



# Complete Genome Sequence of *Burkholderia cenocepacia* K56-2, an Opportunistic Pathogen

 Inmaculada García-Romero,<sup>a</sup>  Miguel A. Valvano<sup>a</sup>

<sup>a</sup>Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, United Kingdom

**ABSTRACT** *Burkholderia cenocepacia* K56-2, an opportunistic bacterium for people with cystic fibrosis (CF), belongs to the *Burkholderia cepacia* complex (Bcc) and is consistently used as a model pathogen. We describe here the closed genome sequence for this strain, which will help advance research in *B. cenocepacia* biology and omics studies.

The *Burkholderia cepacia* complex (Bcc) comprises a set of *Burkholderia* species that cause respiratory infection in people with cystic fibrosis (CF) (1–3). These bacteria are multidrug-opportunistic intracellular pathogens (reviewed in reference 1). *B. cenocepacia*, in particular, can be transmissible (4) and causes a fatal condition known as cepacia syndrome (5, 6). In this whole-genome sequencing (WGS) project, we have closed the genome sequence of *B. cenocepacia* K56-2, which belongs to the ET-12 lineage and was isolated from a CF patient in Canada (2, 3). The strain used in this study was isolated as described in reference 3.

To date, no complete genome sequence for K56-2 has been deposited into GenBank; a draft genome containing 19 contigs was published in 2013 (7). In this work, the genome sequence was assembled *de novo*.

Strain K56-2 was grown in 25 ml LB (Melford) at 37°C and 180 rpm to an optical density (600 nm) of 0.6. Pelleted bacteria (4,000 × *g* for 10 min) were resuspended in a cryopreservative (Microbank; Pro-Lab Diagnostics UK, United Kingdom) and shipped to MicrobesNG (Birmingham, United Kingdom). Two different sequencing approaches were used, Illumina HiSeq technology (a 250-bp paired-end protocol) and a GridION system using a FLO-MIN-106 (R.9.4) flow cell (ONT, United Kingdom). For the former, genomic DNA was purified using a QIAamp DNA minikit (Qiagen, Germany). For the latter, high-molecular-weight genomic DNA was extracted using a Nanobind CCB Big DNA kit (Circulomics, MD, USA). Libraries for Illumina sequencing were prepared using the Nextera XT library prep kit (Illumina, San Diego, USA) according to the manufacturer's protocol. Long-read genomic DNA libraries for GridION sequencing were prepared with the Oxford Nanopore SQK-RBK004 kit (ONT) using 500 ng of high-molecular-weight DNA and in accordance with MicrobesNG protocols ([https://microbesng.com/documents/5/MicrobesNG\\_Methods\\_Document\\_-\\_PDF.pdf](https://microbesng.com/documents/5/MicrobesNG_Methods_Document_-_PDF.pdf)). For Nanopore reads, base calling was performed using Guppy v3.0.6+9999d81 (8), generating 29,635 long reads with an  $N_{50}$  value of 7,943 bp. Illumina reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (9), resulting in 1,734,906 reads of 250 bp. Illumina and Nanopore reads were coassembled using Unicycler v0.4.0 (10), obtaining a final assembly of 4 replicons (with a genome coverage of 107.62×). Default parameters were used for all software unless otherwise specified.

The complete genome sequence of strain K56-2 consists of 3 circular chromosomes of 3,673,077 bp (GC content of 66.89%), 3,211,025 bp (GC content of 67.29%), and 765,157 bp (GC content of 66.92%) and a plasmid of 92,661 bp (GC content of 62.76%). The smaller replicon was defined as a plasmid due to the absence of rRNAs and tRNAs and

**Citation** García-Romero I, Valvano MA. 2020. Complete genome sequence of *Burkholderia cenocepacia* K56-2, an opportunistic pathogen. Microbiol Resour Announc 9:e01015-20. <https://doi.org/10.1128/MRA.01015-20>.

**Editor** J. Cameron Thrash, University of Southern California

**Copyright** © 2020 García-Romero and Valvano. 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 Miguel A. Valvano, [m.valvano@qub.ac.uk](mailto:m.valvano@qub.ac.uk).

**Received** 2 September 2020

**Accepted** 8 October 2020

**Published** 22 October 2020

**TABLE 1** Main features of K56-2 genome annotation

Replicon	Locus tag	AN <sup>a</sup>	No. of:				
			Genes	Proteins	rRNAs	tRNAs	ncRNAs
Chr 1	K562_1	CP053300	3,453	3,337	12	67 <sup>b</sup>	37
Chr 2	K562_2	CP053301	2,902	2,880	3	11	8
Chr 3	K562_3	CP053302	660	651	3	4	2
Plasmid	K562_4	CP053303	106	105			1
Total			7,121	6,973	18	82	48

<sup>a</sup>GenBank accession number.

<sup>b</sup>Including a transfer-messenger RNA (tmRNA).

the high similarity (99.9% identity) with pBCJ2315 (NC\_011003). All the replicons were manually rotated to match those in the *B. cenocepacia* J2315 genome (11). Functional annotation was performed using Prokka 1.12 (12), with a cutoff E value of 1e−05 for coding DNA sequence (CDS) annotation plus adding the noncoding RNA (ncRNA) annotation step using Infernal software v1.1.3 (13) and the Rfam database (14).

According to the automatic annotation, the K56-2 genome comprises 7,121 genes, with 6,973 protein-encoding genes, 82 tRNAs, 18 rRNAs, and 48 noncoding RNAs (Table 1).

**Data availability.** The genome sequence of strain K56-2 was deposited in GenBank under the accession numbers indicated in Table 1. The raw sequence reads were deposited in the SRA under the accession numbers SRR11609092 (GridION) and SRR11609093 (Illumina HiSeq 2500). The associated BioProject and BioSample accession numbers are PRJNA627986 and SAMN14693155, respectively.

## ACKNOWLEDGMENTS

This work was funded by grants from the Medical Research Council (grant MR/P022480/1) and the Biotechnology and Biological Sciences Research Council (BBSRC) (grants BB/S006281/1 and BB/R022607/1) to M.A.V. Genome sequencing was provided by MicrobesNG (Birmingham, UK).

## REFERENCES

- Valvano MA. 2015. Intracellular survival of *Burkholderia cepacia* complex in phagocytic cells. *Can J Microbiol* 61:607–615. <https://doi.org/10.1139/cjm-2015-0316>.
- Mahenthalingam E, Coenye T, Chung JW, Speert DP, Govan JR, Taylor P, Vandamme P. 2000. Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia* complex. *J Clin Microbiol* 38:910–913. <https://doi.org/10.1128/JCM.38.2.910-913.2000>.
- Darling P, Chan M, Cox AD, Sokol PA. 1998. Siderophore production by cystic fibrosis isolates of *Burkholderia cepacia*. *Infect Immun* 66:874–877. <https://doi.org/10.1128/IAI.66.2.874-877.1998>.
- Blanchard AC, Tang L, Tadros M, Muller M, Spilker T, Waters VJ, LiPuma JJ, Tullis E. 2020. *Burkholderia cenocepacia* ET12 transmission in adults with cystic fibrosis. *Thorax* 75:88–90. <https://doi.org/10.1136/thoraxjnl-2019-214098>.
- Isles A, Macluskay I, Corey M, Gold R, Prober C, Fleming P, Levison H. 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *J Pediatr* 104:206–210. [https://doi.org/10.1016/S0022-3476\(84\)80993-2](https://doi.org/10.1016/S0022-3476(84)80993-2).
- Kopp BT, Abdulrahman BA, Khweek AA, Kumar SB, Akhter A, Montione R, Tazi MF, Caution K, McCoy K, Amer AO. 2012. Exaggerated inflammatory responses mediated by *Burkholderia cenocepacia* in human macrophages derived from cystic fibrosis patients. *Biochem Biophys Res Commun* 424:221–227. <https://doi.org/10.1016/j.bbrc.2012.06.066>.
- Varga JJ, Losada L, Zelazny AM, Kim M, McCarrison J, Brinkac L, Sampaio EP, Greenberg DE, Singh I, Heiner C, Ashby M, Nierman WC, Holland SM, Goldberg JB. 2013. Draft genome sequences of *Burkholderia cenocepacia* ET12 lineage strains K56-2 and BC7. *Genome Announc* 1:e00841-13. <https://doi.org/10.1128/genomeA.00841-13>.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Holden M, Seth-Smith H, Crossman L, Sebahia M, Bentley S, Cerdeno-Tarraga A, Thomson N, Bason N, Quail M, Sharp S, Cherevach I, Churcher C, Goodhead I, Hauser H, Holroyd N, Mungall K, Scott P, Walker D, White B, Rose H, Iversen P, Mil-Homens D, Rocha E, Fialho A, Baldwin A, Dowson C, Barrell B, Govan J, Vandamme P, Hart C, Mahenthalingam E, Parkhill J. 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *J Bacteriol* 191:261–277. <https://doi.org/10.1128/JB.01230-08>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Barquist L, Burge SW, Gardner PP. 2016. Studying RNA homology and conservation with infernal: from single sequences to RNA families. *Curr Protoc Bioinformatics* 54:12.13.1–12.13.25. <https://doi.org/10.1002/cpbi.4>.
- Kalvari I, Nawrocki EP, Argasinska J, Quinones-Olvera N, Finn RD, Bateman A, Petrov AI. 2018. Non-coding RNA analysis using the Rfam database. *Curr Protoc Bioinformatics* 62:e51. <https://doi.org/10.1002/cpbi.51>.