











# Complete Genome of *Sinorhizobium meliloti* AK76, a Symbiont of Wild Diploid *Medicago lupulina* from the Mugodgary Mountain Region

Maria E. Vladimirova,<sup>a</sup>  Victoria S. Muntyan,<sup>a</sup>  Alexey M. Afonin,<sup>a</sup>  Alexey N. Muntyan,<sup>a</sup>  Olga A. Baturina,<sup>b</sup> Elena A. Dzyubenko,<sup>c</sup>  Alla S. Saksaganskaya,<sup>a</sup>  Boris V. Simarov,<sup>a</sup>  Marina L. Roumiantseva,<sup>a</sup>  Marsel R. Kabilov<sup>b</sup>

<sup>a</sup>Federal State Budget Scientific Institution All-Russian Research Institute for Agricultural Microbiology (FSBSI ARRIAM), Laboratory of Genetics and Selection of Microorganisms, Saint Petersburg, Russian Federation

<sup>b</sup>Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences (ICBFM SB RAS), Novosibirsk, Russian Federation

<sup>c</sup>Vavilov All-Russian Institute of Plant Genetic Resources (VIR), Saint Petersburg, Russian Federation

**ABSTRACT** *Sinorhizobium meliloti* is a symbiotic bacterial species forming nitrogen-fixing nodules on roots of annual and perennial *Medicago* spp. We report the full genome sequence of *S. meliloti* strain AK76, an effective symbiont of the wild diploid plant *Medicago lupulina* grown in the Mugodgary Mountain region, Kazakhstan.

*Sinorhizobium meliloti* AK76 was isolated from pink nodules of wild-growing annual diploid *Medicago lupulina* native to the Mugodgary Mountains at the northwest of Kazakhstan (48.58°N, 57.7697°E) in 2002 as a strain tolerant to 0.75 M NaCl (1–3). *Medicago lupulina* plants are well nodulated by two closely related rhizobial species, one of which is *Sinorhizobium medicae* (4). Here, we report the genome of *S. meliloti* native isolate AK76.

The AK76 strain was isolated from sterilized crushed well-developed pink nodules and passaged several times on tryptone-yeast (TY) medium (28°C) in order to get pure single colonies, which were further long-term stored at –80°C (1, 5). To rejuvenate AK76, its cells were restarted on solid TY plates, and then a single colony was grown in liquid TY (28°C, 180 rpm) up to an optical density at 600 nm of >0.75. Cells were harvested by centrifugation at 4,000 × *g*. Genomic DNA was isolated using phenol-chloroform extraction (6).

Genomic DNA was isolated with the GeneJET genomic DNA purification kit (Thermo Fisher Scientific) and fragmented to an average size of 600 bp in a microTUBE AFA fiber snap-cap tube (Covaris S2 instrument). The paired-end (PE) library was constructed using dual-index NEBNext multiplex oligonucleotides and a NEBNext Ultra II DNA library prep kit for Illumina (New England BioLabs). The DNA library was sequenced with a reagent kit v3 (600 cycle) on a MiSeq sequencer (Illumina) in the Genomics Core Facility (Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences [ICBFM SB RAS]), and 4.4 million PE reads were generated.

Default parameters were used for all software unless otherwise specified. The reads were quality trimmed, and adapter sequences were removed using Trim Galore v0.4.4 (<https://github.com/FelixKrueger/TrimGalore>). Long reads for isolated AK76 DNA were obtained using a MinION (Oxford Nanopore) sequencer R9.4 (Federal State Budget Scientific Institution All-Russian Research Institute for Agricultural Microbiology [FSBSI ARRIAM]). A barcoded DNA library constructed with the 1D native barcoding genomic DNA protocol (with kits EXP-NBD103 and SQK-LSK108). The raw fast5 files were base called with Albacore v2.3.1 (the run yielded 307,000 reads with an  $N_{50}$  of 7 kb). The resulting reads were demultiplexed using Deepbiner v0.2.0 (7) and cleaned using Porechop v0.2.3 (<https://github.com/rwrick/Porechop>).

**Editor** Leighton Pritchard, SIPBS, University of Strathclyde

**Copyright** © 2022 Vladimirova et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Marsel R. Kabilov, [kabilov@niboch.nsc.ru](mailto:kabilov@niboch.nsc.ru).

The authors declare no conflict of interest.

**Received** 10 November 2021

**Accepted** 13 February 2022

**Published** 28 February 2022

Illumina and Nanopore reads were assembled into 4 replicons—chromosome (3.58 Mb; GC content, 62.85%), pSymA (1.58 Mb; GC content, 60.29%), pSymB (1.72 Mb; GC content, 62.28%), and pAK76 (172.8 kb; GC content, 59.11%)—by Tricycler v0.4.1 (8). The complete genome size of AK76 is 7.05 Mb. Genome annotation of AK76 by NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.0 (9) resulted in 6,650 coding DNA sequences (CDS), including sequences encoding 6,296 proteins, 68 RNAs (tRNA, transfer-messenger RNA [tmRNA], noncoding RNA [ncRNA], and rRNA of the 3 *rrn* operons), and 286 pseudogenes. A single genomic island (40.29 kb; GC content, 58.47%) containing 53 open reading frames (ORFs) was detected on the AK76 chromosome using Islander v0.1 (10). PHASTER web server (accessed 14 September 2021) analysis (11) indicated that 16 of these ORFs were homologous to the intact *Rhizobium* phage 16-3 (GenBank accession number [NC\\_011103](https://doi.org/10.1101/111103)). GTDB-tk v1.7 (12) with reference data for Genome Taxonomy Database (GTDB) R202 were used for taxonomic classifications of the genome. Analysis confirmed that the AK76 strain is *Sinorhizobium meliloti* and the closest strain to AK76 is USDA1002 (GenBank assembly accession number [GCA\\_009601385](https://doi.org/10.1101/1385)).

**Data availability.** The genome sequence of *Sinorhizobium meliloti* AK76 is deposited in GenBank under accession numbers [CP066358](https://doi.org/10.1101/1385), [CP066359](https://doi.org/10.1101/1385), [CP066360](https://doi.org/10.1101/1385), and [CP066361](https://doi.org/10.1101/1385). The raw reads are available in SRA under the accession numbers [SRX13025480](https://doi.org/10.1101/1385) and [SRX13025481](https://doi.org/10.1101/1385). This announcement describes the first version of the genome assembly.

## ACKNOWLEDGMENT

This work was supported by research grant RSF 20-16-00105.

## REFERENCES

- Ibragimova MV, Rumyantseva ML, Onishchuk OP, Belova VS, Kurchak ON, Andronov EE, Simarov BV, Dzyubenko NI. 2006. Symbiosis between the root-nodule bacterium *Sinorhizobium meliloti* and alfalfa (*Medicago sativa*) under salinization conditions. *Microbiology* 75:77–81. <https://doi.org/10.1134/S0026261706010140>.
- Ivanov AI. 1980. Alfalfa. p 350. In Brezhnev D (ed), *Scientific works of the Academy of Agricultural Sciences*. State Publishing House, Kolos, Moscow, Russia. (In Russian.)
- Greene SL, Kisha TJ, Dzyubenko NI. 2008. Conserving alfalfa wild relatives: is past introgression with Russian varieties evident today? *Crop Sci* 48:1853–1864. <https://doi.org/10.2135/cropsci2007.12.0668>.
- Bailly X, Giuntini E, Sexton MC, Ryan PJ, Lower RPJ, Harrison PW, Kumar N, Young JPW. 2011. Population genomics of *Sinorhizobium medicae* based on low-coverage sequencing of sympatric isolates. *ISME J* 5:1722–1734. <https://doi.org/10.1038/ismej.2011.55>.
- Beringer JE. 1974. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* 84:188–198. <https://doi.org/10.1099/00221287-84-1-188>.
- Green MR, Sambrook J. 2017. Isolation of high-molecular-weight DNA using organic solvents. *Cold Spring Harb Protoc* 2017:pdb.prot093450. <https://doi.org/10.1101/pdb.prot093450>.
- Wick RR, Judd LM, Holt KE. 2018. Deepbiner: demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *bioRxiv* <https://doi.org/10.1101/366526>.
- Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *bioRxiv* <https://doi.org/10.1101/2021.07.04.451066>.
- Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>.
- Hudson CM, Lau BY, Williams KP. 2015. Islander: a database of precisely mapped genomic islands in tRNA and tmRNA genes. *Nucleic Acids Res* 43:D48–D53. <https://doi.org/10.1093/nar/gku1072>.
- Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
- Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-TK: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.