Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes

Pâmela Menna,[1,2,3] Fernando Gomes Barcellos[1,3] and Mariangela Hungria[1,2,3]

Correspondence
Mariangela Hungria
hungria@cnpso.embrapa.br *or*
hungria@pq.cnpq.br

[1]Embrapa Soja, Cx Postal 231, 86001-970 Londrina, Paraná, Brazil

[2]Universidade Estadual de Londrina, Department of Microbiology, Cx Postal 60001, 86051-990 Londrina, Paraná, Brazil

[3]Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-MCT), Brasilia, Federal District, Brazil

The genus *Bradyrhizobium* encompasses a variety of bacteria that can live in symbiotic and endophytic associations with legumes and non-legumes, and are characterized by physiological and symbiotic versatility and broad geographical distribution. However, despite indications of great genetic variability within the genus, only eight species have been described, mainly because of the highly conserved nature of the 16S rRNA gene. In this study, 169 strains isolated from 43 different legumes were analysed by rep-PCR with the BOX primer, by sequence analysis of the 16S rRNA gene and the 16S–23S rRNA intergenic transcribed spacer (ITS) and by multilocus sequence analysis (MLSA) of four housekeeping genes, *glnII*, *recA*, *atpD* and *dnaK*. Considering a cut-off at a level of 70 % similarity, 80 rep-PCR profiles were distinguished, which, together with type strains, were clustered at a very low level of similarity (24 %). In both single and concatenated analyses of the 16S rRNA gene and ITS sequences, two large groups were formed, with bootstrap support of 99 % in the concatenated analysis. The first group included the type and/or reference strains of *Bradyrhizobium japonicum*, *B. betae*, *B. liaoningense*, *B. canariense* and *B. yuanmingense* and *B. japonicum* USDA 110, and the second group included strains related to *Bradyrhizobium elkanii* USDA 76^T, *B. pachyrhizi* PAC48^T and *B. jicamae* PAC68^T. Similar results were obtained with MLSA of *glnII*, *recA*, *atpD* and *dnaK*. Greatest variability was observed when the *atpD* gene was amplified, and five strains related to *B. elkanii* revealed a level of variability never reported before. Another important observation was that a group composed of strains USDA 110, SEMIA 5080 and SEMIA 6059, all isolated from soybean, clustered in all six trees with high bootstrap support and were quite distinct from the clusters that included *B. japonicum* USDA 6^T. The results confirm that MLSA is a rapid and reliable way of providing information on phylogenetic relationships and of identifying rhizobial strains potentially representative of novel species.

INTRODUCTION

The genus *Bradyrhizobium* was created to accommodate bacteria capable of establishing N_2-fixing symbioses with several leguminous species and characterized by slow growth and an alkaline reaction in culture media containing mannitol as a carbon source (Jordan, 1982). Based on 16S rRNA gene sequences, the genus *Bradyrhizobium* forms a clade in the *Alphaproteobacteria*, along with oligotrophic soil and aquatic bacteria such as *Rhodopseudomonas palustris*, *Rhodoplanes roseus*, *Nitrobacter winogradskyi* and *Blastobacter denitrificans* and pathogens such as *Afipia* species, among others (Saito *et al.*, 1998; Willems *et al.*, 2001a; van Berkum & Eardly, 2002; Sawada *et al.*, 2003). Symbiotic *Bradyrhizobium* strains have been isolated from

Abbreviations: ITS, intergenic transcribed spacer; MLSA, multilocus sequence analysis.

The GenBank/EMBL/DDBJ accession numbers for the sequences obtained in this study are detailed in Supplementary Table S3.

Details of strains, host plants and sequence accession numbers and BOX-PCR profiles are available as supplementary material with the online version of this paper.

the nodules of highly divergent legume tribes, including herbaceous and woody species of tropical and temperate origin and aquatic legumes such as *Aeschynomene* species, and the non-legume *Parasponia andersonii* (Trinick, 1973; Sprent, 2001; Menna *et al.*, 2006). In addition, bacteria belonging to this genus have been reported as endophytes of wild and modern rice (*Oryza sativa* L.) (e.g. Tan *et al.*, 2001), and *Bradyrhizobium betae* refers to endophytes, isolated from sugar beet (*Beta vulgaris* L.) roots affected by tumour-like deformations, that do not nodulate legumes (Rivas *et al.*, 2004).

In the last few years, much attention has been focused on the genus *Bradyrhizobium*, and high genetic diversity among strains has been demonstrated with several molecular markers (van Berkum & Fuhrmann, 2000; Tan *et al.*, 2001; van Berkum & Eardly, 2002; Willems *et al.*, 2003; Liu *et al.*, 2005; Vinuesa *et al.*, 2005b; Germano *et al.*, 2006; Giongo *et al.*, 2008; Menna *et al.*, 2009). Intriguing questions arise on the evolution and ecology of this versatile genus. However, despite its impressive diversity and worldwide distribution, only eight species have been described to this point: three symbionts of *Glycine* species [*Bradyrhizobium japonicum* (Jordan, 1982), *B. elkanii* (Kuykendall *et al.*, 1992) and *B. liaoningense* (Xu *et al.*, 1995)], one nodulating *Lespedeza cuneata* [*Bradyrhizobium yuanmingense* (Yao *et al.*, 2002)], another nodulating shrubs of the tribes Genisteae and Loteae [*Bradyrhizobium canariense* (Vinuesa *et al.*, 2005a)], the already mentioned *B. betae* (Rivas *et al.*, 2004) and two species that nodulate *Pachyrhizus erosus* (*Bradyrhizobium pachyrhizi* and *B. jicamae*) that have been described recently (Ramírez-Bahena *et al.*, 2009). The new combination '*Bradyrhizobium denitrificans*', proposed by van Berkum *et al.* (2006) as a result of the reclassification of *Blastobacter denitrificans*, is yet to be validly published. A variety of other strains are commonly referred to as *Bradyrhizobium* sp., followed by the name of the host legume. Apparently, the definition of species in the genus is limited by the low diversity of the 16S rRNA gene sequences reported so far (e.g. van Berkum & Fuhrmann, 2000; Willems *et al.*, 2001b; Qian *et al.*, 2003).

DNA–DNA hybridization is generally required in order to define a novel species (Garrity & Holt, 2001); however, an increasing number of arguments against insistence on its use has been raised, including the high cost and intensive work required for its development (Vandamme *et al.*, 1996; Coenye *et al.*, 2005), the existence of more accurate approaches (Konstantinidis & Tiedje, 2004) and doubts about its adequacy (Achtman & Wagner, 2008). Multilocus sequence analysis (MLSA), which consists of examination of the sequences of several conserved housekeeping genes dispersed over at least 100 kb of the genome, has been proposed as a more accessible tool for assessing phylogeny and taxonomy of prokaryotes (Brett *et al.*, 1998; Maiden *et al.*, 1998; Godoy *et al.*, 2003; Cooper & Feil, 2004). Using MLSA, Moulin *et al.* (2004) observed high genetic diversity among *Bradyrhizobium* strains, and Ribeiro *et al.* (2009) found strong concordance between MLSA of five housekeeping genes and DNA–DNA hybridization in *Rhizobium* microsymbionts of common bean (*Phaseolus vulgaris* L.). Nevertheless, at present, the use of a small number of housekeeping genes cannot replace DNA–DNA hybridization in species delineation.

In this study, the genetic diversity of 169 *Bradyrhizobium* strains isolated from various legumes from various countries was analysed using the amplification of repetitive and conserved DNA elements (rep-PCR) and sequence analysis of the 16S rRNA gene, the 16S–23S rRNA intergenic transcribed spacer (ITS) and four additional conserved housekeeping genes (*glnII*, *recA*, *atpD* and *dnaK*). The objective was to apply a polyphasic approach to the appraisal of the genetic diversity, phylogenetic relationships and taxonomic position of these strains, thus gaining a better understanding of the genus *Bradyrhizobium*.

METHODS

Strains. One hundred and sixty-nine *Bradyrhizobium* strains from the Brazilian *Rhizobium* Culture Collection SEMIA, of the FEPAGRO-MIRCEN [Fundação Estadual de Pesquisa Agropecuária (Rio Grande do Sul, Brazil) – Microbiological Resources Center] (IBP World Catalogue of *Rhizobium* Collections no. 443 in the WFCC World Data Center on Microorganisms), were used in this study (Table 1). Additional information about the strains is given in Supplementary Table S1, available in IJSEM Online. The strains were isolated from members of the three subfamilies and 12 tribes of the family Leguminosae (Supplementary Table S2). Fifty of these strains are recommended for use in commercial inoculants in Brazil, and 68 were previously analysed by BOX-PCR (Menna *et al.*, 2009) and 16S rRNA gene sequencing (Menna *et al.*, 2006). Preparation of stock cultures, strain growth conditions and maintenance were as described by Menna *et al.* (2006). Three type strains were used in the studies, *B. japonicum* USDA 6^T, *B. liaoningense* LMG 18230^T and *B. elkanii* USDA 76^T, provided by the USDA (Beltsville, MD, USA) and deposited at the Culture Collection of Diazotrophic and Plant Growth Promoting Bacteria of Embrapa Soja.

DNA extraction and rep-PCR (BOX) genomic fingerprinting. Total genomic DNA of the 169 strains was extracted as described by Kaschuk *et al.* (2006a), and amplification by PCR with the primer BOX A1R (Versalovic *et al.*, 1994; Koeuth *et al.*, 1995) was performed as described by Kaschuk *et al.* (2006a), with a 1 kb DNA marker (Invitrogen) being included on the left and right and in the centre of each gel. The amplified fragments were separated by horizontal electrophoresis on 1.5 % agarose gels (Kaschuk *et al.*, 2006a), which were then stained with ethidium bromide, visualized under UV irradiation and photographed.

Sequencing of the 16S rRNA gene and ITS. The DNA of 80 strains representative of the rep-PCR groups was submitted to amplification with primers for the 16S rRNA gene and ITS region (Table 2). The PCR products were purified with the PureLink PCR Purification kit (Invitrogen), and the reactions were performed as described by Menna *et al.* (2006). Sequencing was performed on a MEGA BACE 1000 (Amersham Biosciences) capillary sequencer, as described by Menna *et al.* (2006).

Sequencing of *glnII*, *recA*, *atpD* and *dnaK* genes. DNA from 40 strains selected after analysis of the 16S rRNA gene and ITS sequences was amplified using primers specific to the regions coding for four housekeeping genes, *glnII*, *recA*, *atpD* and *dnaK*. The primers, amplification conditions and references are listed in Table 2. Purification and sequencing were performed as described above.

Table 1. Strains included in this study

Strain	Host plant	rep-PCR group	16S rRNA+ITS group	MLSA group	Classification based on MLSA
USDA 6^T	*Glycine max*	Group IX	GI-4	GI-7	*B. japonicum*
USDA 76^T	*Glycine max*	Group XVIII	GII-3	GII-1	*B. elkanii*
LMG 18230^T	*Glycine max*	Isolated	GI-7	ND	*B. liaoningense*
SEMIA 501	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 509	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 510	*Glycine max*	Group XXVI	GI-6	ND	ND
SEMIA 511	*Glycine max*	Isolated	GI-isolated	GI-7	*B. japonicum*
SEMIA 512	*Glycine max*	Group IX	GI-4	GI-7	*B. japonicum*
SEMIA 513	Not known	Group XII	ND	ND	ND
SEMIA 515	*Glycine max*	Group XII	ND	ND	ND
SEMIA 516	*Glycine max*	Group XXVI	ND	ND	ND
SEMIA 518	*Glycine max*	Group XI	ND	ND	ND
SEMIA 527	*Glycine max*	Group XIV	ND	ND	ND
SEMIA 528	*Glycine max*	Group XII	ND	ND	ND
SEMIA 538	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 542	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 543	*Glycine max*	Group XII	ND	ND	ND
SEMIA 549	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 556	*Glycine max*	Group IX	ND	ND	ND
SEMIA 560	*Glycine max*	Group XI	GI-3	GI-3	*Bradyrhizobium* sp.
SEMIA 565	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 566	*Glycine max*	Group XII	ND	ND	ND
SEMIA 567	*Glycine max*	Group XII	ND	ND	ND
SEMIA 568	*Glycine max*	Group XII	ND	ND	ND
SEMIA 571	*Glycine max*	Group XII	ND	ND	ND
SEMIA 574	*Glycine max*	Group XXV	ND	ND	ND
SEMIA 576	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 577	*Glycine max*	Group XXI	ND	ND	ND
SEMIA 579	*Glycine max*	Group XII	ND	ND	ND
SEMIA 580	*Glycine max*	Group XII	GI-4	ND	ND
SEMIA 581	*Glycine max*	Group XXV	ND	ND	ND
SEMIA 583	*Glycine max*	Group VI	ND	ND	ND
SEMIA 584	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 587*	*Glycine max*	Group XX	GII-isolated	GII-1	*B. elkanii*
SEMIA 589	*Glycine max*	Group XXV	ND	ND	ND
SEMIA 590	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 591	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 593	*Glycine max*	Group XII	ND	ND	ND
SEMIA 598	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 613	*Vigna unguiculata*	Group VII	GI-3	ND	ND
SEMIA 621	*Lespedeza striata*	Isolated	GII-5	ND	ND
SEMIA 635	*Vigna unguiculata*	Group VII	ND	ND	ND
SEMIA 637	*Acacia mearnsii*	Isolated	GI-5	ND	ND
SEMIA 656*	*Neonotonia wightii*	Isolated	GI-isolated	GI-4	*Bradyrhizobium* sp.
SEMIA 662*	*Vigna unguiculata*	Group II	ND	ND	ND
SEMIA 695*	*Neonotonia wightii*	Group II	GII-2	GII-1	*B. elkanii*
SEMIA 696†	*Desmodium uncinatum*	Group XX	ND	ND	ND
SEMIA 839†	*Lotus pedunculatus*	Isolated	GI-1	ND	ND
SEMIA 905†	*Ornithopus sativus*	Group V	ND	ND	ND
SEMIA 928†	*Lupinus* sp.	Isolated	GI-2	GI-5	*B. canariense*
SEMIA 929†	*Ornithopus sativus*	Group V	GI-9	ND	ND
SEMIA 938*	*Lupinus albus*	Group II	ND	ND	ND
SEMIA 5000	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 5002	*Glycine max*	Group XVIII	GII-3	ND	ND
SEMIA 5003	*Glycine max*	Group VI	GI-10	ND	ND
SEMIA 5005	*Glycine max*	Group XIII	ND	ND	ND

Table 1. cont.

Strain	Host plant	rep-PCR group	16S rRNA + ITS group	MLSA group	Classification based on MLSA
SEMIA 5010	*Glycine max*	Group XXIII	ND	ND	ND
SEMIA 5011	*Glycine max*	Group XIX	GII-3	GII-1	*B. elkanii*
SEMIA 5012	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 5013	*Glycine max*	Group XXII	ND	ND	ND
SEMIA 5014	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5015	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5016	*Glycine max*	Group XIX	ND	ND	ND
SEMIA 5018	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5019*	*Glycine max*	Group XX	GII-3	GII-1	*B. elkanii*
SEMIA 5020	*Glycine max*	Isolated	GI-6	ND	ND
SEMIA 5021	*Glycine max*	Isolated	GI-6	ND	ND
SEMIA 5022	*Glycine max*	Isolated	GI-7	ND	ND
SEMIA 5023	*Glycine max*	Group XXIII	ND	ND	ND
SEMIA 5024	*Glycine max*	Group XXV	ND	ND	ND
SEMIA 5025	*Glycine max*	Group XXIII	GI-7	GI-1	*B. liaoningense*
SEMIA 5026	*Glycine max*	Isolated	GII-3	GII-1	*B. elkanii*
SEMIA 5027	*Glycine max*	Group XXI	GII-2	GII-1	*B. elkanii*
SEMIA 5028	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5029	*Glycine max*	Group X	GI-4	ND	ND
SEMIA 5030	*Glycine max*	Group XXIII	ND	ND	ND
SEMIA 5032	*Glycine max*	Group XXV	ND	ND	ND
SEMIA 5034	*Glycine max*	Isolated	GI-3	ND	ND
SEMIA 5036	*Glycine max*	Group XXV	GI-6	ND	ND
SEMIA 5037	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5038	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 5039	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5042	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 5043	*Glycine max*	Isolated	GI-6	ND	ND
SEMIA 5044	*Glycine max*	Group XVI	ND	ND	ND
SEMIA 5045	*Glycine max*	Isolated	GI-1	GI-7	*B. japonicum*
SEMIA 5046	*Glycine max*	Isolated	GI-1	ND	ND
SEMIA 5048	*Glycine max*	Isolated	GI-1	ND	ND
SEMIA 5051	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5052	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5055	*Glycine max*	Group X	ND	ND	ND
SEMIA 5056	*Glycine max*	Isolated	GI-1	ND	ND
SEMIA 5057	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5058	*Glycine max*	Group XVII	ND	ND	ND
SEMIA 5059	*Glycine max*	Group XVI	ND	ND	ND
SEMIA 5060	*Glycine max*	Isolated	GI-6	ND	ND
SEMIA 5061	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5062	*Glycine max*	Isolated	GI-10	GI-1	*B. liaoningense*
SEMIA 5063	*Glycine max*	Group XI	ND	ND	ND
SEMIA 5064	*Glycine max*	Group VIII	GI-1	ND	ND
SEMIA 5065	*Glycine max*	Group XIV	ND	ND	ND
SEMIA 5066	*Glycine max*	Group XXII	GII-2	ND	ND
SEMIA 5067	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5068	*Glycine max*	Group XIII	GI-4	ND	ND
SEMIA 5069	*Glycine max*	Group XX	ND	ND	ND
SEMIA 5070	*Glycine max*	Group XX	ND	ND	ND
SEMIA 5071	*Glycine max*	Group VIII	ND	ND	ND
SEMIA 5072	*Glycine max*	Group XXVI	ND	ND	ND
SEMIA 5073	*Glycine max*	Group XXVI	ND	ND	ND
SEMIA 5074	*Glycine max*	Group XXVI	ND	ND	ND
SEMIA 5075	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 5079*	*Glycine max*	Group XV	GI-4	GI-7	*B. japonicum*

Table 1. cont.

Strain	Host plant	rep-PCR group	16S rRNA+ITS group	MLSA group	Classification based on MLSA
SEMIA 5080*	*Glycine max*	Group XVI	GI-6	GI-6	*B. japonicum*
SEMIA 5081	*Glycine max*	Group IX	ND	ND	ND
SEMIA 5082	*Glycine max*	Group XV	ND	ND	ND
SEMIA 5083	*Glycine max*	Group XVII	GI-6	ND	ND
SEMIA 5084	*Glycine max*	Group XVI	ND	ND	ND
SEMIA 5085	*Glycine max*	Isolated	GI-4	ND	ND
SEMIA 5086	*Glycine max*	Group XIII	ND	ND	ND
SEMIA 5087	*Glycine max*	Group XVIII	ND	ND	ND
SEMIA 5090	*Glycine max*	Group XVII	ND	ND	ND
SEMIA 6002*	*Vigna unguiculata*	Group III	ND	ND	ND
SEMIA 6014	*Stylosanthes guianensis*	Isolated	GI-isolated	GI-4	*Bradyrhizobium* sp.
SEMIA 6028*	*Desmodium uncinatum*	Isolated	GII-6	GII-3	*Bradyrhizobium* sp.
SEMIA 6053*	*Clitoria ternatea*	Isolated	GII-1	GII-3	*Bradyrhizobium* sp.
SEMIA 6056	*Arachis hypogaea*	Isolated	GII-5	ND	ND
SEMIA 6057	*Psophocarpus tetragonolobus*	Group XXVII	ND	ND	ND
SEMIA 6059*	*Psophocarpus tetragonolobus*	Group XXVII	GI-6	GI-6	*B. japonicum*
SEMIA 6069†	*Leucaena leucocephala*	Isolated	GII-5	GII-3	*Bradyrhizobium* sp.
SEMIA 6071	*Stylosanthes* sp.	Group IX	ND	ND	ND
SEMIA 6077	*Stylosanthes* sp.	Isolated	GI-12	GI-2	*B. yuanmingense*
SEMIA 6093	*Aeschynomene americana*	Isolated	GII-5	GII-3	*Bradyrhizobium* sp.
SEMIA 6099†	*Dimorphandra exaltata*	Isolated	GII-6	GII-3	*Bradyrhizobium* sp.
SEMIA 6100*	*Albizia falcataria*	Group IV	ND	ND	ND
SEMIA 6101*	*Dalbergia nigra*	Isolated	GII-isolated	GII-2	*Bradyrhizobium* sp.
SEMIA 6118	*Stylosanthes* sp.	Isolated	GII-6	ND	ND
SEMIA 6144*	*Arachis hypogaea*	Group III	GI-12	ND	ND
SEMIA 6145*	*Crotalaria juncea*	Isolated	GII-1	ND	ND
SEMIA 6146*	*Centrosema* sp.	Isolated	GII-6	GII-2	*Bradyrhizobium* sp.
SEMIA 6148*	*Neonotonia wightii*	Isolated	GII-6	GII-isolated	*Bradyrhizobium* sp.
SEMIA 6149*	*Galactia striata*	Group IV	ND	ND	ND
SEMIA 6150*	*Acacia mearnsii*	Isolated	GII-6	ND	ND
SEMIA 6152*	*Calopogonium* sp.	Isolated	GII-4	GII-3	*Bradyrhizobium* sp.
SEMIA 6155*	*Stylosanthes* sp.	Isolated	GI-4	ND	ND
SEMIA 6156*	*Crotalaria spectabilis*	Isolated	GI-9	GI-3	*Bradyrhizobium* sp.
SEMIA 6157*	*Cajanus cajans*	Isolated	GII-6	ND	ND
SEMIA 6158*	*Crotalaria spectabilis*	Isolated	GII-6	ND	ND
SEMIA 6160*	*Albizia lebbek*	Group IV	GII-1	GII-3	*Bradyrhizobium* sp.
SEMIA 6163*	*Acacia mearnsii*	Group I	GI-11	GI-isolated	*Bradyrhizobium* sp.
SEMIA 6164*	*Acacia mearnsii*	Isolated	GI-5	GI-8	*Bradyrhizobium* sp.
SEMIA 6169*	*Albizia falcataria*	Group IV	ND	ND	ND
SEMIA 6175*	*Pueraria phaseoloides*	Isolated	GII-1	ND	ND
SEMIA 6179	*Acacia miersii*	Isolated	GI-isolated	GI-8	*Bradyrhizobium* sp.
SEMIA 6181	*Melanoxylum brauna*	Group XXIV	ND	ND	ND
SEMIA 6186	*Acacia miersii*	Group XXIV	GI-8	GI-8	*Bradyrhizobium* sp.
SEMIA 6187	*Acacia miersii*	Isolated	GI-11	GI-8	*Bradyrhizobium* sp.
SEMIA 6189	*Acacia miersii*	Group I	ND	ND	ND
SEMIA 6192*	*Tipuana tipa*	Isolated	GI-12	GI-3	*Bradyrhizobium* sp.
SEMIA 6208*	*Desmodium ovalifolium*	Isolated	GII-4	ND	ND
SEMIA 6319*	*Arachis* sp.	Isolated	GI-10	GI-2	*B. yuanmingense*
SEMIA 6368	*Centrosema plumieri*	Isolated	GI-8	ND	ND
SEMIA 6374	*Arachis pintoi*	Isolated	GI-12	GI-4	*Bradyrhizobium* sp.
SEMIA 6384*	*Mimosa acutistipula*	Isolated	GII-2	ND	ND
SEMIA 6388†	*Acacia podalyriaefolia*	Group XXVIII	ND	ND	ND
SEMIA 6391*	*Acacia auriculiformis*	Group IV	ND	ND	ND
SEMIA 6420*	*Acacia mangium*	Group I	ND	ND	ND
SEMIA 6424*	*Centrosema pubescens*	Isolated	GII-4	ND	ND
SEMIA 6425*	*Centrosema pubescens*	Isolated	GII-6	ND	ND

Table 1. cont.

Strain	Host plant	rep-PCR group	16S rRNA + ITS group	MLSA group	Classification based on MLSA
SEMIA 6426†	*Erythrina poeppigeana*	Group XVIII	ND	ND	ND
SEMIA 6434†	*Inga* sp.	Isolated	GI-isolated	GI-isolated	*Bradyrhizobium* sp.
SEMIA 6440*	*Arachis pintoi*	Isolated	GII-1	GII-3	*Bradyrhizobium* sp.
SEMIA 6443†	*Acosmium nitens*	Group XXVIII	GII-4	ND	ND

ND, Not determined in this study.
*Analysed by Menna *et al.* (2006, 2009).
†Strains recommended as commercial inoculants in Brazil.

Cluster analyses. In the rep-PCR analysis, the sizes of the fragments were normalized according to the sizes of the DNA marker. Cluster analyses were performed with the Bionumerics program version 4.6 (Applied Maths), using the UPGMA algorithm (Sneath & Sokal, 1973) and the Jaccard coefficient (J; Jaccard, 1912), with the optimum values indicated by the Bionumerics program for the tolerance and optimization parameters.

The 16S rRNA gene, ITS, *glnII*, *recA*, *atpD* and *dnaK* sequences generated were analysed with the programs Phred (Ewing & Green, 1998; Ewing *et al.*, 1998), Phrap (http://www.phrap.org) and Consed (Gordon *et al.*, 1998). The sequences obtained and confirmed in the 3′ and 5′ directions were submitted to GenBank under the accession numbers listed in Supplementary Table S3. The single and concatenated sequences obtained were analysed using the MEGA software version 4.0 with the default parameters, K2P distance model (Kimura, 1980) and the neighbour-joining algorithm (Saitou & Nei, 1987). The conserved, variable and parsimony-informative regions were analysed in consensus sequences containing gaps; therefore, the number of analysed sites is always greater because it includes gaps. A site is parsimony-informative if it contains at least two types of nucleotide (or amino acid) and at least two of them occur with a minimum frequency of two, and the parameter was estimated using the MEGA program. Statistical support for tree nodes was evaluated by bootstrap analyses (Felsenstein, 1985) with 1000 samplings (Hedges, 1992).

Sequences of reference and type strains used for alignment and comparison are detailed in Supplementary Table S4. The genome sequence of *Caulobacter crescentus* CB15 (GenBank accession no. AE005673) was used as an outgroup.

RESULTS

Morpho-physiological properties

All 169 strains were characterized by slow growth and an alkaline reaction in a culture medium containing mannitol as carbon source after 5–7 days of growth, typical morpho-physiological properties of *Bradyrhizobium* strains. Mucus production *in vitro* varied from little to profuse after this period, with no correlation with the host plant or the ecosystem of isolation (data not shown).

rep-PCR genomic fingerprinting

DNA profiles after amplification with the BOX primer were obtained for all 169 strains, with a mean of 20 bands differing in intensity and varying in size from 200 to 5000 bp (Supplementary Fig. S1).

Cluster analysis allowed the identification of numerous well-defined groups of bradyrhizobia, with a high level of intraspecific diversity. Considering a cut-off at 70 % similarity, as suggested for studies of diversity of rhizobia using rep-PCR (Grange & Hungria, 2004; Alberton *et al.*, 2006; Kaschuk *et al.*, 2006a, b), 80 profiles were distinguished and, together with the type strains, all strains were clustered at the very low level of 24 % similarity (Supplementary Fig. S1 and Table 1). In some of the groups, strains isolated from the same host legume and in the same country were clustered with similarities higher than 70 %, e.g. SEMIA strains 6163, 6420 and 6189, isolated from *Acacia* species in Brazil. However, in other groups, strains isolated from different legume species and countries showed similarities as high as 100 %, e.g. SEMIA strains 662, 695 and 938, respectively isolated from *Vigna unguiculata*, *Neotonia wightii* and *Lupinus albus* (Supplementary Fig. S1 and Table 1).

Diversity in the 16S rRNA gene and ITS region

When DNA of the 80 strains representative of each rep-PCR group was submitted to sequence analysis of the 16S rRNA gene and ITS region, unique bands were obtained in each amplification, of approximately 1500 and 700 bp, respectively. For the alignment of the 16S rRNA gene, a region of 1480 bp was considered. No statistical differences were detected among the strains in the contents of T, C, A and G, and the estimated mean frequencies for these nucleotides were respectively 20.1, 24.1, 24.3 and 31.5 % (Table 3). In the consensus sequence of all *Bradyrhizobium* strains, containing 1492 analysed sites, there were 1365 conserved positions, 119 variable positions and 68 parsimony-informative sites.

The phylogenetic tree built with the 16S rRNA gene sequences split the strains into two large groups, with respective bootstrap support of 96 and 99 % (Fig. 1a). The first group included 48 SEMIA strains and the following reference strains: *B. japonicum* USDA 6^T, *B. betae* PL7HG1^T, *B. liaoningense* LMG 18230^T, *B. canariense* BC-C2, *B. yuanmingense* CCBAU 10071^T and *B. japonicum*

Table 2. Primers and DNA amplification conditions used in this study

Mixtures of bases used at certain positions are given as: K, T or G; S, G or C; Y, C or T; R, A or G; M, A or C. Primer positions are given according to the corresponding sequence of *B. japonicum* USDA 110.

Primer	Sequence (5′–3′)	Target gene (position)	PCR cycling conditions	Reference
TSrecAf	CAACTGCMYTGCGTATCGTCGAAGG	*recA* (8–32)	2 min 95 °C, 35 × (45 s 95 °C, 30 s 58 °C, 1.5 min 72 °C), 7 min 72 °C	Stepkowski *et al.* (2005)
TSrecAr	CGGATCTGGTTGATGAAGATCACCATG	*recA* (620–594)		
TSatpDf	TCTGGTCCGYGGCCAGGAAG	*atpD* (189–208)	2 min 95 °C, 35 × (45 s 95 °C, 30 s 58 °C, 1.5 min 72 °C), 7 min 72 °C	Stepkowski *et al.* (2005)
TSatpDr	CGACACTTCCGARCCSGCCTG	*atpD* (804–784)		
TSglnIIf	AAGCTCGAGTACATCTGGCTCGACGG	*glnII* (13–38)	2 min 95 °C, 35 × (45 s 95 °C, 30 s 58 °C, 1.5 min 72 °C), 7 min 72 °C	Stepkowski *et al.* (2005)
TSglnIIr	SGAGCCGTTCCAGTCGGTGTCG	*glnII* (681–660)		
BRdnaKf	TTCGACATCGACGCSAACGG	*dnaK* (1411–1430)	2 min 95 °C, 35 × (45 s 95 °C, 30 s 58 °C, 1.5 min 72 °C), 7 min 72 °C	This study
BRdnaKr	GCCTGCTGCKTGTACATGGC	*dnaK* (1905–1885)		
fD1	AGAGTTTGATCCTGGCTCAG	16S rRNA (9–29)	2 min 95 °C, 30 × (15 s 94 °C, 45 s 93 °C, 45 s 55 °C, 2 min 72 °C), 5 min 72 °C	Weisburg *et al.* (1991)
rD1	CTTAAGGAGGTGATCCAGCC	16S rRNA (1474–1494)		
FGPS1490	TGCGGCTGGATCACCTCCTT	16S rRNA (1490–1510)	3 min 94 °C, 35 × (1 min 94 °C, 1 min 55 °C, 2 min 72 °C), 6 min 72 °C	Laguerre *et al.* (1996)
FGPS130	CCGGGTTTCCCCATTCGG	23S rRNA (148–130)		

USDA 110. In this first group, it was possible to observe 11 subgroups, six of which did not include any type or reference strains, in addition to two isolated strains (*B. betae* PL7HG1^T and SEMIA 839). Bootstrap support for these subgroups ranged from 62 to 95 %. The second group comprised 32 SEMIA strains and the reference strains *B. pachyrhizi* PAC48^T, *B. jicamae* PAC68^T and *B. elkanii* USDA 76^T, and was split in four subgroups and three isolated strains (SEMIA strains 695, 6056 and 6093), with 83 % bootstrap support (Fig. 1a).

When compared with the 16S rRNA gene, greater variability was detected in the ITS region (Fig. 1b). Similar to the results obtained with the 16S rRNA gene, no statistical variation was observed in the percentages of T, C, A and G, and, considering the gaps and insertion of nucleotides observed in some strains, the consensus sequence of the *Bradyrhizobium* strains contained 1115 analysed sites, with 590 conserved positions, 377 variable positions and 294 parsimony-informative sites (Table 3).

The phylogenetic tree of the ITS sequences resulted in two large groups with 76 and 99 % bootstrap support, respectively (Fig. 1b). The first group joined 48 SEMIA strains and all type and reference strains except for *B. jicamae* PAC68^T, *B. pachyrhizi* PAC48^T and *B. elkanii* USDA 76^T. It could be split into more subgroups than observed for the 16S rRNA gene, 12, seven of which did not include any type or reference strain, in addition to nine isolated strains (SEMIA strains 6163, 6014, 6187, 656, 6164, 511, 6179 and 6434 and *B. betae* PL7HG1^T). The second large group contained 32 SEMIA strains and the reference strains *B. jicamae* PAC68^T, *B. pachyrhizi* PAC48^T and *B. elkanii* USDA 76^T; it could be split into eight subgroups, with two strains remaining isolated (SEMIA 6158 and *B. jicamae* PAC68^T). Nine SEMIA strains were clustered with *B. elkanii* USDA 76^T in the first subgroup, and SEMIA 587 showed high genetic divergence in comparison with *B. elkanii* USDA 76^T. In addition, four SEMIA strains were clustered with *B. pachyrhizi* PAC48^T in the eighth subgroup. The strains from the remaining six subgroups showed no relation with type or reference strains (Fig. 1b).

In an attempt to improve the phylogenetic information obtainable from the ribosomal regions, the concatenated sequences of the 16S rRNA gene and ITS region were analysed. This was possible because, in a previous study with strains of *Bradyrhizobium* performed by Willems *et al.* (2003), the authors observed high correlation between ITS sequencing and DNA–DNA hybridization, such that strains with more than 95.5 % ITS sequence similarity belonged to the same genospecies, showing more than 60 % DNA–DNA hybridization. The concatenated sequence of the 16S rRNA gene and ITS region of *Bradyrhizobium* strains from our study presented 2271 nucleotides, with the consensus sequence including gaps consisting of 2597 analysed sites, of which 1919 were conserved, 524 variable and 380 parsimony-informative (Table 3). In the resulting neighbour-joining tree, two large groups were again well defined,

Table 3. Sequence information obtained in this study

Forty (16S rRNA gene, ITS) or 80 (other genes) strains were analysed, together with seven, eight or nine type and reference strains (see Methods).

Locus	Strains analysed (*n*)	Nucleotides (%)				Frequency T/C/A/G (%)
		Conserved	Variable	Parsimony-informative	Total*	
ITS	89	590 (52.9)	377 (33.8)	294 (26.3)	791/1115	26.0/24.7/20.1/29.1
16S rRNA	89	1365 (91.5)	119 (7.9)	68 (4.5)	1480/1492	20.1/24.1/24.3/31.5
16S rRNA + ITS	89	1919 (73.9)	524 (20.1)	380 (14.6)	2271/2597	22.2/24.3/22.8/30.7
atpD	48	221 (36.6)	378 (62.6)	352 (58.2)	574/604	18.6/32.2/16.1/33.2
dnaK	47	270 (69.2)	107 (27.4)	74 (18.9)	370/390	10.3/31.1/24.6/34.0
glnII	49	423 (66.3)	181 (28.3)	132 (20.6)	602/638	17.3/33.1/19.9/29.7
recA	47	377 (71.8)	148 (28.2)	108 (20.5)	525/525	15.9/30.9/17.2/36.0
ITS	49	574 (52.3)	380 (34.6)	271 (24.7)	793/1097	26.0/24.8/20.0/29.2
16S rRNA	49	1373 (91.9)	111 (7.4)	65 (4.3)	1481/1493	20.1/24.0/24.3/31.5
Concatenated genes	47	1315 (60.6)	791 (36.5)	658 (30.3)	2075/2167	19.2/27.9/21.0/31.8
Concatenated genes without *atpD*	47	1089 (70.0)	417 (26.8)	308 (19.8)	1498/1555	19.3/27.3/21.8/31.6

*Mean number of nucleotides amplified/number of sites analysed, including gaps.

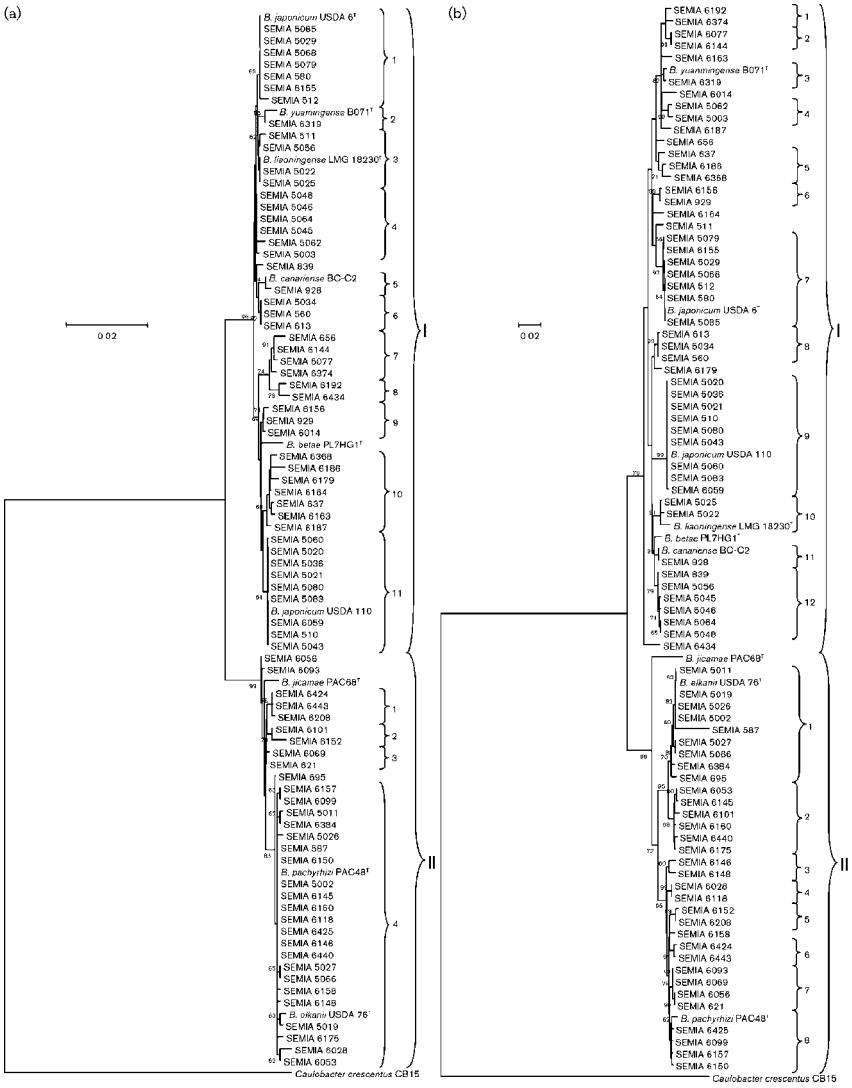

Fig. 1. Phylogenetic relationships of 90 taxa based on the 16S rRNA gene (a) and the ITS region (b). Phylogeny was inferred using the neighbour-joining method. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Evolutionary distances were computed using the maximum composite likelihood method and are in units of the number of base substitutions per site. Positions containing gaps and missing data were eliminated from the dataset (complete deletion option). Phylogenetic analyses were conducted in MEGA4.

both with bootstrap support of 99 % (Fig. 2). The first comprised 12 subgroups, with five SEMIA strains (SEMIA 511, 6179, 6014, 656 and 6434) and *B. betae* PL7HG1^T remaining isolated. All type and reference strains were included in this first group, except *B. elkanii* USDA 76^T, *B. jicamae* PAC68^T and *B. pachyrhizi* PAC48^T. The second group contained 32 strains and the latter three type and reference strains and could be split into six subgroups, with two SEMIA strains 6101 and 587 and *B. jicamae* PAC68^T occupying isolated positions (Fig. 2).

In general, in the trees built individually with the 16S rRNA gene and ITS sequences, it was not possible to establish a clear relation between the clustering of the strains and their host plant species or their geographical origin.

Diversity in the *atpD*, *dnaK*, *glnII* and *recA* gene sequences

The additional housekeeping genes selected to refine the phylogenetic analysis in this study are highly conserved among bacteria of the order *Rhizobiales*, and are dispersed in the genome of *B. japonicum* USDA 110. They encode the following important proteins: *atpD* encodes the ATP synthase beta-chain, *glnII* encodes glutamine synthetase II, *recA* encodes the recombination protein RecA and *dnaK* encodes a 70 kDa chaperone protein. Forty strains representative of the subgroups of the concatenated tree built with the 16S rRNA gene and ITS sequences, in addition to the type and reference strains, were selected for the analysis and are indicated with asterisks in Fig. 2. The amplification resulted in fragments with mean lengths of 574 bp for *atpD*, 370 bp for *dnaK*, 602 bp for *glnII* and 525 bp for *recA*; the sequence characteristics are provided in Table 3. The lowest level of conservation was observed for *atpD*, only 36.6 %; for the other sequences, the levels of conservation were 69.2 % for *dnaK*, 66.3 % for *glnII*, 71.8 % for *recA* and 52.3 % for the ITS region. For each housekeeping gene, phylogenetic trees were constructed using the neighbour-joining method and resulted in distinct groups (Figs 3 and 4). When neighbour-joining trees were built with the deduced amino acid sequences, clustering was not as clear (data not shown).

The greatest variability was observed with the *atpD* gene (Table 3, Fig. 3a), and five strains (SEMIA strains 5026, 5027, 587, 5011 and 5019) formed a distinct cluster. These strains were isolated from soybean plants (*Glycine max* L.) in Brazil, the USA and Thailand (Supplementary Table S1) and, in the 16S rRNA gene and ITS analyses, they showed higher similarity to *B. elkanii* USDA 76^T. Interestingly, the PCR products obtained with *atpD* primers for these five strains were larger (approx. 700 bp) than for the other strains. Although the similarity values were low, BLASTX showed higher similarity (~50 %) of these sequences to a putative insertion sequence transposase protein of *Rhizobium etli* CFN 42^T (Gene ID: 1005217 RHE_PD00002), while all other subgroups showed high *atpD* sequence similarity to type and/or reference strains of *Bradyrhizobium*. In general, the subgroups formed with *atpD* were also observed in the 16S rRNA gene tree, e.g. the clusters with *B. liaoningense* LMG 18230^T, *B. yuanmingense* CCBAU 10071^T and *B. japonicum* USDA 6^T. It is noteworthy that SEMIA strains 5080 and 6059 were clustered with *B. japonicum* USDA 110, in a different group from that of the type strain USDA 6^T (Fig. 3a), similar to what was observed in the 16S rRNA gene and ITS analyses (Figs 1 and 2).

The two large groups observed in the 16S rRNA gene and ITS trees were also detected in the *dnaK* and *glnII* trees, but were not well defined in the *atpD* or *recA* trees. However, three subgroups were observed in all trees, and two included type or reference strains, *B. japonicum* USDA 6^T and SEMIA strains 512 and 5079, both isolated from soybean, and *B. canariense* BC-C2 and SEMIA 928, the latter isolated from *Lupinus* sp. and of unknown geographical origin. Another noteworthy group present in all trees was *B. japonicum* USDA 110 and SEMIA strains 5080 and 6059, all isolated from soybean. Furthermore, the subgroup that included *B. elkanii* USDA 76^T and SEMIA strains 5026, 587, 5027, 5019 and 5011, also isolated from soybean, was present in all trees except that generated with *atpD*. *B. yuanmingense* CCBAU 10071^T was also always clustered with SEMIA 6319, isolated from *Arachis* sp. in Bolivia, except in the tree generated with *glnII* sequences. *B. liaoningense* LMG 18230^T and SEMIA 5025, isolated from soybean in Thailand, were also always clustered together except in the *glnII* tree. Interestingly, SEMIA strains 511 and 5045 were clustered with *B. japonicum* USDA 6^T in the analyses of all four housekeeping genes, but not when the 16S rRNA gene or ITS region were considered.

Analysis of concatenated *atpD*, *dnaK*, *glnII* and *recA* sequences

Assuming that each tree represents a partial view of the information contained in the genome, all four sequences were concatenated in order to gain a fuller understanding of the strains. The concatenation resulted in a consensus sequence of 2075 bp with 2167 analysed sites; 1315 were conserved, 791 were variable and 658 were parsimony-informative (Table 3).

The tree built with the concatenated genes resulted in two large groups, with a bootstrap support of 100 % (Fig. 5). Once more, the first group contained 23 SEMIA strains together with the type or reference strains of *B. japonicum*, *B. liaoningense*, *B. canariense* and *B. yuanmingense* and *B. japonicum* USDA 110. However, within this group, ten subgroups were formed, and some subgroups were clearly distantly related to the type strains, such as subgroups 2, 3, 4, 9 and 10, in addition to three isolated strains, SEMIA strains 656, 6163 and 6186. The second large group comprised 17 SEMIA strains and *B. elkanii* USDA 76^T, with four subgroups and with SEMIA strains 6148 and 695 and *B. elkanii* USDA 76^T remaining isolated. The fourth subgroup, containing SEMIA strains 5027, 587, 5011,

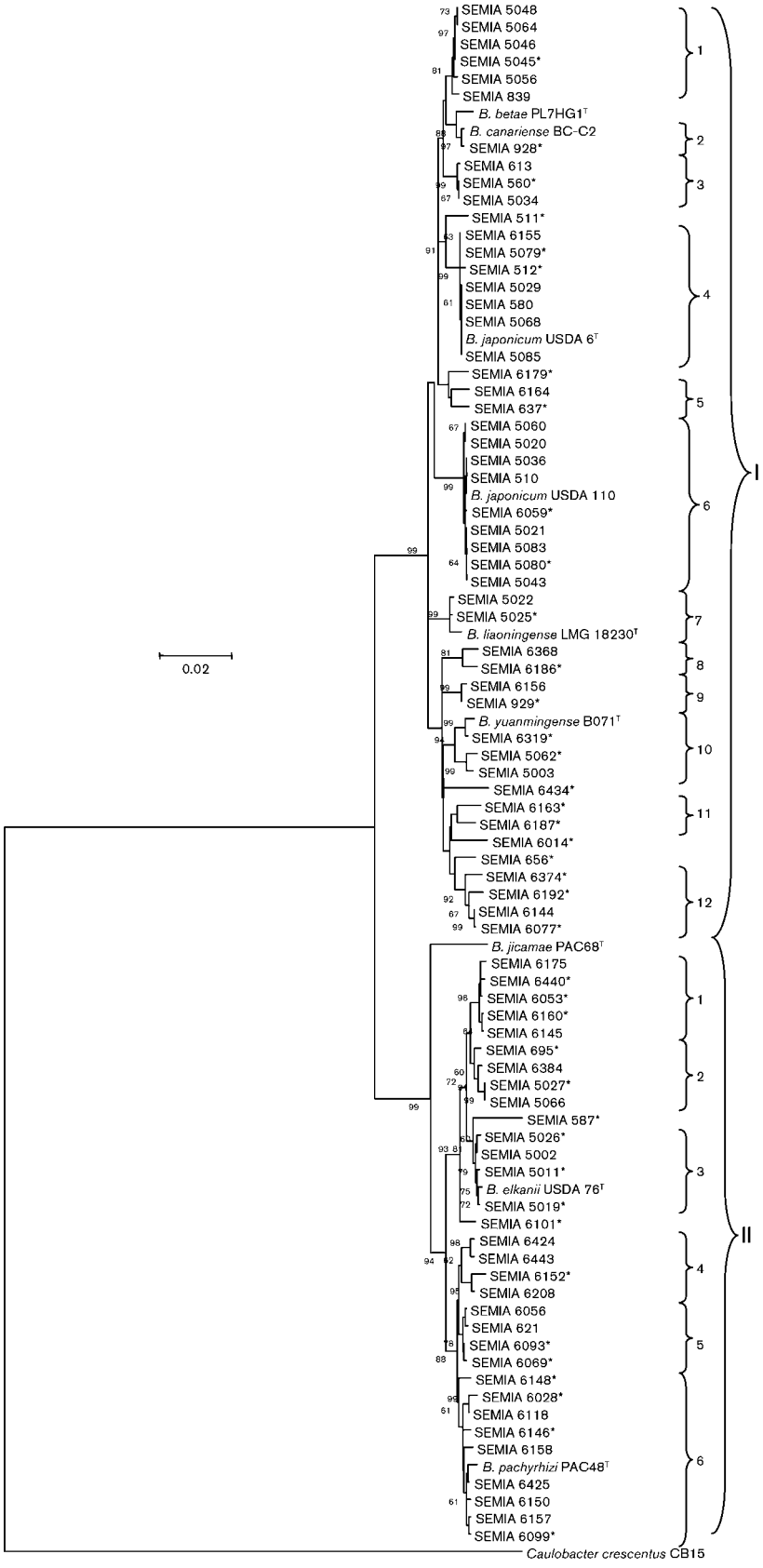

Fig. 2. Phylogenetic relationships of 90 taxa based on concatenated 16S rRNA gene and ITS region sequences. Asterisks (*) indicate strains that were selected for MLSA. See legend to Fig. 1 for further details.

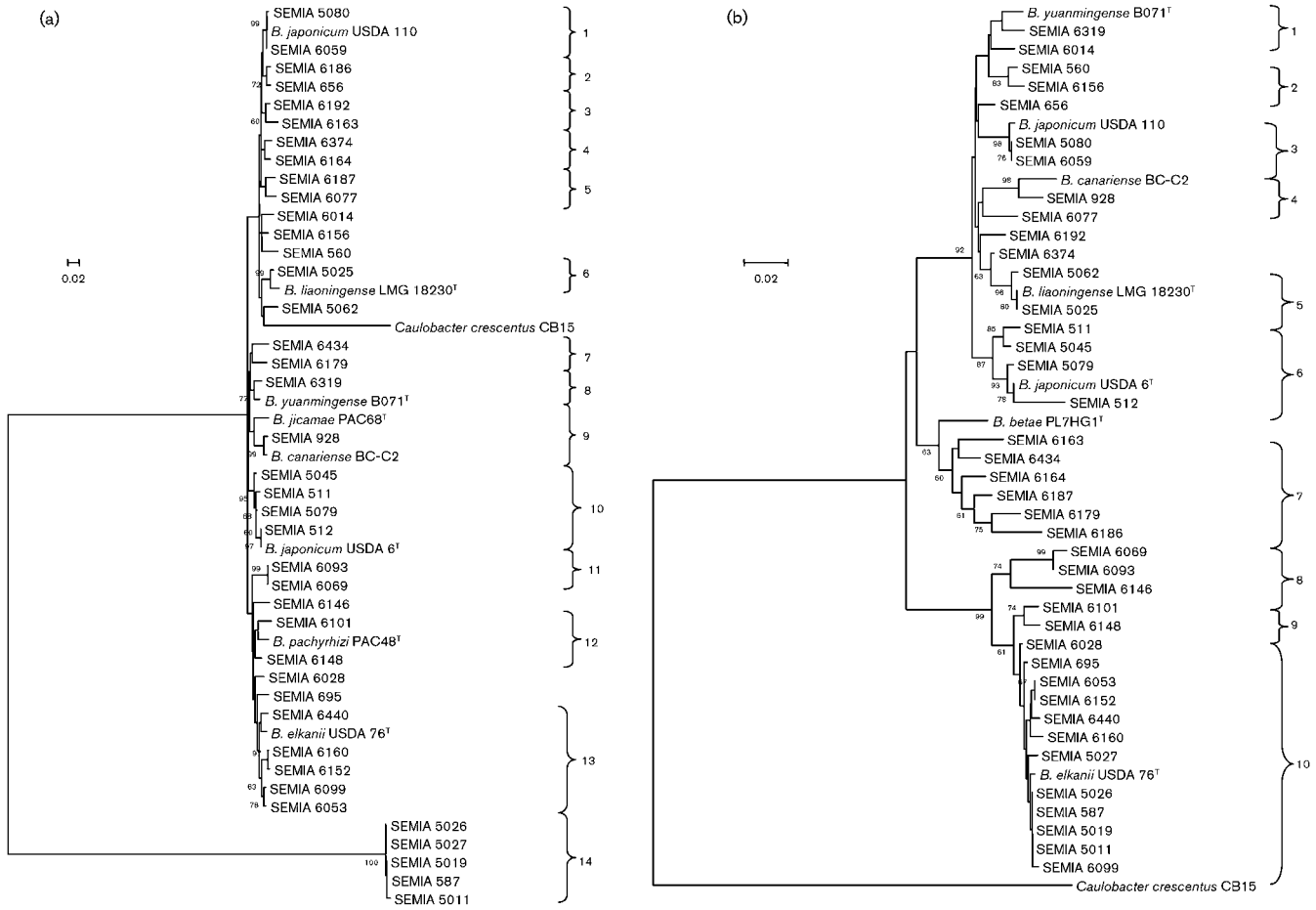

Fig. 3. Phylogenetic trees reconstructed using the neighbour-joining method for 49 strains based on *atpD* (a) and 48 strains based on *dnaK* (b). Codon positions included were first + second + third + non-coding. See legend to Fig. 1 for further details.

5026 and 5019, was more distant from the others (Fig. 5), which might be attributed to the variability in the *atpD* gene.

We then built another tree without the *atpD* gene (Fig. 6). The consensus sequence was 1498 bp long, with 1089 conserved sites, 417 variable but informative sites and 308 parsimony-informative sites (Table 3). The first large group (Fig. 6) was similar to that observed in the tree built with all housekeeping genes (Fig. 5). However, differences were observed in the second large group, in which SEMIA strains 5026, 587, 5027, 5019 and 5011 were clustered with *B. elkanii* USDA 76^T (Fig. 6). The large group, subgroups and isolated strains formed in the analysis of concatenated sequences are detailed in Table 1.

DISCUSSION

The 169 SEMIA strains used in this study had been isolated from 43 distinct host legumes, 104 of them from soybean. Forty of these strains were obtained from selection programmes aimed at identifying effective rhizobial strains, and today are officially recommended for use in commercial inoculants for application to 64 legumes. Previous analysis of these 'elite strains' by Menna *et al.* (2006, 2009) revealed a level of genetic diversity beyond that described previously for the genus *Bradyrhizobium* (e.g. van Berkum & Fuhrmann, 2000; Willems *et al.*, 2001a, b; Qian *et al.*, 2003; Vinuesa *et al.*, 2005b), suggesting the need for further analyses to refine phylogenetic relationships and taxonomic classification. The additional strains and housekeeping genes included in this study confirmed the presence of high intra- and interspecific genetic diversity within the genus *Bradyrhizobium*.

Genetic characterization and evaluation of *Bradyrhizobium* strains has greatly improved with the availability of several molecular techniques (van Berkum & Fuhrmann, 2000; Willems *et al.*, 2003; Rodríguez-Navarro *et al.*, 2004; Stepkowski *et al.*, 2005). One example is the amplification of repetitive regions dispersed in the genome of eubacteria (rep-PCR), such as the sequences known as BOX, which result in highly characteristic bacterial fingerprints

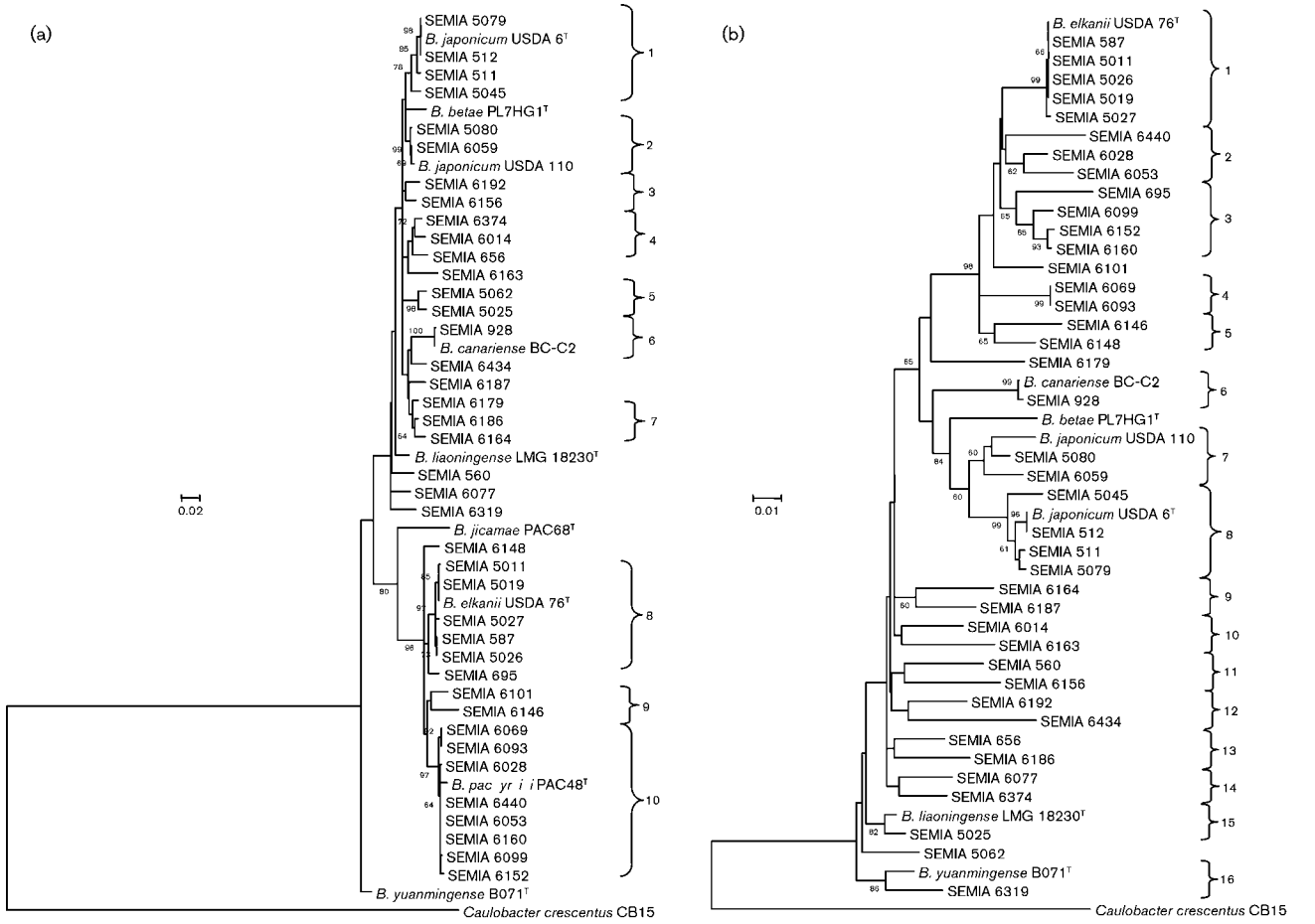

Fig. 4. Phylogenetic trees inferred using the neighbour-joining method for 50 strains based on *glnII* (a) and 48 strains based on *recA* (b). See legends to Figs 1 and 3 for further details.

(Mostasso *et al.*, 2002; Saldaña *et al.*, 2003; Hungria *et al.*, 2006; Wang *et al.*, 2006; Pinto *et al.*, 2007; Menna *et al.*, 2009). rep-PCR has been used to assess genetic diversity of strains isolated from soybean (e.g. Abaidoo *et al.*, 2000; Batista *et al.*, 2007; Loureiro *et al.*, 2007), *Lupinus* species (Barrera *et al.*, 1997), *Acacia albida* (Dupuy & Dreyfus, 1992), *Aeschynomene* species (Wong *et al.*, 1994) and other legumes (Parker & Lunk, 2000; Willems *et al.*, 2003; Wolde-Meskel *et al.*, 2004). Similarly, our BOX-PCR profiles have shown remarkable diversity among *Bradyrhizobium* strains isolated from various legumes and countries, such that they share the low level of similarity of 24 %.

BOX-PCR fingerprinting is particularly useful for identification of elite strains, with application in maintaining quality control of both culture collections and commercial inoculants. However, the current definition of a novel species as well as its classification is based on a polyphasic classification, taking into account phenotypic as well as genetic characteristics (Vandamme *et al.*, 1996), with the sequence of the 16S rRNA gene providing the backbone of the classification (Willems & Collins, 1993; Yanagi & Yamasato, 1993; Young & Haukka, 1996; Garrity & Holt, 2001; Young *et al.*, 2001). In *Bradyrhizobium*, as a result of the limited diversity of the 16S rRNA gene sequence (Olsen & Woese, 1993; Ludwig & Schleifer, 1994; Barrera *et al.*, 1997; Wang & Martínez-Romero, 2000), only six species have been accepted (Jordan, 1982; Kuykendall *et al.*, 1992; Xu *et al.*, 1995; Yao *et al.*, 2002; Rivas *et al.*, 2004; Vinuesa *et al.*, 2005a), in addition to two recently described species (Ramírez-Bahena *et al.*, 2009).

The proposed strategy to improve the definition of species – the analysis of additional conserved ribosomal regions, the 23S rRNA gene and the ITS (Willems *et al.*, 2001b, 2003), which may reveal higher genetic diversity within *Bradyrhizobium* (Doignon-Bourcier *et al.*, 2000; Willems *et al.*, 2001b, 2003; van Berkum & Eardly, 2002; Stepkowski *et al.*, 2005; Germano *et al.*, 2006) – is supported in our study by our analysis of the ITS region (Fig. 1b). Willems *et al.* (2003), using ITS sequencing and DNA–DNA hybridization, observed a strong correlation between the results obtained with the two methods, such that strains with more than 95.5 % ITS sequence similarity belonged to

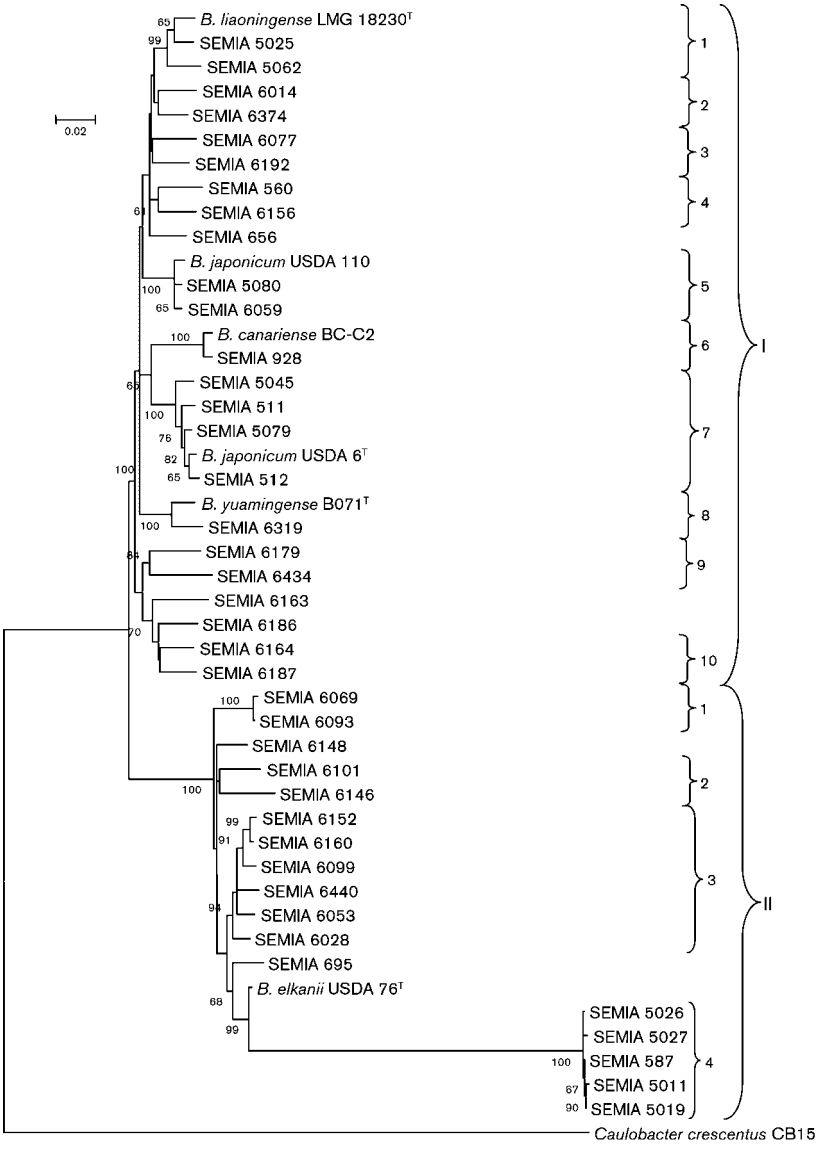

Fig. 5. Evolutionary tree inferred using the neighbour-joining method for 47 strains based on concatenated sequences of the *atpD*, *dnaK*, *glnII* and *recA* genes. See legends to Figs 1 and 3 for further details.

the same genospecies, showing more than 60 % DNA–DNA hybridization. Indeed, based on the results of Willems *et al.* (2003), we have being able to concatenate the 16S rRNA gene and ITS sequences for the *Bradyrhizobium* strains used in this study. However, that may not be possible for other rhizobial species, where the rates of evolution of the two genes might be different and different copies of the ITS are also often observed. Finally, based on the proposal of Willems *et al.* (2003) of a cut-off at 95.5 % similarity, some clusters observed in our study with high bootstrap support in the ITS analysis should represent novel species, particularly SEMIA strains 6163, 6014, 6187, 656, 6164, 511, 6179, 6434 and 6158 and the strains in subgroups 1, 2, 4, 5, 6, 8 and 12 of large group I and in subgroups 2–7 of large group II (Fig. 1b). In conclusion, although controversy remains over the use of the ITS to resolve evolutionary relationships properly among strains of *Bradyrhizobium* (Vinuesa *et al.*, 2005b), in our study, the use of the ITS sequence in association with the 16S rRNA gene sequence undoubtedly improved phylogenetic definition (Fig. 2).

It has been proposed that sequences of at least four housekeeping genes should be used for phylogenetic analysis and taxonomic classification of bradyrhizobia (Stepkowski *et al.*, 2003, 2005; Vinuesa *et al.*, 2005a). This MLSA approach combines congruent genomic data, excluding partitions with a significant level of incongruence, providing reliable genotypic characterization (Bull *et al.*, 1993; Miyamoto & Fitch, 1995). Among the housekeeping genes used in MLSA for phylogenetic reconstructions of the order *Rhizobiales* are *atpD*, *dnaK*, *glnII*, *gltA*, *glnA*, *recA*, *rpoB* and *thrC* (Turner & Young, 2000; Gaunt *et al.*, 2001; Weir *et al.*, 2004; Stepkowski *et al.*, 2005; Vinuesa *et al.*, 2005a; Martens *et al.*, 2007; Ribeiro *et al.*, 2009), and the approach has already been used in

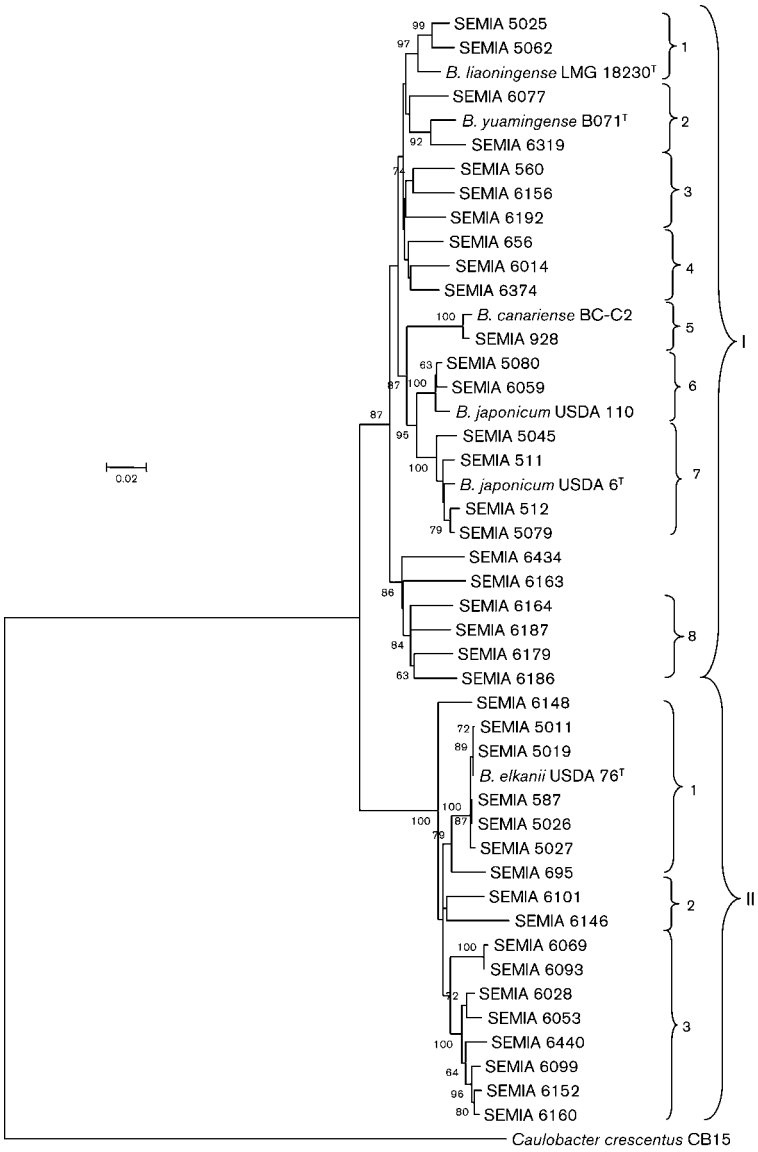

Fig. 6. Evolutionary tree inferred using the neighbour-joining method for 47 strains based on concatenated sequences of the *dnaK*, *glnII* and *recA* genes. See legends to Figs 1 and 3 for further details.

studies with the genus *Bradyrhizobium* (Stepkowski *et al.*, 2003, 2005; Parker, 2004). One important example is the description of *B. canariense*, based on the phylogeny of *atpD*, *glnII*, *recA* and the ITS (Vinuesa *et al.*, 2005a). Likewise, the application of MLSA in our study clearly contributed to a better-defined phylogeny, as well as to the identification of new subgroups highly indicative of novel species. One group that should be emphasized in our study is that composed of *B. japonicum* USDA 110 and SEMIA strains 5080 and 6059, isolated from soybean, which were clustered in all six trees in a different group from that of *B. japonicum* USDA 6^T, clearly indicating the existence of a novel species. Complementary studies are now in progress in our laboratory to describe this novel species.

This study also identified sequences that are related to transposases when the DNA of SEMIA strains 5011, 5019, 587, 5027 and 5026, all symbionts of soybean, was subjected to a PCR with primers specific for the *atpD* gene (Table 3). These strains were isolated in Brazil, USA and Thailand, i.e. they show a broad geographical distribution. In other studies with rhizobia, the use of *atpD* for phylogenetic inference and taxonomic classification has proven to be highly correlated with the phylogeny based on the 16S rRNA gene (Gaunt *et al.*, 2001; Stepkowski *et al.*, 2005; Vinuesa *et al.*, 2005a; Wang *et al.*, 2007; Santillana *et al.*, 2008), and the differences seen in this study have not been detected before. For the *atpD* gene, insertions have been reported of 15 bp in *Azorhizobium caulinodans* and *Brevundimonas bullata* and of 12 bp in *B. japonicum*, providing support for the monophyly of these three species (Gaunt *et al.*, 2001). However, in this study, the high genetic heterogeneity of *atpD* could have resulted from genetic recombination of this gene. Genetic information from such genes can be useful to determine the genetic nature of microbial species

(Achtman & Wagner, 2008). Therefore, we suggest more studies with *atpD*, and recommend that the gene should be used with caution in studies of phylogeny and taxonomy of bacteria belonging to the genus *Bradyrhizobium*.

An important consideration is that, in general, all subgroups formed in the trees built with the housekeeping genes *atpD*, *dnaK*, *glnII* and *recA*, as well as with the ITS, were similar to those obtained with the 16S rRNA gene, confirming the usefulness of these genes for phylogenetic reconstruction of the genus *Bradyrhizobium*. In our study, geographical origin did not affect the patterns of housekeeping genes, thus reinforcing the belief of a common origin for *Bradyrhizobium* with subsequent diffusion of the strains by soil-contaminated seeds (Pérez-Ramírez *et al.*, 1998) or by artificial inoculation (Strijdom, 1998).

In conclusion, concatenation of housekeeping gene sequences has been suggested as a powerful method to improve phylogenetic analysis of *Bradyrhizobium* (Stepkowski *et al.*, 2005; Vinuesa *et al.*, 2005a; Martens *et al.*, 2007) and, indeed, in our study, the tree built with the concatenated *atpD*, *dnaK*, *recA* and *glnII* gene sequences, as well as the tree built with concatenated 16S rRNA gene and ITS sequences, detected high intra- and interspecies genetic variation. Therefore, although the use of a small number of housekeeping genes cannot replace DNA–DNA hybridization in species delineation at present, MLSA has proven to be a reliable technique that is applicable to phylogeny and taxonomic studies of *Bradyrhizobium*, improving the definition of novel species. Finally, the analysis of four housekeeping genes in addition to the 16S rRNA gene and ITS has clearly indicated that there are novel species of *Bradyrhizobium* still to be described.

ACKNOWLEDGEMENTS

The work was partially supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil), projects 552393/2005-3 and 577933/2008-6. The authors thank Professor Fabiano L. Thompson and Professor Manuel Megías for extremely useful suggestions and discussion. The authors also thank Ligia Maria Oliveira Chueire (Embrapa Soja) for help in several steps of this work and Dr Manuel Megias for helpful discussion. P. M. has received a fellowship from CNPq (project 505946/2004-1) and F. G. B. a fellowship from CAPES. M. H. is also a research fellow of CNPq (300698/2007-0).

REFERENCES

Abaidoo, R. C., Keyser, H. H., Singleton, P. W. & Borthakur, D. (2000). *Bradyrhizobium* spp. (TGx) isolates nodulating the new soybean cultivars in Africa are diverse and distinct from bradyrhizobia that nodulate North American soybeans. *Int J Syst Evol Microbiol* **50**, 225–234.

Achtman, M. & Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nat Rev Microbiol* **6**, 431–440.

Alberton, O., Kaschuk, G. & Hungria, M. (2006). Sampling effects on the assessment of genetic diversity of rhizobia associated with soybean and common bean. *Soil Biol Biochem* **38**, 1298–1307.

Barrera, L. L., Trujillo, M. E., Goodfellow, M., Garcia, F. J., Hernandez-Lucas, I., Davila, G., van Berkum, P. & Martinez-Romero, E. (1997). Biodiversity of bradyrhizobia nodulating *Lupinus* spp. *Int J Syst Bacteriol* **47**, 1086–1091.

Batista, J. S. S., Hungria, M., Barcellos, F. G., Ferreira, M. C. & Mendes, I. C. (2007). Variability in *Bradyrhizobium japonicum* and *B. elkanii* seven years after introduction of both the exotic microsymbiont and the soybean host in a Cerrados soil. *Microb Ecol* **53**, 270–284.

Brett, P. J., Deshazer, D. & Woods, D. E. (1998). *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *Int J Syst Bacteriol* **48**, 317–320.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L. & Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analyses. *Syst Biol* **42**, 384–397.

Coenye, T., Gevers, D., Van de Peer, Y., Vandamme, P. & Swings, J. (2005). Towards a prokaryotic genomic taxonomy. *FEMS Microbiol Rev* **29**, 147–167.

Cooper, J. E. & Feil, E. J. (2004). Multilocus sequence typing – what is resolved? *Trends Microbiol* **12**, 373–377.

Doignon-Bourcier, F., Willems, A., Coopman, R., Laguerre, G., Gillis, M. & de Lajudie, P. (2000). Genotypic characterization of *Bradyrhizobium* strains nodulating small Senegalese legumes by 16S–23S rRNA interfgenic gene spacers and amplified length polymorphism fingerprint analyses. *Appl Environ Microbiol* **66**, 3987–3997.

Dupuy, N. C. & Dreyfus, B. L. (1992). *Bradyrhizobium* populations occur in deep soil under the leguminous tree *Acacia albida*. *Appl Environ Microbiol* **58**, 2415–2419.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Garrity, G. M. & Holt, J. G. (2001). The road map to the *Manual*. In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol. 1, pp. 119–166. Edited by D. R. Boone, R. W. Castenholz & G. M. Garrity. New York: Springer.

Gaunt, M. W., Turner, S. L., Rigottier-Gois, L., Lloyd-Macgilp, S. A. & Young, J. P. (2001). Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int J Syst Evol Microbiol* **51**, 2037–2048.

Germano, M. G., Menna, P., Mostasso, F. L. & Hungria, M. (2006). RFLP analysis of the rRNA operon of a Brazilian collection of bradyrhizobial strains from 33 legume species. *Int J Syst Evol Microbiol* **56**, 217–229.

Giongo, A., Ambrosini, A., Vargas, L. K., Freire, J. R. J., Bodanese-Zanettini, M. H. & Passaglia, L. M. P. (2008). Evaluation of genetic diversity of bradyrhizobia strains nodulating soybean [*Glycine max* (L.) Merrill] isolated from South Brazilian fields. *Appl Soil Ecol* **38**, 261–269.

Godoy, D., Randle, G., Simpson, A. J., Aanensen, D. M., Pitt, T. L., Kinoshita, R. & Spratt, B. G. (2003). Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. *J Clin Microbiol* **41**, 2068–2079.

Gordon, D., Abajian, C. & Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Res* **8**, 195–202.

Grange, L. & Hungria, M. (2004). Genetic diversity of indigenous common bean (*Phaseolus vulgaris*) rhizobia in two Brazilian ecosystems. *Soil Biol Biochem* **36**, 1389–1398.

Hedges, S. B. (1992). The number of replications needed for accurate estimation of the bootstrap p-value in phylogenetic studies. *Mol Biol Evol* **9**, 366–369.

Hungria, M., Chueire, L. M. O., Megías, M., Lamrabet, Y., Probanza, A., Guttierrez-Mañero, F. J. & Campo, R. J. (2006). Genetic diversity of indigenous tropical fast-growing rhizobia isolated from soybean nodules. *Plant Soil* **288**, 343–356.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytol* **11**, 37–50.

Jordan, D. C. (1982). Transfer of *Rhizobium japonicum* Buchan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing root nodule bacteria from leguminous plants. *Int J Syst Bacteriol* **32**, 136–139.

Kaschuk, G., Hungria, M., Andrade, D. S. & Campo, R. J. (2006a). Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. *Appl Soil Ecol* **32**, 210–220.

Kaschuk, G., Hungria, M., Santos, J. C. P. & Berton-Junior, J. F. (2006b). Differences in common bean rhizobial populations associated with soil tillage management in southern Brazil. *Soil Tillage Res* **87**, 205–217.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Koeuth, T., Versalovic, J. & Lupski, J. R. (1995). Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. *Genome Res* **5**, 408–418.

Konstantinidis, K. T. & Tiedje, J. M. (2004). Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc Natl Acad Sci U S A* **101**, 3160–3165.

Kuykendall, L. D., Saxena, B., Devine, T. E. & Udell, S. E. (1992). Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. *Can J Microbiol* **38**, 501–505.

Laguerre, G., Mavingui, P., Allard, M. R., Charnay, M. P., Louvrier, P., Mazurier, S. I., Rigottier-Gois, L. & Amarger, N. (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.

Liu, J., Wang, E. T. & Chen, W. X. (2005). Diverse rhizobia associated with woody legumes *Wisteria sinensis*, *Cercis racemosa* and *Amorpha fruticosa* grown in the temperate zone of China. *Syst Appl Microbiol* **28**, 465–477.

Loureiro, M. F., Kaschuk, G., Alberton, O. & Hungria, M. (2007). Soybean [*Glycine max* (L.) Merrill] rhizobial diversity in Brazilian oxisols under various soil, cropping and inoculation managements. *Biol Fertil Soils* **43**, 665–674.

Ludwig, W. & Schleifer, K. H. (1994). Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol Rev* **15**, 155–173.

Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K. & other authors (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* **95**, 3140–3145.

Martens, M., Delaere, M., Coopman, R., De Vos, P., Gillis, M. & Willems, A. (2007). Multilocus sequence analysis of *Ensifer* and related taxa. *Int J Syst Evol Microbiol* **57**, 489–503.

Menna, P., Hungria, M., Barcellos, F. G., Bangel, E. V., Hess, P. N. & Martinez-Romero, E. (2006). Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains used in Brazilian commercial inoculants. *Syst Appl Microbiol* **29**, 315–332.

Menna, P., Pereira, A. A., Bangel, E. V. & Hungria, M. (2009). rep-PCR of tropical rhizobia for strain fingerprinting, biodiversity appraisal and as a taxonomic and phylogenetic tool. *Symbiosis* **48**, 120–130.

Miyamoto, M. M. & Fitch, W. M. (1995). Testing species phylogenies and phylogenetic methods with congruence. *Syst Biol* **44**, 64–76.

Mostasso, L., Mostasso, F. L., Dias, B. G., Vargas, M. A. T. & Hungria, M. (2002). Selection of bean (*Phaseolus vulgaris* L.) rhizobial strains for the Brazilian Cerrados. *Field Crops Res* **73**, 121–132.

Moulin, L., Béna, G., Boivin-Masson, C. & Stepkowski, T. (2004). Phylogenetic analysis of symbiotic nodulation genes support vertical and lateral gene co-transfer within the *Bradyrhizobium* genus. *Mol Phylogenet Evol* **30**, 720–732.

Olsen, G. J. & Woese, C. R. (1993). Ribosomal RNA: a key to phylogeny. *FASEB J* **7**, 113–123.

Parker, M. A. (2004). rRNA and *dnaK* relationships of *Bradyrhizobium* sp. nodule bacteria from four Papilionoid legume trees in Costa Rica. *Syst Appl Microbiol* **27**, 334–342.

Parker, M. A. & Lunk, A. (2000). Relationships of bradyrhizobia from *Platypodium* and *Machaerium* (Papilionoideae: tribe Dalbergieae) on Barro Colorado Island, Panama. *Int J Syst Evol Microbiol* **50**, 1179–1186.

Pérez-Ramírez, N. O., Rogel, M. A., Wang, E., Castellanos, J. Z. & Martínez-Romero, E. (1998). Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. *FEMS Microbiol Ecol* **26**, 289–296.

Pinto, F. G. S., Hungria, M. & Mercante, F. M. (2007). Polyphasic characterization of Brazilian *Rhizobium tropici* strains effective in fixing N_2 with common bean (*Phaseolus vulgaris* L.). *Soil Biol Biochem* **39**, 1851–1864.

Qian, J., Kwon, S. W. & Parker, M. A. (2003). rRNA and *nifD* phylogeny of *Bradyrhizobium* from sites across the Pacific Basin. *FEMS Microbiol Lett* **219**, 159–165.

Ramírez-Bahena, M.-H., Peix, A., Rivas, P., Camacho, M., Rodríguez-Navarro, D. N., Mateos, P. F., Martínez-Molina, E., Willems, A. & Velázquez, E. (2009). *Bradyrhizobium pachyrhizi* sp. nov. and *Bradyrhizobium jicamae* sp. nov., isolated from effective nodules of *Pachyrhizus erosus*. *Int J Syst Evol Microbiol* **59**, 1929–1934.

Ribeiro, R. A., Barcellos, F. G., Thompson, F. L. & Hungria, M. (2009). Multilocus sequence analysis of Brazilian *Rhizobium* microsymbionts of common bean (*Phaseolus vulgaris* L.) reveals unexpected taxonomic diversity. *Res Microbiol* **160**, 297–306.

Rivas, R., Willems, A., Palomo, J. L., García-Benavides, P., Mateos, P. F., Martínez-Molina, E., Gillis, M. & Velázquez, E. (2004). *Bradyrhizobium betae* sp. nov., isolated from roots of *Beta vulgaris* affected by tumour-like deformations. *Int J Syst Evol Microbiol* **54**, 1271–1275.

Rodríguez-Navarro, D. N., Camacho, M., Leidi, E. O., Rivas, R. & Velázquez, E. (2004). Phenotypic and genotypic characterization of rhizobia from diverse geographical origin that nodulate *Pachyrhizus* species. *Syst Appl Microbiol* **27**, 737–745.

Saito, A., Mitsui, H., Hattori, R., Minamisawa, K. & Hattori, T. (1998). Slow-growing and oligotrophic soil bacteria phylogenetically close to *Bradyrhizobium japonicum*. *FEMS Microbiol Ecol* **25**, 277–286.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Saldaña, G., Martinez-Alcántara, V., Vinardell, J. M., Bellogín, R., Ruíz-Sainz, J. E. & Balatti, P. A. (2003). Genetic diversity of fast-growing rhizobia that nodulate soybean (*Glycine max* L. Merr.). *Arch Microbiol* **180**, 45–52.

Santillana, N., Ramírez-Bahena, M. H., García-Fraile, P., Velázquez, E. & Zúñiga, D. (2008). Phylogenetic diversity based on *rrs*, *atpD*, *recA* genes and 16S–23S intergenic sequence analyses of rhizobial strains isolated from *Vicia faba* and *Pisum sativum* in Peru. *Arch Microbiol* **189**, 239–247.

Sawada, H., Kuykendall, L. D. & Young, J. M. (2003). Changing concepts in the systematics of bacterial nitrogen-fixing legume symbionts. *J Gen Appl Microbiol* **49**, 155–179.

Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical Taxonomy*. San Francisco: W. H. Freeman.

Sprent, J. I. (2001). *Nodulation in Legumes*. Kew, UK: Royal Botanic Gardens.

Stepkowski, T., Czaplinska, M., Miedzinska, K. & Moulin, L. (2003). The variable part of the *dnaK* gene as an alternative marker for phylogenetic studies of rhizobia and related alpha *Proteobacteria*. *Syst Appl Microbiol* **26**, 483–494.

Stepkowski, T., Moulin, L., Krzyzanska, A., McInnes, A., Law, I. J. & Howieson, J. (2005). European origin of *Bradyrhizobium* populations infecting lupins and serradella in soils of western Australia and South Africa. *Appl Environ Microbiol* **71**, 7041–7052.

Strijdom, B. W. (1998). South African studies on biological nitrogen-fixing systems and the exploitation of the nodule bacterium-legume symbiosis. *S Afr J Sci* **94**, 11–23.

Tan, Z., Hurek, T., Vinuesa, P., Muller, P., Ladha, J. K. & Reinhold-Hurek, B. (2001). Specific detection of *Bradyrhizobium* and *Rhizobium* strains colonizing rice (*Oryza sativa*) roots by 16S–23S ribosomal DNA intergenic spacer-targeted PCR. *Appl Environ Microbiol* **67**, 3655–3664.

Trinick, M. J. (1973). Symbiosis between *Rhizobium* and the non-legume, *Trema aspera*. *Nature* **244**, 459–460.

Turner, S. L. & Young, J. P. W. (2000). The glutamine synthetase of rhizobia: phylogenetics and evolutionary implications. *Mol Biol Evol* **17**, 309–319.

van Berkum, P. & Eardly, B. D. (2002). The aquatic budding bacterium *Blastobacter denitrificans* is a nitrogen-fixing symbiont of *Aeschynomene indica*. *Appl Environ Microbiol* **68**, 1132–1136.

van Berkum, P. & Fuhrmann, J. J. (2000). Evolutionary relationships among the soybean bradyrhizobia reconstructed from 16S rRNA gene and internally transcribed spacer region sequence divergence. *Int J Syst Evol Microbiol* **50**, 2165–2172.

van Berkum, P., Leibold, J. M. & Eardly, B. D. (2006). Proposal for combining *Bradyrhizobium* spp. (*Aeschynomene indica*) with *Blastobacter denitrificans* and to transfer *Blastobacter denitrificans* (Hirsh and Muller, 1985) to the genus *Bradyrhizobium* as *Bradyrhizobium denitrificans* (comb. nov.). *Syst Appl Microbiol* **29**, 207–215.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. & Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**, 407–438.

Versalovic, J., Schneider, M., de Bruijn, F. J. & Lupski, J. R. (1994). Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol* **5**, 25–40.

Vinuesa, P., León-Barrios, M., Silva, C., Willems, A., Jarabo-Lorenzo, A., Pérez-Galdona, R., Werner, D. & Martínez-Romero, E. (2005a). *Bradyrhizobium canariense* sp. nov., an acid-tolerant endosymbiont that nodulates endemic genistoid legumes (Papilionoideae: Genisteae) from the Canary Islands, along with *Bradyrhizobium japonicum* bv. *genistearum*, *Bradyrhizobium* genospecies alpha and *Bradyrhizobium* genospecies beta. *Int J Syst Evol Microbiol* **55**, 569–575.

Vinuesa, P., Silva, C., Werner, D. & Martinez-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. T. & Martínez-Romero, E. (2000). Phylogeny of root- and stem-nodule bacteria associated with legumes. In *Prokaryotic Nitrogen Fixation: a Model System for Analysis of a Biological Process*, pp. 177–186. Edited by E. W. Triplett. Wymondham, UK: Horizon Scientific.

Wang, F. Q., Wang, E. T., Zhang, Y. F. & Chen, W. X. (2006). Characterization of rhizobia isolated from *Albizia* spp. in comparison with microsymbionts of *Acacia* spp. and *Leucaena leucocephala* grown in China. *Syst Appl Microbiol* **29**, 502–517.

Wang, F. Q., Wang, E. T., Liu, J., Chen, Q., Sui, X. H., Chen, W. F. & Chen, W. X. (2007). *Mesorhizobium albiziae* sp. nov., a novel bacterium that nodulates *Albizia kalkora* in a subtropical region of China. *Int J Syst Evol Microbiol* **57**, 1192–1199.

Weir, B. S., Turner, S. J., Silvester, W. B., Park, D. C. & Young, J. M. (2004). Unexpectedly diverse *Mesorhizobium* strains and *Rhizobium leguminosarum* nodulate native legume genera of New Zealand, while introduced legume weeds are nodulated by *Bradyrhizobium* species. *Appl Environ Microbiol* **70**, 5980–5987.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**, 697–703.

Willems, A. & Collins, M. D. (1993). Phylogenetic analysis of rhizobia and agrobacteria based on 16S rRNA gene sequences. *Int J Syst Bacteriol* **43**, 305–313.

Willems, A., Coopman, R. & Gillis, M. (2001a). Phylogenetic and DNA–DNA hybridization analyses of *Bradyrhizobium* species. *Int J Syst Evol Microbiol* **51**, 111–117.

Willems, A., Doignon-Bourcier, F., Goris, J., Coopman, R., de Lajudie, P., De Vos, P. & Gillis, M. (2001b). DNA–DNA hybridization study of *Bradyrhizobium* strains. *Int J Syst Evol Microbiol* **51**, 1315–1322.

Willems, A., Munive, A., de Lajudie, P. & Gillis, M. (2003). In most *Bradyrhizobium* groups sequence comparison of 16S–23S rDNA internal transcribed spacer regions corroborates DNA–DNA hybridizations. *Syst Appl Microbiol* **26**, 203–210.

Wolde-Meskel, E., Terefework, Z., Lindström, K. & Frostegard, A. (2004). Rhizobia nodulating African *Acacia* spp. and *Sesbania sesban* trees in southern Ethiopian soils are metabolically and genomically diverse. *Soil Biol Biochem* **36**, 2013–2025.

Wong, F. Y. K., Stackebrandt, E., Ladha, J. K., Fleischman, D. E., Date, R. A. & Fuerst, J. A. (1994). Phylogenetic analysis of *Bradyrhizobium japonicum* and photosynthetic stem-nodulating bacteria from *Aeschynomene* species grown in separated geographical regions. *Appl Environ Microbiol* **60**, 940–946.

Xu, L. M., Ge, C., Cui, Z., Li, J. & Fan, H. (1995). *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybeans. *Int J Syst Bacteriol* **45**, 706–711.

Yanagi, M. & Yamasato, K. (1993). Phylogenetic analysis of the family *Rhizobiaceae* and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer. *FEMS Microbiol Lett* **107**, 115–120.

Yao, Z. Y., Kan, F. L., Wang, E. T., Wei, G. H. & Chen, W. X. (2002). Characterization of rhizobia that nodulate legume species of the genus *Lespedeza* and description of *Bradyrhizobium yuanmingense* sp. nov. *Int J Syst Evol Microbiol* **52**, 2219–2230.

Young, J. P. W. & Haukka, K. E. (1996). Diversity and phylogeny of rhizobia. *New Phytol* **133**, 87–94.

Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A. & Sawada, H. (2001). A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie *et al.* 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int J Syst Evol Microbiol* **51**, 89–103.