



Draft Genome Sequence of *Pantoea agglomerans* CPHN2, a Potential Plant-Growth-Promoting Endophyte

Pradeep Kumar,^a Varsha Chauhan,^a Amita Suneja Dang,^b Ajit Kumar,^c  Pooja Suneja^a

^aDepartment of Microbiology, M.D. University, Rohtak, India

^bCentre for Medical Biotechnology, M.D. University, Rohtak, India

^cCentre for Bioinformatics, M.D. University, Rohtak, India

ABSTRACT Here, we report the draft genome sequence of *Pantoea agglomerans* CPHN2, an endophyte isolated from nodules of *Cicer arietinum* (Chickpea) from Hisar, Haryana, India. The genome was 4,839,532 bp and exhibited a GC content of 55.2% and 4,508 genes with 4,468 coding sequences, 1 rRNA, 71 tRNAs, and 1 CRISPR.

Members of *Pantoea*, a genus of Gram-negative bacteria, have been isolated from diverse habitats and been found to exhibit an association with a variety of hosts, such as plants, animals, and humans. Many species of this genus are associated with plants, either on their surfaces or as endophytes. *Pantoea agglomerans*, one of the most common species associated with plants, is able to fix atmospheric nitrogen and possess multiple plant-growth-promoting (multi-PGP) traits. Numerous studies have reported the endophytic occurrence of *P. agglomerans* in various crops, such as rice (1), chickpea (2), sugarcane (3), sweet potato (4), and wheat (5).

P. agglomerans CPHN2, a nonrhizobial endophyte, was isolated from nodules of *Cicer arietinum* (chickpea) from Hisar, Haryana, India (29.25°N, 76.05°E), with a motive to access its ability to promote plant growth. Healthy nodules of *C. arietinum* were detached from the plants and surface sterilized (6). The uncrushed sterilized nodules were placed on tryptone soya agar (TSA) plates as a control to check the efficacy of surface sterilization. Macerated samples were streaked onto TSA plates and examined for the growth of endophytic bacteria after an incubation ($28 \pm 2^\circ\text{C}$) of 3 days. The colonies were selected and purified. The genomic DNA was isolated using a modified cetyltrimethylammonium bromide (CTAB) protocol (7) and amplified using 16S rRNA universal primers (1541R and 8F) (8). The amplicon was sequenced, a similarity analysis was performed at the BLAST server, and the sequence was submitted to GenBank (MH298522). The isolate was observed to possess plant-growth-promoting traits, such as the solubilization of phosphate and the production of ammonia, siderophore, and a large amount of indole-3-acetic acid (9). The presence of these prominent plant-growth-promoting traits encouraged us to go for whole-genome sequencing of this strain which further confirmed it as *Pantoea agglomerans*.

A paired-end (PE) sequencing library was prepared from the DNA sample using the Illumina TruSeq nano DNA library prep kit. The PE Illumina library was loaded into the NextSeq 500 instrument for cluster generation and sequencing. Paired-end sequencing on the NextSeq 500 instrument helps in reading template fragments in both forward and backward orientations. Whole-genome sequencing (Illumina NextSeq 500, 2×150 bp) yielded 2,898,434,450 reads representing $\sim 100\times$ genome coverage. Fastp v0.201 (10), Shovill v1.0.4 (11), and Prokka v1.14.5 (12) tools of the Galaxy server were used for genome analysis and annotation. Fastp v0.201 was used to filter and trim the final assembly, and the genome sequence was assembled with Shovill v1.0.4. Default parameters were used for the software. After the filtering and trimming steps, a total of 9,731,611 paired-end reads were obtained. Shovill combined the reads into 32 contigs, resulting in a total assembly size of 4,839,532 bp and

Editor David A. Baltrus, University of Arizona

Copyright © 2022 Kumar 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 Ajit Kumar, akumar.cbt.mdu@gmail.com, or Pooja Suneja, poojapavit@gmail.com.

The authors declare no conflict of interest.

Received 14 April 2022

Accepted 5 July 2022

Published 21 July 2022

an N_{50} value of 558,390 bp. Prokka v1.14.5 was used for the annotation, yielding 4,468 coding sequences. The genome was also observed to carry 1 rRNA gene, 71 tRNA genes, and 1 CRISPR gene. The GC content of the genome was found to be 55.2%. The genome was annotated automatically using Galaxy (<https://usegalaxy.org/>) for a further comparative analysis and genome annotation and RAST (<https://rast.nmpdr.org/>) for a better graphical analysis (13).

The reported draft genome sequence of *P. agglomerans* CPHN2 will help us improve strain-specific fingerprinting, which is essential for the registration and protection of commercial bioinoculant products. Also, this draft genome sequence of *P. agglomerans* CPHN2 will provide a platform for elucidating various plant-growth-promoting mechanisms and their regulation, which can be used to further improve the practical and commercial value of the strain by optimization of formulation technology, performance, and efficacy of the *P. agglomerans* CPHN2 strain.

Data availability. The *P. agglomerans* CPHN2 whole-genome sequence has been deposited at GenBank under the accession no. [CP098412](#) and [CP098413](#) for plasmids and [CP098414](#) for the chromosome and annotated by the prokaryotic genome annotation pipeline (PGAP) (14). The BioProject accession no. is [PRJNA811747](#) (Sequence Read Archive accession no. [SRR18189611](#)).

ACKNOWLEDGMENTS

P.K. acknowledges the Council of Scientific & Industrial Research, New Delhi, India, for a CSIR fellowship. We also acknowledge the Department of Science and Technology, Govt. of India for providing the FIST grant (grant no. 1196 SR/FST/LS-I/2017/4) to the Department of Microbiology, Maharshi Dayanand University, Rohtak, Haryana, India.

REFERENCES

- Feng Y, Shen D, Song W. 2006. Rice endophyte *Pantoea agglomerans* YS19 promotes host plant growth and affects allocations of host photosynthates. *J Appl Microbiol* 100:938–945. <https://doi.org/10.1111/j.1365-2672.2006.02843.x>.
- Mishra A, Chauhan PS, Chaudhry V, Tripathi M, Nautiyal CS. 2011. Rhizosphere competent *Pantoea agglomerans* enhances maize (*Zea mays*) and chickpea (*Cicer arietinum* L.) growth, without altering the rhizosphere functional diversity. *Antonie Van Leeuwenhoek* 100:405–413. <https://doi.org/10.1007/s10482-011-9596-8>.
- Quecine MC, Araújo WL, Rossetto PB, Ferreira A, Tsui S, Lacava PT, Mondin M, de Azevedo JL, Pizzirani K, Aparecida A. 2012. Sugarcane growth promotion by the endophytic bacterium *Pantoea agglomerans* 33.1. *Appl Environ Microbiol* 78:7511–7518. <https://doi.org/10.1128/AEM.00836-12>.
- Asis CA, Adachi K. 2004. Isolation of endophytic diazotroph *Pantoea agglomerans* and nondiazotroph *Enterobacter asburiae* from sweetpotato stem in Japan. *Lett Appl Microbiol* 38:19–23. <https://doi.org/10.1046/j.1472-765x.2003.01434.x>.
- Cherif-Silini H, Thissera B, Bouket AC, Saadaoui N, Silini A, Eshelli M, Alenezi FN, Vallat A, Luptakova L, Yahiaoui B, Cherrad S, Vacher S, Rateb ME, Belbahri L. 2019. Durum wheat stress tolerance induced by endophyte *Pantoea agglomerans* with genes contributing to plant functions and secondary metabolite arsenal. *Int J Mol Sci* 20:3989. <https://doi.org/10.3390/ijms20163989>.
- Vincent JM. 1970. A manual for the practical study of the root-nodule bacteria. *In* A manual for the practical study of the root-nodule bacteria. Blackwell Publishing Ltd, Oxford, United Kingdom.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. 1991. Current protocols in molecular biology, vol 1. Wiley Interscience, New York, NY.
- Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703. <https://doi.org/10.1128/jb.173.2.697-703.1991>.
- Maheshwari R, Bhutani N, Bhardwaj A, Suneja P. 2019. Functional diversity of cultivable endophytes from *Cicer arietinum* and *Pisum sativum*: bio-prospecting their plant growth potential. *Biocatal Agric Biotechnol* 20: 101229. <https://doi.org/10.1016/j.bcab.2019.101229>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Seemann T. 2017. Shovill: faster SPAdes assembly of Illumina reads (v0. 9.0). <https://github.com/tseemann/shovill>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Chauhan V, Shrivastava K, Anand S, Kumar C, Singh A, Varma-Basil M. 2020. Draft genome sequence of *Mycobacterium simiae*, a potential pathogen isolated from the normal human oral cavity. *Microbiol Resour Announc* 9: e01185-20. <https://doi.org/10.1128/MRA.01185-20>.
- 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>.