



# Methylome Characterization of *Burkholderia pseudomallei* Strain 982 at Single-Base Resolution

 Kar-Wai Hong,<sup>a</sup> Kok Keng Tee,<sup>b</sup> Wai-Fong Yin,<sup>a</sup>  Richard J. Roberts,<sup>c</sup>  Kok-Gan Chan<sup>a,d</sup>

<sup>a</sup>International Genome Centre, Jiangsu University, Zhenjiang, China

<sup>b</sup>Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

<sup>c</sup>New England Biolabs, Ipswich, Massachusetts, USA

<sup>d</sup>Division of Genetics and Molecular Biology, Institute of Biological Sciences, University of Malaya, Kuala Lumpur, Malaysia

**ABSTRACT** *Burkholderia pseudomallei* is the etiological agent of melioidosis, which has been studied by transcriptome and secretome analyses. However, little is known about the methylome of this pathogen. Here, we present the complete genome and methylome of melioidosis-causing *B. pseudomallei* strain 982.

**B***urkholderia pseudomallei* is a Gram-negative bacterium from the class *Betaproteobacteria* that causes melioidosis with a case fatality rate of 10 to 50% (1, 2).

Here, we have sequenced *B. pseudomallei* strain 982 in order to understand its methylome. This strain was originally isolated from the pus of a 13-year-old male patient on 21 March 2015 in Pahang, Malaysia. The pus sample was serially diluted prior to being spread on blood agar followed by incubation at 37°C to obtain a single colony (3). The bacterial sample was identified using the analytical profile index (API) approach and a Microflex benchtop matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometer (Bruker Daltonik GmbH, Germany) (4). This isolate was cultivated in lysogeny broth, and its genomic DNA was extracted using a MasterPure DNA purification kit (Epicentre, Madison, WI, USA). The quantity and quality of the extracted genomic DNA were measured using a Qubit 2.0 fluorometer and a NanoDrop 2000 spectrophotometer (both Thermo Fisher Scientific, Waltham, MA, USA) and a Pippin Pulse electrophoresis power supply (Sage Science, Beverly, MA, USA), respectively (5). Genomic DNA (10 μg) was sheared using a g-TUBE (Covaris, Woburn, Massachusetts, USA). A 20-kb template library was prepared using the BluePippin size-selection system and sequenced using a Pacific Biosciences RS II platform in two runs of the same batch by using P5-C3 chemistry following the MagBead loading protocol (all Pacific Biosciences, Menlo Park, CA) (6, 7), as this increases sequencing coverage, which will facilitate downstream analysis.

*De novo* assembly of the genome was performed using the Hierarchical Genome Assembly Process (HGAP) version 2.0 (8). The circularity of the assembled genome was studied using Gepard (9) and Contiguity (10). The overlapping ends were trimmed using Minimus2 of the AMOS software package (11). The annotation and detection of prophages were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) and the PHAge Search Tool Enhanced Release (PHASTER) (12, 13). For base modification analysis, the RS\_Modification\_and\_Motif\_Analysis.1 protocol (<https://github.com/sanger-pathogens/Bio-PacbioMethylation/blob/master/lib/Bio/PacbioMethylation/RSModificationRunner.pm>) with a default modification quality value score of 30 was used, and the motif analysis summary was deposited in REBASE (14). Default parameters were used for all software unless otherwise specified.

A total of 527,043,402 reads and 458,304,859 reads were generated from two individual SMRT Cell v2 instances with  $N_{50}$  values of 9,087 bp and 9,162 bp, respectively,

**Citation** Hong K-W, Tee KK, Yin W-F, Roberts RJ, Chan K-G. 2019. Methylome characterization of *Burkholderia pseudomallei* strain 982 at single-base resolution. *Microbiol Resour Announc* 8:e00898-19. <https://doi.org/10.1128/MRA.00898-19>.

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

**Copyright** © 2019 Hong 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 Kok-Gan Chan, [kokgan@um.edu.my](mailto:kokgan@um.edu.my).

**Received** 18 August 2019

**Accepted** 25 September 2019

**Published** 24 October 2019

**TABLE 1** MTase and its predicted DNA recognition sequences in *B. pseudomallei* strain 982

RM system gene annotated	Chromosome	Type (subtype)	Locus tag (start bp–stop bp)	Predicted recognition motif <sup>a</sup>	Methylated type
M.Bps982II	1	I (gamma)	AMS56_03820 (799640–801286)	GTC <sup>m6</sup> <u>ANNNNNNTGG</u>	m <sup>6</sup> A
S.Bps982II	1	I	AMS56_03825 (801276–802589)	GTCANNNNNNTGG	
Bps982IIP	1	I	AMS56_03830 (802586–805717)	GTCANNNNNNTGG	
M.Bps982ORF1205P	1	II (beta)	AMS56_01205 (247616–248392) <sup>b</sup>		
M.Bps982ORF17150P	1	II	AMS56_17150 (3830654–3831883)		
M.Bps982I	1	III (beta)	AMS56_01845 (377543–379576) <sup>b</sup>	CAC <sup>m6</sup> <u>AG</u>	m <sup>6</sup> A
Bps982IP	1	III	AMS56_01840 (374473–377502) <sup>b</sup>	CACAG	
M.Bps982IV	1	III (beta)	AMS56_02205 (546377–547897)	CAG <sup>m6</sup> <u>ATG</u>	
Bps982IVP	1	III	AMS56_02210 (547907–550591)	CAGATG	
M.Bps982III	2	II (beta)	AMS56_(1628287–1629135)	GT <u>WW</u> <sup>m6</sup> <u>AC</u>	m <sup>6</sup> A
M.Bps982ORF25235P	2	II	AMS56_(1703198–1703761)		
M.Bps982ORF25241P	2	II	AMS56_(1705313–1706002) <sup>b</sup>	YGG <sup>m5</sup> <u>CCR</u>	m <sup>5</sup> C

<sup>a</sup> Modified bases are highlighted in bold and underlined when on the complementary strand.

<sup>b</sup> A complementary strand sequence.

giving an average sequencing coverage of 116.91×. This bacterium consisted of two circular chromosomes of 4,028,032 bp (67.87% G+C content) and 3,156,645 bp (68.56% G+C content). It contained 6,199 genes, 6,035 of which are protein-coding genes and 61 of which are tRNA genes, with 4 rRNA operons (5S, 16S, and 23S) and 1 noncoding RNA. PHASTER indicated that chromosome 1 harbored one intact prophage region, while chromosome 2 possessed two incomplete prophage regions.

Four methylation motifs were detected, namely, GTC<sup>m6</sup>ANNNNNNTGG, GTWW<sup>m6</sup>AC, CAG<sup>m6</sup>ATG, and CAC<sup>m6</sup>AG (Table 1). By performing a comprehensive search in REBASE, the DNA methylase (MTase) which is responsible for the motif GTWWAC was found to be conserved across the *Burkholderiaceae* family (14, 15).

**Data availability.** The whole-genome sequence of *B. pseudomallei* strain 982 has been deposited at DDBJ/EMBL/GenBank under the accession numbers [CP012576](#) and [CP012577](#) for chromosomes 1 and 2, respectively. The version described in this paper is available under NCBI BioProject number [PRJNA293915](#) and NCBI BioSample number [SAMN04011951](#). The raw reads of sequenced genomic DNA of this bacterium and the output of base modification and motif analysis were deposited in the SRA under accession numbers [SRX6717522](#) and [SRZ189982](#), respectively. The description of genes related to the restriction modification (RM) system of this strain is available in REBASE under the organism number 16827 (<http://rebase.neb.com/cgi-bin/onumget?16827>).

## ACKNOWLEDGMENTS

We thank the University of Malaya for research grants (FRGS grant FP022-2018A and HIR grant H-50001-A000027 to K.-G.C. and grant PG085-2015B awarded to K.-W.H.).

We are grateful to Yuet-Meng Cheong for providing *Burkholderia pseudomallei* strain 982.

All authors listed made substantial direct and indirect contributions to the work and approved it for publication.

We declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. R.J.R. works for New England Biolabs, a company that sells research reagents, including restriction enzymes and DNA methyltransferases, to the scientific community.

## REFERENCES

- Wiersinga WJ, Currie BJ, Peacock SJ. 2012. Melioidosis. *N Engl J Med* 367:1035–1044. <https://doi.org/10.1056/NEJMr1204699>.
- Wiersinga WJ, Virk HS, Torres AG, Currie BJ, Peacock SJ, Dance DAB, Limmathurotsakul D. 2018. Melioidosis. *Nat Rev Dis Primers* 4:17107. <https://doi.org/10.1038/nrdp.2017.107>.
- Limmathurotsakul D, Wuthiekanun V, Chantratita N, Wongsuvan G, Thanwisai A, Biaklang M, Tumapa S, Lee S, Day NPJ, Peacock SJ. 2007. Simultaneous infection with more than one strain of *Burkholderia pseudomallei* is uncommon in human melioidosis. *J Clin Microbiol* 45:3830–3832. <https://doi.org/10.1128/JCM.01297-07>.
- Wang H, Chen YL, Teng SH, Xu ZP, Xu YC, Hsueh PR. 2016. Evaluation of the Bruker Biotyper matrix-assisted laser desorption/ionization time-of-flight mass spectrometry system for identification of clinical and environmental isolates of *Burkholderia pseudomallei*. *Front Microbiol* 7:415. <https://doi.org/10.3389/fmicb.2016.00415>.
- Hong KW, Asmah Hani AW, Nurul Aina Murni CA, Pusparani RR, Chong

- CK, Verasahib K, Yusoff WNW, Noordin NM, Tee KK, Yin WF, Yu CY, Ang GY, Chan KG. 2017. Comparative genomic and phylogenetic analysis of a toxigenic clinical isolate of *Corynebacterium diphtheriae* strain B-D-16–78 from Malaysia. *Infect Genet Evol* 54:263–270. <https://doi.org/10.1016/j.meegid.2017.07.015>.
6. Lim YL, Yong D, Ee R, Tee KK, Yin WF, Chan KG. 2015. Complete genome sequence of *Serratia multitudinisentens* RB-25<sup>T</sup>, a novel chitinolytic bacterium. *J Biotechnol* 207:32–33. <https://doi.org/10.1016/j.jbiotec.2015.04.027>.
7. Ee R, Lim YL, Yin WF, See-Too WS, Roberts RJ, Chan KG. 2016. Novel methyltransferase recognition motif identified in *Chania multitudinisentens* RB-25<sup>T</sup> gen. nov., sp. nov. *Front Microbiol* 7:1362. <https://doi.org/10.3389/fmicb.2016.01362>.
8. Chin C-S, 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. <https://doi.org/10.1038/nmeth.2474>.
9. Krumsiek J, Arnold R, Rattei T. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23:1026–1028. <https://doi.org/10.1093/bioinformatics/btm039>.
10. Sullivan MJ, Ben Zakour NL, Forde BM, Stanton-Cook M, Beatson SA. 2015. Contiguity: contig adjacency graph construction and visualisation. *PeerJ PrePrint* 3:e1037v1. <https://peerj.com/preprints/1037/>.
11. Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M. 2011. Next generation sequence assembly with AMOS. *Curr Protoc Bioinformatics Chapter 11:Unit 11.8*. <https://doi.org/10.1002/0471250953.bi1108s33>.
12. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
13. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
14. Roberts RJ, Vincze T, Posfai J, Macelis D. 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 43:D298–D299. <https://doi.org/10.1093/nar/gku1046>.
15. Erill I, Puigvert M, Legrand L, Guarischi-Sousa R, Vandecasteele C, Setubal JC, Genin S, Guidot A, Valls M. 2017. Comparative analysis of *Ralstonia solanacearum* methylomes. *Front Plant Sci* 8:504. <https://doi.org/10.3389/fpls.2017.00504>.