A peer-reviewed version of this preprint was published in PeerJ on 16 May 2018.

View the peer-reviewed version (peerj.com/articles/4776), which is the preferred citable publication unless you specifically need to cite this preprint.

https://doi.org/10.7717/peerj.4776
Novel, non-symbiotic isolates of *Neorhizobium* from a dryland agricultural soil

Amalia Soenens, Juan Imperial

1 Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid-Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Pozuelo de Alarcón, Madrid, Spain
2 Instituto de Ciencias Agrarias, Consejo Superior de Investigaciones Científicas, Madrid, Spain

Corresponding Author: Juan Imperial
Email address: juan.imperial@upm.es

Semi-selective enrichment, followed by PCR screening, resulted in the successful direct isolation of fast-growing Rhizobia from a dryland agricultural soil. Over 50% of these isolates belong to the genus *Neorhizobium*, as concluded from partial *rpoB* and near-complete 16S rDNA sequence analysis. Further genotypic and genomic analysis of five representative isolates confirmed that they form a coherent group within *Neorhizobium*, closer to *N. galegae* than to the remaining *Neorhizobium* species, but clearly differentiated from the former, and constituting at least one new genomospecies within *Neorhizobium*. All the isolates lacked *nod* and *nif* symbiotic genes but contained a *repABC* replication / maintenance region, characteristic of rhizobial plasmids, within large contigs from their draft genome sequences. These *repABC* sequences were related, but not identical, to *repABC* sequences found in symbiotic plasmids from *N. galegae*, suggesting that the non-symbiotic isolates have the potential to harbor symbiotic plasmids. This is the first report of non-symbiotic members of *Neorhizobium* from soil.
Novel, non-symbiotic isolates of *Neorhizobium* from a dryland agricultural soil

Amalia Soenens¹, Juan Imperial¹,²

¹ Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid – Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 28223 Pozuelo de Alarcón (Madrid), Spain

² Instituto de Ciencias Agrarias, Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain

Corresponding Author:

Juan Imperial

CBGP, Campus de Montegancedo de la UPM, Ctra. M-40, km 38, 28223 Pozuelo de Alarcón (Madrid), Spain

Email address: juan.imperial@csic.es
ABSTRACT

Semi-selective enrichment, followed by PCR screening, resulted in the successful direct isolation of fast-growing Rhizobia from a dryland agricultural soil. Over 50% of these isolates belong to the genus *Neorhizobium*, as concluded from partial *rpoB* and near-complete 16S rDNA sequence analysis. Further genotypic and genomic analysis of five representative isolates confirmed that they form a coherent group within *Neorhizobium*, closer to *N. galegae* than to the remaining *Neorhizobium* species, but clearly differentiated from the former, and constituting at least one new genomospecies within *Neorhizobium*. All the isolates lacked *nod* and *nif* symbiotic genes but contained a *repABC* replication / maintenance region, characteristic of rhizobial plasmids, within large contigs from their draft genome sequences. These *repABC* sequences were related, but not identical, to *repABC* sequences found in symbiotic plasmids from *N. galegae*, suggesting that the non-symbiotic isolates have the potential to harbor symbiotic plasmids. This is the first report of non-symbiotic members of *Neorhizobium* from soil.
INTRODUCTION

A group of α-proteobacteria from the Rhizobiales order, especially within the Rhizobiaceae and Bradyrhizobiaceae families, are collectively known as Rhizobia because they have the ability to establish root-nodule symbioses with legumes. Within these nodules, Rhizobia fix atmospheric nitrogen, and this fixed nitrogen is assimilated by the plant. This makes most legumes uniquely independent from the need of any exogenous nitrogen fertilizer, an important ecological and agricultural trait that is at the basis of any effort aimed at sustainable agriculture.

Since their first isolation from legume root nodules (Beijerinck, 1888), it has been known that Rhizobia are present in the soil, wherefrom they can colonize emerging roots of their legume host. However, they have been rarely isolated directly from the soil, given that the use of trap legume plants provides a facile method for Rhizobia isolation from root nodules. This is very convenient, especially because rhizobial soil populations have been often estimated as ranging between $10^2$-$10^5$ cells per gram of soil, depending on soil type and host plant (Singleton & Tavares, 1986). However, the use of legume trap plants allows a very limited glimpse at rhizobial populations in soil, for two reasons. First, legume-rhizobial symbioses are usually very specific, and a specific legume can only be nodulated by a specific type of Rhizobium, a phenomenon whose molecular bases have been intensively studied in recent years. Second, the genetic determinants for a successful symbiosis with a host legume are but a small fraction of the genetic complement of a Rhizobium. These determinants are often present in plasmids or mobile genomic islands that can be transferred, exchanged or lost (López-Guerrero et al., 2012; Andrews & Andrews, 2017). Therefore, it is possible that, for any given Rhizobium, a large non-symbiotic subpopulation co-exists with the symbiotic subpopulation and represents an ill-studied
reservoir of genetic diversity. That this is indeed the case has been demonstrated in very few instances. In a pioneering study, (Sullivan et al., 2002) were able to show that *Mesorhizobium loti*, the microsymbiont of *Lotus* spp. harbors its symbiotic determinants within a symbiotic genetic island integrated in its chromosome, and that this island can excise and be transferred to other cells. Furthermore, after inoculating a soil with a strain containing a marked symbiotic island, they were able to recover it in different genomic backgrounds, thus proving that the island undergoes cell-to-cell transfer in the soil, and that non-symbiotic *M. loti* strains are present in the soil that can receive the marked symbiotic island and thus acquire the ability to nodulate *Lotus*. Despite the above, the direct isolation of Rhizobia from soil has received little attention, and even in those few cases, the interest was placed on symbiotically-competent Rhizobia. For example, Louvrier et al. (Louvrier, Laguerre & Amarger, 1995), in ground-breaking work, devised a semi-selective culture medium to enrich *Rhizobium leguminosarum* from soil. They were interested in isolating symbiotic strains that had not been selected by the plant host, in order to test the hypothesis that the different plant host species this bacterium colonizes, select specific genotypes among those present in the soil (Louvrier, Laguerre & Amarger, 1996). Likewise, Tong and Sadowsky optimized a *Bradyrhizobium japonicum* and *B. elkanii* – enriching medium (Tong & Sadowsky, 1994) that was later used to isolate symbiotic and non-symbiotic *Bradyrhizobia* from soil (Pongsilp et al., 2002). In our lab, we have recently built upon Laguerre and Amarger’s observations by carrying out a population genomics study of genotype selection by the legume host in the same agricultural soil in Burgundy (Jorrín & Imperial, 2015). In this study, many non-symbiotic Rhizobia were isolated (Jorrín, 2016; Jorrín & Imperial, unpublished data). The above results are in line with the recent identification of non-symbiotic Rhizobia as abundant components of plant microbiomes (Lundberg et al., 2012; Shakya et al., 2013;
Chaparro, Badri & Vivanco, 2014; Ofek-Lalzar et al., 2014; De Souza et al., 2016) and the isolation of non-symbiotic Rhizobia from plant material (Segovia et al., 1991; Sullivan et al., 1996; Van Insberghe et al., 2015; Jones et al., 2016).

In this work, we set out to directly isolate Rhizobia from a dryland agricultural soil in Southern Spain where no record of legume cultivation is available. A large fraction of those isolates was found to constitute a hitherto unsuspected, non-symbiotic clade within the recently described genus *Neorhizobium* (Mousavi et al., 2014), whose known members were, up to this work, legume symbionts: *N. galegae*, isolated as symbionts of the cold-climate legume *Galega* sp. (Lindström, 1989) as well as of many other legumes; *N. alkalisoli*, from nodules of *Caragana intermedia* in Northern China (Li Lu et al., 2009); and *N. huautlense* from nodules of *Sesbania herbacea* (Wang et al., 1998).

**MATERIALS & METHODS**

**Bacterial strains and growth conditions**

Bacterial strains used in this work are listed in Table 1. Rhizobial strains were grown in Yeast Mannitol Broth (YMB; Vincent, 1970) at 28 °C, either in liquid culture or on solid media supplemented with 1.5% agar. For long-term maintenance, strains were grown at 28°C in YMB and preserved in 20% glycerol at -80°C.

**Soil**

Tomejil soil is a dryland agricultural soil from the Las Torres-Tomejil Experimental Agricultural Station of the Instituto de Investigación y Formación Agraria y Pesquera de Andalucía (IFAPA) in Seville, Spain, where no record of previous legume cultivation exists. A
103 10 m x 10 m plot (37°24'33.10" N, 5°34'51.91" W; 77 m above sea level) was selected and fenced for our research. For soil collection, portions of soil down to 30-50 cm depth were collected with a shovel from several points within the plot. Soil samples were maintained at 4 °C and -20 °C. Physicochemical characterization was done externally in Laboratorio de Edafología y Técnicas Analíticas Instrumentales, EUIT Agrícola. UPM, Madrid. Physicochemical properties of the soil are described in Table S1.

109 Direct isolation of Rhizobia from soil

For *Neorhizobium* isolation we have used a modification of the Louvrier et al. protocol (Louvrier, Laguerre & Amarger, 1995) for enrichment of fast-growing Rhizobia (Jorrín, 2016; Jorrín & Imperial, unpublished data; Fig. S1), as follows. In an erlenmeyer flask a 10^-1 soil dilution was made using a salt buffer (0.1 g NaCl, 0.5 g K2HPO4, 0.2 g MgSO4·7 H2O, pH 6.8). This was shaken overnight at 28 ºC at 200 rpm. Serial dilutions up to 10^-6 depending on the soil, were made in the same salt buffer and 100 μl of each dilution was plated in the semi-selective media MNBP (Louvrier, Laguerre & Amarger, 1995) supplemented with 25 ppm Congo red. Plates were left for four days at 28 ºC. After the incubation period, white-pink colonies were picked with a sterile toothpick into Yeast Mannitol Broth (YMB), Luria-broth (LB; Bertani, 1951), and MNBP agar plates and incubated at 28 ºC for two days. Negative LB colonies were then streaked out a second time into YMB and LB agar plates. Confirmed negative LB colonies were grown on YMB agar plates until pure cultures were obtained. Genomic DNA was extracted from pure cultures using the alkaline lysis method (Baele et al., 2000) and tested for *fnrN* by PCR amplification. Positive *fnrN* isolates were characterized phylogenetically by PCR.
amplification and sequencing of the housekeeping genes 16S rDNA and rpoB. Presence of symbiotic genes was determined by amplification and sequencing (if present) of nodC and nifH.

**Genotypic characterization**

Bacterial DNA previously isolated by alkaline lysis (Baele et al., 2000) was used as substrate for PCR amplifications. Full-length 16S rDNA (Weisburg et al., 1991), partial nodC (Sarita et al., 2005) and partial nifH (Ando et al., 2005) sequences were amplified with primers described in the references. Partial rpoB was amplified with: F_rpoB (5′-GARTTCGACGCCAAGGAYAT-3′) and R_rpoB (5′-GAAGAACACGCGAGTTGAACAT-3′). Amplifications were carried out in 25 μl solution containing DNA (5-10ng), 2.5 μl 10x PCR buffer containing magnesium chloride (Roche Applied Science, Penzberg, Germany), 10mM of each dNTP, 10 μM of each primer, 1μl DMSO and 1 U of Taq DNA polymerase (Roche Applied Science). Unincorporated primers and dNTPs were removed from PCR products with the NucleoSpin®Extract II Kit (Macherey-Nagel, Düren, Germany) or, when needed, by gel electrophoresis followed by band purification with the same kit. Sanger sequencing was carried out externally (STAB Vida, Lisbon, Portugal).

**Genomic characterization**

For bacterial genome sequencing of Tomejil isolates: T20_22, T7_12, T25_27, T25_13, T6_25, and type Neorhizobium strains not available in public databases, bacteria were grown in Tryptone Yeast (TY; Beringer, 1974). The bacterial pellet obtained after centrifugation was used to extract total DNA using the CTAB method (Feil, Feil & Copeland, 2012). DNA quantity and quality was assessed by spectrophotometry (Nanodrop, NanoDrop Technologies, Wilmington, Delaware).
DE, USA) and fluorescence (Qubit, Invitrogen by Life Technologies, Singapore), and integrity
and purity were checked by electrophoresis in a 0.8% agarose gel. Draft genomic sequences of
bacterial strains were obtained externally (MicrobesNG, Birmingham, UK) with Illumina
technology (MiSeq v3, PE 2x300 bp), and reads were assembly using SPAdes (Bankevich et al.,
2012) and annotated with Prokka (Seemann, 2014).

Bioinformatics

For phylogenetic and sequence analyses, nucleotide sequences obtained from PCR products
were corrected and assembled if necessary with SerialCloner2-6
(http://serialbasics.free.fr/Serial_Cloner.html) and 4peaks (http://nucleobytes.com/4peaks/).
Sequences were aligned with the ClustalW algorithm (Chenna et al., 2003) in MEGA 6.0
(Tamura et al., 2013). Randomized Axelerated Maximum Likelihood (RAxML) (Stamatakis,
2014) or MEGA 6.0 (Tamura et al., 2013) were used for the construction of phylogenetic trees.
Phylogenetic trees were visualized with FigTree v.1.4.3
(http://tree.bio.ed.ac.uk/software/figtree/) and edited with Adobe Illustrator CS5 (Adobe
Systems, San José, CA, USA). In order to study genomic identity among strains, Average
Nucleotide Identity (either based on MUMmer alignments, ANIm, or based on BLAST
alignments, ANIb) was calculated with the JSpeciesWS online server (Richter et al., 2016). A
distance dendrogram was generated by hierarchical cluster analysis of 100 - % ANI matrices
(Chan et al., 2012) with StataSE v.14.0 (StataCorp, College Station, TX, USA) after
computation of Euclidean distances with the Average Linked method.

DNA sequences
GenBank accession and Bioproject numbers for sequences and genomes, respectively, obtained in this work are listed in Supplementary Table S2, together with accession numbers for reference sequences used in the analyses.

RESULTS

Direct isolation of Neorhizobium sp. from soil

A total of 28 independent isolations were carried out from Tomejil soil. After four days, pink-white colonies appearing on MNBP plates were chosen as putative fast-growing rhizobia. From those, we chose to discard all that were able to grow on LB agar. In our hands, this includes most Ensifer spp. and many non-rhizobial isolates. However, most Rhizobium spp. were unable to grow on the 5 g.l\(^{-1}\) NaCl present in LB agar, and thus this step resulted in a further enrichment of Rhizobium spp. DNA from 342 of these putative Rhizobium spp. was amplified for the fnrN gene by PCR. This gene is important for microaerobic metabolism (Gutiérrez et al., 1997) and has been found in all sequenced Rhizobium spp. and Agrobacterium spp., but not in other Rhizobia. Ninety-nine of the DNAs resulted in good amplification of a major band of a size similar to what was expected. In order to reduce the number of isolates for further characterization, a fragment of the rpoB gene was amplified and sequenced from all of them. The rpoB marker has been shown to be very effective in discriminating phylogenetically close bacteria (Mollet, Drancourt & Raoult, 1997; Khamis et al., 2003; Case et al., 2006), including Rhizobia (Jorrín, 2016; Jorrín & Imperial, unpublished data). Forty-six of the 99 partial rpoB sequences corresponded, as expected, to Rhizobium spp. (data not shown), while the remaining 53, surprisingly, clustered with sequences from genus Neorhizobium (Fig. 1B). These sequences were compared and
classified into groups (twenty in total) if they differed in at least one nucleotide. In order to facilitate further studies, a representative strain of each of these \textit{rpoB} sequence types was chosen. Near complete 16S \textit{rDNA} sequences were obtained from these representative strains and compared against those in databanks, confirming that they formed a diverse clade that was closely related to known members of the genus \textit{Neorhizobium}, but distant from the \textit{Agrobacterium-Rhizobium} group (Fig. 1A).

**Genotypic and genomic characterization of \textit{Neorhizobium} sp. isolates**

Since all previously characterized \textit{Neorhizobium} isolates had been obtained from nitrogen-fixing legume root nodules, we tried to amplify \textit{nodC} and \textit{nifH}, markers related to symbiosis and nitrogen fixation, respectively, from our Tomejil isolates. Negative results were obtained in all cases, suggesting that these isolates were non-symbiotic. Since direct tests of symbiotic ability with legume plants were hampered by the fact that the known hosts of \textit{Neorhizobium} spp. are very diverse and are not found in the Tomejil area, we reasoned that obtaining and characterizing the genome sequence of some of the Tomejil isolates, even at draft level, would be worthwhile.

We chose five representatives of the most abundant / diverse \textit{rpoB} groups, so as to try to obtain a picture of the genomic diversity within this \textit{Neorhizobium} sp. clade. Genomic DNA was sequenced and assembled to draft level and this assembly was used for subsequent analysis. All five isolates had similar genome size and G+C (%) composition (Table 2).

We first searched for \textit{fnrN} genes. This was important because no \textit{fnrN} gene has been described in \textit{Neorhizobium} spp. isolates, and it was absent from the \textit{Neorhizobium} spp. genomic sequences
available in databanks (data not shown). Using the well-characterized *R. leguminosarum fnrN*

gene for BLAST comparisons, we were unable to find any relevant hit against genomic
sequences from Tomejil isolates, despite the fact that the isolates were chosen because they
showed clear amplification of a unique band with fnrN primers (Fig. S2A). Amplified PCR
bands from some of the isolates were sequenced and compared with databanks. Part of the
amplified region showed similarity with genes encoding a poly (3-hydroxybutyrate)
depolymerase from Rhizobia, including *Neorhizobium*. This gene, together with an upstream
ORF, was present in genome assemblies of the Tomejil strains and contained sequences partially
complementary to those of fnrN primers that may explain the successful amplifications observed
(Fig. S2B).

We then tried to find *nod* and *nif* genes by running BLAST searches of known *Neorhizobium nod*
and *nif* genes against draft genome sequences, with negative results in all cases (data not shown).
Since *nod* and *nif* genes are harbored on large megaplasmids in *N. galegae* (Österman et al.,
2014), we searched for similar plasmids in the genomes of Tomejil strains. All five genomes
contained a set of *repABC* sequences characteristic of rhizobial plasmids (Cevallos et al., 2008;
Pinto et al., 2012), highly similar (85-86%) to the *repABC* cluster from the 1.8 Mb megaplasmid
from *N. galegae* bv. orientalis (Table 3) and that are probably responsible for the replication a
large plasmid in each of these strains. This is supported by the fact that *repABC* homologues are
present in large contigs (83-523 kb) in the Tomejil strains, despite the draft level quality of their
genome sequences (27-69 contigs, Table 2). Multiple sequence alignment of these *repABC*
sequences showed that plasmids from the Tomejil strains were highly related, and separated from
the symbiotic plasmids from *N. galegae* bv. orientalis and *N. galegae* bv. officinalis (Fig. 2). The
completely sequenced strain *N. galegae* bv. officinalis HAMBI 1140 harbors an additional megaplasmid (175 kb) that appeared to be highly unrelated to both *N. galegae* symbiotic plasmids and plasmids from Tomejil strains (Table 3, Fig. 2).

In order to ascertain the phylogenetic relationships among the sequenced Tomejil strains, and between this clade and the type strains of the three *Neorhizobium* species: *N. galegae*, *N. alkalisoli*, *N. huautlense*, we pulled out the complete sequences of the genes *atpD*, *glnII*, *recA*, *rpoB* and *thrC* from the genome assemblies. These genes had been used successfully to define the *Neorhizobium* genus and its species (Mousavi et al., 2014). Figure 3 shows a maximum-likelihood phylogenetic tree derived from a multiple alignment of the concatenated genes, rooted by using the *Agrobacterium tumefaciens* Ach5 sequences. Very similar results were obtained when the phylogenetic tree was derived by the neighbor-joining method (data not shown). There was very strong support for two clades within the *Neorhizobium* genus; one formed by *N. alkalisoli* and *N. huautlense*, and the other by *N. galegae* and the Tomejil strains. Four of the Tomejil strains (T7_12, T20_22, T25_13, T25_27) formed a well-supported group clearly separated from the fifth strain, T6_25, and all five, in turn, clearly separated from *N. galegae*.

Genome-wide comparisons were carried out by calculating pairwise Average Nucleotide Identities (Richter & Rosselló-Mora, 2009; Richter et al., 2016) with the above genomes. Both ANIm and ANIb scores were calculated, and the ANIb matrix is shown in Table S3. A distance matrix was generated from the ANIb matrix using 100-ANIb (Chan et al., 2012) and Euclidean distances calculated. These were represented in the dendrogram shown in Fig. 4. The dendrogram faithfully reproduced the topology of the multilocus phylogenetic tree (Fig. 3).

Using a 95% ANI value as a widely accepted delimiter for genomic species (Richter & Rosselló-
Mora, 2009), our results are consistent with Tomejil strains representing at least a clearly
differentiated genospecies within *Neorhizobium*, and probably two, represented by strains T6_25
on one hand, and T7_12, T20_22, T25_13, and T25_27, on the other.

**DISCUSSION**

Direct isolation of Rhizobia from soil without resorting to trapping them inside their legume host
is complicated by the fact that Rhizobial populations in soil can be small (Singleton & Tavares,
1986) and because of the paucity of selective characters that can be used. This is especially true
for non-symbiotic variants lacking any symbiotic marker. As a result, very few studies have
aimed at isolating non-symbiotic Rhizobia from soil. Groundbreaking work by Sullivan and
Ronson with *Mesorhizobium loti* (Sullivan et al., 1995), did not only reveal the existence of an
abundant population of non-symbiotic variants in the soil, but also showed that the genetic
determinants for symbiosis could be readily transferred to these variants within the soil. Some of
these non-symbiotic variants were further characterized and suggested to be representatives of
species hitherto undescribed. Using a semi-selective culture medium, Tong and Sadowsky (Tong
& Sadowsky, 1994) were able to enrich in soybean-specific Rhizobia of the species
*Bradyrhizobium japonicum* and *B. elkanii*. Using this medium, the same group later described
that about half of the Bradyrhizobia isolated directly from Thai soils were non-symbiotic,
although they did not characterize them further (Pongsilp et al., 2002). Louvrier et al (Louvrier,
Laguerre & Amarger, 1995) were also able to enrich members of the genus *Rhizobium* using a
specific medium. Although they focused on symbiotic isolates (Louvrier, Laguerre & Amarger,
1996) they also isolated a fraction of non-symbiotic Rhizobia (Laguerre, pers. comm, 2011).
Our group had previously used the *Rhizobium* semi-selective medium, together with the *fnrN* gene marker, to isolate Rhizobia from soil (Jorrín 2016; Jorrín & Imperial, unpublished data) in order to study, at the genomic level, the legume host selection of specific *R. leguminosarum* genotypes (Jorrín & Imperial, 2015). This phenomenon had been described by the Amarger-Laguerre (Laguerre et al., 2003) and the Young (Young & Wexler, 1988) groups. During this study, we had been able to isolate a number of non-symbiotic members of the genus *Rhizobium* (Jorrín, 2016; Jorrín & Imperial, unpublished data). Therefore, it was surprising that the same enrichment methodology resulted in the isolation of a large (over 50%) proportion of non-symbiotic members of the genus *Neorhizobium* from soil samples from the IFAPA Tomejil Experimental Station in Carmona (Seville, Spain), first because non-symbiotic *Neorhizobium* had not been previously isolated, and second because our genetic screening with *fnrN* should have left them behind. The serendipitous reason why these particular *Neorhizobium* were chosen as putative *Rhizobium* has been presented above. Clearly, they grow well on MNBP semi-selective medium, with cultural characteristics similar to those of *Rhizobium*. However, it is possible that, among the 243 colonies that tested negative for *fnrN* amplification, other *Neorhizobium* that do not have a high enough conservation of the poly (3-hydroxybutyrate) depolymerase region exist in the Tomejil soil.

At any rate, *Neorhizobium* appears to be at least as abundant in the particular sampled soil as *Rhizobium* and, given the high rates of horizontal gene transfer that can take place in soil (Sullivan et al., 2002), they may become relevant in the establishment of symbioses with native legumes after receiving the appropriate symbiotic genes. Acquisition of these genes might be facilitated by the fact that *N. galegae* has been shown to harbor symbiotic genes on a megaplasmid (Radeva et al., 2001; Österman et al., 2014). As shown above, all the sequenced
Tomejil strains contain a set of *repABC* genes characteristic of plasmids from the Rhizobiales (Cevallos et al., 2008; Pinto, Pappas & Winans, 2012) in large contigs, suggesting that these strains also harbor a megaplasmid and are, thus, probably able to receive similar, large plasmids containing symbiotic genes.

A final consideration is whether the existence of populations of non-symbiotic *Neorhizobium* in soils reflects a normal situation in agricultural soils. This would require a large screening that is beyond the scope of this work. However, our preliminary studies suggest that in at least two other soils from Southern Spain that we have tested, members of *Neorhizobium* are not present in detectable numbers, and that the predominant isolates resulting from our semi-selective screening are *Rhizobium* spp. (Soenens & Imperial, unpublished results). This would then shift the question to why the Tomejil soil harbors a *Neorhizobium* population, a question that would require a better understanding of the ecology of this group in soil.

**CONCLUSIONS**

In conclusion, our work has allowed, for the first time, the isolation and identification of non-symbiotic members of the genus *Neorhizobium* from soil. Genotypic and genomic characterization of these isolates suggests that they are representatives of one, or perhaps two, new genospecies within *Neorhizobium*. It also suggests that soils harbor a large diversity of Rhizobial diversity in the form of non-symbiotic variants that have traditionally escaped characterization and that may play an important role in the biology of these organisms.

**ACKNOWLEDGEMENTS**
We thank Francisco Temprano, Dulce Rodríguez-Navarro, and Francisco Perea, from IFAPA “Las Torres-Tomejil,” for facilitating our soil sampling in the Tomejil experimental farm.

REFERENCES


Table 1 (on next page)

Bacterial strains used in this study.
1. Bacterial strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristics</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. alkalisoli</td>
<td>Type strain, Nod⁺ Fix⁺, LB⁻</td>
<td>(Mousavi et al. 2014)</td>
</tr>
<tr>
<td>DSM 21826ᵀ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. galegae</td>
<td>Type strain, Nod⁺ Fix⁺, LB⁻</td>
<td>(Mousavi et al. 2014)</td>
</tr>
<tr>
<td>HAMBI 540ᵀ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. huautlense</td>
<td>Type strain, Nod⁺ Fix⁺, LB⁻</td>
<td>(Mousavi et al. 2014)</td>
</tr>
<tr>
<td>DSM 21817ᵀ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neorhizobium sp.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4_1</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T4_8</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T5_2</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T5_26</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T5_27</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T6_1</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T6_21</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T6_23</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T6_25</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T7_1</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T7_7</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T7_8</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T7_9</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T7_11</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁻, LB⁻</td>
<td>This study</td>
</tr>
</tbody>
</table>
T7_12  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_20 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T8_5  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T9_24 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T11_12 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T13_2  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T16_1  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T16_2  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T16_4  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T16_9  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T16_12 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_4  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_6  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_14 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_15 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T17_26 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T18_15 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T20_10 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T20_15 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T20_22 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T20_25 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T21_1  Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
T21_15 Soil isolate, Tomejil, Nod³ Fix⁻, LB⁻  This study
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>T21_19</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T22_7</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T22_11</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T22_47</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T23_12</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T23_26</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T24_19</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T24_25</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_4</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_5</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_7</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_13</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_19</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_20</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_27</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_28</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T25_30</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
<tr>
<td>T28_6</td>
<td>Soil isolate, Tomejil, Nod⁻ Fix⁺, LB⁻</td>
<td>This study</td>
</tr>
</tbody>
</table>

1Nod: nodulation phenotype; Fix: nitrogen fixation phenotype; LB: growth on LB medium.
Table 2 (on next page)

Genomic features of *Neorhizobium* genomes sequenced in this work.
Genomic features of *Neorhizobium* genomes sequenced in this work.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Number of contigs</th>
<th>Largest contig (bp)</th>
<th>Total genome length (bp)</th>
<th>G+C (%)</th>
<th>N50</th>
</tr>
</thead>
<tbody>
<tr>
<td>T20_22</td>
<td>37</td>
<td>1,052,711</td>
<td>6,608,977</td>
<td>61.47</td>
<td>508,270</td>
</tr>
<tr>
<td>T7_12</td>
<td>52</td>
<td>1,197,185</td>
<td>6,627,103</td>
<td>61.44</td>
<td>347,929</td>
</tr>
<tr>
<td>T25_27</td>
<td>27</td>
<td>1,446,028</td>
<td>6,462,352</td>
<td>61.49</td>
<td>734,252</td>
</tr>
<tr>
<td>T25_13</td>
<td>42</td>
<td>721,675</td>
<td>6,322,993</td>
<td>61.56</td>
<td>419,328</td>
</tr>
<tr>
<td>T6_25</td>
<td>69</td>
<td>488,027</td>
<td>6,750,064</td>
<td>61.35</td>
<td>186,129</td>
</tr>
</tbody>
</table>
Table 3 (on next page)

Presence of repABC regions in the genomes of Tomejil strains.

DNA regions similar to the 3,628 bp region containing repABC genes from the symbiotic megaplasmid from N. galegae bv. orientalis HAMBI 540 were located in genome sequences by BLAST, extracted and compared by multiple alignment (ClustalW).
Presence of *repABC* regions in the genomes of Tomejil strains.

DNA regions similar to the 3,628 bp region containing *repABC* genes from the symbiotic megaplasmid from *N. galegae* bv. orientalis HAMBI 540 were located in genome sequences by BLAST, extracted and compared by multiple alignment (ClustalW).

<table>
<thead>
<tr>
<th>Genome</th>
<th>Number of <em>repABC</em> regions</th>
<th>(%) identity to HAMBI 540</th>
<th>size of contig (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N. galegae</em> bv. orientalis HAMBI 540</td>
<td>1</td>
<td>100</td>
<td>1,807,065</td>
</tr>
<tr>
<td><em>N. galegae</em> bv. officinalis HAMBI 1141</td>
<td>2</td>
<td>95</td>
<td>1,638,739</td>
</tr>
<tr>
<td>T20_22</td>
<td>1</td>
<td>85</td>
<td>376,046</td>
</tr>
<tr>
<td>T7_12</td>
<td>1</td>
<td>85</td>
<td>523,062</td>
</tr>
<tr>
<td>T25_27</td>
<td>1</td>
<td>85</td>
<td>263,545</td>
</tr>
<tr>
<td>T25_13</td>
<td>1</td>
<td>86</td>
<td>721,675</td>
</tr>
<tr>
<td>T6_25</td>
<td>1</td>
<td>85</td>
<td>82,656</td>
</tr>
</tbody>
</table>
Figure 1 (on next page)

Phylogenetic tree of representative Tomejil soil isolates based on PCR amplified, near-complete 16S rDNA (1,234 bp, panel A) and partial rpoB (356 bp, panel B) sequences.

Maximum likelihood trees (RAxML) were derived from ClustalW alignments. T: Tomejil soil isolates representative of the different rpoB genotypes (see text). The number of strains within each genotype group is indicated within parentheses. Trees include sequences from type strains of *Neorhizobium* species, as the closest taxonomic relatives, and of *A. tumefaciens*, as outgroup. Bootstrap support (1,000 replications) for the different nodes is indicated. Bars represent the number of substitutions per base. Genbank accession numbers are listed on Supplemental Table S1.
Phylogenetic tree based on repABC sequences from Tomejil genome sequences and from Neorhizobium megaplasmids.

Sequences similar to the *N. galegae* bv. orientalis HAMBI 540 1.8 Mb megaplasmid repABC region (3,628 bp) were extracted from Tomejil draft genomes sequences and from the *N. galegae* bv. officinalis HAMBI 1141 genome sequence, aligned with ClustalW, and a Neighbor-Joining consensus tree derived. Bootstrap support (1,000 replications) for the different nodes is indicated. Bar represents the number of substitutions per base.
**Figure 3** (on next page)

Phylogenetic tree of sequenced Tomejil strains and of *Neorhizobium* type strains based on a concatenation of complete *atpD, glnII, recA, rpoB*, and *thrC* genes (8,949 bp).

Maximum likelihood trees (RAxML) were derived from ClustalW alignments. The tree includes the *A. tumefaciens* type strain as outgroup. Bootstrap support (1,000 replications) for the different nodes is indicated. Bar represents the number of substitutions per base.
Neorhizobium galegae HAMBI 540\textsuperscript{T}

Neorhizobium huautlense DSM 21817\textsuperscript{T}

Neorhizobium alkalisoli DSM 21826\textsuperscript{T}

Agrobacterium tumefaciens Ach5
Dendrogram representation of a Euclidean distance matrix derived from pairwise ANIb distances among Tomejil and *Neorhizobium* type strain genomes.

The vertical red line indicates the 95% ANI threshold.