

# Complete genome sequence of Curtobacterium sp. strain SGAir0471, isolated from Singapore air samples

Hlaing, Phu Pwint Thin; Junqueira, Ana Carolina M.; Uchida, Akira; Purbojati, Rikky Wenang; Wong, Anthony; Kushwaha, Kavita K.; Putra, Alexander; Kee, Carmon; Gaultier, Nicolas Eugene; Premkrishnan, Balakrishnan N. V.; Heinle, Cassie Elizabeth; Lim, Serene Boon Yuean; Vettath, Vineeth Kodengil; Drautz-Moses, Daniela Isabel; Schuster, Stephan Christoph

2019

Hlaing, P. P. T., Junqueira, A. C. M., Uchida, A., Purbojati, R. W., Wong, A., Kushwaha, K. K., . . . Schuster, S. C. (2019). Complete genome sequence of Curtobacterium sp. strain SGAir0471, isolated from Singapore air samples. *Microbiology Resource Announcements*, 8(47), e00960-19-. doi:10.1128/MRA.00960-19

<https://hdl.handle.net/10356/142213>

<https://doi.org/10.1128/MRA.00960-19>



---

© 2019 Hlaing et al. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

*Downloaded on 28 Jan 2021 02:15:24 SGT*



# Complete Genome Sequence of *Curtobacterium* sp. Strain SGAir0471, Isolated from Singapore Air Samples

Phu Pwint Thin Hlaing,<sup>a</sup>  Ana Carolina M. Junqueira,<sup>b</sup> Akira Uchida,<sup>a</sup>  Rikky W. Purbojati,<sup>a</sup> Anthony Wong,<sup>a</sup> Kavita K. Kushwaha,<sup>a</sup> Alexander Putra,<sup>a</sup> Carmon Kee,<sup>a</sup> Nicolas E. Gaultier,<sup>a</sup> Balakrishnan N. V. Premkrishnan,<sup>a</sup> Cassie E. Heinle,<sup>a</sup> Serene B. Y. Lim,<sup>a</sup> Vineeth Kodengil Vettath,<sup>a</sup> Daniela I. Drautz-Moses,<sup>a</sup> Stephan C. Schuster<sup>a</sup>

<sup>a</sup>Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore

<sup>b</sup>Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

**ABSTRACT** *Curtobacterium* sp. strain SGAir0471 was isolated from tropical air samples collected in Singapore. The genome was assembled using PacBio RS II long reads and Illumina MiSeq short paired-end reads. The complete genome measures 3.53 Mb and consists of 3,151 protein-coding genes, 49 tRNAs, and 12 rRNAs.

The genus *Curtobacterium* contains rod-shaped Gram-positive bacteria and was established as a new genus in 1972 in a taxonomic study carried out on coryneform bacteria (1). *Curtobacterium* spp. mainly inhabit plants and often soil (2, 3). Some *Curtobacterium* spp. are well-established plant pathogens (4) and have also been isolated from clinical samples from patients with respiratory distress, soft tissue infection, conjunctivitis, and lymphadenopathy (5). Thus, a complete genome of a *Curtobacterium* sp. may help us to understand the possible mechanisms used by this genus to cause infections.

*Curtobacterium* sp. strain SGAir0471 was isolated from outdoor air in Singapore (1.2233N, 103.84472E) using an Andersen single-stage impactor (SKC, Inc., USA). Airborne particles were impacted onto potato dextrose agar (Sigma-Aldrich, USA), and further isolation was done on Trypticase soy agar (Becton, Dickinson, USA) at 30°C until an axenic culture was obtained. Genomic DNA was purified using the Wizard Genomic DNA purification kit (Promega, USA) following the manufacturer's protocol. The sequencing library was prepared with the SMRTbell template prep kit 1.0 (Pacific Biosciences, USA), followed by single-molecule real-time (SMRT) sequencing conducted on the PacBio RS II platform. Additionally, short reads were generated on a MiSeq (Illumina, USA) 300-bp paired-end run using whole-genome shotgun libraries constructed with the TruSeq Nano DNA library preparation kit (Illumina, USA).

Default parameters were used for all software unless otherwise specified. A total of 44,706 subreads were used for *de novo* assembly with the Hierarchical Genome Assembly Process (HGAP) version 3 included in the PacBio SMRT Analysis version 2.3.0 package (6). Quality control of long reads was performed using the PreAssembler filter version 1 protocol from HGAP, and Cutadapt version 1.8.1 (7) was used to remove adapter sequences for short reads. Further improvements to genome quality were achieved by polishing the draft assembly via Quiver (6) and correcting errors using Pilon version 1.16 (8) using 706,319 MiSeq paired-end reads. The consensus assembly generated one chromosomal contig of 3,531,991 bp (124.2-fold coverage) with an average G+C content of 71.7%.

Genome annotation was done using NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) version 4.4 (9). The complete genome was predicted to contain a total of 3,294 genes, including 3,151 protein-coding genes (PCGs), 12 ribosomal subunits (5S, 16S, 23S), 49 tRNAs, 3 noncoding RNAs, and 79 pseudogenes.

**Citation** Hlaing PPT, Junqueira ACM, Uchida A, Purbojati RW, Wong A, Kushwaha KK, Putra A, Kee C, Gaultier NE, Premkrishnan BNV, Heinle CE, Lim SBY, Vettath VK, Drautz-Moses DI, Schuster SC. 2019. Complete genome sequence of *Curtobacterium* sp. strain SGAir0471, isolated from Singapore air samples. *Microbiol Resour Annot* 8:e00960-19. <https://doi.org/10.1128/MRA.00960-19>.

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

**Copyright** © 2019 Hlaing 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 Stephan C. Schuster, [SCSchuster@ntu.edu.sg](mailto:SCSchuster@ntu.edu.sg).

P.P.T.H. and A.C.M.J. contributed equally to this work.

**Received** 2 September 2019

**Accepted** 28 October 2019

**Published** 21 November 2019

Taxonomical identification was performed using average nucleotide identity (ANI) utilizing the Microbial Species Identifier (MiSI) method (10), which was run using ANICalculator with default parameters against a database of 6,387 bacterial RefSeq genomes created using a text filter for “type, synonym type, and proxytype” and subsequently using the getorf program with the -find 3 option. This revealed 78.4% identity with *Curtobacterium ammoniigenes* NBRC 101786 with an alignment fraction value of 0.22. The 16S rRNA analysis using Barrnap version 0.7 (11) and BLASTn (12) was run against the SILVA database (13) and resulted in a 99.7% identity with *Curtobacterium* sp. strain B18. As the ANI result is below the threshold for species-level identification, the *Curtobacterium* genus was assigned to the isolate based on the combined ANI and 16S rRNA sequence similarity results.

**Data availability.** The genome sequence of *Curtobacterium* sp. strain SGAir0471 has been deposited in DDBJ/EMBL/GenBank under the accession number [CP027869](https://doi.org/10.1093/nar/gkw569). The SRA database accession numbers are [SRR8894900](https://doi.org/10.1093/nar/gkw569) and [SRR8894901](https://doi.org/10.1093/nar/gkw569).

## ACKNOWLEDGMENTS

This work was supported by Singapore Ministry of Education Academic Research Fund tier 3 grant MOE2013-T3-1-013.

We thank Anjali Bansal Gupta for the constructive review of the manuscript.

## REFERENCES

1. Yamada K, Komagata K. 1972. Taxonomic studies on coryneform bacteria. V. Classification of coryneform bacteria. *J Gen Appl Microbiol* 18: 417–431. <https://doi.org/10.2323/jgam.18.417>.
2. Francis MJ, Doherty RR, Patel M, Hamblin JF, Ojaimi S, Korman TM. 2011. *Curtobacterium flaccumfaciens* septic arthritis following puncture with a Coxspur Hawthorn thorn. *J Clin Microbiol* 49:2759–2760. <https://doi.org/10.1128/JCM.00340-11>.
3. Dunleavy JM. 1989. *Curtobacterium plantarum* sp. nov. Is ubiquitous in plant leaves and is seed transmitted in soybean and corn. *Int J Syst Evol Microbiol* 39:240–249. <https://doi.org/10.1099/00207713-39-3-240>.
4. Chase AB, Arevalo P, Polz MF, Berlemont R, Martiny J. 2016. Evidence for ecological flexibility in the cosmopolitan genus *Curtobacterium*. *Front Microbiol* 7:1874. <https://doi.org/10.3389/fmicb.2016.01874>.
5. Funke G, Aravena-Roman M, Frodl R. 2005. First description of *Curtobacterium* spp. isolated from human clinical specimens. *J Clin Microbiol* 43:1032. <https://doi.org/10.1128/JCM.43.3.1032-1036.2005>.
6. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Turner SW, Korfach J, Clum A, Copeland A, Huddleston J, Eichler EE. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
7. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10. <https://doi.org/10.14806/ej.17.1.200>.
8. 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>.
9. Tatusova T, DiCuccio M, Badretdin A, Chetverin V, Nawrocki EP, Zaslavsky L, Pruitt KD, Ostell J, Lomsadze A, Borodovsky M. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
10. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpidis NC, Pati A. 2015. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 43:6761–6771. <https://doi.org/10.1093/nar/gkv657>.
11. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
12. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
13. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and Web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.