Fig. 4. STRUCTURE analyses of *Kummerowia*-associated *Bradyrhizobium* populations from different areas. Concatenated taxonomic markers, (a) housekeeping genes (b) and symbiosis-related genes (c) were analyzed. The inferred ancestries are sequentially designated as I, II, III and IV shown in bars and filled with different grayness or dots. The horizontal axis represents current *Bradyrhizobium* individuals (in the same order in all panels), and the bar for each individual is filled according to the inferred proportions of single-nucleotide alleles that were derived from each of the ancestry. The vertical axis represents the percentage of the strains. Sou. or Nor. indicates that the strains were isolated from southern or northern ecoregions of China respectively.

Table 3
Recombination analysis by DNASP and CLONALFRAME.

Genes	Length (bp)	Rm	r/m	ρ/θ
Concatenated taxonomic markers	1212	118	1.39	0.29
Exurban	1212	102	0.72	0.12
Urban	1212	40	0.62	0.07
Concatenated housekeeping genes	1371	93	2.52	0.67
Exurban	1371	85	2.30	0.54
Urban	1371	18	0.39	0.062
Concatenated symbiosis-related genes	1083	82	0.17	0.01
Exurban	1083	61	0.047	0.0029
Urban	1083	10	0.012	0.00017

Note: Rm, observed minimum number of recombination events; r/m, the relative impact of recombination compared with that of point mutation in the genetic diversification of the lineage; ρ/θ, the relative frequency of the occurrence of recombination compared with that of point mutation in the history of the lineage.

and symbiosis-related genes, respectively) (Table 2). Nevertheless, the Dxy values between southern and northern China (0.08296, 0.04931 and 0.03823 for the taxonomic markers, housekeeping and symbiosis-related genes respectively) were close to the mean π values (0.07347, 0.04330 and 0.03113 for the taxonomic markers, housekeeping and symbiosis-related genes, respectively; Table S6).

Recombination and evolutionary lineages of Bradyrhizobium *of* Kummerowia *in two segregated ecoregions*

A total of 118, 93 and 82 putative recombination events (Rm) were inferred for the concatenated taxonomic markers, housekeeping and symbiosis-related genes, respectively, with DNASP software (Table 3). The symbiosis-related genes had the lower Rm value than other genes and lower levels of recombination rates in evolutionary history.

Recombination rates in evolutionary history were further assessed by calculating the r/m and ρ/θ ratios using the CLONALFRAME programme. Compared with the occurrence of point mutations, the relative frequency of the occurrence of recombination (ρ/θ) was less than 1.0 for all the test genes in the history of the lineage (Table 3). However, compared with that of the point mutation (r/m), the relative impact of the recombination was greater than 1.0 for the taxonomic markers and housekeeping genes but not for the symbiosis-related genes. In addition, the Rm, ρ/θ and r/m values of *Kummerowia* bradyrhizobia isolated from urban areas were much lower than those isolated from exurban areas for all of the test genes.

Compared with *SMc00019, truA* and *thrA*, the topological structures were significantly incongruent with the phylogeny of the inferred species for all the test genes except for *thrA* ($p < 0.05$), as revealed by the Shimodaira–Hasegaewa test performed using the Modeltest 3.7 programme integrated in PAUP software (Table S7).

Discussion

Kummerowia, a legume, usually grows wildly in nature; however, human have recently started cultivating it. *Bradyrhizobium* is the main symbiotic species of *Kummerowia*; it has high nitrogen fixation abilities. *Bradyrhizobium* can form nodules on the roots of *Kummerowia* in the soil. We systematically analysed the evolutionary relationship of *Bradyrhizobium* based on three taxonomic markers, three housekeeping genes and two symbiosis-related genes. The results showed that low levels of biodiversity, infrequent gene exchange, infrequent recombination and independent evolutionary lineage of *Bradyrhizobium* strains nodulating *Kummerowia* gradually evolved and emerged with urban segregation.

Low level of biodiversity for Bradyrhizobium *strains nodulating* Kummerowia *with urban segregation*

The *Bradyrhizobium* strains nodulating *Kummerowia* obtained and isolated in previous studies [19,8] were all covered here. The high genetic diversity of the 14 genospecies is indicated on the phylogenetic tree based on three genes *SMc00019*, *thrA* and *truA*, which were also previously introduced as taxonomic markers to differentiate the relationship between intraspecific and interspecific genomes among rhizobial strains [40]. The high *Hd* (0.981) and π (0.11631) values presented in Table S3 also indicate the high biodiversity of *Kummerowia* bradyrhizobia. In addition, the phylogenetic trees of the concatenated sequences of the housekeeping genes *atpD*, *glnII* and *recA*, which are also taxonomic markers [33], were similar to those of *SMc00019*, *thrA* and *truA*, which mirrored the high genetic diversity of *Kummerowia* bradyrhizobia. Nonetheless, many strains in species *Bradyrhizobium* sp. I and *B. elkanii* were separated, and the strains *B. diazoefficiens* CCBAU 33037 and *B. yuanmingense* CCBAU 61035 were located far away from their own groups (Fig. 1 and S2). For the symbiosis-related genes, the genetic diversity was relatively lower and the *Bradyrhizobium* strains belonging to different genospecies were embraced on the phylogenetic trees (Fig. 2); this was inferred through the selective communication between legumes and rhizobia. The incongruences of the phylogenetic trees simultaneously revealed by the SH test of single genes compared with the phylogeny of the three concatenated taxonomic markers indicated that the evolutionary direction of each gene significantly differed although they belong to the same genospecies.

Interestingly, many *Bradyrhizobium* strains isolated from exurban areas, including those isolated from southern (six provinces) and northern (one province) China, were closely located and mingled together based on three kinds of test genes, which was highlighted an inconsistency in the biogeographical patterns studied previously for legume-nodulating *Rhizobium*, *Ensifer*, *Burkholderia* and other species [2,16,1,17]. However, an independent group was formed by distributing the strains isolated from urban areas, thereby indicating that the genetic relationships of *Bradyrhizobium* strains were greatly affected by the local environmental variables of wild or cultivated *Kummerowia* with human intervention (known as urban segregation) than by geographic isolation [28] and that compared with strains isolated from exurban areas, those isolated from urban areas will be more easily involved in convergent evolution in the future. However, the strain *B. yuanmingense* CCBAU 101063 isolated from urban areas embraced the strains isolated from exurban areas on the phylogenetic trees for all the test genes and was inferred as one of the few persistent native rhizobia without radical genetic divergence. Thus, the biodiversity of *Bradyrhizobium* strains being driven to lower levels might be attributable to the legume *Kummerowia*, which has suffered urban segregation and artificial cultivation by humans.

Infrequent gene flow and recombination among Kummerowia Bradyrhizobium *with urban segregation*

The gene flow among rhizobial strains isolated from the roots of *Glycine max* [41,6], *Arachis hypogaea* [1], *Medicago sativa* [9], *Caragana* [14] and other legumes was frequent for the core genes or symbiosis-related genes, regardless of the genospecies and geographical distribution. However, bradyrhizobial strains of *Kummerowia* isolated from urban areas in this study had less genetic communication with the isolates from exurban areas, as revealed by the lower Nm values between exurban and urban areas in southern and northern China. Furthermore, artificial cultivation and urban segregation could critically lower the gene flow frequencies among *Kummerowia* bradyrhizobia for all the test genes. This finding is supported by previous studies that have investigated rhizobia associated with *Astragalus* [37,36]. Moreover, due to the high Dxy values between exurban and urban areas (Table 2) in southern and northern China, genetic divergence among *Kummerowia* bradyrhizobia was strongly affected by urban segregation, rather than by geographical distribution (Table S6).

Apart from gene flow, recombination is an important impetus of evolutionary history, and has recently been shown to contribute the diversity of rhizobia comparable with or greater than that of mutation. For example, the higher Rm and r/m (>1.0) values indicate that recombination, and not mutation, plays a vital role in the evolutionary history of the taxonomic markers and housekeeping genes. Nevertheless, all the ρ/θ and r/m values calculated in urban areas were far lower than those calculated in exurban areas, particularly for symbiosis-related genes, thereby indicating a lower frequency of recombination among *Kummerowia* bradyrhizobia in urban areas. This observation is consistent with the significantly lower Rm values observed in urban areas. Therefore, recombination and genetic drift events frequently occurred among the strains under stronger natural selection in exurban areas [23]; however, neutral mutation (as the critical force) occurred in urban areas [25,4].

Independent evolutionary lineage of Kummerowia *bradyrhizobia evolved with urban segregation*

Rhizobial strains isolated from the root nodules of *Kummerowia* in the present study consisted of four lineages (I–IV) for the three kinds of genes (Fig. 4a–c). For the taxonomic markers and symbiosis-related genes, lineages I–III were primarily distributed in exurban areas, although the patterns between them largely differed. Lineage IV, an independent branch, was almost exclusively located in urban areas, with the exception of strain CCBAU 101063. Horizontal gene transfer was also rarely observed between *Kummerowia* bradyrhizobia isolated from exurban and urban areas in the results of the STRUCTURE analyses (Fig. 4). This indicates that the evolutionary lineages of *Kummerowia* bradyrhizobia independently evolved for the taxonomic markers and symbiosis-related genes. Strain CCBAU 101063 might be one of the rhizobial populations that has preserved inherent efficient symbiosis and nitrogen fixation on host plants and strong viability, whereas other strains isolated from urban areas as the dominant populations in lineage IV might have had higher efficient symbiosis and nitrogen fixation or stronger viability [15]. By contrast, lineages I–III showed very similar results for the taxonomic markers and housekeeping genes, regardless of the representative strains isolated from southern or northern China (Fig. 4). This indicates that an independent lineage of the rhizobia would gradually evolve in the history, primarily due to urban segregation rather than biogeographic separation, which has already been established to be a critical factor related to microbial distribution [14,2,16].

Conclusion

In conclusion, our results confirm that the diversity of rhizobial strains associated with *Kummerowia* has drastically reduced in urban environmental conditions, although the formation of rhizobial communities is generally affected not only by biogeographic separation but also by their host legumes in natural wild environments. Gene exchange and recombination were frequent within the exurban or urban areas separately but seldom occurred between the two ecoregions based on the analyses of three types of genes. Furthermore, a single evolutionary lineage gradually formed for the *Bradyrhizobium* species of cultivated *Kummerowia* with urban segregation.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (31500001, 31860001 and 31770039) and the Natural Science Foundation of Inner Mongolia Autonomous Region of China (2017MS(LH)0306).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.syapm.2018.10.003.

References

[1] Chen, J., Hu, M., Ma, H., Wang, Y., Wang, E.T., Zhou, Z., Gu, J. (2016) Genetic diversity and distribution of bradyrhizobia nodulating peanut in acid-neutral soils in Guangdong Province. Syst. Appl. Microbiol. 39, 418–427.

[2] Chidebe, I.N., Jaiswal, S.K., Dakora, F.D. (2018) Distribution and phylogeny of microsymbionts associated with cowpea (*Vigna unguiculata*) nodulation in three agroecological regions of Mozambique. Appl. Environ. Microbiol. 84.

[3] Didelot, X., Falush, D. (2007) Inference of bacterial microevolution using multilocus sequence data. Genetics 175, 1251–1266.

[4] Fraser, C., Hanage, W.P., Spratt, B.G. (2005) Neutral microepidemic evolution of bacterial pathogens. Proc. Natl. Acad. Sci. U S A 102, 1968–1973.

[5] Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

[6] Gyogluu, C., Jaiswal, S.K., Kyei-Boahen, S., Dakora, F.D. (2018) Identification and distribution of microsymbionts associated with soybean nodulation in Mozambican soils. Syst. Appl. Microbiol. 41, 506–515.

[7] Han, L.L., Wang, E.T., Han, T.X., Liu, J., Sui, X.H., Chen, W.F., Chen, W.X. (2009) Unique community structure and biogeography of soybean rhizobia in the saline-alkaline soils of Xinjiang, China. Plant Soil 324, 291–305.

[8] Hao, W., ChaoXin, M., EnTao, W., WenXin, C. (2009) Diversity of rhizobia and interactions among the host legumes and rhizobial genotypes in an agricultural-forestry ecosystem. Plant Soil 314, 169–182.

[9] Harrison, T.L., Wood, C.W., Heath, K.D., Stinchcombe, J.R. (2017) Geographically structured genetic variation in the *Medicago lupulina-Ensifer* mutualism. Evol. Int. J. Org. Evol. 71, 1787–1801.

[10] Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K. (2009) Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9, 1322–1332.

[11] Hudson, R.R., Kaplan, N.L. (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111, 147–164.

[12] Huson, D.H., Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.

[13] Jakobsson, M., Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics (Oxford, England) 23, 1801–1806.

[14] Ji, Z.J., Yan, H., Cui, Q.G., Wang, E.T., Chen, W.F., Chen, W.X. (2015) Genetic divergence and gene flow among *Mesorhizobium* strains nodulating the shrub legume *Caragana*. Syst. Appl. Microbiol. 38, 176–183.

[15] Ji, Z.J., Yan, H., Cui, Q.G., Wang, E.T., Chen, W.F., Chen, W.X. (2017) Competition between rhizobia under different environmental conditions affects the nodulation of a legume. Syst. Appl. Microbiol. 40, 114–119.

[16] Lemaire, B., Chimphango, S.B., Stirton, C., Rafudeen, S., Honnay, O., Smets, E., Chen, W.M., Sprent, J., James, E.K., Muasya, A.M. (2016) Biogeographical patterns of legume-nodulating *Burkholderia* spp.: from African fynbos to continental scales. Appl. Environ. Microbiol. 82, 5099–5115.

[17] Li, Y., Yan, J., Yu, B., Wang, E.T., 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, 5294–5300.

[18] Librado, P., Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics (Oxford, England) 25, 1451–1452.

[19] Lin, D.X., Man, C.X., Wang, E.T., Chen, W.X. (2007) Diverse rhizobia that nodulate two species of *Kummerowia* in China. Arch. Microbiol. 188, 495–507.

[20] Liu, T.Y., Li, Y., Jr., Liu, X.X., Sui, X.H., Zhang, X.X., Wang, E.T., Chen, W.X., Chen, W.F., Pulawska, J. (2012) *Rhizobium cauense* sp. nov., isolated from root nodules of the herbaceous legume *Kummerowia stipulacea* grown in campus lawn soil. Syst. Appl. Microbiol. 35, 415–420.

[21] Nei, M., Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.

[22] Park, K.Y., Lee, S.H., Min, B.K., Lee, K.S., Choi, J.S., Chung, S.R., Min, K.R., Kim, Y. (1999) Inhibitory effect of luteolin 4′-O-glucoside from *Kummerowia striata* and other flavonoids on interleukin-5 bioactivity. Planta Med. 65, 457–459.

[23] Petren, K., Grant, P.R., Grant, B.R., Clack, A.A., Lescano, N.V. (2010) Multilocus genotypes from Charles Darwin's finches: biodiversity lost since the voyage of the Beagle. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365, 1009–1018.

[24] Posada, D., Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics (Oxford, England) 14, 817–818.

[25] Sella, G., Petrov, D.A., Przeworski, M., Andolfatto, P. (2009) Pervasive natural selection in the *Drosophila* genome?PLoS Genet. 5, e1000495.

[26] Shimodaira, H., Hasegawa, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

[27] Somasegaran, P., Hoben, H.J. 1994 Handbook for Rhizobia, Methods in Legume-Rhizobium Technology, Springer-Verlag, New York.

[28] Stefan, A., Van Cauwenberghe, J., Rosu, C.M., Stedel, C., Labrou, N.E., Flemetakis, E., Efrose, R.C. (2018) Genetic diversity and structure of *Rhizobium leguminosarum* populations associated with clover plants are influenced by local environmental variables. Syst. Appl. Microbiol. 41, 251–259.

[29] Swofford, D.L. 1998 PAUP*. Phylogenetic Analysis Using Parsimony, version 4, Sinauer Associates, Sunderland, MA.

[30] Tampakaki, A.P., Fotiadis, C.T., Ntatsi, G., Savvas, D. (2017) A novel symbiovar (aegeanense) of the genus *Ensifer* nodulates *Vigna unguiculata*. J. Sci. Food Agric. 97, 4314–4325.

[31] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

[32] Terefework, Z., Kaijalainen, S., Lindstrom, K. (2001) AFLP fingerprinting as a tool to study the genetic diversity of *Rhizobium galegae* isolated from *Galega orientalis* and *Galega officinalis*. J. Biotechnol. 91, 169–180.

[33] Tong, W., Li, X., Huo, Y., Zhang, L., Cao, Y., Wang, E., Chen, W., Tao, S., Wei, G. (2018) Genomic insight into the taxonomy of *Rhizobium* genospecies that nodulate *Phaseolus vulgaris*. Syst. Appl. Microbiol. 41, 300–310.

[34] Wang, H.K., Xia, Y., Yang, Z.Y., Natschke, S.L., Lee, K.H. (1998) Recent advances in the discovery and development of flavonoids and their analogues as antitumor and anti-HIV agents. Adv. Exp. Med. Biol. 439, 191–225.

[35] Wei, G.H., Wang, E.T., Tan, Z.Y., Zhu, M.E., Chen, W.X. (2002) *Rhizobium indigoferae* sp. nov. and *Sinorhizobium kummerowiae* sp. nov., respectively isolated from *Indigofera* spp. and *Kummerowia stipulacea*. Int. J. Syst. Evol. Microbiol. 52, 2231–2239.

[36] Wei, G.H., Zhang, Z.X., Chen, C., Chen, W.M., Ju, W.T. (2008) Phenotypic and genetic diversity of rhizobia isolated from nodules of the legume genera *Astragalus*, *Lespedeza* and *Hedysarum* in northwestern China. Microbiol. Res. 163, 651–662.

[37] Yan, H., Ji, Z.J., Jiao, Y.S., Wang, E.T., Chen, W.F., Guo, B.L., Chen, W.X. (2016) Genetic diversity and distribution of rhizobia associated with the medicinal legumes *Astragalus* spp. and *Hedysarum polybotrys* in agricultural soils. Syst. Appl. Microbiol. 39, 141–149.

[38] Yao, L.J., Shen, Y.Y., Zhan, J.P., Xu, W., Cui, G.L., Wei, G.H. (2012) *Rhizobium taibaishanense* sp. nov., isolated from a root nodule of *Kummerowia striata*. Int. J. Syst. Evol. Microbiol. 62, 335–341.

[39] Yoo, H.H., Kim, T., Ahn, S., Kim, Y.J., Kim, H.Y., Piao, X.L., Park, J.H. (2005) Evaluation of the estrogenic activity of Leguminosae plants. Biol. Pharm. Bull. 28, 538–540.

[40] Zhang, Y.M., Tian, C.F., Sui, X.H., Chen, W.F., Chen, W.X. (2012) Robust markers reflecting phylogeny and taxonomy of rhizobia. PLoS One 7, e44936.

[41] Zhang, X.X., Guo, H.J., Wang, R., Sui, X.H., Zhang, Y.M., Wang, E.T., Tian, C.F., Chen, W.X. (2014) Genetic divergence of *Bradyrhizobium* strains nodulating soybeans as revealed by multilocus sequence analysis of genes inside and outside the symbiosis island. Appl. Environ. Microbiol. 80, 3181–3190.