Complete genome sequence of the *Medicago* microsymbiont *Ensifer (Sinorhizobium) medicae* strain WSM419

Wayne Reeve1*, Patrick Chain2,3, Graham O’Hara1, Julie Ardley1, Kemanthi Nandesena1, Lambert Bräu1, Ravi Tiwari1, Stephanie Malfatti2,3, Hajnalka Kiss2,3, Alla Lapidus2, Alex Coppeland2, Matt Nolan2, Miriam Land2,4, Loren Hauser2,4, Yun-Juan Chang2,4, Natalia Ivanova2, Konstantinos Movromatis2, Victor Markowitz2, Nikos Kyrpides2, Margaret Gollagher6, Ron Yates1,7, Michael Dilworth1 & John Howieson1,7.

1 Centre for Rhizobium Studies, Murdoch University, Perth, Australia  
2 DOE Joint Genome Institute, Walnut Creek, California, USA  
3 Lawrence Livermore National Laboratory, Livermore, California, USA  
4 Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA  
5 Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA  
6 Institute for Sustainability and Technology Policy, Murdoch University, Perth, Australia  
7 Department of Agriculture and Food, South Perth, Australia

*Corresponding author:* Wayne Reeve

**Keywords:** microsymbiont, non-pathogenic, aerobic, Gram-negative rod, root-nodule bacteria, nitrogen fixation, Alphaproteobacteria

*Ensifer (Sinorhizobium) medicae* is an effective nitrogen fixing microsymbiont of a diverse range of annual *Medicago* (medic) species. Strain WSM419 is an aerobic, motile, non-spore forming, Gram-negative rod isolated from a *M. murex* root nodule collected in Sardinia, Italy in 1981. WSM419 was manufactured commercially in Australia as an inoculant for annual medics during 1985 to 1993 due to its nitrogen fixation, saprophytic competence and acid tolerance properties. Here we describe the basic features of this organism, together with the complete genome sequence, and annotation. This is the first report of a complete genome sequence for a microsymbiont of the group of annual medic species adapted to acid soils. We reveal that its genome size is 6,817,576 bp encoding 6,518 protein-coding genes and 81 RNA only encoding genes. The genome contains a chromosome of size 3,781,904 bp and 3 plasmids of size 1,570,951 bp, 1,245,408 bp and 219,313 bp. The smallest plasmid is a feature unique to this medic microsymbiont.

**Editorial note** - Readers are advised that in Opinion 84 the Judicial Commission of the International Committee on Systematics of Prokaryotes ruled that the genus name *Ensifer* Casida 1982 has priority over *Sinorhizobium* Chen et al. 1988 and the names are synonyms [1]. It was further concluded that the transfer of members of the genus *Sinorhizobium* to the genus *Ensifer*, as proposed by Young [2] would not cause confusion.

**Introduction**

Agricultural systems are nearly always nitrogen deficient, a factor which grossly limits their productivity. In fact, each year some 50 Tg of nitrogen is harvested globally in food crops [3], and must be replaced. External inputs of nitrogen to agriculture may come from mineral fertilizers, the production of which is heavily dependent on fossil fuels. Alternatively, nitrogen can be obtained from symbiotic nitrogen fixation (SNF) by root nodule bacteria (rhizobia) on nodulated legumes [4]. SNF is therefore considered a key biological process on the planet. The commonly accepted figure for global SNF in agriculture is 50-70 million metric tons annually, worth in excess of U.S. $10 billion [5]. Rhizobia associated with forage legumes contribute a substantial proportion of this fixed nitro-
gen across 400 million ha [5]. The amount fixed annually by the Ensifer (Sinorhizobium)-Medicago symbiosis is estimated to be worth $250 million.

A particular constraint to the formation of this symbiosis is acidity, due mainly to the acid-sensitive nature of the microsymbionts [6]. In laboratory culture, the medic microsymbionts fail to grow below pH 5.6 and are considered to be the most acid-sensitive of all the commercial root nodule bacteria [7]. Many agricultural regions have moderately acidic soils (typically in the pH range of 4.0 to 6.0) and this has prevented the Ensifer-Medicago symbiosis reaching its full potential [8]. Consequently, an effort was initiated in the 1980s to discover more acid-tolerant medic microsymbionts from world regions with acidic soils upon which annual medics had evolved. A particular suite of strains isolated from acidic soils on the Italian island of Sardinia proved to be acid soil tolerant [9], an attribute we now know is related to the presence of a unique set of genes required for acid adaptation [10]. Characterization of these acid-tolerant isolates revealed that they belonged to the species E. medicae and could be symbiotically distinguished from the related species E. meliloti by their unique capacity to fix nitrogen in association with annual acid soil adapted Medicago hosts of worldwide agronomic value [11], as well as with the perennial forage legume M. sativa (alfalfa) [12].

One of the acid-tolerant isolates, E. medicae strain WSM419, was isolated in 1981 from a nodule recovered from the roots of an annual medic (M. murex) growing south of Tempio in Sardinia. WSM419 is of particular interest because it is saprophytically competent in the acidic, infertile soils of southern Australia [9,13], and it is also a highly effective nitrogen fixing microsymbiont of a broad range of annual medics of Mediterranean origin [11,12]. These attributes contributed to the commercialization of the strain in Australia as an inoculant for acid soil medics between 1985 and 1993 [14,15]. Here we present a summary classification and a set of features (Table 1) for E. medicae strain WSM419, together with the description of a complete genome sequence and annotation.

**Classification and features**

E. medicae strain WSM419 forms mucoid colonies that may appear as donut shaped (Figure 1, left) on specific media such as YMA [13]. It is a Gram-negative, non-spore-forming rod (Figure 1, center) that has peritrichous flagellae (Figure 1, right). In minimal media E. medicae WSM419 has a mean generation time of 4.1 h when grown at 28°C [33]. It is a member of the Rhizobiaceae family of the class Alphaproteobacteria based on phylogenetic analysis. Figure 2 shows the phylogenetic neighborhood of E. medicae strain WSM419 inferred from a 16S rRNA based phylogenetic tree. An intragenic fragment of 1,440 bp was chosen since the 16S rRNA gene has not been completely sequenced in many type strains. A comparison of the entire 16S rRNA gene of WSM419 to completely sequenced 16S rRNA genes of other sinorhizobia revealed 4 and 18 bp mismatches to the reported sequences of E. meliloti (Sm1021) and E. fredii (YcS2, 15067 and SjzZ4), respectively.

In minimal media E. medicae WSM419 has a mean generation time of 4.1 h when grown at 28°C [33]. It is a member of the Rhizobiaceae family of the class Alphaproteobacteria based on phylogenetic analysis. Figure 2 shows the phylogenetic neighborhood of E. medicae strain WSM419 inferred from a 16S rRNA based phylogenetic tree. An intragenic fragment of 1,440 bp was chosen since

---

**Figure 1.** Unique colony morphology (Left) and scanning (Center) and transmission (Right) electron micrographs of E. medicae strain WSM419.
Ensifer (Sinorhizobium) medicae strain WSM419

Figure 2. Phylogenetic tree showing the relationships of E. medicae strain WSM419 to type strains in the Rhizobiaceae based on aligned sequences of the 16S rRNA gene (1,440 bp internal region). All sites were informative and there were no gap-containing sites. Phylogenetic analyses were performed using MEGA, version 3.1 [34]. Kimura two-parameter distances were derived from the aligned sequences [35] and a bootstrap analysis [36] as performed with 500 replicates in order to construct a consensus unrooted tree using the neighbor-joining method [37] for each gene alignment separately. Genera in this tree include Bradyrhizobium (B); Mesorhizobium (M); Rhizobium (R); Ensifer (Sinorhizobium) (S). Type strains are indicated with a superscript T. Strains with a genome sequencing project registered in GOLD [31] are in bold red print. Published genomes are designated with an asterisk.
**Symbiotaxonomy**

*E. medicae* and *E. meliloti* are traditionally separated on the basis of the effective nodulation (Nod+, Fix+) by *E. medicae* on *M. polymorpha* [38]. Specific symbiotic characteristics that further distinguish *E. medicae* WSM419 from *E. meliloti* include its ability to nodulate and fix nitrogen effectively with a wide range of annual Mediterranean medics, including *M. polymorpha*, *M. arabica*, *M. murex* and *M. sphaerocarpos*. WSM419 is symbiotically competent with these species when grown in acidic soils [39]. In contrast, WSM419 is Fix+ with the alkaline soil species of annual medics such as *M. littoralis*, *M. tornata* and hybrids of *M. littoralis/M. truncatula* [11,40]. WSM419 is also Nod+, Fix+ with the perennial forage legume *M. sativa* [11,12] but is less effective with this species than are some *E. meliloti* isolates. However, WSM419 is more effective at fixing nitrogen with *M. truncatula* than the previously sequenced *E. meliloti* Sm1021, making it an ideal candidate for inoculation of this model legume [12].

| MIGS ID | Property               | Term                                      | Evidence code                  |
|---------|------------------------|-------------------------------------------|-------------------------------|
|         | Current classification  |                                          |                               |
|         | Domain                 | *Bacteria*                                | TAS [17]                      |
|         | Phylum                  | *Proteobacteria*                          | TAS [18]                      |
|         | Class                   | *Alphaproteobacteria*                     | TAS [19,20]                   |
|         | Order                   | *Rhizobiales*                             | TAS [20,21]                   |
|         | Family                  | *Rhizobiaceae*                            | TAS [22,23]                   |
|         | Genus                   | *Ensifer*                                 | TAS [1,2,24-27]               |
|         | Species                 | *Ensifer medicae* strain WSM419            | TAS [1,2,11,24-28]            |
|         | Gram stain              | negative                                  | TAS [29]                      |
|         | Cell shape              | rod                                       | TAS [29]                      |
|         | Motility                | motile                                    | TAS [29]                      |
|         | Sporulation             | non-sporulating                           | TAS [29]                      |
|         | Temperature range       | mesophile                                 | TAS [29]                      |
|         | Optimum temperature     | 28°C                                      | TAS [29]                      |
|         | Salinity                | unknown                                   |                               |
|         | Oxygen requirement      | aerobic                                   | TAS [29]                      |
| MIGS-22 | Carbon source           | galactose, arabinose, glutamate           | TAS [9,13]                    |
|         | Energy source           | chemooorganotroph                         | TAS [9,13]                    |
| MIGS-6  | Habitat                 | Soil, root nodule, host                   | TAS [9]                       |
| MIGS-15 | Biotic relationship     | Free living or symbiotic                  | TAS [9]                       |
| MIGS-14 | Pathogenicity           | none                                      | TAS [16]                      |
|         | Biosafety level         | 1                                         | TAS [30]                      |
|         | Isolation               | *Medicago murex* root nodule              | TAS [9]                       |
| MIGS-4  | Geographic location     | Forestry Station 7 km south of Tempio, Sardinia, Italy | TAS [9] |
| MIGS-5  | Nodule collection date  | May 1st, 1981                             | TAS [31]                      |
| MIGS-4.1| Longitude               | 9.101915                                  | TAS [31]                      |
| MIGS-4.2| Latitude                | 40.888925                                 | TAS [31]                      |
| MIGS-4.3| Depth                   | <10 cm                                    | TAS [31]                      |
| MIGS-4.4| Altitude                | 350m                                      | TAS [31]                      |

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [32]. If the evidence code is IDA, then the property was directly observed for a living isolate by one of the authors or an expert mentioned in the acknowledgements.
Ensifer (Sinorhizobium) medicae strain WSM419

Genome sequencing and annotation
Genome project history
E. medicae WSM419 was selected for sequencing on the basis of its importance as a symbiotic nitrogen fixing bacterium in agriculture, and its tolerance for acidic soils [9,14]. This strain was selected for sequencing as part of the Community Sequencing Program of the Joint Genome Institute (JGI) in 2005. The genome project is deposited in the Genomes OnLine Database [31] and the complete genome sequence in GenBank. A summary of the project information is shown in Table 2.

| MIGS ID | Property               | Term                                                                 |
|---------|------------------------|----------------------------------------------------------------------|
| MIGS-31 | Finishing quality      | Finished                                                             |
|         | Libraries used         | Four Sanger libraries – 3 kb pUC18, 2 kb pTH1522, 8 kb pMCL200 and fosmid pCC1Fos |
| MIGS-28 | Sequencing platforms   | ABI3730xl; MegaBACE4500                                              |
| MIGS-29 | Sequencing coverage    | ~13× Sanger                                                          |
| MIGS-30 | Assemblers             | PHRED/PHRAP/CONSED                                                  |
| MIGS-32 | Gene calling method    | Critica, Generation and Glimmer                                    |
|         | Genbank ID             | CP000739 (pSMED01 or pSymB)<sup>a</sup>                            |
|         |                       | CP000740 (pSMED02 or pSymA)<sup>c</sup>                            |
|         |                       | CP000741 (pSMED03 or accessory plasmid)<sup>d</sup>                 |
| Genbank Date of Release | June 29, 2007               |
| GOLD ID |                       | Gc00590<sup>e</sup>                                              |
| NCBI project ID | 16304                        |
| Database: IMG | 640753051<sup>f</sup>                        |
| Project relevance | Symbiotic nitrogen fixation, agriculture                      |

a [http://www.ncbi.nlm.nih.gov/nuccore/150026743](http://www.ncbi.nlm.nih.gov/nuccore/150026743)

b [http://www.ncbi.nlm.nih.gov/nuccore/150030273](http://www.ncbi.nlm.nih.gov/nuccore/150030273)

c [http://www.ncbi.nlm.nih.gov/nuccore/150031715](http://www.ncbi.nlm.nih.gov/nuccore/150031715)

d [http://www.ncbi.nlm.nih.gov/nuccore/150032810](http://www.ncbi.nlm.nih.gov/nuccore/150032810)

e [http://genomesonline.org/GOLD_CARDS/Gc00590.html](http://genomesonline.org/GOLD_CARDS/Gc00590.html)

Growth conditions and DNA isolation
E. medicae strain WSM419 was grown to mid logarithmic phase in TY medium (a rich medium) [41] on a gyratory shaker at 28°C. DNA was isolated from 60 ml of cells using a CTAB (Cetyl trimethylammonium bromide) bacterial genomic DNA isolation method [JGI general information].

Genome sequencing and assembly
The genome was sequenced using a Sanger platform. All general aspects of library construction and sequencing performed at the JGI can be found at the JGI website. Sequence data statistics from the trace archive for this project are presented in Table 3.

All reads were assembled using the phrap assembler. Possible mis-assemblies were corrected and gaps between contigs were closed by custom primer walks from sub-clones or PCR products. Processing of sequence traces and base calling and assessment of data quality and assembly were performed with the PHRED/PHRAP/CONSED package [42-44]. The initial draft assembly was produced from 84,192 high-quality reads and consisted of 30 contigs (each with at least 20 reads per contig). Gaps in the sequence were primarily identified by mate-pair sequences and then closed by primer walking on gap-spanning library clones or genomic DNA amplified PCR products. True physical gaps were closed by combinatorial and
multiplex PCR. All repeated sequences were addressed using mate-pair sequences and PCR data. Sequence finishing and polishing added 638 reads. The final assembly of the main chromosome and 3 plasmids from 84,830 reads produced approximately 13-fold coverage across the genome. Assessment of final assembly quality was completed as described previously [45].

Table 3. Production sequence data for the *E. medicae* WSM419 genome project (JGI project 4001622)

| Library | Vector | Insert size (kb) | Reads | Mb   | q20 (Mb) |
|---------|--------|-----------------|-------|------|----------|
| BICH    | pMCL200 | 5.9             | 37,091| 36.3 | 25.7     |
| BICG    | pUC18c  | 2.6             | 33,520| 36.8 | 26.1     |
| BICI    | pCC1Fos | 38.8            | 13,929| 13.9 | 8.9      |
| FAUT    | pTH1522 | 2.1             | 7,376 | 6.4  | 5.4      |

## Genome annotation

Automated gene prediction was completed by assessing congruence of gene call results from three independent programs, the Critica [46], Generation, and Glimmer [47] modeling packages, and by comparing the translations to the GenBank non-redundant database using the basic local alignment search tool for proteins (BLASTP). Product description annotations were obtained using searches against the KEGG, InterPro, TIGRfams, PROSITE, and Clusters of Orthologous Groups of protein (COGs) databases. The tRNAscanSE tool [48] was used to find tRNA genes, whereas ribosomal RNAs were found by using BLASTN vs. the 16S and 23S ribosomal RNA databases. Initial comparative analyses of bacterial genomes and gene neighborhoods were completed using the JGI Integrated Microbial Genomes web-based interface. Additional gene prediction analysis and functional annotation was performed within the Integrated Microbial Genomes (IMG-ER) platform [49].

## Genome properties

The genome is 6,817,576 bp long with 61.15% GC content and comprised of four replicons (Table 4); one circular chromosome of size 3,781,904 bp (Figure 3) and three plasmids of size 1,570,951 bp, 1,245,408 bp and 219,313 bp (Figure 4). Of the 6,599 genes predicted, 6,518 were protein-coding genes, and 81 RNA only encoding genes. In addition, 305 pseudogenes were identified. The majority of the genes (70.4%) were assigned a putative function while those remaining were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 5.

Table 4. Genome Statistics for *E. medicae* WSM419.

| Attribute                        | Value   | % of Total |
|----------------------------------|---------|------------|
| Genome size (bp)                 | 6,817,576 | 100.00     |
| DNA coding region (bp)           | 6,001,805 | 88.03      |
| DNA G+C content (bp)             | 4,168,935 | 61.15      |
| Number of replicons              | 4       | 100.00     |
| Extrachromosomal elements        | 3       | 75.00      |
| Total genes                      | 6,599   | 100.00     |
| RNA genes                        | 81      | 1.23       |
| rRNA operons                     | 3       |            |
| Protein-coding genes             | 6,518   | 98.77      |
| Pseudo genes                     | 305     | 4.62       |
| Genes with function prediction   | 4,646   | 70.40      |
| Genes in paralog clusters        | 4,138   | 62.71      |
| Genes assigned to COGs           | 4,999   | 75.75      |
| Genes assigned Pfam domains      | 5,051   | 76.54      |
| Genes with signal peptides       | 2,170   | 32.88      |
| Genes with transmembrane helices | 1,481   | 22.44      |
| CRISPR repeats                   | 0       |            |
Figure 3. Graphical circular map of the chromosome and plasmids of *E. medicae* WSM419. From outside to the center: Genes on forward strand (color by COG categories as denoted in the IMG platform), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, sRNAs red, other RNAs black), GC content, GC skew. The replicons are not drawn to scale.

Table 5. Number of genes associated with the general COG functional categories.

| Code | Value | % age | Description                                           |
|------|-------|-------|-------------------------------------------------------|
| J    | 182   | 2.79  | Translation, ribosomal structure and biogenesis       |
| A    | 0     | 0.00  | RNA processing and modification                       |
| K    | 501   | 7.69  | Transcription                                         |
| L    | 250   | 3.84  | Replication, recombination and repair                 |
| B    | 1     | 0.02  | Chromatin structure and dynamics                      |
| D    | 36    | 0.55  | Cell cycle control, mitosis and meiosis               |
| Y    | 0     | 0.00  | Nuclear structure                                     |
| V    | 56    | 0.86  | Defense mechanisms                                    |
| T    | 247   | 3.79  | Signal transduction mechanisms                        |
| M    | 287   | 4.40  | Cell wall/membrane biogenesis                         |
| N    | 66    | 1.01  | Cell motility                                         |
| Z    | 0     | 0.00  | Cytoskeleton                                          |
| W    | 1     | 0.02  | Extracellular structures                              |
| U    | 106   | 1.63  | Intracellular trafficking and secretion               |
| O    | 178   | 2.73  | Posttranslational modification, protein turnover, chaperones |
| C    | 336   | 5.15  | Energy production and conversion                      |
Table 5 (cont.) Number of genes associated with the general COG functional categories.

| Code | value | % age | Description                                           |
|------|-------|-------|-------------------------------------------------------|
| G    | 582   | 8.93  | Carbohydrate transport and metabolism                 |
| E    | 622   | 9.54  | Amino acid transport and metabolism                   |
| F    | 109   | 1.67  | Nucleotide transport and metabolism                   |
| H    | 196   | 3.01  | Coenzyme transport and metabolism                     |
| I    | 209   | 3.21  | Lipid transport and metabolism                        |
| P    | 296   | 4.54  | Inorganic ion transport and metabolism                |
| Q    | 159   | 2.44  | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 687   | 10.54 | General function prediction only                      |
| S    | 528   | 8.10  | Function unknown                                      |
| -    | 1,519 | 23.30 | Not in COGs                                           |

Acknowledgements
This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396. We would like to gratefully acknowledge the funding received from Murdoch University Strategic Research Fund through the Crop and Plant Research Institute (CapRI), and the Grains Research and Development Corporation (GRDC), to support the National Rhizobium Program (NRP) and the Centre for Rhizobium Studies (CRS) at Murdoch University.

References
1. Judicial Commission of the International Committee on Systematics of Prokaryotes. The genus name Sinorhizobium Chen et al. 1988 is a later synonym of Ensifer Casida 1982 and is not conserved over the latter genus name, and the species name 'Sinorhizobium adhaerens' is not validly published. Opinion 84. Int J Syst Evol Microbiol 2008; 58: 1973. PubMed doi:10.1099/ijs.0.2008/005991-0

2. Young JM. The genus name Ensifer Casida 1982 takes priority over Sinorhizobium Chen et al. 1988, and Sinorhizobium morelense Wang et al. 2002 is a later synonym of Ensifer adhaerens Casida 1982. Is the combination Sinorhizobium adhaerens (Casida 1982) Willems et al. 2003 legitimate? Request for an Opinion. Int J Syst Evol Microbiol 2003; 53:2107-2110. PubMed doi:10.1099/ijs.0.02665-0

3. Peoples MB, Hauggaard-Nielsen H, Jensen EE. Chapter 13. The potential environmental benefits and risks derived from legumes in rotations. In: Emerich, DW & Krishnan HB (Eds.), Agronomy Monograph 52. Nitrogen Fixation in Crop Production Am Soc Agron, Crop Sci Soc Am & Soil Sci Soc Am 2009, pp. 349-385 Madison, Wisconsin, USA.

4. Sprent JI. Legume nodulation: a global perspective. 2009. Oxford, Wiley-Blackwell.

5. Herridge DF, Peoples MB, Boddey RM. Global inputs of biological nitrogen fixation in agricultural systems. Marschner Review. Plant Soil 2008; 311:1-18. doi:10.1007/s11104-008-9668-3

6. Robson AD, Loneragan JF. Nodulation and growth of Medicago truncatula on acid soils. II Colonization of acid soils by Rhizobium meliloti. Aust J Agric Res 1970; 21:435-445. doi:10.1071/AR9700435

7. Graham PH, Parker CA. Diagnostic features in the characterization of the root nodule bacteria of legumes. Plant Soil 1964; 20:383-396. doi:10.1007/BF00009567

8. Howieson JG. Characteristics of an ideotype acid tolerant pasture legume symbiosis in Mediterranean agriculture. Plant Soil 1995; 171:71-76. doi:10.1007/BF00009567

9. Howieson JG, Ewing MA. Acid tolerance in the Rhizobium meliloti-Medicago symbiosis. Aust J Agric Res 1986; 37:55-64. doi:10.1071/AR9860055

10. Reeve WG, Brau L, Castelli J, Garau G, Sohlenkamp C, Geiger O, Dilworth MJ, Glenn AR, Ho-
20. List editor. Validation List No. 107. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol 2006; 56: 1-6. PubMed doi:10.1099/ijs.0.64188-0

21. Kuykendall LD. Order VI. Rhizobiales ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 324.

22. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. Int J Syst Bacteriol 1980; 30: 225-420.

23. Conn HJ. Taxonomic relationships of certain non-sporeforming rods in soil. J Bacteriol 1938; 36: 320-321.

24. Chen WX, Yan GH, Li JL. Numerical taxonomic study of fast-growing soybean rhizobia and a proposal that Rhizobium fredii be assigned to Sinorhizobium gen. nov. Int J Syst Bacteriol 1988; 38: 392-397.

25. De Lajudie P, Willems A, Pot B, Dewettinck D, Maestrojuan G, Neyra M, Collins MD, Dreyfus B, Kersters K, Gillis M. Polyphasic taxonomy of Rhizobia: emendation of the genus Sinorhizobium and description of Sinorhizobium meliloti comb. nov., Sinorhizobium saheli sp. nov., and Sinorhizobium teranga sp. nov. Int J Syst Bacteriol 1994; 44: 715-733.

26. Willems A, Fernández-López M, Muñoz-Adelantado E, Goris J, De Vos P, Martínez-Romero E, Toro N, Gillis M. Description of new Ensifer strains from nodules and proposal to transfer Ensifer adhaerens Casida 1982 to Sinorhizobium as Sinorhizobium adhaerens comb. nov. Request for an opinion. Int J Syst Evol Microbiol 2003; 53: 1207-1217. PubMed doi:10.1099/ijs.0.02264-0

27. Lindström K, Martínez-Romero ME. International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of Agrobacterium and Rhizobium. Minutes of the meeting, 4 July 2001, Hamilton, Canada. Int J Syst Evol Microbiol 2002; 52: 2337. doi:10.1099/ijs.0.02524-0

28. Rome S, Fernandez MP, Brunel B, Normand P, Cleyet-Marel JC. Sinorhizobium medicae sp. nov., isolated from annual Medicago spp. Int J Syst Bacteriol 1996; 46: 972-980. PubMed
29. Kuykendall LD, Hashem F, Wang ET. Genus VII. Sinorhizobium, 2005, pp 358-361. In: Bergey’s Manual of Systematic Bacteriology. Second Edition. Volume 2 The Proteobacteria. Part C The Alpha-, Delta-, and Epsilonproteobacteria. Brenner DJ, Krieg NR, Staley JT (Eds.), Garrity GM (Editor in Chief) Springer Science and Business Media Inc, New York, USA.

30. Biological Agents. Technical rules for biological agents www.baua.de TRBA 466.

31. Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes OnLine Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res 2008; 36:D475-D479. PubMed doi:10.1093/nar/gkm884

32. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nat Genet 2000; 25: 5-29. PubMed doi:10.1038/75556

33. Reeve WG, Tiwari RP, Dilworth MJ, Glenn AR. Calcium affects the growth and survival of Rhizobium meliloti. Soil Biol Biochem 1993; 25:581-586. doi:10.1016/0038-0717(93)90197-J

34. Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004; 5:150-163. PubMed doi:10.1093/bib/bih52.150

35. Kimura M. A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980; 16:111-120. doi:10.1007/BF01731581 PubMed

36. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985; 39:783-791. doi:10.2307/2408678

37. Saitou N, Nei M. Reconstructing phylogenetic trees. Mol Biol Evol 1987; 4:406-425. PubMed

38. Rome S, Fernandez MP, Brunel B, Normand P, Cleyet-Marel JC. Sinorhizobium medicae sp. nov., isolated from annual Medicago spp. Int J Syst Bacteriol 1996; 46:972-980. PubMed

39. Howieson JG, Ewing MA. Annual species of Medicago differ greatly in their ability to nodulate on acid soils. Aust J Agric Res 1989; 40:843-850. doi:10.1071/AR9890843

40. Howieson JG, Evans P, Nutt B. Estimation of host-strain compatibility for symbiotic N-fixation between Rhizobium meliloti, several annual species of Medicago and Medicago sativa. Plant Soil 2000; 219:49-55. doi:10.1023/A:1004795617375

41. Reeve WG, Tiwari RP, Worsely PS, Dilworth MJ, Glenn AR, Howieson JG. Constructs for insertional mutagenesis, transcriptional signal localisation and gene regulation studies in root nodule and other bacteria. Microbiology 1999; 145:1307-1316. PubMed doi:10.1099/13500872-145-6-1307

42. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998; 8:186-194. PubMed

43. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998; 8:175-185. PubMed

44. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res 1998; 8:195-202. PubMed

45. Chain P, Larimer F, Regala W, Lao V, Delcher AL, Land M, Hauser L, Hooper A, Klotz M, Norton J, et al. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph Nitrosomonas europaea. J Bacteriol 2003; 185:2759-2773. doi:10.1128/JB.185.9.2759-2773.2003 PubMed

46. Badger JH, Olsen G. CRITICA: coding region identification tool invoking comparative analysis. Mol Biol Evol 1999; 16:512-524. PubMed

47. Delcher AL, Hamon D, Kasif S, White O, Salzberg SL. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 1999; 27:4636-4641. PubMed doi:10.1093/nar/27.23.4636

48. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997; 25:955-964. PubMed doi:10.1093/nar/25.5.955

49. Markowitz VM, Szeto E, Palaniappan K, Grechkin Y, Chu K, Chen IMA, Duschak I, Anderson I, Lykidis A, Mavromatis K, et al. The Integrated Microbial Genomes (IMG) system in 2007: data content and analysis tool extensions. Nucleic Acids Res 2008; 36:D528-D533. PubMed doi:10.1093/nar/gkm846