



Draft Genome Sequences of Two Glycoalkaloid-Degrading *Arthrobacter* Strains Isolated from Green Potato Peel

Rosanna C. Hennessy,^a Byeonghyeok Park,^b Duleepa Pathiraja,^b In-Geol Choi,^b Mikkel Schultz-Johansen,^a Peter Stougaard^a

^aDepartment of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg C, Denmark

^bDepartment of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, South Korea

ABSTRACT Here, we report the genome sequences of two *Arthrobacter* sp. strains isolated from potato and capable of degrading the toxic potato-derived glycoalkaloids (GAs) α -chaconine and α -solanine. Information from the genome sequences will provide insight into the genetic mechanism of GA degradation by these isolates.

Glycoalkaloids (GAs) are chemical compounds consisting of an alkaloid attached to a carbohydrate moiety (1). The two main GAs produced in potato are α -chaconine and α -solanine. Potato juice is a by-product of starch processing and a potential protein source for human nutrition. To convert starch side streams from feed to food, antinutritional factors, including GAs, must be removed. While fungal pathogens are known to degrade potato-derived GAs, little information on bacterial enzymes is available (2). This announcement reports the draft genome sequences of two *Arthrobacter* strains (S39 and S41) which are capable of degrading α -chaconine and α -solanine.

Previously, the GA-degrading strains *Arthrobacter* sp. S39 and *Arthrobacter* sp. S41 were isolated from green potato peel (2). Genomic DNA was purified from liquid cultures grown at 20°C in lysogenic broth (LB) using the Gentra Puregene Yeast/Bact. kit (Qiagen, Germany) as previously described (3). Library preparation for whole-genome sequencing was performed using the Nextera DNA library prep kit (Illumina, Inc., San Diego, CA). Sequencing was carried out with a MiSeq platform using the MiSeq reagent kit v3 (600 cycles) (Illumina, Inc.). A total of 979,725 paired-end reads (589.8 Mbp) with an average read size of 301 bp and an average insert size of 525 bp were retrieved for *Arthrobacter* sp. S39, and a total of 1,376,771 paired-end reads (828.8 Mbp) with an average read size of 301 bp and an average insert size of 407 bp were obtained for *Arthrobacter* sp. S41. Adapter sequences and low-quality bases (<Q20) were trimmed from raw reads using TrimGalore! software v0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). A total of 962,128 reads (405.8 Mbp, 79.3 \times coverage) for *Arthrobacter* sp. S39 and 1,363,873 reads (596.9 Mbp, 171.9 \times coverage) for *Arthrobacter* sp. S41 passed quality control and were used for assembly using the SPAdes genome assembler v3.7.1 with default settings (4). Genome annotation was conducted using the Prokaryotic Contig Annotation Pipeline Server (P-CAPS) (5) and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (6). The annotation results of P-CAPS were not incorporated into the results of PGAP. P-CAPS was used to inspect the quality of the assembly prior to submission to NCBI.

The draft genome sequence of *Arthrobacter* sp. S39 is 5,118,715 bp long, and its contigs are contained in 48 scaffolds (L_{50} , 5; N_{50} , 391,678 bp) with a mean G+C content of 65.4%. The draft genome sequence of *Arthrobacter* sp. S41 is 3,472,078 bp long, and its contigs are set up in 7 scaffolds (L_{50} , 1; N_{50} , 2,465,108 bp) with a mean G+C content of 55.7%. The annotation of the *Arthrobacter* sp. S39 genome revealed the presence of 4,587 predicted coding sequences (CDS) and 72 predicted RNA coding sequences (18

Citation Hennessy RC, Park B, Pathiraja D, Choi I-G, Schultz-Johansen M, Stougaard P. 2019. Draft genome sequences of two glycoalkaloid-degrading *Arthrobacter* strains isolated from green potato peel. *Microbiol Resour Anounc* 8:e00226-19. <https://doi.org/10.1128/MRA.00226-19>.

Editor Christina Cuomo, Broad Institute of MIT and Harvard University

Copyright © 2019 Hennessy 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 Rosanna C. Hennessy, hennessy@plen.ku.dk.

Received 5 March 2019

Accepted 21 March 2019

Published 18 April 2019

rRNAs, 51 tRNAs, 3 noncoding RNAs [ncRNAs]), compared to 3,119 CDS and 72 RNA genes (4 rRNAs, 65 tRNAs, 3 ncRNAs) for *Arthrobacter* sp. S41.

Carbohydrate-active enzymes (CAZymes) were identified using a local version of dbCAN v6.0 (E value cutoff, $1.0e-5$) (7). A total of 283 predicted CAZy domains were identified, including 141 glycoside hydrolases (GHs) in *Arthrobacter* sp. S39 and 39 GHs in *Arthrobacter* sp. S41. The genomes of both strains encode GHs organized in gene clusters containing α -L-rhamnosidases, β -D-glucosidases, and β -D-galactosidases, which are candidate enzymes for GA degradation with potential application in the removal of toxic GAs present in potato side streams for conversion into protein suitable for human nutrition.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers [SIHX00000000](https://doi.org/10.1093/mpmi/09-16-0186-R) (*Arthrobacter* sp. S39) and [SIHY00000000](https://doi.org/10.1093/mpmi/09-16-0186-R) (*Arthrobacter* sp. S41). The versions described in this paper are SIHX01000000 and SIHY01000000, respectively. Raw sequence reads are available under SRA accession numbers [SRX5464543](https://doi.org/10.1093/mpmi/09-16-0186-R) (*Arthrobacter* sp. S39) and [SRX5464544](https://doi.org/10.1093/mpmi/09-16-0186-R) (*Arthrobacter* sp. S41).

ACKNOWLEDGMENT

This work was supported by Innovation Fund Denmark grant 5158-00001A (pro-Potato: Potato Proteins—Challenges and Industrial Possibilities).

REFERENCES

- Dahlin P, Muller MC, Ekengren S, McKee LS, Bulone V. 2017. The impact of steroidal glycoalkaloids on the physiology of *Phytophthora infestans*, the causative agent of potato late blight. *Mol Plant Microbe Interact* 30: 531–542. <https://doi.org/10.1094/MPMI-09-16-0186-R>.
- Hennessy RC, Jørgensen NOG, Scavenius C, Enghild JJ, Greve-Poulsen M, Sørensen OB, Stougaard P. 2018. A screening method for the isolation of bacteria capable of degrading toxic steroidal glycoalkaloids present in potato. *Front Microbiol* 9:2648. <https://doi.org/10.3389/fmicb.2018.02648>.
- Dichmann SI, Park B, Pathiraja D, Choi I-G, Stougaard P, Hennessy RC. 2018. Draft genome of a novel *Serratia* sp. strain with antifungal activity. *Microbiol Resour Announc* 7:e01340-18. <https://doi.org/10.1128/MRA.01340-18>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
- Park B, Baek M-J, Min B, Choi I-G. 2017. Prokaryotic Contig Annotation Pipeline Server: Web application for a prokaryotic genome annotation pipeline based on the Shiny app package. *J Comput Biol* 24:917–922. <https://doi.org/10.1089/cmb.2017.0066>.
- 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>.
- Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46:W95–W101. <https://doi.org/10.1093/nar/gky418>.