Draft Genome Sequence of *Bradyrhizobium* sp. Strain LVM 105, a Nitrogen-Fixing Symbiont of *Chamaecrista fasciculata* (Michx.) Greene

Hari B. Krishnan,[a,b] Won-Seok Kim,[b] Scott A. Givan[c]

[a]Plant Genetics Research, USDA-Agricultural Research Service, Columbia, Missouri, USA
[b]Plant Science Division, University of Missouri, Columbia, Missouri, USA
[c]Informatics Research Core Facility, Molecular Microbiology and Immunology, University of Missouri, Columbia, Missouri, USA

ABSTRACT Here, we present the draft genome sequence of *Bradyrhizobium* sp. strain LVM 105, a soil bacterium that forms nitrogen-fixing nodules on the roots of partridge pea. The genome size was 8,386,213 bp, with a GC content of 63.8%. The draft assembly contained 97 scaffolds with an N_{50} value of 415,026 bp.

Partridge pea, *Chamaecrista fasciculata*, is an annual legume belonging to the subfamily Caesalpinioideae. This legume enters into symbiosis with rhizobia, resulting in the formation of nitrogen-fixing nodules. The nodulation process in *Chamaecrista* is believed to have evolved independently from that of the subfamily Papilionoideae, representing around 60 million years of independent evolution (1, 2). The characterization of the genomes of the symbionts will afford insights into the origin and evolution of nitrogen-fixing nodules. Here, we present the sequence assembly and annotation of a rhizobium that forms nitrogen-fixing nodules in partridge pea.

Bradyrhizobium sp. strain LVM 105 was isolated from nodules of partridge pea growing in a wooded region in Columbia, Missouri (USA), following established protocol (3). A single colony was grown in liquid yeast extract mannitol (YEM) medium (55 mM mannitol, 0.1% yeast extract, 3 mM K_2HPO_4, 0.8 mM $MgSO_4$, and 1.7 mM NaCl; pH 6.8) for 4 days at 30°C. Genomic DNA was isolated using DNAzol reagent (Thermo Fisher Scientific, Waltham, MA) and sheared to generate average fragment sizes of 550 bp. A library was constructed following the manufacturer's protocol with reagents supplied in the TruSeq DNA PCR-Free sample preparation kit (Illumina, San Diego, CA). The library was diluted and sequenced according to Illumina's standard sequencing protocol for the MiSeq platform. To assemble the draft genome sequence, we started with a single 150-mer paired-end Illumina MiSeq data set derived from a single shotgun library. Using relatively light settings (minimum base quality of 15 and minimum read length of 125), 1,305,481 forward and reverse reads (2,610,962 total reads) were trimmed and filtered to remove low-quality bases using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Reads were also filtered to remove PhiX using Bowtie version 1.2.2 (4). Subsequent to the quality control regimen, 1,070,398 properly paired reads and 188,147 unpaired reads were assembled using SPAdes version 3.11.1 using the "-k auto" option flag (5, 6) and the BayesHammer error-correction algorithm. Actual kmer lengths selected during the assembly were 21, 33, 55, and 77. After removing contigs shorter than 500 nt in length, the final assembly contained 97 scaffolds with a total length of 8,386,213 nt, with an N_{50} value of 415,026 nt and a GC content of 63.81%. We used a variety of software to evaluate the resulting assembly, including QUAST (7), REAPR (8), FRC (9), and custom Bash scripts. The mean read coverage for the assembly was 37.82×. The protein-coding open

reading frames were predicted and annotated using the NCBI Prokaryotic Genome Annotation Pipeline (10).

A total of 7,453 protein-coding sequences, 3 rRNA genes, and 51 tRNA genes were detected. A BLAST search analysis of the 16S rRNA gene sequence of LVM 105 revealed close phylogenetic relationships with different strains of *Bradyrhizobium diazoefficiens*. Genes involved in motility, root colonization, and siderophore production were found in the genome. Genes involved in nitrogen fixation and nodulation were also detected. A cluster of genes coding for a type III secretion system (11), which is involved in the delivery of effector proteins directly into the host plants, was present in the genome.

Data availability. The draft genome sequence of *Bradyrhizobium* sp. strain LVM 105 has been deposited in GenBank under the accession number QZMV00000000. Raw sequence reads were deposited in the Sequence Read Archive under BioProject number PRJNA492229. The version described in this paper is the first version, QZMV01000000.

REFERENCES

1. Singer SR, Maki SL, Farmer AD, Ilut D, May GD, Cannon SB, Doyle JJ. 2009. Venturing beyond beans and peas: what can we learn from *Chamaecrista*? Plant Physiol 151:1041–1047. https://doi.org/10.1104/pp.109.144774.
2. Doyle JJ, Luckow MA. 2003. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiol 131:900–910. https://doi.org/10.1104/pp.102.018150.
3. Somasegaran P, Hoben HJ. 1994. Handbook for rhizobia. Springer, New York, NY.
4. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
6. Center for Algorithmic Biotechnology. 2016. SPAdes. http://cab.spbu.ru/software/spades/.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
8. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD. 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biol 14:R47. https://doi.org/10.1186/gb-2013-14-5-r47.
9. Vezzi F, Narzisi G, Mishra B. 2012. Feature-by-feature— evaluating *de novo* sequence assembly. PLoS One 7:e31002. https://doi.org/10.1371/journal.pone.0031002.
10. 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.
11. Tsukui T, Eda S, Kaneko T, Sato S, Okazaki S, Kakizaki-Chiba K, Itakura M, Mitsui H, Yamashita A, Terasawa K, Minamisawa K. 2013. The type III secretion system of *Bradyrhizobium japonicum* USDA122 mediates symbiotic incompatibility with *Rj2* soybean plants. Appl Environ Microbiol 79:1048–1051. https://doi.org/10.1128/AEM.03297-12.