Population genetic structure of *Sinorhizobium meliloti* and *S. medicae* isolated from nodules of *Medicago* spp. in Mexico

Claudia Silva, Feng Ling Kan & Esperanza Martínez-Romero

Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Correspondence: Claudia Silva, Centro de Ciencias Genómicas, UNAM, Programa de Ecología Genómica, AP 565A, Cuernavaca, Morelos, Mexico. Tel.: +52 777 3131697; fax: +52 777 3175581; e-mail: csilva@ccg.unam.mx

Received 19 June 2006; revised 11 December 2006; accepted 11 January 2007.
First published online 26 March 2007.

DOI:10.1111/j.1574-6941.2007.00301.x

Editor: Kornelia Smalla

Keywords
genetic structure; *Sinorhizobium*; *Medicago*; alfalfa; migration; recombination.

Abstract

We studied the genetic structure of 176 bacterial isolates from nodules of *Medicago sativa*, *M. lupulina* and *M. polymorpha* in fifteen sites distributed in three localities in Mexico. The strains were characterized by multilocus enzyme electrophoresis, plasmid profiles, PCR restriction fragment length polymorphism of 16S rRNA genes and of the intergenic spacer between 16S and 23S rRNA genes, and partial sequences of *glnII*, *recA* and *nodB*. Most of the strains were classified as *Sinorhizobium meliloti*, and a high genetic diversity was recorded. Six strains were classified as *Sinorhizobium medicae*, with no genetic variation. Phylogenetic and population genetic analyses revealed evidence of frequent recombination and migration within species.

Introduction

The genus *Medicago* comprises more than 60 plant species, two-thirds of which are annuals and one-third perennials (Lesins & Lesins, 1979). Members of this genus, native to Eurasia and Africa, especially in the Mediterranean area, are now widespread in temperate regions (Allen & Allen, 1981). The most widely cultivated perennial species is *Medicago sativa* (alfalfa), which is mainly used as forage for domestic animals. Alfalfa was introduced to the Americas by Europeans several hundred years ago. Spaniards took from Mexico many products, such as maize and beans, and introduced others, such as alfalfa, as forage for their cows and horses. Nowadays in Mexico extensive areas are grown with alfalfa, which is the main forage crop in the semi-arid regions of north and central Mexico (www.fao.org). Large colonization and trading enterprises have always been accompanied by the dispersal of many useful plants; beyond the intended transport of valuable plants, humans unknowingly carry with them organisms such as plants, animals and microorganisms, frequently as contaminants of their products. In addition to alfalfa, other annual *Medicago* species are reported in the list of alien flowering plants recorded in Mexico (Villaseñor & Espinosa-García, 2004).

Medicago species establish nitrogen-fixing symbioses with fast-growing rhizobia, and a moderate to high degree of host specificity has been recorded (Brockwell & Hely, 1966; Sprent, 2001). Using multilocus enzyme electrophoresis (MLEE), Eardly *et al.* (1990) analysed the genetic structure of a large collection of bacteria isolated from nodules from perennial and annual species of *Medicago* coming from several distinct geographic regions distributed on four continents. This pioneer study revealed the existence of two deeply diverging phylogenetic divisions, one of which appears to be adapted to annual medic species. Moreover, the analyses of linkage disequilibrium showed evidence of genetic exchange within each group but not across groups, suggesting the existence of a genetic barrier between lineages. Further studies established that the two genetic groups correspond to distinct species, namely *Sinorhizobium meliloti* and *S. medicae* (Rome *et al.*, 1996a, b), the latter being adapted to annual medics.

In Mexico there is a diversity of native *Sinorhizobium* species. Novel species have been recognized that nodulate legume trees in the genera *Leucaena* (Wang *et al.*, 2002) and *Acacia* (Toledo *et al.*, 2003; Lloret *et al.*, 2007); there is no information, however, about the diversity of rhizobia nodulating *Medicago* species. The work presented here is the first study to analyse the nitrogen-fixing bacteria associated with *Medicago* in Mexico. This information is relevant to programs directed towards improving crop productivity through biofertilization with locally adapted or genetically

modified strains. Because *Medicago* species are not native to Mexico we hypothesized that they could be nodulated by (1) native *Sinorhizobium* populations, (2) introduced *S. meliloti* and *S. medicae*, or (3) native *Sinorhizobium* populations that acquired symbiotic genes from *S. meliloti* or *S. medicae*. In this work we used phylogenetic and population genetic analyses to analyse the genetic structure of a large collection of bacteria isolated from cultivated *M. sativa* nodules. In addition, we compared the alfalfa strains with isolates from the annual species *M. lupulina* and *M. polymorpha*, which grow as weeds in Mexico. Fifteen sites distributed in three localities of central Mexico were sampled. The genetic variation was hierarchically assessed using MLEE, plasmid profiles, PCR restriction fragment length polymorphism (PCR-RFLP) of 16S rRNA genes and of the intergenic spacer (ITS) between the 16S and 23S rRNA genes, including partial sequences of two housekeeping genes (*glnII* and *recA*) and the symbiotic gene *nodB*. We discuss our results in the light of biogeographical hypotheses about the origin of the strains and the impact of the evolutionary forces in shaping their genetic structure.

Materials and methods

Sampling and strain isolation

Fifteen sites distributed in three localities of central Mexico (Guanajuato, Texcoco and Cuernavaca) were sampled. Strains from Guanajuato were isolated from nodules of field-grown alfalfa plants (*M. sativa*) or alfalfa plants grown in the laboratory using soil from the different sites. The twelve sites sampled in Guanajuato (designated G1 to G12) were agricultural fields where alfalfa has been grown for more than 12 years; the plants were about $3-5$ years old, and no inoculation and low levels of nitrogen fertilization ($20–40\,kg\,ha^{-1}$) had been applied. An agricultural plot in Texcoco was sampled (named Texcoco site 1) to address the importance of geographic distance on the diversity and genetic structure of alfalfa symbionts, the site being *c.* 400 km from Guanajuato. In order to compare the structure of the symbionts from a cultivated perennial *Medicago* with those of an annual weedy species, a population of *M. polymorpha* located *c.* 2 km from Texacoco site 1, at a site named Texacoco 2, was sampled. These sites are *c.* 100 km apart from the third sampling location in Cuernavaca, where isolates were from *M. lupulina* (black medic) growing as a weed. At each site at least five plants were sampled. Isolates were recovered using a standard method (Vincent, 1970) on peptone-yeast extract (PY) medium.

MLEE, genetic diversity and cluster analysis

Cell lysates were obtained as described previously (Silva *et al.*, 1999), and electrophoretic separation on 12.6% starch gels and staining of the enzymes was performed according to the procedures described by Selander *et al.* (1986). Nine metabolic enzymes were evaluated: aldolase (ALD), aconitase (ACO), glucose-6-phosphate dehydrogenase (G6PDH), hexokinase (HEX), isocitrate dehydrogenase (IDH), malate dehydrogenase (MDH), malic enzyme (ME), phosphoglucose isomerase (PGI) and phosphoglucomutase (PGM).

Distinctive mobility variants of each enzyme were equated with alleles at the corresponding locus. The combined allele profiles for the nine enzymes were defined as multilocus genotypes (electrophoretic types or ETs). Based on allele frequencies for ETs, the genetic diversity for an enzyme locus was calculated as $H = (1 - \Sigma x_i^2)\ (n/(n–1))$, where x_i is the frequency of the ith allele and n is the number of ETs. The total mean genetic diversity (H) is the arithmetic mean of H values across loci (Selander *et al.*, 1986). To compute the diversity values the program ETDIV version 2.2 was used (Whittam, 1990). The genetic distance between each pair of distinct ETs was estimated as mean character differences, a similarity matrix was constructed using PAUP*4.0v10b (Swofford, 2002) and clustered by the unweighted pair group method with arithmetic mean (UPGMA) method using MEGA2 version 2.1 (Kumar *et al.*, 2001).

Genetic differentiation and linkage disequilibrium analyses

To estimate the relative genetic differentiation at different levels, we used indices related to Nei's $G_{ST} = (Ht - Hs)/Ht$ (Nei, 1987), where Ht is the expected diversity in an equivalent randomly mating total population and Hs is the average diversity of the subpopulations. To compute the G_{ST} values we used ETDIV version 2.2 (Whittam, 1990). These indices range from 0, if there is no genetic differentiation, to 1 if there is maximal genetic differentiation (Nei, 1987). To test if the G_{ST} indices were significantly different from 0, we performed a chi-square test of independence as $\chi^2 = nG_{ST}\ (a - 1)$, where n is the number of individuals and a is the total number of alleles. There are $(k - 1)\ (a - 1)$ degrees of freedom, where k is the number of subdivisions, and significance was examined at $P < 0.05$ (Hagen & Hamrick, 1996; Silva *et al.*, 1999).

To determine the extent to which populations exhibit nonrandom associations of alleles between loci, we used a multilocus index based on the distribution of allelic mismatches between pairs of isolates over all loci. The ratio of the variance in mismatches observed in a population (Vo) to the expected variance of the corresponding population at linkage equilibrium (random association of alleles, Ve) provides a measure of linkage disequilibrium. If there is no linkage disequilibrium $Vo/Ve = 1$. The significance of the difference between Vo and Ve was calculated using a Monte Carlo procedure with 1000 iterations, carried out with the LDV program (Souza *et al.*, 1992).

Visualization and cluster analysis of plasmid profiles

The plasmid content of the isolates was visualized following the Eckhardt procedure (Eckhardt, 1978). Plasmid mobilities were determined in 0.7% agarose gels using plasmids from *R. etli* CFN42^T and *R. tropici* CIAT899^T as molecular size references. A database for the presence/absence of the various plasmid sizes was constructed. The genetic distance between each pair of distinct plasmid profiles was estimated as mean character differences, and a similarity matrix was constructed using PAUP* (Swofford, 2002), and clustered by the UPGMA method using MEGA2 (Kumar *et al.*, 2001).

PCR amplification and RFLP of 16S rRNA genes and the ITS between the 16S and 23S rRNA genes

Genomic DNA was purified using a GenomicPrep kit (Amersham). Amplifications were performed with *Taq* polymerase (Roche) using the reaction mixture and conditions described in Silva *et al.* (2003). To amplify the 16S rRNA gene, primers fD1 and rD1 were used (Weisburg *et al.*, 1991) with an annealing temperature of 58 °C. A single band of about 1500 bp was obtained, and five restriction endonucleases (HinfI, HhaI, MspI, RsaI and Sau3AI) were used to digest the amplification product. Primers FGPS1490 (Navarro *et al.*, 1992) and FGPS132 (Pansonnet & Nesme, 1994) were used to amplify the ITS between the 16S and 23S rRNA genes, with an annealing temperature of 55 °C, as described elsewhere (Vinuesa *et al.*, 1998). All strains produced a single band of *c.* 1350 bp, and four restriction endonucleases (AluI, CfoI, HaeIII and MspI) were used to digest the amplification product. Restriction products were resolved by agarose (2%) gel electrophoresis.

PCR amplification and sequencing of *glnII*, *recA* and *nodB*

Partial sequences of the housekeeping genes *glnII* and *recA* were amplified using primers *glnII* 12F and *glnII* 689R, and *recA* 41F and *recA* 640R, respectively (Vinuesa *et al.*, 2005b). To amplify the symbiotic gene *nodB*, primers *nodB*3F and *nodB*RR were used (Silva *et al.*, 2003). Amplifications were performed with *Taq* polymerase (Roche) using the reaction mixture and conditions described in Silva *et al.* (2003). Amplification products were purified using the PCR product purification system of Roche and automatically sequenced using the same primers. The accession numbers generated in this study are DQ767659 – DQ767682. All accession numbers used in this study are listed in Table S1 of the supplementary material (available online).

Phylogenetic analyses based on sequence data

Multiple nucleotide sequence alignments were generated using CLUSTALX (Thompson *et al.*, 1997). Phylogenetic trees were inferred under the maximum likelihood (ML) optimality criterion using PHYML2.4.4 (Guindon & Gascuel, 2003) and the nucleotide substitution model selected by the Akaike information criterion (AIC), as implemented in MODELTEST3.06 (Posada & Crandall, 1998) as described elsewhere (Vinuesa *et al.*, 2005b). The robustness of ML topologies was inferred by nonparametric bootstrap tests using PHYML and 100 data replicates.

Results

Chromosomal genetic diversity based on MLEE

The electrophoretic mobility variants of nine metabolic enzymes were determined for a total of 176 isolates obtained from *Medicago* spp. nodules. The multilocus genotypes (ETs) for the *S. meliloti* type strain USDA1002^T and the *S. medicae* type strain A3217^T were also determined. Five isolates displayed the same ET as A3217^T, and, based on the results of the other molecular markers analysed (see below), they were classified as *S. medicae*. The *S. medicae* strains comprised one isolate from Cuernavaca, one from Guanajuato, and three isolates from Texcoco site 2. Twenty-six strains were initially isolated from *M. polymorpha* from Texcoco site 2; however, because all displayed the same plasmid profile, only three isolates were included in further analyses. Because no genetic diversity was detected in the *S. medicae* population ($H = 0$), no further population genetics analysis could be performed. The remaining 171 isolates were classified as *S. meliloti*, and more than half of them (59%) presented the same ET as the type strain USDA 1002^T. The high diversity displayed by the *S. meliloti* population ($H = 0.40$) permitted us to analyse the genetic structure of the 14 sampled populations (Table 1). Two of the Guanajuato populations were composed of only one ET and hence their diversity was 0. Likewise, only two ETs were present in the Texcoco site 1 population, which displayed a very low diversity ($H = 0.11$). The highest diversity estimates were recorded for the Guanajuato population. No new genotypes appeared in the Texcoco site 1 and Cuernavaca populations, indicating that Guanajuato contains the entire *S. meliloti* diversity of the sample (Table 1). The genetic relatedness among the 31 ETs from the *S. meliloti* population and the single ET from the *S. medicae* population is depicted in Fig. 1.

Genetic differentiation and linkage disequilibrium

The most frequent *S. meliloti* genotype (ET1) was found in all 14 sampling sites, and was the predominant genotype in most of them. The most abundant genotypes were recovered both from plants grown in the field and from plants grown

Table 1. Sampling scheme and population genetic parameters for *Sinorhizobium meliloti* isolated in Mexico

Population	*Medicago* species	No. of isolates	No. of Ets	Mean no. of alleles	*H**	ITS pattern
G1	*M. sativa*	12	5	1.8	0.28	A, B
G2	*M. sativa*	3	1	1.0	0.00	A
G3	*M. sativa*	17	6	2.0	0.35	A, B
G4	*M. sativa*	17	3	1.2	0.15	A, B
G5	*M. sativa*	8	4	1.6	0.26	A, B
G6	*M. sativa*	6	1	1.0	0.00	B
G7	*M. sativa*	12	3	1.6	0.37	A, B
G8	*M. sativa*	7	3	1.2	0.15	A, B
G9	*M. sativa*	21	3	1.4	0.22	A, B
G10	*M. sativa*	16	3	1.6	0.37	A, B
G11	*M. sativa*	12	8	1.7	0.21	A, B
G12	*M. sativa*	16	4	1.8	0.39	A, B
Guanajuato	*M. sativa*	147	30	3.7	0.40	A, B
Texcoco site 1	*M. sativa*	13	2	1.1	0.11	A, B
Cuernavaca	*M. lupulina*	11	4	1.3	0.19	A
Total		171	30	3.7	0.40	

*Mean genetic diversity.
ETs, electrophoretic types.

in the laboratory, and no genetic differentiation ($G_{ST} = 0.00$) was detected between these two sources of strains. The genetic differentiation estimates revealed a low and nonsignificant differentiation among the 12 Guanajuato populations. The same result was found for the populations from Guanajuato, Cuernavaca and Texcoco (Table 2). Moreover, no significant differentiation was found for all the pairwise comparisons between the 14 populations (data not shown). The lack of genetic differentiation at local and wider geographic distances is indicative of a very high migration rate among sampling sites.

The multilocus index used to measure the nonrandom association of alleles within populations showed linkage disequilibrium when the Guanajuato isolates and the total population were analysed (Table 2). However, when only the different ETs were included in the analyses, no significant linkage disequilibrium was detected both in the Guanajuato and in the total populations (Table 2). Taken together, these results point towards an epidemic genetic structure (Maynard Smith *et al.*, 1993), in which genetic exchange is frequent among the members of the *S. meliloti* population, and some abundant genotypes produce epidemic clones within the population.

Plasmid profiles

The five *S. medicae* strains displayed the same plasmid profile, which was composed of two megaplasmids of about 1600 and 1300 kb, and a plasmid of about 150 kb, which was very similar to that of USDA A321^T (1600, 1300 and 400 kb). For the 171 *S. meliloti* strains, 62 distinct plasmid profiles were detected; the most abundant one (P1) was found in most of the populations and was identical to that of USDA 1002^T, which is composed of one megaplasmid of about 1600 kb and a plasmid of about 230 kb. All the strains presented the 1600-kb megaplasmid, and in some cases it was possible to visualize a second megaplasmid of *c*. 1300 kb: these two megaplasmids are similar to pSymA and pSymB of strain 1021 (Galibert *et al.*, 2001). For nine strains (P4) it was not possible to visualize any plasmids in the Eckhardt gels. Various explanations are possible for this result; the most likely one is that the cells were not broken during the lysis procedure. We know that they are symbiotic *S. meliloti* strains because all the other molecular markers analysed (including *nodB* sequences) placed them within the population, and in reinoculation experiments they formed nodules on alfalfa plants.

Although the majority of the strains (63%) contained only two or three plasmids, some strains presented up to seven plasmids, and in many cases these included small plasmids (< 100 kb). If the small plasmids are eliminated from the analysis, the number of profiles is reduced from 63 to 38, indicating that they contribute strongly to plasmid profile diversity. The distribution of the plasmid profiles in the chromosomal genotypes is depicted in Fig. 1. Because ET1 contained more than half of the isolates, we decided to group the 101 strains according to their plasmid profiles in order to visualize the internal dynamics of this abundant and widely distributed genotype (Fig. 2). It is evident from Figs 1 and 2 that a given plasmid profile can be found in different chromosomal backgrounds and in different localities. This result suggests that the portion of the genome held in plasmids is subject to genetic exchange among chromosomal backgrounds.

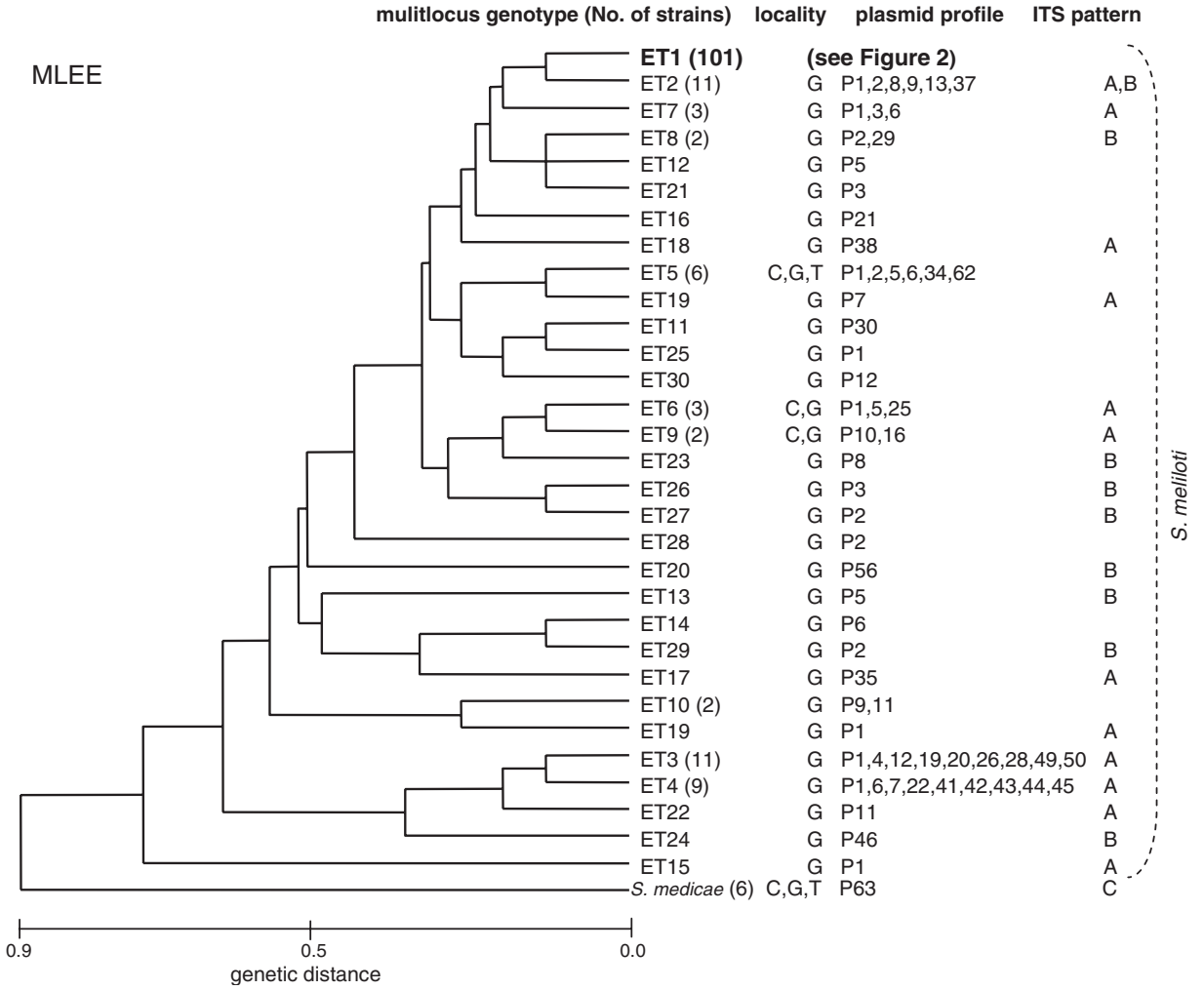

Fig. 1. Dendrogram depicting the genetic relatedness of 176 *Sinorhizobium* strains isolated from three *Medicago* species in Mexico, based on their multilocus genotypes (ETs). The code in the terminal nodes describes the ET number designation, and in parentheses the number of strains for the ETs with multiple strains. In the second column the locality of isolation is described as G, Guanajuato; C, Cuernavaca; T, Texcoco. The third column indicates the plasmid profiles found for the strains contained in the ET. The fourth column shows the ITS patterns displayed by the strains. The most abundant ET is highlighted in boldface and its description is presented in Fig. 2.

Restriction analyses of 16S and ITS rRNA gene regions

Twenty-three strains were selected as representatives of the 15 sampling sites and host species, and were subject to RFLP 16S rRNA gene analysis. Twenty strains from 14 sampling sites and isolated from *M. sativa* and *M. lupulina* showed identical patterns to that of *S. meliloti* USDA 1002^{T}. Three strains differed in the restriction pattern of RsaI and were identical to the pattern displayed by *S. medicae* USDA $A321^{T}$. These strains were isolates from *M. lupulina*, *M. sativa* and *Medicago* sp. from Cuernavaca, Guanajuato and Texcoco site 2, respectively.

Owing to the lack of variability within species detected by 16S rRNA gene analysis, we performed restriction analyses for the more variable 16S–23S ITS region for 118 of our strains and for 16 reference strains (nine *S. meliloti*, five *S.*

Table 2. Genetic differentiation (G_{ST}) and linkage disequilibrium estimates (*Vo*/*Ve*) for *S. meliloti* from Mexico

Population[†]	No. of isolates	No. of ETs	*Gst*[‡]	Isolates *Vo*/*Ve*[§]	ETs *Vo*/*Ve*
Guanajuato (12)	147	30	0.15^{NS}	1.81*	1.13^{NS}
Total (3)	171	30	0.00^{NS}	1.18*	1.16^{NS}

[†]The number of populations is indicated in parentheses.
[‡]Genetic differentiation index.
[§]Observed/expected variance of the mismatch distribution.
*Significant.
[NS]Nonsignificant.
ETs, electrophoretic types.

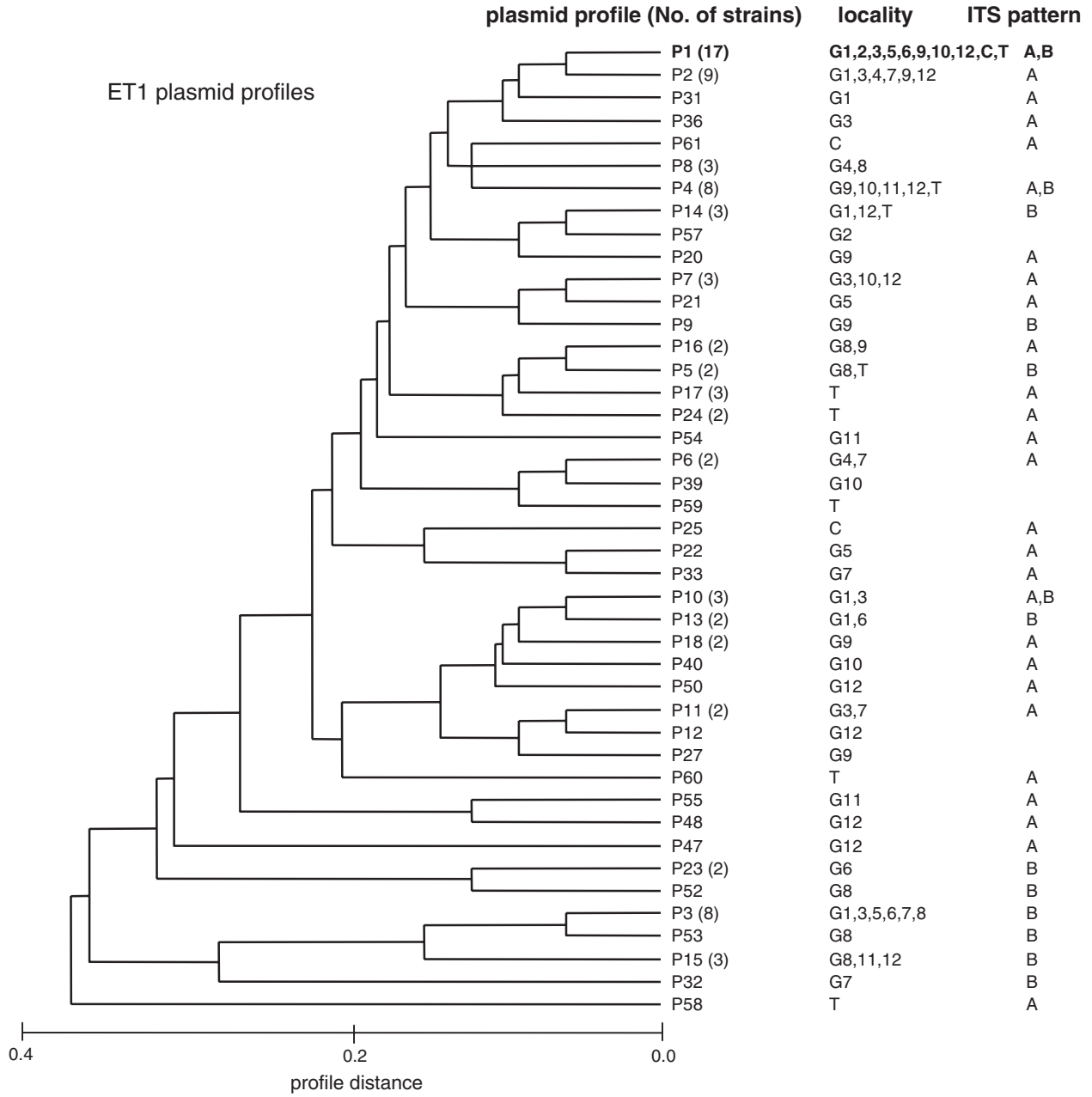

Fig. 2. Dendrogram depicting the plasmid profile distances for the 101 *S. meliloti* strains corresponding to ET1 (see Fig. 1). In the second column the sampling sites are indicated as G1–G12, Guanajuato; C, Cuernavaca; T, Texcoco. The third column shows the ITS patterns displayed by the strains. The most abundant plasmid profile is highlighted in boldface.

medicae, and the type strains of *S. americanum* and *S. morelense*). Among the 115 *S. meliloti* strains from Mexico, two ITS patterns were recorded (Table 3). Eighty strains showed the same pattern (A) as USDA 1002^T and most of the *S. meliloti* reference strains, while 35 strains displayed a pattern (B) that was also detected in two of the reference strains. The three *S. medicae* strains from Mexico displayed the same pattern (C) as USDA $A321^T$ and two of the reference strains; two of the *S. medicae* reference strains displayed a different restriction pattern (D) that was not found in the Mexican strains. *Sinorhizobium americanum* and *S. morelense* each displayed completely distinct patterns (Table 3). The results of the ITS restriction patterns differentiated *S. meliloti* from *S. medicae*, and showed the presence of two genotypes within each species. The two restriction patterns of the *S. meliloti* population did not

Table 3. Restriction patterns of the intergenic spacer (ITS) between 16S and 23S rRNA genes displayed by 118 Mexican isolates and 16 *Sinorhizobium* reference strains

	Restriction enzymes				ITS
Species/strain	AluI	CfoI	HaeIII	MspI	pattern
S. meliloti					
80 Mexican isolates	a	a	a	a	A
35 Mexican isolates	b	b	b	a	B
USDA 1002^T	a	a	a	a	A
56A14*	b	b	b	a	B
M119*	a	a	a	a	A
CC2013*	a	a	a	a	A
RCR2011	a	a	a	a	A
15A5*	a	a	a	a	A
102F51*	b	b	b	a	B
Rm41	a	a	a	a	A
1021	a	a	a	a	A
S. medicae					
Three Mexican isolates	c	c	c	b	C
USDA 1037^T	c	c	c	b	C
M3*	c	c	c	b	C
M102*	c	c	c	b	C
M104-1*	c	d	d	b	D
M104-2*	c	d	d	b	D
S. americanum CFNEI 156^T	d	e	e	c	E
S. morelense Lc04^T	e	f	F	d	F

*Strains provided by Bertrand Eardly (Pennsylvania State University).

show evidence of geographic or genotypic differentiation, because both were found in most of the populations and no association with multilocus genotype or plasmid profile was evident (Table 1, Figs 1 and 2).

Sequence-based phylogenies

The DNA sequences obtained were used as queries in BLASTN searches to retrieve related sequences from the public databases. We included in the phylogenies sequences showing high identities with our sequences, and those available for other *Sinorhizobium* species. Sequences from representative species of the fast-growing rhizobia in the genera *Rhizobium* and *Mesorhizobium* were also included. *Bradyrhizobium japonicum* sequences were used as outgroups, with the exception of *glnII* for which Turner & Young (2000) had shown an anomalous phylogenetic placement.

The sequences of the chromosomal genes *recA* and *glnII* confirmed the taxonomic assignment of our strains as either *S. meliloti* or *S. medicae* (Figs 3 and 4). In both phylogenetic reconstructions the *Sinorhizobium*, *Rhizobium* and *Mesorhizobium* groups were well supported. Sequences from several *Sinorhizobium* species could be included for the *recA* phylogeny, while for the *glnII* gene fewer *Sinorhizobium* sequences could be retrieved; however, many sequences related to our isolates were found. These strains were rhizobia studied by Wernegreen *et al.* (1999) (Wernegreen & Riley, 1999) isolated in diverse geographic localities from *Medicago*, *Melilotus* and *Trigonella* species. In the *glnII* phylogeny the *S. meliloti* strains formed two groups; the first one is represented by strain 1021 for which the complete genome is available (Galibert *et al.*, 2001), while the second group is formed by three strains isolated in Australia from *Trigonella suavissima*.

The phylogeny inferred from the symbiotic gene *nodB* showed that the *S. meliloti* and *S. medicae* strains conformed a well-supported clade (Fig. 5), which is related to a clade formed by strains isolated from *M. ruthenica* and other legumes in China that were reclassified as '*R. gallicum* bv. *orientale*' by Silva *et al.* (2005). Within the *S. meliloti*/*S. medicae* clade four subclades can be identified. The first one is represented by strain 1021 and the Mexican strains from Guanajuato. The second subclade is formed by most of the *S. medicae* strains, with the exception of strain C × 26, which is part of the third subclade, which includes *S. meliloti* strains from Cuernavaca and Texcoco site 1. The *T. suavissima* strains form the fourth subclade. The presence of this substructure in the sequence of *nodB* could be indicative of the existence of distinct biovars within the *S. meliloti*/*S. medicae* group. Recently, two biovars were described for *S. meliloti*: bv. *medicaginis* for *Sinorhizobium* effective on *M. laciniata* and bv. *meliloti* for the known *S. meliloti* strains effective on *M. sativa*, such as strain 1021 (Villegas *et al.*, 2006). The biovar status of the other three subclades remains to be studied.

It is noteworthy that for the three genes the Mexican strains displayed sequences identical to those of strains isolated in distant geographic regions (Figs 3–5), such as Australia or Syria, suggesting a low genetic differentiation among the worldwide populations of *S. meliloti* and *S. medicae*.

Discussion

Genetic diversity of *S. meliloti* and *S. medicae* and the origin of the Mexican populations

In this work we found that the naturalized populations of *Medicago* species in Mexico are nodulated by *S. meliloti* and *S. medicae*. No evidence of nodulation by native *Sinorhizobium* species or of the lateral transfer of symbiotic genes to other *Sinorhizobium* species was detected. We found that *S. meliloti* is the main symbiont of cultivated alfalfa plants in Mexico. *Sinorhizobium medicae* was detected in all the nodules of *M. polymorpha* from Texcoco site 2, in one nodule of *M. lupulina* from Cuernavaca, and in one nodule of alfalfa from Guanajuato. The six *S. medicae* strains characterized in this work were homogeneous for all the genetic markers analysed, indicating that they belong to a single clone that is widespread in central Mexico. The fact

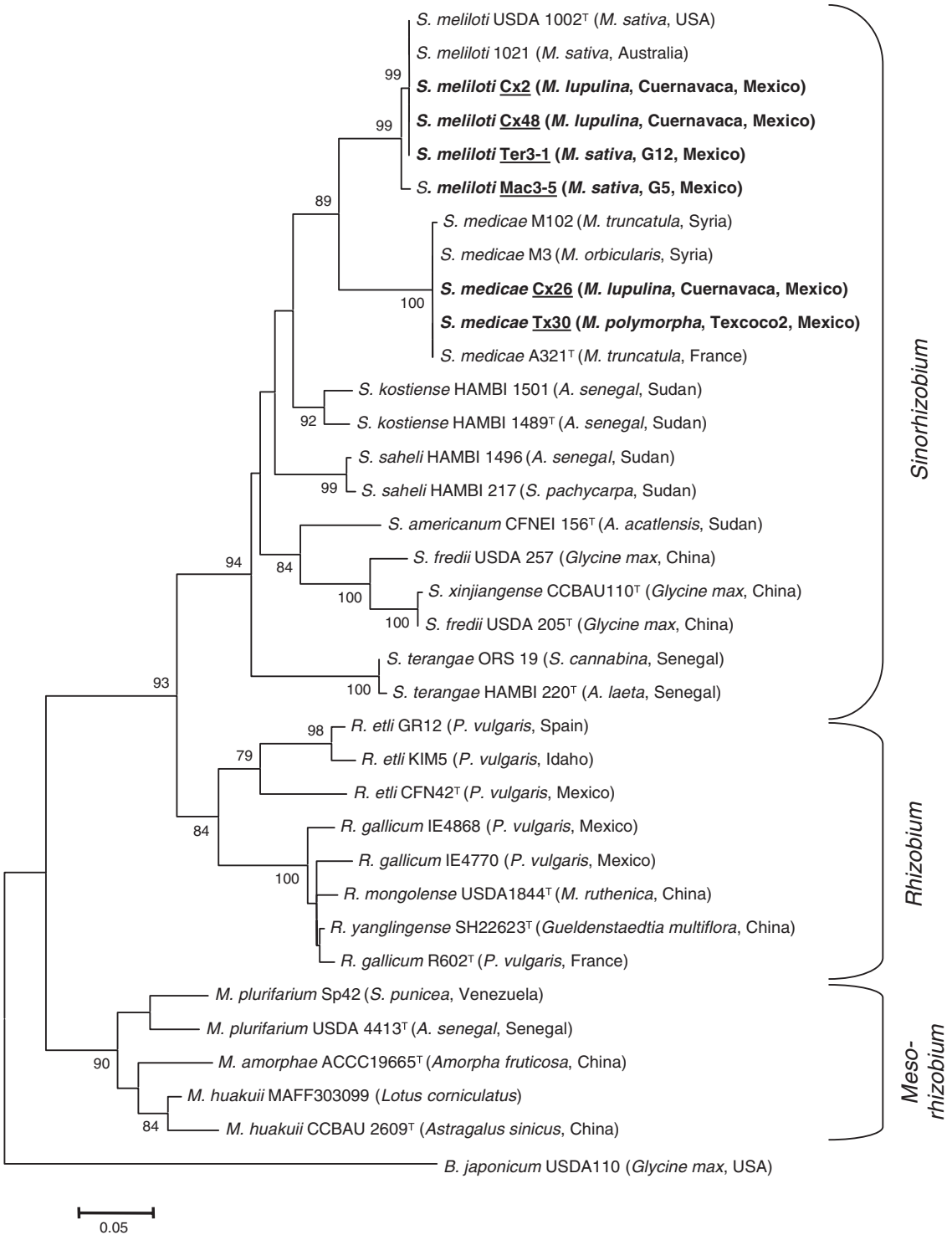

Fig. 3. Phylogenetic relationships inferred from *recA* gene sequences. The phylogeny was estimated under the maximum-likelihood optimality criterion using the GTR+I+G model with the gamma shape parameter equal to 2.6 and the proportion of invariant sites equal to 0.5. Bootstrap support values $> 75\%$ are indicated at the corresponding nodes. The host and geographic origin for each strain are shown in parentheses. The strains described in this study are underlined and in boldface. Brackets to the right indicate the generic affiliation of the strains.

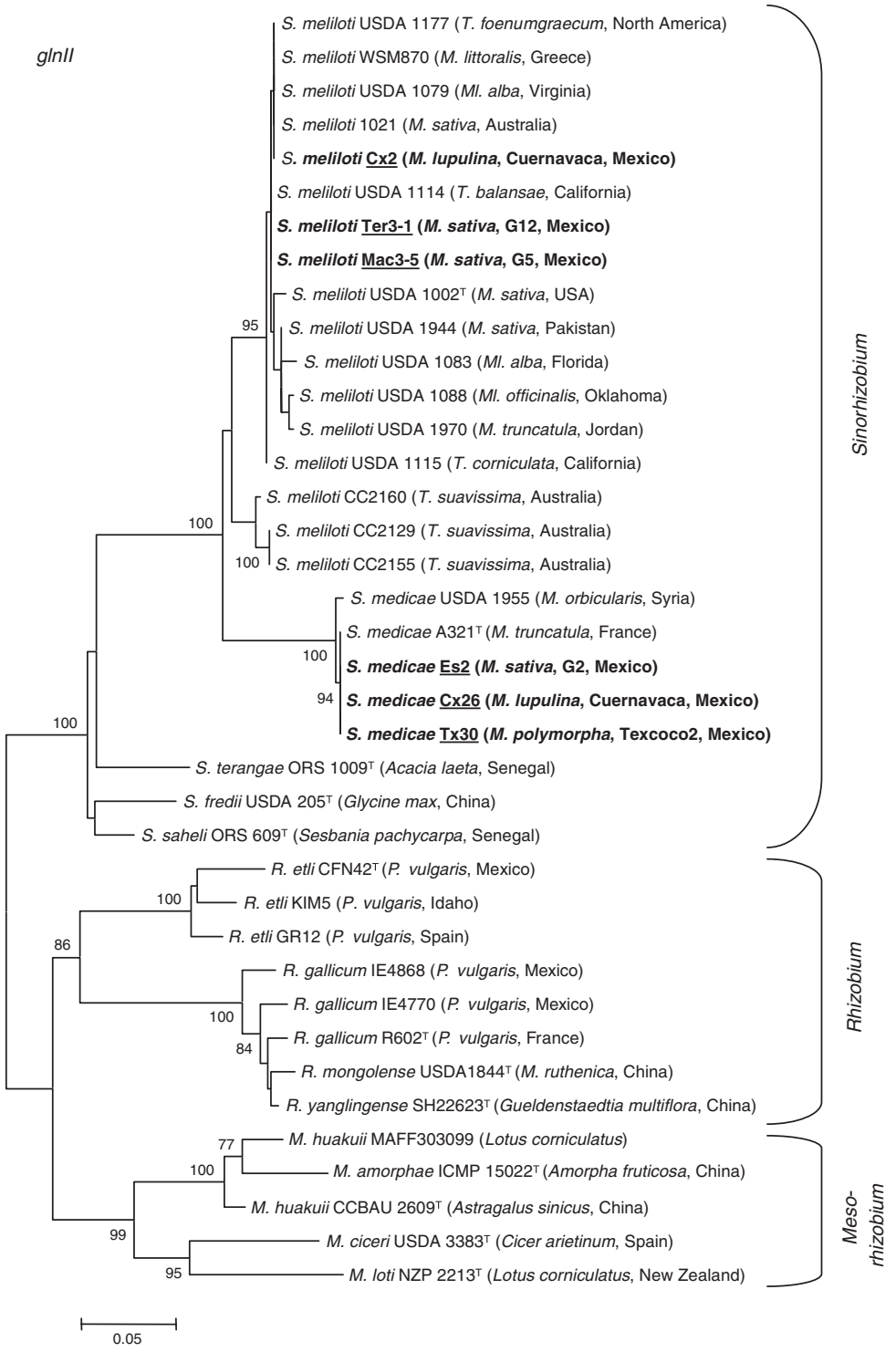

Fig. 4. Phylogenetic relationships inferred from *glnII* gene sequences. The phylogeny was estimated under the maximum-likelihood optimality criterion using the GTR+I+G model with the gamma shape equal to 2.7 and the proportion of invariant sites equal to 0.5. Bootstrap support values $> 75\%$ are indicated at the corresponding nodes. The host and geographic origin for each strain are shown in parentheses. The strains described in this study are underlined and in boldface. Brackets to the right indicate the generic affiliation of the strains.

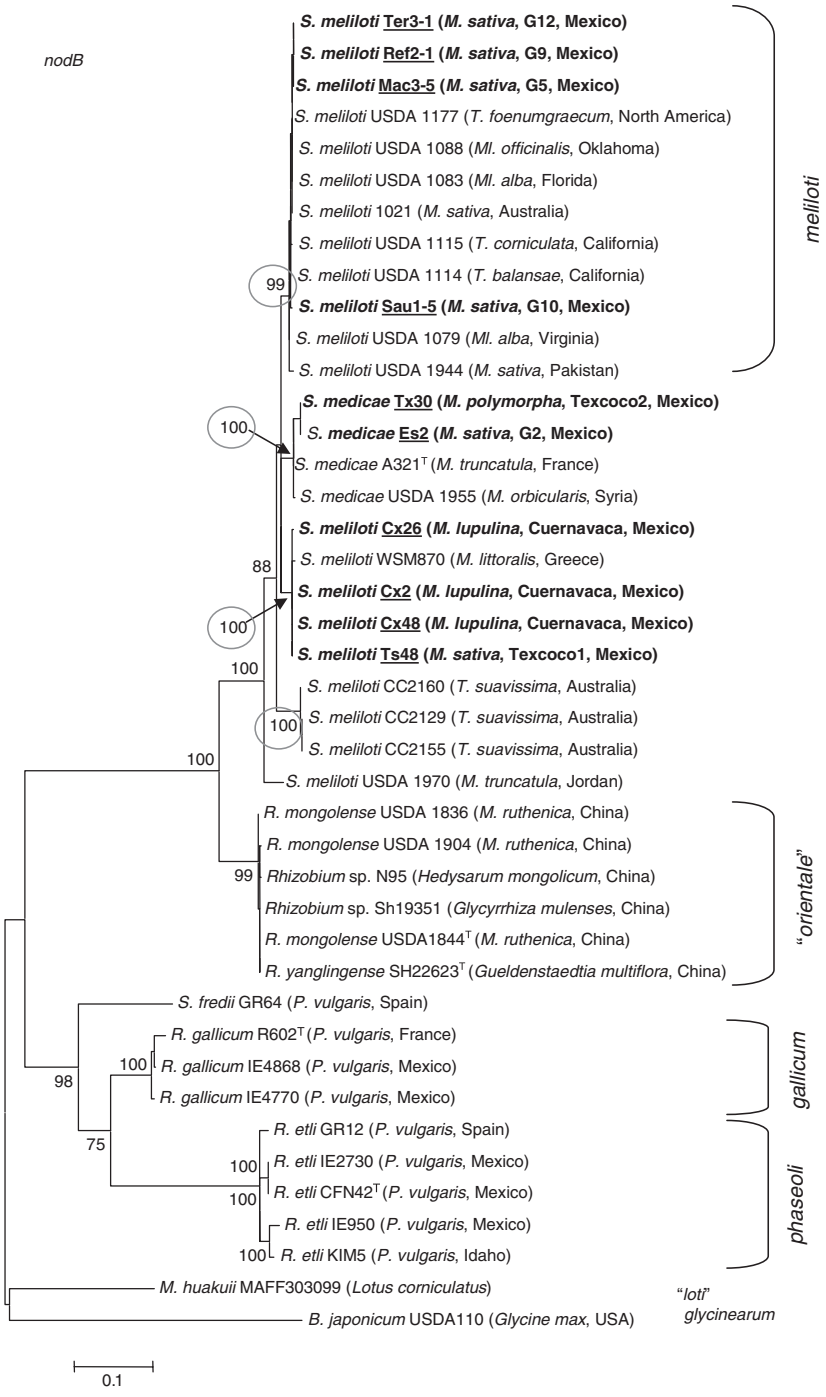

Fig. 5. Maximum-likelihood *nodB* phylogeny inferred under the HKY+I+G model with the gamma shape parameter equal to 2.0 and the proportion of invariant sites equal to 0.2. Bootstrap support values $> 75\%$ are indicated at the corresponding nodes. The host and geographic origin for each strain are in parentheses. The strains described in this study are underlined and in boldface. The nodes supporting the four subclades within the *S. meliloti*/*S. medicae* group are indicated by circles. The recognized biovars are indicated to the right.

that the genetic variation of the *S. medicae* population is lower than that of the *S. meliloti* population is in agreement with other studies (Eardly *et al.*, 1990; Rome *et al.*, 1996a; Biondi *et al.*, 2003). These results support the proposed hypotheses that *S. medicae* originated from an ancestral *S. meliloti* lineage (Biondi *et al.*, 2003), and that these sister species diverged later than the other species of the genus (Rome *et al.*, 1996b; Turner & Young, 2000).

The multilocus genetic diversity displayed by the Mexican *S. meliloti* population ($H = 0.40$) is higher than the one reported by Eardly *et al.* (1990) for a collection of strains isolated from several *Medicago* species at localities on five continents ($H = 0.25$), indicating that a flourishing *S. meliloti* population is present in Mexican soils. A high genetic diversity has been recorded for *S. meliloti* in many other studies using different genetic markers (Eardly *et al.*, 1990;

Bromfield *et al*., 1995, 1998; Paffetti *et al*., 1996), even for places that did not have a history of alfalfa cultivation (Paffetti *et al*., 1998; Carelli *et al*., 2000; Bromfield *et al*., 2001). Thus, what is the origin of the world *S. meliloti* and *S. medicae* populations? It is difficult to determine whether the rhizobial populations are native to an area or were introduced to it. The coevolution hypothesis states that rhizobia and host legumes coevolved locally (Martínez-Romero & Caballero-Mellado, 1996). If this is true for *Medicago* species, *S. meliloti* and *S. medicae* should have evolved in Eurasia and Africa and then dispersed to other parts of the world. Roumiantseva *et al*. (2002) studied the diversity of *S. meliloti* from the Central Asian alfalfa gene center to challenge the hypothesis that in the centre of origin of alfalfa a wider range of rhizobial genetic variants would be expected than elsewhere. However, their results do not support the idea that Central Asia is a repository of alfalfa symbionts with many divergent endemic genotypes, because the levels of genetic diversity were surprisingly modest (Roumiantseva *et al*., 2002). The dispersion of rhizobia could involve natural and human-mediated mechanisms. In the case of the Mexican populations, we are aware that inoculants for alfalfa were introduced in some areas. Although we have information that the sampled plots were not inoculated, we cannot rule out the possibility that they were inoculated when the crop was introduced several decades ago. Another human-mediated dispersion mechanism could be the involuntary transportation of rhizobia on the seed-coats of introduced plants, as suggested for the common bean rhizobia (Pérez-Ramírez *et al*., 1998). In the particular case of alfalfa, this would apply not only to *S. meliloti* and *S. medicae*, but also to acid-tolerant rhizobia related to strain Or191, which inefficiently nodulates alfalfa in acidic soils of distant geographic locations in North and South America (del Papa *et al*., 1999). Nevertheless, the presence of *S. meliloti* populations in places with no history of alfalfa cultivation could be the result of natural dispersion mechanisms, such as wind-driven dust or running water (Griffin *et al*., 2002; Wang & Chen, 2004; Silva *et al*., 2005; Vinuesa *et al*., 2005b).

The influence of recombination and migration on the population structure

The absence of genetic exchange between *S. meliloti* and *S. medicae* and the evidence for genetic exchange within each species have been reported in several studies (Eardly *et al*., 1990; Rome *et al*., 1996a; Roumiantseva *et al*., 2002; Biondi *et al*., 2003), suggesting that although these sympatric species share ecological niches there is a genetic barrier that maintains the distinctiveness of each species.

In this work we sampled the nodules of perennial and annual *Medicago* species distributed in three distant locations of central Mexico with the aim of assessing the importance of geographic distance on the composition and structure of rhizobia. It was surprising that the same genotypes were found nodulating different *Medicago* species in sites separated by as much as 400 km. The population genetic analyses for *S. meliloti* showed that recombination and migration may be important evolutionary forces contributing to the generation of genetic diversity. The lack of genetic differentiation among sampling sites indicates that the distribution of the genetic variation is homogeneous at the geographic scales tested, and suggests high levels of gene flow among populations. In concordance, the linkage disequilibrium analyses provided evidence of frequent recombination within the whole *S. meliloti* population. The epidemic genetic structure is produced by the presence of some very abundant clones that are selected by the plants to form nodules (Bromfield *et al*., 1995; Silva *et al*., 1999; McInnes *et al*., 2004). The most abundant clone was found in all three localities and occupied more than half of the nodules. This clone displayed the same genotype as the *S. meliloti* type strain USDA 1002^{T}, which was isolated in the USA. The sequence data revealed that the same *S. meliloti* and *S. medicae* haplotypes were present in strains from diverse geographic origins, suggesting that migration occurs at very wide geographic distances. Similar findings have been reported for other rhizobial populations in the genera *Rhizobium*, *Mesorhizobium* and *Bradyrhizobium* nodulating diverse legumes in several continents (Vinuesa & Silva, 2004; Silva *et al*., 2005; Vinuesa *et al*., 2005a, b), suggesting that migration and recombination are forces with a strong influence on the evolutionary history of rhizobia at a global scale.

Acknowledgements

We are grateful to En Tao Wang for his help in the collection and identification of the Texcoco and Cuernavaca plants, and to Oscar Olalde for his help in the sampling of the Guanajuato plants and soils. Pablo Vinuesa is acknowledged for supervision of phylogenetic analyses and a critical review of the manuscript. Three anonymous reviewers are acknowledged for important comments that improved the final version of the manuscript. We thank Bertrand Eardly for providing reference strains, Lourdes Lloret for providing some of the *recA* sequences, Michael Dunn for reviewing the manuscript, and Marco Antonio Rogel for his help with the laboratory techniques. This work was supported by PAPIIT IN201106 from DGAPA, UNAM, Mexico.

References

Allen ON & Allen EK (1981) *The Leguminosae. A Source Book of Characteristics, Uses and Nodulation*. Macmillan Publishers LTD, London.

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.

Brockwell J & Hely FW (1966) Symbiotic characteristics of *Rhizobium meliloti*: an appraisal of systematic treatment of nodulation and nitrogen fixation interaction between hosts and rhizobia of diverse origins. *Aust J Agric Res* **17**: 885–899.

Bromfield ES, Butler G & Barran LR (2001) Temporal effects on the composition of a population of *Sinorhizobium meliloti* associated with *Medicago sativa* and *Melilotus alba*. *Can J Microbiol* **47**: 567–573.

Bromfield ESP, Barran LR & Wheatcroft R (1995) Relative genetic structure of a population of *Rhizobium meliloti* isolated directly from soil and from nodules of alfalfa (*Medicago sativa*) and sweet clover (*Melilotus alba*). *Mol Ecol* **4**: 183–188.

Bromfield ESP, Behara AMP, Singh RS & Barran LR (1998) Genetic variation in local populations of *Sinorhizobium meliloti*. *Soil Biol Biochem* **30**: 1707–1716.

Carelli M, Gnocchi S, Fancelli S, Mengoni A, Paffetti D, Scotti C & Bazzicalupo M (2000) Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa cultivars in Italian soils. *Appl Environ Microbiol* **66**: 4785–4789.

del Papa MF, Balague LJ, Sowinski SC *et al.* (1999) Isolation and characterization of alfalfa-nodulating rhizobia present in acidic soils of central Argentina and Uruguay. *Appl Environ Microbiol* **65**: 1420–1427.

Eardly BD, Materon LA, Smith NH, Johnson DA, Rumbaugh MD & Selander RK (1990) Genetic structure of natural populations of the nitrogen fixing bacterium *Rhizobium meliloti*. *Appl Environ Microbiol* **56**: 187–194.

Eckhardt T (1978) A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. *Plasmid* **1**: 584–588.

Galibert F, Finan TM, Long SR *et al.* (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.

Griffin DW, Kellogg CA, Garrison VH & Shinn EA (2002) The global transport of dust – an intercontinental river of dust, microorganisms and toxic chemicals flows through the Earth's atmosphere. *Am Sci* **90**: 228–235.

Guindon S & Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.

Hagen MJ & Hamrick JL (1996) A hierarchical analysis of population genetic structure in *Rhizobium leguminosarum* bv. *trifolii*. *Mol Ecol* **5**: 177–186.

Kumar S, Tamura K, Jakobsen IB & Nei M (2001) MEGA2: molecular evolutionary genetics analysis (software). *Bioinformatics* **17**: 1244–1245.

Lesins KA & Lesins I (1979) *Genus Medicago (Leguminosae): A Taxogenetic Study*, Dr W. Junk bv Publishers, The Hague.

Lloret L, Ormeño-Orrillo E, Rincón R, Martínez-Romero J & Martínez-Romero E (2007) *Sinorhizobium mexicanum* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in México. *Int J Syst Evol Bacteriol*, in press.

Martínez-Romero E & Caballero-Mellado J (1996) *Rhizobium* phylogenies and bacterial genetic diversity. *Critical Rev Plant Sci* **15**: 113–140.

Maynard Smith J, Smith NH, O'Rourke M & Spratt BG (1993) How clonal are bacteria? *Proc Natl Acad Sci USA* **90**: 4384–4388.

McInnes A, Thies JE, Abbot JG & Howeieson JG (2004) Structure and diversity among rhizobial strains, populations and communities – a review. *Soil Biol Biochem* **36**: 1295–1308.

Navarro E, Simonet P, Normand P & Bardin R (1992) Characterization of natural population of *Nitrobacter* spp. using PCR/RFLP analysis of the ribosomal intergenic spacer. *Arch Microbiol* **157**: 107–115.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Paffetti D, Scotti C, Gnocchi S, Fancelli S & Bazzicalupo M (1996) Genetic diversity of an Italian *Rhizobium meliloti* population from different *Medicago sativa* varieties. *Appl Environ Microbiol* **62**: 2279–2285.

Paffetti D, Daguin F, Fancelli S, Gnocchi S, Lippi F, Scotti C & Bazzicalupo M (1998) Influence of plant genotype on the selection of nodulating *Sinorhizobium meliloti* strains by *Medicago sativa*. *Antonie Van Leeuwenhoek* **73**: 3–8.

Pansonnet C & Nesme X (1994) Identification of *Agrobacterium* strains by PCR-RFLP analysis of pTi and chromosomal regions. *Arch Microbiol* **161**: 300–309.

Pérez-Ramírez NO, Rogel MA, Wang E, Castellanos JZ & Martínez-Romero E (1998) Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. *FEMS Microbiol Ecol* **26**: 289–296.

Posada D & Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Rome S, Brunel B, Normand P, Fernandez M & Cleyet Marel JC (1996a) Evidence that two genomic species of *Rhizobium* are associated with *Medicago truncatula*. *Arch Microbiol* **165**: 285–288.

Rome S, Fernandez MP, Brunel B, Normand P & Cleyet-Marel JC (1996b) *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. *Int J Syst Bacteriol* **46**: 972–980.

Roumiantseva ML, Andronov EE, Sharypova LA, Dammann-Kalinowski T, Keller M, Young JPW & Simarov BV (2002) Diversity of *Sinorhizobium meliloti* from the Central Asian alfalfa gene center. *Appl Environ Microbiol* **68**: 4694–4697.

Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN & Whittam TS (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol* **51**: 873–884.

Silva C, Eguiarte LE & Souza V (1999) Reticulated and epidemic population genetic structure of *Rhizobium etli* biovar *phaseoli* in a traditionally managed locality in Mexico. *Mol Ecol* **8**: 277–287.

Silva C, Vinuesa P, Eguiarte LE, Martínez-Romero E & Souza V (2003) *Rhizobium etli* and *Rhizobium gallicum* nodulate common bean (*Phaseolus vulgaris*) in a traditionally managed

milpa plot in Mexico: population genetics and biogeographic implications. *Appl Environ Microbiol* **69**: 884–893.

Silva C, Vinuesa P, Eguiarte LE, Souza V & Martínez-Romero E (2005) Evolutionary genetics and biogeographic structure of *Rhizobium gallicum sensu lato*, a widely distributed bacterial symbiont of diverse legumes. *Mol Ecol* **14**: 4033–4050.

Souza V, Nguyen TT, Hudson RR, Piñero D & Lenski RE (1992) Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? *Proc Natl Acad Sci USA* **89**: 8389–8393.

Sprent JI (2001) *Nodulation in Legumes*. Royal Botanic Gardens, Kew.

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods (Software)*. Sinuauer Associates, Sunderland, MA.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882.

Toledo I, Lloret L & Martínez-Romero E (2003) *Sinorhizobium americanum* sp. nov., a new *Sinorhizobium* species nodulating native *Acacia* spp. in Mexico. *Syst Appl Microbiol* **26**: 54–64.

Turner SL & Young JP (2000) The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. *Mol Biol Evol* **17**: 309–319.

Villaseñor JL & Espinosa-García J (2004) The alien flowering plants of Mexico. *Diversity Distrib* **10**: 113–123.

Villegas MD, Rome S, Maure L, Domergue O, Gardan L, Bailly X, Cleyet-Marel JC & Brunel B (2006) Nitrogen-fixing sinorhizobia with *Medicago laciniata* constitute a novel biovar (bv. *medicaginis*) of *S. meliloti*. *Syst Appl Microbiol* **29**: 526–538.

Vincent JM (1970) *A Manual for the Practical Study of the Root Nodule Bacteria*. Blackwell, Oxford.

Vinuesa P & Silva C (2004) Species delineation and biogeography of symbiotic bacteria associated with cultivated and wild legumes. *Biological Resources and Migration* (Werner D, ed), pp. 143–155. Springer Verlag, Berlin.

Vinuesa P, Rademaker JLW, de Bruijn FJ & Werner D (1998) Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary Islands by PCR-restriction fragment length polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rDNA intergenic spacers, repetitive extragenic palindromic PCR genomic fingerprinting and partial 16S rDNA sequencing. *Appl Environ Microbiol* **64**: 2096–2104.

Vinuesa P, Silva C, Lorite MJ, Izaguirre-Mayoral ML, Bedmar EJ & Martínez-Romero E (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.

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.

Wang E & Chen WX (2004) Estimation of biotic and abiotic factors that affect migration of rhizobia. *Biological Resources and Migration* (Werner D, ed), pp. 167–171. Springer Verlag, Berlin.

Wang E, Tan ZY, Willems A, Fernández-Lopez M, Reinhold-Hurek B & Martínez-Romero E (2002) *Sinorhizobium morelense* sp. nov., a *Leucaena leucocephala*-associated bacterium that is highly resistant to multiple antibiotics. *Int J Syst Evol Bacteriol* **52**: 1687–1693.

Weisburg WG, Barns SM, Pelletie DA & Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**: 697–703.

Wernegreen J & Riley MA (1999) Comparison of the evolutionary dynamics of symbiotic and housekeeping loci: a case for the genetic coherence of rhizobial lineages. *Mol Biol Evol* **16**: 98–113.

Whittam TS (1990). ETDIV and ETCLUS programs, Pennsylvania State University, Philadelphia, PA.

Supplementary material

The following supplementary material is available for this article:

Table S1. Accession numbers for the sequences used in this study. The sequences generated in this study are highlighted in boldface.

This material is available as part of the online article from: http://www.blackwellsynergy.com/doi/abs/10.1111/j.1742-4658.2006.05-00301.x (This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.