genome**A**nnouncements™

Check for updates

Complete Genome Sequence of *Mesorhizobium sophorae* ICMP 19535^T, a Highly Specific, Nitrogen-Fixing Symbiont of New Zealand Endemic *Sophora* spp.

Sofie E. De Meyer,[a,b] **Dung Tuan Nguyen,**[c] **Penghao Wang,**[a] **Mitchell Andrews**[c]

Centre for Rhizobium Studies, Murdoch University, Murdoch, Western Australia, Australia[a]; Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium[b]; Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln, New Zealand[c]

ABSTRACT We report here the complete genome sequence of *Mesorhizobium sophorae* ICMP 19535^T. This strain was isolated from *Sophora microphylla* root nodules and can nodulate and fix nitrogen with this host and also with *Sophora prostrata*, *Sophora longicarinata*, and *Clianthus puniceus*. The genome consists of 8.05 Mb.

Most legume species (plant family *Leguminosae*) can fix atmospheric nitrogen via symbiotic bacteria ("rhizobia") in root nodules, which can give them an advantage under low-nitrogen soil conditions (1, 2). New Zealand is geographically isolated, and this separation has given rise to a unique native flora (3). There are four genera of indigenous legumes on the main New Zealand islands, and all of them are in the subfamily *Papilionoideae* and all nodulate. These genera are *Carmichaelia*, *Clianthus*, and *Montigena* in the tribe Galegeae, and *Sophora* is within the tribe Sophoreae (4, 5).

Forty-eight rhizobial isolates from four New Zealand native *Sophora* spp. sampled at eight different field sites were separated into eight groups and three individual isolates on the basis of their concatenated *recA*, *glnII*, and *rpoB* gene sequences, while showing almost identical unique *nodA* and *nodC* gene sequences (6). Seven of these groups have been formally identified as new species, one of which is *Mesorhizobium sophorae* ICMP 19535^T (7, 8). This relationship between New Zealand native *Sophora* spp. and *Mesorhizobium* spp. with specific nodulation gene sequences is highly specific (9).

Genomic DNA was isolated using the GENTRA Pure-Gene kit (Qiagen). The genome sequence was determined using the PacBio RS II (Pacific Biosciences) platform at ChunLab, Inc. (South Korea). Raw sequences were assembled with PacBio SMRT Analysis version 2.0 software (Pacific Biosciences). Gene prediction was performed using tRNA-scan version 1.3.1 (10) for the tRNA search, Cmsearch (INFERNAL version 1.0.2) and Rfam version 12.0 for the rRNA and noncoding RNA searches, and gPilerCR version 1.06 (11) and CRT1.2 (12) for the clustered regularly interspaced short palindromic repeat (CRISPR) searches. Functional annotation was performed by a homology search against the KEGG (13), SEED (14), Swissprot (15), and eggNOG (16) databases. In total, 8,049,106 bp of sequence information was obtained with a total of 24 contigs. The average contig size was 334,479 bp, with the largest being 1,626,386 bp and the shortest 14,820 bp. The genome features an average GC content of 62.22%, and the N_{50} value was 666,711 bp. A total of 7,659 protein-coding sequences and 6 rRNAs and 58 tRNAs were predicted.

The genomic sequence comparison against other sequenced *Mesorhizobium* strains indicated that around 70% of the sequences of ICMP 19535^T are highly conserved. However, ICMP 19535^T demonstrated segments of unique sequences and genomics rearrangements. The closest strains that were matched to ICMP 19535^T include the

rhizobia *Mesorhizobium ciceri* bv. *biserrulae* WSM1284 and *Mesorhizobium loti* NZP2037. Preliminary sequence analysis revealed that the ICMP 19535^T genome contains many key genes known to be associated with legume symbiosis. A number of clusters of nodulation genes (*nod*) were identified, and, interestingly, one cluster of *nod* genes is located on a genomic island that contains multiple *nif* genes, which are known to encode enzymes involved in nitrogen fixation. Moreover, a cluster of *fix* genes was discovered. Further analysis of the genome will lead to a better understanding of the strain's genomic mechanisms that regulate nodulation and nitrogen fixation associated with native New Zealand legume symbioses.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under accession number NNRI00000000. The version described here is the first version, NNRI01000000.

REFERENCES

1. The Legume Phylogeny Working Group (LPWG). 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. Taxon 66:44–77. https://doi.org/10.12705/661.3.
2. Andrews M, Raven JA, Lea PJ. 2013. Do plants need nitrate? The mechanisms by which nitrogen form affects plants. Ann Appl Biol 163:174–199. https://doi.org/10.1111/aab.12045.
3. McGlone MS, Duncan RP, Heenan PB. 2001. Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand. J Biogeogr 28:199–216. https://doi.org/10.1046/j.1365-2699.2001.00525.x.
4. Wagstaff SJ, Heenan PB, Sanderson MJ. 1999. Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). Am J Bot 86:1346–1356. https://doi.org/10.2307/2656781.
5. Heenan PB, De Lange PJ, Wilton AD. 2001. *Sophora* (Fabaceae) in New Zealand: taxonomy, distribution, and biogeography. N Z J Bot 39:17–53. https://doi.org/10.1080/0028825X.2001.9512715.
6. Tan HW, Heenan PB, De Meyer SE, Willems A, Andrews M. 2015. Diverse novel mesorhizobia nodulate New Zealand native *Sophora* species. Syst Appl Microbiol 38:91–98. https://doi.org/10.1016/j.syapm.2014.11.003.
7. De Meyer SE, Tan HW, Heenan PB, Andrews M, Willems A. 2015. *Mesorhizobium waimense* sp. nov. isolated from *Sophora longicarinata* root nodules and *Mesorhizobium cantuariense* sp. nov. isolated from *Sophora microphylla* root nodules. Int J Syst Evol Microbiol 65:3419–3426. https://doi.org/10.1099/ijsem.0.000430.
8. De Meyer SE, Tan HW, Andrews M, Heenan PB, Willems A. 2015. *Mesorhizobium calcicola* sp. nov., *Mesorhizobium waitakense* sp. nov., *Mesorhizobium sophorae* sp. nov., *Mesorhizobium newzealandense* sp. nov. and *Mesorhizobium kowhaii* sp. nov. isolated from *Sophora* root nodules in New Zealand. Int J Syst Evol Microbiol 66:786–795. https://doi.org/10.1099/ijsem.0.000796.
9. Andrews M, Andrews ME. 2017. Specificity in legume-rhizobia symbioses. Int J Mol Sci 18:705. https://doi.org/10.3390/ijms18040705.
10. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res 33:W686–W689. https://doi.org/10.1093/nar/gki366.
11. Edgar RC. 2007. PILER-CR: fast and accurate identification of CRISPR repeats. BMC Bioinformatics 8:18. https://doi.org/10.1186/1471-2105-8-18.
12. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, Hugenholtz P. 2007. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 8:209. https://doi.org/10.1186/1471-2105-8-209.
13. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42:D199–D205. https://doi.org/10.1093/nar/gkt1076.
14. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691–5702. https://doi.org/10.1093/nar/gki866.
15. UniProt Consortium. 2015. UniProt: a hub for protein information. Nucleic Acids Res 43:D204–D212. https://doi.org/10.1093/nar/gku989.
16. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldón T, Rattei T, Creevey C, Kuhn M, Jensen LJ, von Mering C, Bork P. 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res 42:D231–D239. https://doi.org/10.1093/nar/gkt1253.