



Complete Genome Sequence of an Efficient *Rhizobium leguminosarum* bv. *viciae* Strain, A1

 Alexey M. Afonin,^a Emma S. Gribchenko,^a Anton S. Sulima,^a Vladimir A. Zhukov^a

^aLaboratory of Genetics of Plant-Microbe Interactions, All-Russia Research Institute for Agricultural Microbiology, St. Petersburg, Russia

ABSTRACT *Rhizobium leguminosarum* strain A1 is used in inoculation experiments with a wide range of pea (*Pisum sativum* L.) lines. In this study, we report the genome sequence of strain A1, consisting of a 5.06-Mbp circular chromosome and circular plasmids ranging from 804,800 bp to 154,738 bp long.

Strain A1 was first isolated in 1986 from the nodules of a pea plant cultivated in a field adjacent to the All-Russia Research Institute for Agricultural Microbiology (ARRIAM) (1). It was shown to possess useful qualities, such as being able to nodulate a wide range of pea varieties, including those carrying the *Sym2A* allele inherent in line NGB2150 (JI1357, WBH 2150) with the Afghan or Afghanistan phenotype, and it was capable of overcoming competitive nodulation blocking, making it an efficient strain for agricultural use (1, 2). Additionally, it was shown to produce large quantities of various lipochitin oligosaccharides (LCOs), including Nod factors (3). The strain was observed to lose symbiotic properties much faster than other strains used in inoculation experiments (e.g., RCAM1026) (data not shown). The genetic factors responsible for the multiple unique features remained undiscovered.

Plants of *Pisum sativum* line NGB2150 were inoculated with the A1 strain, pink nodules were harvested, bacteria were isolated as described previously (4), and the strain was preserved in 10% glycerol at -80°C . For DNA isolation, the strain was revived on solid tryptone-yeast extract (TY) medium. One colony was chosen for subsequent procedures. The strain was cultivated in 50 ml of liquid TY medium in a 100-ml flask at 28°C and 200 rpm (5). The culture was harvested after 48 h of incubation. DNA was isolated using the phenol-chloroform method (6) and quantified with a spectrophotometer (BioSpec-mini; Shimadzu, Japan). The required library absorption parameters were an A_{260}/A_{280} ratio of ~ 2 and an A_{260}/A_{230} ratio of >1.8 .

Long-read whole-genome sequencing was performed using a MinION sequencer (Oxford Nanopore, United Kingdom) in the ARRIAM. The SQK-LSK109 ligation sequencing kit and the EXP-NBD104 native barcoding expansion 1-12 kit were used to prepare the library according to the manufacturer's instructions, omitting the DNA-shearing step. The reads were base called and demultiplexed using Guppy base caller (v. 3.3.0). The resulting read N_{50} value was 31,627 bp, with a total read length of 0.2 Gbp and estimated coverage of $25\times$. The Flye pipeline (v. 2.6) (7) was used to assemble the Nanopore reads. The resulting assembly was corrected four times using Racon (v. 1.3.2) (8) (with the modifiers $-m 8 -x -6 -g -8 -w 500$), followed by a single polish using the medaka program (v. 0.10.0) with default parameters.

Short-read whole-genome sequencing of the strain was carried out on an Illumina system with the TruSeq DNA PCR-free kit; in total, 8,721,349 bp of 2×150 -bp sequence reads were generated. The reads were quality trimmed and adapter sequences and possible contaminants were removed as described previously (9); after filtering, the expected coverage was about $218\times$. The short reads were used to polish the assem-

Citation Afonin AM, Gribchenko ES, Sulima AS, Zhukov VA. 2020. Complete genome sequence of an efficient *Rhizobium leguminosarum* bv. *viciae* strain, A1. Microbiol Resour Announc 9:e00249-20. <https://doi.org/10.1128/MRA.00249-20>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2020 Afonin 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 Alexey M. Afonin, aafonin@arriam.ru.

Received 11 March 2020

Accepted 14 April 2020

Published 7 May 2020

TABLE 1 Characteristics of the replicons of strain *Rhizobium leguminosarum* A1

Amplicon	Length (bp)	No. of CDSs	No. of tRNAs	GC content (%)	Accession no.
A1 chromosome	5,067,161	4,943	52	60.94	CP049730
pRL12	804,800	745	0	60.71	CP049731
pRL11	661,136	656	2	60.70	CP049732
pRL10	632,304	588	0	60.87	CP049733
pRLa11	328,507	321	0	57.67	CP049734
pRLa12	154,738	185	0	56.48	CP049735

bled genome using the Pilon (v. 1.22) algorithm (10). PGAP was used to annotate the assembled transcripts (11).

The genome of strain A1 consists of 6 replicons, including 1 chromosome and 5 plasmids. The statistics for the amplicons are presented in Table 1.

The circularity of all of the assembled fragments was reported by the Flye assembly pipeline and verified by mapping the long reads to the assembled fragments using Minimap2 (12), with the map-ont mapping mode, and inspecting the coverage uniformity. The chromosome was rotated so that the *dnaA* gene was placed at the start of the sequence; for each plasmid, a *repABC* operon was located and placed at the start of the sequence.

The relation to other *Rhizobium leguminosarum* strains was determined using the average nucleotide identity (ANI) method (13). The strain closest to the A1 strain was strain RCAM1026 (97.2%), which is used for inoculation studies in the ARRIAM (14), placing the strain in genospecies C, according to reference 15.

The coding sequences (CDSs) predicted by the PGAP were annotated using eggNOG mapper (v. 2) (16) with the eggNOG (v. 5.0) database (17). Additionally, the CDSs were compared to the latest version of the UniProt Swiss-Prot curated database (18). BLASTp (v. 2.9.0+) was used to search the database with the value $1e-10$, and the identity threshold was set at 60% (19).

The plasmid pRLa11 contains 13 predicted Nod factor-associated genes (*nodA* [NCBI accession number [WP_017958626.1](#)], *nodJ* [[WP_017958630.1](#)], *nodN* [[WP_138333862.1](#)], *nodM* [[WP_138333863.1](#)], *nodL* [[WP_138333865.1](#)], *nodE* [[WP_138333867.1](#)], *nodF* [[WP_138333869.1](#)], *nodD1* [[WP_138333871.1](#)], *nodB* [[WP_138333873.1](#)], *nodC* [[WP_138333874.1](#)], *nodI* [[WP_138333928.1](#)], *nodT* [[WP_165586599.1](#)], and *nodX* [[WP_138333876.1](#)]), as described previously (3). Additionally, 5 nodulation genes were found on the chromosome (*nodT* [[WP_018068951.1](#)], *nodG* [[WP_018070560.1](#)], *nodT* [[WP_130672970.1](#)], *nodN* [[WP_130673101.1](#)], and *nodL* [[WP_130673413.1](#)]), 1 on pRL10 (*nodT* [[WP_018071823.1](#)]), and 1 on pRLa12 (*tolC* family protein [[WP_165586630.1](#)]). The gene present in four distinct copies was *nodT*, which was previously reported to be involved in the secretion of small molecules and, presumably, nodulation factors (20). Multiple clusters of genes annotated as *vir* genes belonging to type IV secretion systems were found on the pRLa11 and pRLa12 plasmids.

The large number of *nod* genes found in the genome is probably responsible for the previously observed high variability of the LCOs produced (3). Multiple genes encoding secretion systems, including the four copies of the *nodT* gene, may be the cause of the increased Nod factor excretion by the strain described previously (3). The full-genome sequence of this strain will be useful for further investigation of the symbiotic properties of this strain.

Data availability. The assemblies and sequence data have been deposited in the NCBI database. The BioProject number is [PRJNA609819](#), the BioSample number is [SAMN14260269](#), and the assembly accession numbers are [CP049730](#) to [CP049735](#). The raw Illumina data can be found under number [SRR11216745](#), and the demultiplexed fastq file, with barcodes removed, from the MinION runs can be found under number [SRR11216744](#). This announcement describes the first version of the genome assembly.

ACKNOWLEDGMENTS

This work was supported by RFBR grant 18-34-00844 for A.M.A. and A.S.S. (Nanopore sequencing and data processing) and RSF grant 17-76-30016 for E.S.G. and V.A.Z. (Illumina sequencing and library preparation).

REFERENCES

- Chetkova SA, Tikhonovich IA. 1986. Isolation and study of *Rhizobium leguminosarum* strains effective on peas from Afghanistan. *Mikrobiologiya* 55:143–147.
- Sulima AS, Zhukov VA, Afonin AA, Zhernakov AI, Tikhonovich IA, Lutova LA. 2017. Selection signatures in the first exon of paralogous receptor kinase genes from the *Sym2* region of the *Pisum sativum* L. genome. *Front Plant Sci* 8:1957. <https://doi.org/10.3389/fpls.2017.01957>.
- Ovtsyna AO, Rademaker GJ, Esser E, Weinman J, Rolfe BG, Tikhonovich IA, Lugtenberg BJ, Thomas-Oates JE, Spaink HP. 1999. Comparison of characteristics of the *nodX* genes from various *Rhizobium leguminosarum* strains. *Mol Plant Microbe Interact* 12:252–258. <https://doi.org/10.1094/MPMI.1999.12.3.252>.
- Novikova N, Safronova V. 1992. Transconjugants of *Agrobacterium radiobacter* harbouring *sym* genes of *Rhizobium galegae* can form an effective symbiosis with *Medicago sativa*. *FEMS Microbiol Lett* 93:261–268. <https://doi.org/10.1111/j.1574-6968.1992.tb05107.x>.
- Beringer JE. 1974. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* 84:188–198. <https://doi.org/10.1099/00221287-84-1-188>.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. 2002. Short protocols in molecular biology: A compendium of methods from Current Protocols in Molecular Biology, 5th ed. John Wiley & Sons, New York, NY.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>.
- Muntyan VS, Baturina OA, Afonin AM, Cherkasova ME, Laktionov YV, Saksaganskaya AS, Kabilov MR, Roumiantseva ML. 2019. Draft genome sequence of *Sinorhizobium meliloti* AK555. *Microbiol Resour Announc* 8:e01567-18. <https://doi.org/10.1128/MRA.01567-18>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
- Afonin A, Sulima A, Zhernakov A, Zhukov V. 2017. Draft genome of the strain RCAM1026 *Rhizobium leguminosarum* bv. *viciae*. *Genom Data* 11:85–86. <https://doi.org/10.1016/j.gdata.2016.12.003>.
- Kumar N, Lad G, Giuntini E, Kaye ME, Udomwong P, Shamsani NJ, Young JPW, Bailly X. 2015. Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biol* 5:140133. <https://doi.org/10.1098/rsob.140133>.
- Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol Biol Evol* 34:2115–2122. <https://doi.org/10.1093/molbev/msx148>.
- Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309–D314. <https://doi.org/10.1093/nar/gky1085>.
- Brouns S. 2012. Molecular biology. A Swiss army knife of immunity. *Science* 337:808–809. <https://doi.org/10.1126/science.1227253>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Rivilla R, Sutton JM, Downie JA. 1995. *Rhizobium leguminosarum* NodT is related to a family of outer-membrane transport proteins that includes TolC, PrtF, CyaE and AprF. *Gene* 161:27–31. [https://doi.org/10.1016/0378-1119\(95\)00235-x](https://doi.org/10.1016/0378-1119(95)00235-x).