AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Draft Genome Sequence of *Rhodococcus* sp. Strain 66b

Louise F. Thatcher, Cindy A. Myers, Cathryn A. O'Sullivan, Margaret M. Roper

CSIRO Agriculture and Food, Centre for Environment and Life Sciences, Wembley, Western Australia, Australia

ABSTRACT We report here the draft genome sequence and annotation of *Rhodococcus* sp. strain 66b isolated from the soil of southwest Western Australia. This strain exhibits a range of bioactivities, including plant growth promotion, biosurfactant production, and wax degradation. Whole-genome sequencing was conducted to uncover the underlying mechanisms.

Received 25 March 2017 **Accepted** 4 April 2017 **Published** 25 May 2017

Citation Thatcher LF, Myers CA, O'Sullivan CA, Roper MM. 2017. Draft genome sequence of *Rhodococcus* sp. strain 66b. Genome Announc 5:e00229-17. https://doi.org/10.1128/genomeA.00229-17.

Address correspondence to Louise F. Thatcher, Louise.Thatcher@csiro.au.

Within the *Actinobacteria* phylum, species from the genus *Rhodococcus* are significant for their bioremediation and industrial applications, for example, the production of enzymes or metabolites involved in degradation of organic compounds or the production of biosurfactants or bioflocculants (1). We isolated and purified *Rhodococcus* sp. strain 66b from naturally water-repellent soils based on its wax-degrading ability and further assessed it for biosurfactant production based on surface tension assays (2). The production of biosurfactants was associated with its potential to degrade waxes that cause water repellency in sandy soils. Sequencing of 16S rRNA designated strain 66b a *Rhodococcus* species (2, 3).

DNA for whole-genome sequencing was extracted from mycelia and coccoid/rod-shaped fragments using a Mo Bio PowerLyzer UltraClean microbial DNA isolation kit, followed by preparation of an indexed Illumina TruSeq library (350-bp insert) and sequencing using 150-bp paired-end reads on an Illumina MiSeq instrument by the Australian Genome Research Facility (AGRF), Melbourne, Australia. A total of 0.48 Gbp of raw data were generated from a sequence run using approximately 1/10 of a sequencing lane. Reads were assessed for quality, trimmed [CutAdapt (4)], and sorted as per Thatcher et al. (5), and overlapping reads were merged using FLASH (version 1.2.11) (6). *De novo* assembly of reads (paired-end, singletons, merged) was performed using SPAdes (version 3.9.0) (7) with the "--careful" option and k-mer lengths of 21, 33, 55, and 77. Contigs less than 1,000 bp were removed. The 66b genome was assembled into 6.68 Mbp (57 scaffolds; N_{50} count, 7 scaffolds; N50 length, 0.38 Mb), with a G+C content of 62.3%. Coding sequences, functional annotation, and secondary metabolite biosynthesis gene clusters were predicted by Prokka (version 1.11) (8) [incorporating Prodigal (version 2.6.3) (9)], Blast2GO (version 1.0.2) (10), and antiSMASH (version 3.0.5.1) (11), respectively.

Blast2GO (10) best BLAST hits analysis for species comparisons revealed the closest neighbor strain for 66b to be *Rhodococcus erythropolis*. Genome annotation by Prokka (8) allowed for the identification of 6,359 coding sequences and 54 tRNAs within the 66b genome. Prediction of secondary metabolite clusters by antiSMASH (11) suggested the 66b genome harbors at least 18 biosynthetic gene clusters, of which 7 are nonribosomal peptide synthetases. Other biosynthetic gene clusters include 2 polyketide synthases and a bacteriocin, ectoine, terpene, and butyrolactone cluster, suggesting its potential to produce metabolites contributing to its bioactivities and potential to improve soil wettability.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. MUZB00000000. The version described in this paper is version MUZB01000000.

ACKNOWLEDGMENTS

This research was supported by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and was undertaken with the assistance of resources from the Australian Genome Research Facility (AGRF), which is supported by the Australian Government. The sequenced strain was obtained from an Australian Grains Research and Development Corporation (GRDC)-funded project to M.M.R. The GRDC had no role in the study design, data collection, or interpretation, or the decision to submit the work for publication.

We thank Ondrej Hlinka and Angela Williams for assistance in running genome assembly and metabolite gene cluster analysis scripts.

REFERENCES

1. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk HP, Clément C, Ouhdouch Y, van Wezel GP. 2016. Taxonomy, physiology, and natural products of *Actinobacteria*. Microbiol Mol Biol Rev 80:1–43. https://doi.org/10.1128/MMBR.00019-15.
2. Roper MM. 2004. The isolation and characterisation of bacteria with the potential to degrade waxes that cause water repellency in sandy soils. Aust J Soil Res 42:427–434. https://doi.org/10.1071/SR03153.
3. Roper MM, Myers CA, Lee J, O'Sullivan CA. 2015. Suppression of *Fusarium* crown rot in wheat by endophytic *Actinobacteria*, abstr 2015. *In* Australasian Plant Pathology Society Conference, Fremantle, Australia, 14 to 16 September 2015.
4. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
5. Thatcher LF, Kamphuis LG, Hane JK, Oñate-Sánchez L, Singh KB. 2015. The Arabidopsis KH-domain RNA-binding protein ESR1 functions in components of jasmonate signalling, unlinking growth restraint and resistance to stress. PLoS One 10:e0126978. https://doi.org/10.1371/journal.pone.0126978.
6. Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507.
7. 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.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
9. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.
10. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. https://doi.org/10.1093/bioinformatics/bti610.
11. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.