



Genome Sequence of the Historical Clinical Isolate *Burkholderia pseudomallei* PHLS 6

 Patrik D'haeseleer,^a  Shannon L. Johnson,^b Karen W. Davenport,^b Patrick S. Chain,^b Joe Schoeniger,^c Debjit Ray,^c Anupama Sinha,^c Kelly P. Williams,^c José Peña,^a Steven S. Branda,^c Sahar El-Etr^a

Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory (LLNL), Livermore, California, USA^a; Los Alamos National Laboratory (LANL), Los Alamos, New Mexico, USA^b; Sandia National Laboratories (SNL), Livermore, California, USA^c

Here, we present the draft genome sequence of *Burkholderia pseudomallei* PHLS 6, a virulent clinical strain isolated from a melioidosis patient in Bangladesh in 1960. The draft genome consists of 39 contigs and is 7,322,181 bp long.

Received 17 May 2016 Accepted 18 May 2016 Published 30 June 2016

Citation D'haeseleer P, Johnson SL, Davenport KW, Chain PS, Schoeniger J, Ray D, Sinha A, Williams KP, Peña J, Branda SS, El-Etr S. 2016. Genome sequence of the historical clinical isolate *Burkholderia pseudomallei* PHLS 6. *Genome Announc* 4(3):e00649-16. doi:10.1128/genomeA.00649-16.

Copyright © 2016 D'haeseleer 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 Steven S. Branda, sbranda@sandia.gov.

Burkholderia pseudomallei is a Gram-negative pathogenic saprophyte endemic to Southeast Asia, India, and northern Australia (1, 2). It is the causative agent of melioidosis, a potentially fatal infectious disease with a wide spectrum of symptoms (3). Melioidosis mortality rates range from 14% to 40% (4, 5), exceeding 70% without treatment with effective antimicrobials (6, 7). The global number of melioidosis cases per year has been estimated to be around 165,000, including 89,000 deaths (2). *B. pseudomallei* is on the CDC category B list (8, 9) and is considered to be a potential bioweapon (10) due to its prevalence in soil, multiple routes of infection, low infectious dose, high mortality rates, native resistance to a wide range of antibiotics and disinfectants, and lack of an effective vaccine.

Like other *Burkholderia* species, *B. pseudomallei* strains exhibit a remarkable level of genomic plasticity (11–16). The high abundance of simple sequence repeats (SSRs) in *Burkholderia mallei* and *B. pseudomallei* strains drives rapid genomic adaptation (17), even during the course of acute infection in a single patient (18–20) or short-term passaging *in vitro* (20, 21). *B. pseudomallei* PHLS 6 was chosen for sequencing because it is a virulent clinical strain that has undergone minimal lab manipulation. It was originally isolated from a melioidosis patient in Bangladesh in 1960. The PHLS 6 strain was also found to be highly efficient at entering and surviving inside amoebae, a mechanism thought to play a role in the prevalence of *B. pseudomallei* in soils (22).

Genomic DNA was isolated using the Norgen bacterial genomic DNA isolation kit and initially sequenced independently using Illumina technology at Sandia National Laboratories (SNL) and Los Alamos National Laboratory (LANL). SNL sequence data consisted of 10,692,593 NextSeq 150-bp paired-end (PE) reads of short inserts (Nextera DNA library prep kit) and 3,500,043 reads of long inserts (Nextera mate-pair prep kit) trimmed with NxTrim (23). Assembly with SPAdes, Bridger, and Gap Filler (24–26) yielded 40 scaffolds with 220 contigs. LANL sequence data consisted of 15,757,418 MiSeq 250-bp PE reads of short inserts (300 ± 70 bp) (NEBNext Ultra DNA library prep), which were trimmed and filtered for quality and reduced to a total of 300×

genome coverage for use in the assemblies. Assemblies from IDBA and Velvet were merged with parallel Phrap (27–30), producing 70 contigs. Because of the high number of repeats in the genome, it was decided to combine the SNL long-insert data and LANL short-insert data using ALLPaths (31) and merge the resulting assembly with the prior SNL and LANL assemblies in parallel Phrap to generate the final improved draft assembly, with 39 contigs totaling 7,322,181 bp (G+C content, 68.1%). Annotation utilized an LANL in-house Ergatis workflow with minor manual curation (32) and found 6,156 coding regions, 11 rRNA sequences, and 59 tRNA sequences. The genome contains 118 genes associated with antibiotic and toxic compound resistance, as well as 16 genes associated with cellular invasion and intracellular resistance (33).

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. [LWRR000000000](https://www.ncbi.nlm.nih.gov/nuclink/LWRR000000000/). The version described in this paper is version LWRR01000000.

ACKNOWLEDGMENTS

This research was supported by LLNL Laboratory-directed Research and Development (LDRD) Project 15-ERD-017 and SNL Project 165767. Work at LLNL was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. SNL is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

This paper is approved by LLNL for unlimited release (LLNL-JRNL-684377).

FUNDING INFORMATION

This work, including the efforts of Patrik D'haeseleer, José Peña, Steven S. Branda, and Sahar El-Etr, was funded by DOE | Laboratory Directed Research and Development (LDRD) (15-ERD-017).

REFERENCES

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

2. Limmathurotsakul D, Golding N, Dance DA, Messina JP, Pigott DM, Moyes CL, Rolim DB, Bertherat E, Day NP, Peacock SJ, Hay SI. 2016. Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. *Nat Microbiol* 1:15008. <http://dx.doi.org/10.1038/nmicrobiol.2015.8>.
3. Wiersinga WJ, van der Poll T, White NJ, Day NP, Peacock SJ. 2006. Melioidosis: insights into the pathogenicity of *Burkholderia pseudomallei*. *Nat Rev Microbiol* 4:272–282. <http://dx.doi.org/10.1038/nrmicro1385>.
4. Limmathurotsakul D, Wongratnacheewin S, Teerawattanasook N, Wongsuvan G, Chaisuksant S, Chetchotisakd P, Chaowagul W, Day NP, Peacock SJ. 2010. Increasing incidence of human melioidosis in northeast Thailand. *Am J Trop Med Hyg* 82:1113–1117. <http://dx.doi.org/10.4269/ajtmh.2010.10-0038>.
5. Currie BJ, Ward L, Cheng AC. 2010. The epidemiology and clinical spectrum of melioidosis: 540 cases from the 20 year Darwin prospective study. *PLoS Negl Trop Dis* 4:e900. <http://dx.doi.org/10.1371/journal.pntd.0000900>.
6. White NJ, Dance DA, Chaowagul W, Wattanagoon Y, Wuthiekanun V, Pitakwatchara N. 1989. Halving of mortality of severe melioidosis by ceftazidime. *Lancet* 2:697–701. [http://dx.doi.org/10.1016/S0140-6736\(89\)90768-X](http://dx.doi.org/10.1016/S0140-6736(89)90768-X).
7. Lipsitz R, Garges S, Aurigemma R, Baccam P, Blaney DD, Cheng AC, Currie BJ, Dance D, Gee JE, Larsen J, Limmathurotsakul D, Morrow MG, Norton R, O'Mara E, Peacock SJ, Pesik N, Rogers LP, Schweizer HP, Steinmetz I, Tan G, Tan P, Wiersinga WJ, Wuthiekanun V, Smith TL. 2012. Workshop on treatment of and postexposure prophylaxis for *Burkholderia pseudomallei* and *B. mallei* infection. *Emerg Infect Dis* 18:e2. <http://dx.doi.org/10.3201/eid1812.120638>.
8. Rotz LD, Khan AS, Lillibrige SR, Ostroff SM, Hughes JM. 2002. Public health assessment of potential biological terrorism agents. *Emerg Infect Dis* 8:225–230. <http://dx.doi.org/10.3201/eid0802.010164>.
9. National Institute of Allergy and Infectious Diseases. 2003. NIAID biodefense research agenda for category B and C priority pathogens. National Institute of Allergy and Infectious Diseases, Bethesda, MD. <http://www.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/Documents/categorybandc.pdf>.
10. Thavaselvam D, Vijayaraghavan R. 2010. Biological warfare agents. *J Pharm Bioallied Sci* 2:179–188. <http://dx.doi.org/10.4103/0975-7406.68499>.
11. Holden MT, Titball RW, Peacock SJ, Cerdeño-Tárraga 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, Crosssett 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>.
12. Ong C, Ooi CH, Wang D, Chong H, Ng KC, Rodrigues F, Lee MA, Tan P. 2004. Patterns of large-scale genomic variation in virulent and avirulent *Burkholderia* species. *Genome Res* 14:2295–2307. <http://dx.doi.org/10.1101/gr.1608904>.
13. Ou K, Ong C, Koh SY, Rodrigues F, Sim SH, Wong D, Ooi CH, Ng KC, Jikuya H, Yau CC, Soon SY, Kesuma D, Lee MA, Tan P. 2005. Integrative genomic, transcriptional, and proteomic diversity in natural isolates of the human pathogen *Burkholderia pseudomallei*. *J Bacteriol* 187:4276–4285. <http://dx.doi.org/10.1128/JB.187.12.4276-4285.2005>.
14. Tumapa S, Holden MT, Vesaratchavest M, Wuthiekanun V, Limmathurotsakul D, Chierakul W, Feil EJ, Currie BJ, Day NP, Nierman WC, Peacock SJ. 2008. *Burkholderia pseudomallei* genome plasticity associated with genomic island variation. *BMC Genomics* 9:190. <http://dx.doi.org/10.1186/1471-2164-9-190>.
15. Sim SH, Yu Y, Lin CH, Karuturi RK, Wuthiekanun V, Tuanyok A, Chua HH, Ong C, Paramalingam SS, Tan G, Tang L, Lau G, Ooi EE, Woods D, Feil E, Peacock SJ, Tan P. 2008. The core and accessory genomes of *Burkholderia pseudomallei*: implications for human melioidosis. *PLoS Pathog* 4:e1000178. <http://dx.doi.org/10.1371/journal.ppat.1000178>.
16. Chain PSG. 2011. Genomic versatility in the *Burkholderia* genus: from strains to species. Ph.D. thesis. Michigan State University, East Lansing, MI.
17. Song H, Hwang J, Myung J, Seo H, Yi H, Sim HS, Kim BS, Nierman WC, Kim HS. 2009. Simple sequence repeat (SSR)-based gene diversity in *Burkholderia pseudomallei* and *Burkholderia mallei*. *Mol Cells* 27:237–241. <http://dx.doi.org/10.1007/s10059-009-0029-8>.
18. Price EP, Hornstra HM, Limmathurotsakul D, Max TL, Sarovich DS, Vogler AJ, Dale JL, Ginther JL, Leadem B, Colman RE, Foster JT, Tuanyok A, Wagner DM, Peacock SJ, Pearson T, Keim P. 2010. Within-host evolution of *Burkholderia pseudomallei* in four cases of acute melioidosis. *PLoS Pathog* 6:e1000725. <http://dx.doi.org/10.1371/journal.ppat.1000725>.
19. Limmathurotsakul D, Holden MT, Coupland P, Price EP, Chantratita N, Wuthiekanun V, Amornchai P, Parkhill J, Peacock SJ. 2014. Microevolution of *Burkholderia pseudomallei* during an acute infection. *J Clin Microbiol* 52:3418–3421. <http://dx.doi.org/10.1128/JCM.01219-14>.
20. Romero CM, DeShazer D, Feldblyum T, Ravel J, Woods D, Kim HS, Yu Y, Ronning CM, Nierman WC. 2006. Genome sequence alterations detected upon passage of *Burkholderia mallei* ATCC 23344 in culture and in mammalian hosts. *BMC Genomics* 7:228. <http://dx.doi.org/10.1186/1471-2164-7-228>.
21. U'Ren JM, Schupp JM, Pearson T, Hornstra H, Friedman CL, Smith KL, Daugherty RR, Rhoton SD, Leadem B, Georgia S, Cardon M, Huynh LY, DeShazer D, Harvey SP, Robison R, Gal D, Mayo MJ, Wagner D, Currie BJ, Keim P. 2007. Tandem repeat regions within the *Burkholderia pseudomallei* genome and their application for high resolution genotyping. *BMC Microbiol* 7:23. <http://dx.doi.org/10.1186/1471-2180-7-23>.
22. Moore EA. 2010. Interactions of *Burkholderia pseudomallei* and *Acanthamoeba castellanii* and their effects on virulence in human monocytes. Masters' thesis. Brigham Young University, Provo, UT.
23. O'Connell J, Schulz-Trieglaff O, Carlson E, Hims MM, Gormley NA, Cox AJ. 2015. NxTrim: optimized trimming of Illumina mate pair reads. *Bioinformatics* 31:2035–2037. <http://dx.doi.org/10.1093/bioinformatics/btv057>.
24. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyskhin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
25. Chang Z, Li G, Liu J, Zhang Y, Ashby C, Liu D, Cramer CL, Huang X. 2015. Bridger: a new framework for *de novo* transcriptome assembly using RNA-seq data. *Genome Biol* 16:30. <http://dx.doi.org/10.1186/s13059-015-0596-2>.
26. Nadalin F, Vezzi F, Policriti A. 2012. GapFiller: a *de novo* assembly approach to fill the gap within paired reads. *BMC Bioinformatic* 13(Suppl 14):S8. <http://dx.doi.org/10.1186/1471-2105-13-S14-S8>.
27. Peng Y, Leung HCM, Yiu SM, Chin FYL. 2010. IDBA—a practical iterative de Bruijn graph *de novo* assembler. *Lect Notes Comput Sci* 6044:426–440. http://dx.doi.org/10.1007/978-3-642-12683-3_28.
28. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
29. Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* 8:175–185. <http://dx.doi.org/10.1101/gr.8.3.175>.
30. Ewing B, Green P. 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* 8:186–194. <http://dx.doi.org/10.1101/gr.8.3.186>.
31. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB. 2008. ALLPATHS: *de novo* assembly of whole-genome shotgun microreads. *Genome Res* 18:810–820. <http://dx.doi.org/10.1101/gr.7337908>.
32. Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. 2010. An Ergatis-based prokaryotic genome annotation Web server. *Bioinformatics* 26:1122–1124. <http://dx.doi.org/10.1093/bioinformatics/btq090>.
33. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 46:D206–D214. <http://dx.doi.org/10.1093/nar/gkt1226>.