

Finished Annotated Genome Sequence of *Burkholderia pseudomallei* Strain Bp1651, a Multidrug-Resistant Clinical Isolate

Julia V. Bugrysheva,^a David Sue,^a Janetta Hakovirta,^a Vladimir N. Loparev,^a Kristen Knipe,^a Scott A. Sammons,^a Satishkumar Ranganathan-Ganakammal,^a Shankar Changayil,^a Ganesh Srinivasamoorthy,^a Michael R. Weil,^b Roman L. Tatusov,^{a,c} Jay E. Gee,^a Mindy G. Elrod,^a Alex R. Hoffmaster,^a Linda M. Weigel^a

Office of Infectious Diseases^a and Office of the Chief Information Officer,^b Centers for Disease Control and Prevention, Department of Health and Human Services, Atlanta, Georgia, USA; SRA International, Inc, Atlanta, Georgia, USA^c

***Burkholderia pseudomallei* strain Bp1651, a human isolate, is resistant to all clinically relevant antibiotics. We report here on the finished genome sequence assembly and annotation of the two chromosomes of this strain. This genome sequence may assist in understanding the mechanisms of antimicrobial resistance for this pathogenic species.**

Received 14 October 2015 Accepted 17 October 2015 Published 3 December 2015

Citation Bugrysheva JV, Sue D, Hakovirta J, Loparev VN, Knipe K, Sammons SA, Ranganathan-Ganakammal S, Changayil S, Srinivasamoorthy G, Weil MR, Tatusov RL, Gee JE, Elrod MG, Hoffmaster AR, Weigel LM. 2015. Finished annotated genome sequence of *Burkholderia pseudomallei* strain Bp1651, a multidrug-resistant clinical isolate. *Genome Announc* 3(6):e01427-15. doi:10.1128/genomeA.01427-15.

Copyright © 2015 Bugrysheva 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 Julia V. Bugrysheva, vol5@cdc.gov.

Burkholderia pseudomallei is a Gram-negative bacillus and the etiologic agent of the potentially fatal, infectious disease called melioidosis (1). This saprophytic organism is considered endemic in Southeast Asia and Northern Australia. The infection usually occurs following contact with contaminated soil or water (1). *B. pseudomallei* strain Bp1651 was isolated from sputum of a hospital patient in the United States. This infection is likely to have been acquired in Australia, where the patient traveled several years prior. Antimicrobial susceptibility testing revealed that the strain was resistant to amoxicillin-clavulanic acid, ceftazidime, doxycycline, imipenem, and sulfamethoxazole/trimethoprim.

The *B. pseudomallei* Bp1651 genome was sequenced with the PacBio RSII system (Pacific Biosciences, USA). Eight single-molecule real-time (SMRT) cells were sequenced, and the data were assembled with the RS_HGAP_Assembly.3 protocol implemented in SMRT Portal version 2.3.0. There were a total of 243,100 subreads with an average length of 3,279 bp after filtering. A 30× long-read cutoff (6,316 bp) was used for preassembly, and the longest 15× of the corrected reads were assembled using Celera Assembler. The assembly resulted in two contigs (average depth of coverage 95.2) that correspond to the two chromosomes of this bacterium. We were able to close the contig corresponding to chromosome 1 into a circular chromosome based on the PacBio sequence assembly. The contig for chromosome 2 was closed into a circular chromosome using a paired-end option without addition of “N”s (see https://www.ncbi.nlm.nih.gov/assembly/aggp/AGP_Specification for definition) by the National Center for Biotechnology Information (NCBI) Genome Submission group during submission of the sequence.

Even though the error rate for a single read with PacBio could be 12% or more, combining the multiple reads of the same region by the hierarchical genome-assembly process (HGAP) increases accuracy to 99.999%, as reported previously (2). The accuracy of the current HGAP3 assembly was verified by comparing the assembled genome to the whole-genome map generated by whole-

genome mapping technology (OpGen Inc., USA) with *Bam*HI digest. We designated the sequence of the Bp1651 genome as finished (noncontiguous finished quality) because the ends of chromosome 2 were closed artificially (3).

Chromosome 1 is 4,112,209-bp long with a G+C content of 67.8%, while chromosome 2 is 3,147,816-bp long with a G+C content of 68.5%. These features are consistent with previously published *B. pseudomallei* genomes (4, 5).

The genome of *B. pseudomallei* Bp1651 was annotated using Prokka software (6). Chromosome 1 was predicted to contain 3,521 protein-coding sequences, 9 rRNA genes, and 59 tRNA genes. Chromosome 2 was predicted to contain 2,411 protein-coding sequences, 3 rRNA genes, and 26 tRNA genes. A predicted function was assigned to 75% of protein-coding genes. These numbers are similar to previously published genome annotations of *B. pseudomallei* (4, 5). We are currently analyzing sequences that may be associated with the antibiotic resistance of this strain.

Nucleotide sequence accession numbers. The finished annotated sequences of the two chromosomes of *B. pseudomallei* Bp1651 have been deposited at NCBI GenBank under the accession numbers [CP012041](https://www.ncbi.nlm.nih.gov/nuccore/CP012041) and [CP012042](https://www.ncbi.nlm.nih.gov/nuccore/CP012042).

ACKNOWLEDGMENTS

We thank Nancy Watz and Ellen Jo Baron, Stanford Medical Center, for providing the Bp1651 strain of *B. pseudomallei*. We also thank Mark Borodovsky, Georgia Institute of Technology, and Tom de Man, CDC, for assistance with genome annotation.

This work was supported by the Office of Infectious Diseases, CDC. The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. The mention of company names or products does not constitute endorsement by the CDC.

REFERENCES

1. Wiersinga WJ, Currie BJ, Peacock SJ. 2012. Melioidosis. *N Engl J Med* 367:1035–1044. [http://dx.doi.org/10.1056/NEJMra1204699](https://doi.org/10.1056/NEJMra1204699).

2. Chin C, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <http://dx.doi.org/10.1038/nmeth.2474>.
3. Chain PS, Grafham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, Highlander S, Hugenholtz P, Khouri HM, Kodira CD, Kolker E, Kyrpides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Genomic Standards Consortium Human Microbiome Project Jumpstart Consortium, Detter JC. 2009. Genomics: genome project standards in a new era of sequencing. *Science* 326:236–237. <http://dx.doi.org/10.1126/science.1180614>.
4. Johnson SL, Bishop-Lilly KA, Ladner JT, Daligault HE, Davenport KW, Jaissle J, Frey KG, Koroleva GI, Bruce DC, Coyne SR, Broomall SM, Li PE, Teshima H, Gibbons HS, Palacios GF, Rosenzweig CN, Redden CL, Xu Y, Minogue TD, Chain PS. 2015. Complete genome sequences for 59 *Burkholderia* isolates, both pathogenic and near neighbor. *Genome Announc* 3(2):e00159-15. <http://dx.doi.org/10.1128/genomeA.00159-15>.
5. Holden MTG, Titball RW, Peacock SJ, Cerdeno-Tarraga AM, Atkins T, Crossman LC, Pitt T, Churcher C, Mungall K, Bentley SD, Sebaihia M, Thomson NR, Bason N, Beacham IR, Brooks K, Brown KA, Brown NF, Challis GL, Cherevach I, Chillingworth T, Cronin A, Crossett B, Davis P, DeShazer D, Feltwell T, Fraser A, Hance Z, Hauser H, Holroyd S, Jagels K, Keith KE, Maddison M, Moule S, Price C, Quail MA, Rabinowitsch E, Rutherford K, Sanders M, Simmonds M, Songsivilai S, Stevens K, Tumapa S, Vesaratchavest M, Whitehead S, Yeats C, Barrell BG, Oyston PC, Parkhill J. 2004. Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc Natl Acad Sci USA* 101:14240–14245. <http://dx.doi.org/10.1073/pnas.0403302101>.
6. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.