Phylogenetic multilocus sequence analysis identifies seven novel *Ensifer* genospecies isolated from a less-well-explored biogeographical region in East Africa

Tulu Degefu,1,2 Endalkachew Wolde-meskel1,2 and Åsa Frostegård1

Correspondence
Tulu Degefu
tulu.degefu@umb.no or tukanovic97@yahoo.com

1Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, PO Box 5003, NO-1432 Aas, Norway
2School of Plant and Horticultural Sciences, Hawassa University, PO Box 5, Hawassa, Ethiopia

The diversity of 71 rhizobial strains belonging to the genus *Ensifer*, isolated from root nodules of woody legumes growing in southern Ethiopia, was studied using multilocus sequence analysis (MLSA) and phenotypic approaches. Phylogenetic analyses based on core genes revealed that 43 strains were clustered in seven distinct and consistent positions (genospecies I–VII), while another 25 strains were also distinct but were discrepant in their placement on the different gene trees. The remaining three strains occupied the same phylogenetic branches as defined *Ensifer* species and thus were not distinct. Irrespective of their chromosomal background, the majority of the test strains were highly related with respect to their *nifH* and *nodC* gene sequences, suggesting that these symbions might have acquired these genes recently from a common origin. On the *nifH* phylogenetic tree, the branch containing the test strains and reference species isolated from woody legumes in Africa was clearly separate from those isolated outside the continent, suggesting that these symbions have a long history of separate evolution within *Ensifer* for this gene. A cross-inoculation study showed that our strains were capable of eliciting effective nodulation on the homologous host and on other host species. This suggests a potential to improve nitrogen fixation by selecting for broad-host-range inoculants. Our study confirms the presence of a wide diversity of *Ensifer* in East Africa and, while contributing to the general knowledge of the biodiversity within the genus, also highlights the need to focus on previously less-well-explored biogeographical regions to unravel as-yet-unidentified rhizobial resources.

INTRODUCTION

The large and economically important plant family Leguminosae comprises about 19 400 species in 727 genera and includes both short-lived herbs and woody perennials, which are widely distributed on all habitable continents (Martin *et al.*, 2004). Their successful spread and establishment in a wide variety of biotopes is attributed largely to their ability to form nitrogen-fixing symbiosis with rhizobia. Rhizobia, which are Gram-negative bacteria, team up with legumes and trigger the formation of nodules in which the bacteria fix atmospheric nitrogen. This makes their use in sustainable agriculture an interesting alternative to chemical fertilizers, since it is both cost-effective and eco-friendly. They also have a large potential to rehabilitate degraded land by improving soil fertility. However, these benefits from legume–rhizobium symbioses can only be realized when the legume is nodulated with a compatible rhizobial partner, a reflection of symbiotic specificity governed by complex molecular signalling between the two symbionts (Sharma *et al.*, 1993).

In the recent classification of rhizobia, there are 97 species with validly published names within 13 genera, isolated from a broad spectrum of host plants and geographical areas (http://www.bacterio.cict.fr/). Most of the described species in the rhizobial genera belong to the subclass *Alphaproteobacteria*, but there are also a few representatives within the subclass *Betaproteobacteria*, such as *Burkholderia* species (Sawada *et al.*, 2003), and within the *Gammaproteobacteria*, such as *Pseudomonas* and *Klebsiella* species (Shiraishi *et al.*, 2010). Among the groups within the *Alphaproteobacteria*, the genus *Ensifer* currently consists of 16 species with validly published names, isolated from a wide range of legume species. Most members of this genus were previously

**Abbreviations:** AIC, Akaike information criterion; ANI, average nucleotide identity; MLSA, multilocus sequence analysis.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this study are detailed in Table S1.

Five supplementary tables and a supplementary figure are available with the online version of this paper.
assigned to the genus Sinorhizobium, but the latter genus is now amalgamated with the genus Ensifer (Willems et al., 2003; Young, 2003). Among the 16 species defined in the genus to date, six species, including Ensifer arboris and Ensifer kostiensis (Nick et al., 1999), Ensifer saheli and Ensifer terangae (de Lajudie et al., 1994) and Ensifer garamanticus and Ensifer numidicus (Merabet et al., 2010), were isolated from different herbaceous and woody legume species growing in African soils.

In previous investigations based on phenotypic and modern molecular methods, we reported the existence of a hitherto unknown, wide taxonomic diversity of Ethiopian isolates belonging to the genera Rhizobium, Ensifer, Mesorhizobium and Bradyrhizobium, with several of the isolates being related only distantly to defined species (Degefu et al., 2011; Woldem eskel et al., 2004, 2005). Among the strains represented in these main phylogenetic branches in our collection, 71 were found to belong to the Ensifer branch. All were isolated from root nodules of three Acacia species (Acacia abyssinica, Acacia seyal and Acacia tortilis) and Sesbania sesban, growing at different locations in southern Ethiopia. In the present investigation, we examined the phylogeny and taxonomic positions of those isolates further by multilocus sequence analysis (MLSA) of nine genes, including seven core and two symbiosis-related genes. Phylogenetic analyses of the core genes revealed that the majority of the test strains (~96 %) occupied seven unique, well-supported phylogenetic positions, designated genospecies I–VII, which clustered separately from previously defined Ensifer species.

METHODS

Bacterial strains and growth conditions. The test strains are part of a larger collection of rhizobial strains isolated from leguminous trees and herbs in southern Ethiopia (Wolde meskel et al., 2005). Seventy-one of those strains were classified as belonging to the genus Ensifer (Wolde meskel et al., 2004, 2005). The phylogeny and taxonomic positions of the strains were investigated further in the present study using MLSA of seven core and two symbiosis-related genes. Table S1 (available in IJSEM Online) summarizes the test strains, the plant host from which they were isolated, their geographical origin as well as their 16S rRNA-RFLP genotype grouping, based on earlier studies, the genospecies grouping based on the results from the present study and the GenBank accession numbers of all genes sequenced here. The strains were kept in 15 % (v/v) glycerol at −80 °C and cultured in YMB at 28 °C for 4–5 days.

Molecular techniques. Isolation of total genomic DNA followed Boom et al. (1990) with a slight modification, using diatomaceous earth or Celite analytical filter aid as a DNA-binding solid support (Terefework et al., 2001). The 16S rRNA gene (Weisburg et al., 1991) and internal fragments of the six housekeeping genes recA, rpoB, gyrB, atpD, gap and pnp (Gaunt et al., 2001; Martens et al., 2008) as well as the symbiosis-related genes nodC (Laguette et al., 2001) and nifH (Rivas et al., 2002) were amplified and sequenced using the conditions described in the references indicated.

Phylogenetic data analysis. Multiple nucleotide alignments for the test isolates and reference species in the genus Ensifer were carried out using the Clustal W program (Thompson et al., 1994) from MEGA version 4.0 (Tamura et al., 2007). Phylogenetic trees were reconstructed using maximum-likelihood methods. The best-fit model (the optimal model of nucleotide substitution) was estimated by Modeltest 3.7 (Posada & Crandall, 1998) using the Akaike information criterion (AIC) and the hierarchical likelihood ratio test (Posada & Buckley, 2004). Using the two approaches, the same best-fit model was obtained for the 16S rRNA, gyrB, rpoB, recA, atpD and nodC genes. However, for the gap, pnp and nifH genes, the two options generated different best-fit models for explaining our dataset. In those cases, we reconstructed the tree using both models. However, since we found no difference in topology, the model selected with AIC was used for phylogenetic inference. Sequence information is presented in Table S2. The robustness of the tree topology was calculated from bootstrap analysis with 100 replications.

Cross-inoculation tests. All test isolates were evaluated for their ability to produce nodules on nine different legume host species (Acacia tortilis, Acacia seyal, Acacia nilotica, Acacia abyssinica, Sesbania sesban, Sesbania auclata, Leucanea leucocephala, Vigna unguiculata and Vigna radiata). Seedlings were grown in modified Leonard jars filled with washed and sterilized river sand. Four surface-sterilized and pre-germinated seedlings were planted aseptically into each of three replicate pots. The seedlings were later thinned to one by snipping on the top, as described previously (Chen et al., 2008). A single colony of each isolate was picked from yeast extract-mannitol agar (YMA) plates and multiplied in yeast extract-mannitol broth (YMB). Each seedling was inoculated with 1 ml bacterial culture at exponential growth phase (Somasegaran & Hoben, 1994). Non-inoculated seedlings of the different plant species, either supplied with mineral nitrogen (as 0.1 % KNO₃ in nutrient solution) or grown without nitrogen, served respectively as positive and negative controls. Seedlings were grown under natural sunlight and temperature and were fertilized with one-quarter-strength modified Jensen’s N-free medium (Somasegaran & Hoben, 1994). Nodule assessment and harvesting took place 8 and 5 weeks after inoculation for woody and herbaceous legume species, respectively (Maatallah et al., 2002; Somasegaran & Hoben, 1994). At harvest, plants were scored for nodulation as described in Table S3.

RESULTS

The nucleotide sequences of each core gene from the test strains and Ensifer reference strains were aligned to reconstruct the phylogenetic trees (Figs. 1, 2 and S1). All the test strains were found on the Ensifer branch of the Alphaproteobacteria. The nucleotide sequences of all genes generated in this study were deposited in the NCBI public database (for accession numbers see Table S1). At the time of writing, strains of 16 Ensifer reference species with validly published names were available. Ensifer xinjiangensis was not included in our analyses since this name was reported to be a later synonym of the earlier described Ensifer fredii (Martens et al., 2008). In addition to the phylogenetic trees of individual genes, we also concatenated the sequences of the seven core genes to reconstruct a composite tree (Fig. 3). This was used, together with average nucleotide identity (ANI) values, for further exploration of the phylogenetic positions of our test isolates.

Comparative grouping of the test strains based on different core genes

All test strains and strains of 16 currently defined Ensifer species were clearly differentiated based on MLSA of seven
core genes (Figs 1, 2 and S1). Of the 71 test strains, 68 (~96%) represented novel genomic species, occupying seven well-supported and distinct positions on the trees. The remaining three strains (AC14C, AC50a and AC50e) were not distinct in their placement; thus strain AC14C was identified as Sinorhizobium americanum (this species has not yet been reclassified in Ensifer), while AC50a and AC50e were closely related to Ensifer meliloti (Fig. 3 and Table S4). Forty-three of the 68 test strains were consistent with respect to their placement on the different gene trees and grouped into distinct clusters (Figs 1 and 2). Hereafter, these are referred to as genospecies I–VII for convenience. The remaining 25 test strains shared a higher degree of sequence relatedness (~100%) to genospecies I–VII than to any of the reference species. Although intermingled within the branches comprising the designated genospecies, their placement on the trees varied depending on the gene analysed. These 25 strains were therefore designated inconsistent (Table S4).

Based on analyses of 16S rRNA gene sequences, strains from the seven designated genospecies, but also those with inconsistent grouping, were affiliated to defined species in the genus Ensifer with similarity ranging between 97 and 100% (Table S5). The highest values (99–100%) were obtained between the seven genospecies and one of the eight defined species Ensifer kummerowiae, E. saheli, E. fredii, E. meliloti, E. medicae, E. adhaerens, E. garamanticus and S. americanum. All genospecies except genospecies IV occupied distinct and highly supported branches (Figs 1 and S1). The test isolates belonging to genospecies I and II were closely related to reference strains of Ensifer mexicanus (99.8% similarity) and E. saheli (99.8% similarity), respectively. Genospecies III comprised isolates from S. sesban only, and these isolates were related to strains of E. kummerowiae and E. meliloti with sequence similarities of about 99.4%. The closest relative of the test strains in genospecies IV was E. garamanticus (100% sequence similarity to the type strain). The test isolates belonging to genospecies V and VI were highly related to strains of E. fredii and S. americanum. Strains belonging to genospecies VII were tightly and distinctly clustered (100% bootstrap support) on branches containing the different genomovars of E. adhaerens, with E. adhaerens gv. C as the closest described species (Table S5).

### Analyses of individual housekeeping genes

The phylogenetic trees constructed from sequence alignments of individual housekeeping genes are shown in Fig. 2. In almost all cases, the test strains within each of the designated genospecies formed tight but also distinct clusters of their own, consistently excluding any of the defined Ensifer species. There were, however, a few discrepancies. Genospecies II and V were highly related in their rpoB gene sequences and thus occupied the same branch (Fig. 2c). Similarly, strains in genospecies IV were tightly clustered with E. garamanticus on the recA and rpoB gene trees, in agreement with their grouping based on 16S rRNA gene sequences. The groupings of the test isolates within genospecies I–VII were tight (>87% bootstrap support) for all individual genes; thus, sequence variation of the same gene within each of the genospecies was low, ranging between 0 and 1%. In contrast, comparative analyses of individual
Fig. 2: Maximum-likelihood phylogenetic trees based on recA (a), gyrB (b), pppGpp (c), atpD (d), gapA (e) and pnp (f) gene sequences showing relationships among the different Acacia and Sesbania microsymbionts (i.e., bold and recognized species of the genus Ensifer) per site. Numbers in parentheses indicate the total number of test strains in the designated genospecies represented on each branch. Bootstrap values above 50 (based on 100 replications) are shown at each node. Bars, 0.02 estimated nucleotide substitutions.
Analysis of concatenated gene sequences

Concatenated sequences were produced from the seven housekeeping genes analysed in this study. The strains within each of the defined genospecies were distinctly, but also tightly, placed on the composite tree (Fig. 3), where all the genospecies designated were clearly differentiated from each other and from defined Ensifer species. This is in agreement with the grouping based on individual core genes. The test strains representing genospecies V, I and II were placed on a well-supported separate clade (97 % bootstrap value), where each genospecies occupied a distinct position, excluding any of the Ensifer reference species. E. fredii, S. americanum and E. saheli were the closest neighbours to genospecies V, I and II, respectively, but ANI values were not more than 92 %. Similarly, genospecies VI, III, IV and VII occupied distinct branches, and ANI values showed only distant relatedness to defined species in the genus. E. fredii (93 % ANI), E. kostiensis (88 % ANI), E. terangae (93 % ANI) and E. adhaerens (95 % ANI) were the closest possible but distant phylogenetic neighbours to genospecies VI, III, IV and VII, respectively. Pairwise comparisons between the novel genospecies and reference species based on concatenated gene sequences displayed higher sequence variation, ranging between 4.2 and 12 %. In contrast, there were high sequence similarities (≥ 99 %) among the test strains within each of the respective genospecies.

Phylogenetic analysis of host-determinant genes

Phylogenetic trees were also generated from the nodC and nifH genes (Fig. 4a, b). The grouping of the test strains on
trees generated from these genes was different from that for the core genes. For example, the \textit{nifH}-based phylogenetic tree placed all the Ethiopian strains into one well-supported major cluster (88 % bootstrap value), on which two main conspicuous subclusters were recognized. The majority of the test isolates (83 %), representing genospecies I, II, III, V, VI and VII, were grouped on one of the subclusters (Fig. 4a). The remaining isolates, which comprised genospecies IV, occupied the other subcluster. Similarly, sequence analysis of the \textit{nodC} gene grouped all of the test isolates into one well-supported major clade (100 % bootstrap value), on which three main tight subgroups were distinguished in this clade (Fig. 4b). The majority of the test strains comprising genospecies I, III, V, VI and VII were found in one of the subgroups, while strains in genospecies II and IV occupied the other two tight subclusters on a major clade. In both gene analyses, a major branching separated strains isolated in Africa (\textit{nifH} tree), or Africa and India (\textit{nodC} tree), from those originating outside the continent. Thus, the \textit{nifH} phylogenetic tree contained one highly supported branch (bootstrap value 88 %) where all the test strains clustered, along with reference species isolated from woody legumes in Africa. The latter included \textit{Ensifer} sp. HAMBI 1480 from \textit{A. senegal} in Sudan, \textit{Ensifer} sp. DW0607 from Kenya nodulating \textit{A. seyal} and two strains from Senegal (\textit{E. terangae} ORS 1073 and ORS 1009 from \textit{A. senegal} and \textit{A. laeta}, respectively), which were most closely related to our test isolates, with which they shared $\geq 98$ % sequence similarity.

Based on analysis of the \textit{nodC} gene, \textit{E. arboris} strain LMG 14919 (Iglesias \textit{et al}., 2007) and \textit{Mesorhizobium} strain ORS 3324 (Diouf \textit{et al}., 2010), both from Africa (Senegal), and the uncultured \textit{Rhizobiales} bacterium clone S25 Cl210 (Sarita \textit{et al}., 2005), isolated from Asia (India), were the closest neighbour sequences available in the database, but they were related only distantly to our test isolates, with sequence similarities ranging between 87 and 96 %.

**Infectiveness and cross-nodulation**

The microsymbionts of the different \textit{Acacia} species induced effective root nodules with only five of the woody legumes tested (Table S3). Microsymbionts of \textit{S. sesban} had an even narrower host range, and induced effective nodules only with their homologous host. \textit{S. sesban}, on the other hand, nodulated with all strains tested and was therefore the most promiscuous of the host plants, followed by \textit{A. tortilis}, \textit{A. seyal} and \textit{A. nilotica}, which nodulated with at least 66 test strains. None of the isolates formed nodules on the crop legumes tested.

**DISCUSSION**

**Diverse \textit{Ensifer} species from root nodules of \textit{Acacia} species and \textit{S. sesban} tree legumes**

All 71 test strains that were examined here were defined in an earlier study (Wolde-meskel \textit{et al}., 2005) as belonging to the genus \textit{Ensifer}. In that study, we demonstrated that they formed a separate branch within a cluster representing various AFLP patterns and 16S rRNA PCR-RFLP genotypes (Wolde-meskel \textit{et al}., 2005). In addition, partial 16S rRNA gene sequences were analysed for 13 of the strains, which separated into five genospecies (Wolde-meskel \textit{et al}., 2005). All of the representative strains in those five genospecies were affiliated phylogenetically to one of five defined species (\textit{E. fredii}, \textit{E. meliloti}, \textit{E. medicae}, \textit{E. saheli} and \textit{E. adhaerens}). This is in agreement with the results in the present investigation based on MLSA analysis of seven core and two symbiosis-related genes, where we found that our isolates were affiliated to these and three other species (\textit{E. garamanticus}, \textit{E. kummerowiae} and \textit{S. americanum}) (Table S5).

Phylogenetic analysis of full 16S rRNA gene sequences confirmed that all 71 test strains belonged to the genus \textit{Ensifer}. The majority of the test isolates occupied seven distinct and tight phylogenetic positions (genospecies I–VII; Figs 1 and S1). Of the remaining three isolates, AC50a and AC50e were closely associated with \textit{E. meliloti} and AC14C was grouped with \textit{S. americanum}. Strains AC50a and AC50e were previously found to be closely associated with \textit{E. meliloti} based on AFLP and 16S rRNA PCR-RFLP genotyping (Wolde-meskel \textit{et al}., 2005), in agreement with the present result. Twenty-five test isolates were discrepant in their groupings (when their placement on the 16S RNA gene tree is compared with trees reconstructed from the other housekeeping genes), but still occupied the branches comprising the genospecies designated here (Fig. S1). Comparatively, the MLSA-based groupings of the test isolates in the present study were generally in agreement with the AFLP grouping determined earlier (Table S1). Taken together, the wide diversity of the rhizobial isolates investigated in this study, while strengthening earlier findings on the diversity of rhizobia from Ethiopian soils (Wolde-meskel \textit{et al}., 2005), also indicates that the country represents a biogeographical region with large potential for unearthng as-yet-unidentified rhizobial biodiversity in the tropics.

MLSA has been indicated to be a powerful alternative to the laborious DNA–DNA hybridization technique for species delineation (Martens \textit{et al}., 2008; Rossello-Móra, 2011; Vinuesa \textit{et al}., 2005). The rationale behind using several housekeeping genes in bacterial systematic has already been demonstrated (Martens \textit{et al}., 2008). Hence, we sequenced and analysed six protein-coding genes used previously for systematic studies of the genus \textit{Ensifer} (Martens \textit{et al}., 2007, 2008). The analyses clearly revealed that the test strains within each of the seven designated genospecies were consistently and distinctly placed on each gene tree (Fig. 2). The inconsistent strains were found to be discrepant in their placement among the different gene trees, but still grouped within the clusters comprising the seven genospecies (Fig. S1). The discrepant clustering of these isolates might be attributed to several factors, including differences in evolutionary histories of the genes,
horizontal gene transfer and subsequent recombination events (Charles et al., 2005).

The gyrB gene is one of the molecular markers reported previously to have the best capacity to delineate Ensifer species, and a cut-off point of 97.2% similarity in gyrB gene sequences was reported for genomic species identity in the genus (Martens et al., 2008). The gyrB gene sequence similarity between the Ethiopian Ensifer strains and members of defined species in the genus was not more than 94%, supporting the distinctness of the test isolates. Similarly, comparative analyses of the remaining housekeeping genes showed a low degree of similarity.
genes demonstrated that the test strains within each of the defined genospecies were unique; thus, they occupied seven well-supported phylogenetic positions that clustered distinctly from each other and the reference species (Fig. 2). The exception to this finding was the grouping based on rpoB genes, for which members of genospecies II and V occupied the same, well-supported branch. However, with respect to other housekeeping gene sequences, these genospecies were distantly related to each other. Another exception was the tight grouping of genospecies IV with E. garamanticus in both its recA and rpoB genes (Fig. 2a, c); however, they were only distantly related with respect to their atpD genes (Fig. 2d). Our unsuccessful attempts to amplify the gyrB, gap and pnp genes of E. garamanticus using PCR primer sets employed for the test strains can probably be explained by differences in nucleotide sequences at the primer-annealing sites.

A polyphasic taxonomic approach such as DNA–DNA hybridization and phenotyping methods is reported to be a consensus tool for bacterial systematics (Stackebrandt & Goebel, 1994). However, DNA–DNA hybridization has been criticized for various reasons (Vandamme et al., 1996), and several studies have demonstrated that concatenation of carefully selected housekeeping gene sequences is more powerful and reliable for defining species (Konstantinidis et al., 2006; Martens et al., 2008). Therefore, we concatenated the sequences of seven core genes to construct a composite phylogenetic tree (Fig. 3). This grouped the test isolates into seven well-resolved and distinct genospecies, further consolidating the grouping based on individual core genes. The ANI values between the genospecies and the reference species did not exceed 95.8%, which is far less than the ANI values reported earlier by other authors for Ensifer species (Martens et al., 2008), further supporting the distinctness of the test isolates reported here.

The accessory and core genes revealed discordant phylogeny

Symbiosis-related genes in rhizobia confer differences in host range (Young et al., 2006). Discordant phylogenies have been well documented when trees generated from core and symbiotic genes are compared (Haukka et al., 1998; Lloret et al., 2007; Toledo et al., 2003), a reflection of differences in evolutionary histories and the occurrence of horizontal gene transfer and recombination events (Charles et al., 2005). Therefore, distantly related species can carry accessory genes of identical/similar sequences, for example in Ensifer and Mesorhizobium strains nodulating A. tortilis (Bala et al., 2002), in Azorhizobium and Ensifer strains nodulating Sesbania rostrata (Terefework et al., 2000), in E. mexicanum and ‘Sinorhizobium chiapanecum’ strains nodulating Acacia (Rincón-Rosales et al., 2009) and in different Mesorhizobium strains nodulating the shrub legume Anagyris latifolia (Donate-Correa et al., 2007). Similarly, accessory genes can be different among rhizobial strains of the same species. We found highly related nodC (≥97% similarity) and nifH (≥98% similarity) gene sequences and similar patterns of grouping among the majority of the test strains, irrespective of their chromosomal background, suggesting that these symbionts might have acquired these genes recently from a common origin. The close similarity in host range displayed by these test isolates (especially those from the different Acacia species) could be a reflection of high relatedness in the sequences of their symbiotic genes and may be attributed to the dry ecoclimatic environment to which the host species are generally adapted. It is known that different Acacia species are adapted to and grow in many of the world’s arid and semi-arid areas, including the African continent (Zahrn, 1999). Similar to our results, Haukka et al. (1998) also reported that phylogenetic trees generated from symbiotic genes show some correlation with host plant range. The exception to this observation was that, unlike the test strains from the three Acacia species, those from S. sesban were specific in nodulation and formed effective symbiosis only with their homologous host. Here, it is to be noted that the nodC gene sequences of isolates from S. sesban comprised a distinct subgroup within the cluster containing test strains from Acacia species, which might explain their specificity.

On the nifH phylogenetic tree, the branch containing the test strains, along with reference species isolated from woody legumes in Africa, was clearly separated from those outside the continent. This suggests that these symbionts have a long history of separate evolution within Ensifer for this particular gene under specific ecoclimatic conditions to which these legume hosts are adapted. Similar studies by Haukka et al. (1998) on Ensifer from Prosopis species of two continents (Africa and Latin America) demonstrated that the nodA and nifH gene sequences of Ensifer strains from Latin American Prosopis were highly related within the continent (for example, Brazil and Mexico), but were quite different from those of the African symbionts of Prosopis chilensis. Overall, this convergence between symbiotic genes and nodulation phenotypes of the test strains that harbour distinct chromosomal profiles further supports the discordant phylogeny between housekeeping and symbiotic genes.

Host range of Ensifer species from Ethiopia and its implication

The legume–rhizobium symbiosis contributes a substantial amount of the nitrogen budget (70 million tons of nitrogen per year) in terrestrial ecosystems (Zahrn, 1999). This is affected by the ability of the hosts to enter into symbioses with different rhizobia. Some legumes have specific rhizobial requirements (e.g. Sesbania spp.), while others are capable of interacting with diverse rhizobia. In our present cross-inoculation undertakings on nine legume species (Table S3), the test isolates from different Acacia species had broader host ranges than the microsymbionts from S. sesban. Depending on the legume species, the strains exhibited variations in their symbiotic performance (effectiveness), but all of them had accumulated a high nitrogen content that was comparable to the positive control (plant
supplied with chemical fertilizer; data not shown). The ability of the test isolates from nodules of the different Acacia species to nodulate a range of legumes has both silvicultural and ecological importance, and provides useful information for applied research aimed at enhancing nitrogen fixation in agriculture/afroforestry.

**Conclusions**

MLSA proved to be a useful tool for resolving the taxonomic and phylogenetic positions of our test strains, and demonstrated the existence of diverse genospecies only distantly affiliated to known Ensifer species. Several of the isolates could potentially be assigned to novel species. While contributing to our knowledge of the biodiversity within the genus Ensifer, our results also emphasize the need to focus on previously less-well-explored biogeographical regions to investigate hitherto-unidentified rhizobial resources. A practical outcome can be the development of improved inocula by using native strains that are both effective nitrogen fixers and competitive enough for nodulation. Several of the isolates in our collection were capable of eliciting effective nodules on a number of legumes and are thus interesting candidates for development of inoculants with the potential to increase the soil nitrogen status and crop yields via rhizobium–legume symbiosis in Ethiopia and other areas with similar biogeoclimatic conditions.

**ACKNOWLEDGEMENTS**

The authors gratefully acknowledge financial support from the Norwegian Universities Committee for Development Research and Education (NUFU) for PhD research undertaken at the Norwegian University of Life Sciences (UMB), Department of Chemistry, Biotechnology and Food Sciences, Environmental Microbiology unit. T.D. is thankful to the Norwegian State Educational Loan Fund (Laånekassen) for a PhD stipend.

**REFERENCES**


Rincón-Rosales, R., Lloret, L., Ponce, E. & Martínez-Romero, E. (2009). Rhizobia with different symbiotic efficiencies nodulate *Acacia* angustissima in Mexico, including *Sinorhizobium* chiapense sp. nov. which has common symbiotic genes with *Sinorhizobium* mexicanum. *FEBS Microbiol Ecol* 67, 103–117.


