

Complete Genome Sequence of the Exopolysaccharide-Producing *Burkholderia caribensis* Type Strain MWAP64

Yanling Pan, Ka Fai Kong, Jimmy S. H. Tsang

Molecular Microbiology Laboratory, School of Biological Sciences, University of Hong Kong, Hong Kong Special Administrative Region, China

We report the complete genome sequence of *Burkholderia caribensis* MWAP64 (LMG 18531), which was isolated from soil for its proficiency in producing large amounts of exopolysaccharide that help form microaggregates in a vertisol. There are four replicons with a total size of 9,032,119 bp.

Received 26 November 2015 Accepted 7 December 2015 Published 28 January 2016

Citation Pan Y, Kong KF, Tsang JSH. 2016. Complete genome sequence of the exopolysaccharide-producing *Burkholderia caribensis* type strain MWAP64. *Genome Announc* 4(1):e01636-15. doi:10.1128/genomeA.01636-15.

Copyright © 2016 Pan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jimmy S. H. Tsang, jshtsang@hku.hk.

Burkholderia caribensis MWAP64 (LMG 18531) was isolated from a fraction of a vertisol in the southeast of the island of Martinique in the French West Indies (1). It is a motile, Gram-negative, rod-shaped bacterium that is pleomorphic in actively growing cultures. It produces large amounts of exopolysaccharide in sugar-enriched agar media and forms mucoid colonies. It is the type strain of this species. Here, we describe the complete genome sequence of this bacterium.

Previous studies on a haloacid-degrading bacterium strain MBA4 using DNA-DNA hybridization showed that it has a DNA homology value of 74% with that of *Burkholderia caribensis* MWAP64 (2). The complete genome of MBA4 was recently determined and deposited in GenBank as CP012746, CP012747, and CP012748 (3). For strain MWAP64, a whole-genome shotgun approach was used with the Roche 454 GS FLX Titanium and PacBio RS II systems. The raw sequencing reads of MWAP64 obtained from the 454 sequencing platform were filtered and analyzed with CLC Genomic Workbench version 6.5 (CLC bio, Aarhus, Denmark). There were 138,033 clean reads with an average length of 355 bp. For the PacBio long-read sequencing system, SMRT Analysis version 2.3.0 and HGAP version 2 were used for filtering, trimming, and assembly. There were 73,629 clean reads with an average length of 11,521 bp. The depth of coverage is around 100-fold. The genome of *B. caribensis* MBA4 was used as a reference sequence. More than 99.94% of the 48,990,523 bp from the 454 clean reads were mapped to the assembled contigs. Four circular replicons, two chromosomes, and two plasmids were assembled. The sizes of these four replicons are, respectively, 3,688,127; 2,890,272; 2,011,268; and 442,452 bp. The total size of the genome is 9,032,119 bp, with a tabulated GC content of 62.58%.

The completed genome was annotated automatically with the Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) from NCBI (4) and the Rapid Annotations using Subsystems Technology (RAST) server (5). The complete genome contains 7,911 predicted genes, including 18 rRNAs, 62 tRNAs, and 1 ncRNA. Apart from 157 pseudogenes, the genome contains 7,673 protein-coding genes. Among these protein-coding genes, 5,768 were assigned with functions, while 1,905 were annotated as hy-

pothetical. Furthermore, 634 tandem repeats were identified by Tandem Repeats Finder (6). The online CRISPR finder (7) predicted the presence of 7 questionable CRISPR regions in the genome. The online prophage search tool PHAST (8) identified one intact, one incomplete, and one questionable prophage region in the four replicons. The metagenomic analysis server WebMGA (9) assigned 5,946 genes into 23 clusters of orthologous groups (COG) categories. Pathway Tools version 19.5 predicted 283 pathways, including 1,647 enzymatic reactions and 17 transport reactions (10).

Nucleotide sequence accession numbers. The genome sequence of *Burkholderia caribensis* MWAP64 has been deposited in GenBank under the accession numbers [CP013102](https://ncbi.nlm.nih.gov/nuccore/CP013102), [CP013103](https://ncbi.nlm.nih.gov/nuccore/CP013103), [CP013104](https://ncbi.nlm.nih.gov/nuccore/CP013104), and [CP013105](https://ncbi.nlm.nih.gov/nuccore/CP013105).

ACKNOWLEDGMENTS

We thank A. S. Chan for advice and the University Centre for Genomic Sciences for 454 sequencing.

This work has been supported by grants from the General Research Fund (project number HKU 780511 and 17120514) of the Research Grants Council of the Hong Kong Special Administrative Region, China.

FUNDING INFORMATION

Research Grants Council, University Grants Committee (RGC, UGC) provided funding to Jimmy S. H. Tsang under grant numbers HKU780511 and 17120514.

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

- Achouak W, Christen R, Barakat M, Martel MH, Heulin T. 1999. *Burkholderia caribensis* sp. nov., an exopolysaccharide-producing bacterium isolated from vertisol microaggregates in Martinique. *Int J Syst Bacteriol* 49:787–794. <http://dx.doi.org/10.1099/00207713-49-2-787>.
- Chan YP. 2005. Taxonomic analysis of a haloacid degrading *Burkholderia* species MBA4. M.Ph. thesis. University of Hong Kong, Hong Kong SAR, China.
- Pan Y, Kong KF, Tsang JSH. 2014. Draft genome sequence of the haloacid-degrading *Burkholderia caribensis* strain mba4. *Genome Announc* 2(1):e00047-14. <http://dx.doi.org/10.1128/genomeA.00047-14>.

4. Angiuoli S, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira C, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures. (SOPs) for (meta)genomic annotation. *OMICS* 12:137–141. <http://dx.doi.org/10.1089/omi.2008.0017>.
5. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
6. Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580. <http://dx.doi.org/10.1093/nar/27.2.573>.
7. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <http://dx.doi.org/10.1093/nar/gkm360>.
8. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <http://dx.doi.org/10.1093/nar/gkr485>.
9. Wu S, Zhu Z, Fu L, Niu B, Li W. 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 12:444. <http://dx.doi.org/10.1186/1471-2164-12-444>.
10. Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, Kothari A, Weaver D, Lee T, Subhraveti P, Spaulding A, Fulcher C, Keseler IM, Caspi R. 2015. Pathway tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief Bioinform* [Epub ahead of print.] <http://dx.doi.org/10.1093/bib/bbv079>.