Near-Complete Whole-Genome Sequencing of Two *Burkholderia pseudomallei* Strains Harbouring Novel Molecular Class D Beta-Lactamase Genes, Isolated from Malaysia

Ummu Afeera Zainulabid,[a,e] Norirwanisyam Mohamad Zain,[b,d] Janaki Arumugam,[c] Norhidayah Kamarudin,[a] Muhammad 'Adil Zainal Abidin,[a] Abdullah Shamshir Abd Mokti,[c] Noorlina Nordin,[c] Ahmad Faizal Rakawi,[c] Aman Shah Abdul Majid,[d] Gnanasekaran Ashok,[d] Alex Lourdes Francis,[d] Darren Dean Tay,[e] Natesan Vijayalakshami,[d] How Soon Hin,[a] ⓘ Hajar Fauzan Ahmad[e]

ᵃKulliyyah of Medicine, International Islamic University of Malaysia, Jalan Sultan Ahmad Shah, Kuantan, Pahang, Malaysia
ᵇHospital Raja Permaisuri Bainun, Jalan Raja Ashman Shah, Ipoh, Perak, Malaysia
ᶜHospital Tengku Ampuan Afzan, Jalan Tanah Putih, Kuantan, Pahang, Malaysia
ᵈFaculty of Medicine, Quest International University, Ipoh, Perak, Malaysia
ᵉFaculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, Gambang, Pahang, Malaysia

Ummu Afeera Zainulabid and Norirwanisyam Mohamad Zain contributed equally.

ABSTRACT   Here, we present two 7.1- and 7.3-bp near-complete genome sequences of *Burkholderia pseudomallei* strains of HTAA077 and HRPB058, isolated from a pus culture from a confirmed melioidosis patient at Hospital Tengku Ampuan Afzan, Kuantan, Malaysia, and from blood culture from a patient at Hospital Raja Permaisuri Bainun, Ipoh, Malaysia, using a Nanopore MinION instrument.

*B*urkholderia pseudomallei is a pathogenic saprophyte that causes melioidosis (1). The disease is endemic in Malaysia, with various clinical presentations (2, 3), and its management includes 6 months of antibiotic regimen administration (4). Clinical specimens were collected from patients suspected of bacteremia in accordance with ethical approvals (International Islamic University Malaysia [IIUM] Research Ethics Committee [IREC] and National Medical Research Register [NMRR] via IREC 2021-193, NMRR-17-3238-39448, and NMRR-21-360-58751). The blood was cultured using a BD Bactec FX blood culture system (Becton, Dickinson and Company, NJ), and both samples were subcultured on blood agar media (Bio-Rad, CA) at 37°C to obtain a pure culture. The genomic DNA of the strains was extracted as described previously using the GenElute bacterial genomic DNA kit (Sigma-Aldrich Co., St. Louis, USA) (5). For the library preparation, DNA size selection using SPRIselect (Beckman Coulter Inc., IN) was performed to retain fragments larger than 6 kb. Approximately 1,000 ng of size-selected DNA as measured by a QFX fluorometer (Denovix Inc., Wilmington, USA) was used directly for library preparation using the ligation sequencing kit SQK-LSK110 (Oxford Nanopore Technologies [ONT], Oxford, UK). The prepared library was sequenced on a Flongle flow cell R9.4.1 (Oxford Nanopore Technologies) and base-called used the Guppy v5.0.7 "super accuracy model." The raw fastq reads were length filtered to retain reads longer than 2 kb, followed by *de novo* assembly using the Flye v2.9 –nano-hq model (6). The assembled genome was subsequently polished with one round of RACON (https://github.com/isovic/racon) and then one round of Medaka (https://github.com/nanoporetech/medaka) (7), generating the final consensus assembly that consists of 2 circular sequences, which the genome automatically circularized with Flye, but not rotated and closed contigs. Genome completeness was subsequently assessed using BUSCO5 (https://gitlab.com/ezlab/busco) based on the Burkholderiales_odb10 lineage database (8). Genome annotation was carried out using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) with the GeneMarkS2-v.1.14_1.25

method (9), and the presence of acquired antibiotic resistance genes (ARGs), virulence factors, and human pathogenicity was analyzed using the program VirulenceFinder 2.0 (10) and ResFinder 2.1 (11). The HTAA077 strain has a total length of 7,276,308 bp, an average GC content of 68%, and ONT reads with a mean length of 67,923 and an N_{50} length of 4,082,678 bp. It contains 3 contigs, comprising 6,193 coding sequences with 76 RNAs. In addition, the HRPB058 strain has a total length of 7,094,056 bp, an average GC content of 68.1%, and ONT reads with a mean length of 32,1475 and an N_{50} length of 3,961,310 bp. It contains 2 contigs, comprising 5,897 coding sequences with 76 RNAs and 5,897 encoded proteins. Both virulence factors found in this isolate are related to antibiotic resistance genes of bla_{OXA-43}, bla_{OXA-57}, and bla_{OXA-59} that encoded an unknown beta-lactam group.

Data availability. The complete genome sequences of *Burkholderia pseudomallei* are available in the NCBI SRA database under accession numbers SRR18556255 (HTAA077) and SRR18556256 (HRPB056), and the BioProject number is PRJNA821438. The BioSample accession numbers for HTAA077 and HRPB058 are SAMN27097645 and SAMN27097644, respectively. The NCBI PGAP genome annotation numbers are GCA_024506485.1 (JANIED010000001) and GCA_024506415.1 (JANIED010000002) for HTAA077 and HRPB058, respectively.

ACKNOWLEDGMENTS

REFERENCES

1. Janaki A, Howsoon H, Zainulabid UA, Mahmud MIAM, Rahman JA, Kashfi A, Rahman A, Chunn Kuan Y. 2020. Seroconversion of indirect immunofluorescent antibody IgM and IgG in melioidosis. Bangladesh J Infect Dis 7:61–66. https://doi.org/10.3329/bjid.v7i2.51514.

2. Sia TLL, Mohan A, Ooi MH, Chien SL, Tan LS, Goh C, Pang DCL, Currie BJ, Wong JS, Podin Y. 2021. Epidemiological and clinical characteristics of melioidosis caused by gentamicin-susceptible Burkholderia pseudomallei in Sarawak. Open Forum Infect Dis 8:ofab460. https://doi.org/10.1093/ofid/ofab460.

3. Chowdhury S, Barai L, Afroze SR, Ghosh PK, Afroz F, Rahman H, Ghosh S, Hossain MB, Rahman MZ, Das P, Rahim MA. 2022. The epidemiology of melioidosis and its association with diabetes mellitus: a systematic review and meta-analysis. Pathogens 11:149. https://doi.org/10.3390/pathogens11020149.

4. Karunanayake P. 2022. Melioidosis: clinical aspects. Clin Med (Lond) 22: 6–8. https://doi.org/10.7861/clinmed.2022-0014.

5. Ahmad HF, Schreiber L, Marshall IPG, Andersen PJ, Castro-Mejía JL, Nielsen DS. 2019. Draft genome sequence of Streptococcus anginosus strain CALM001, isolated from the gut of an elderly dane. Microbiol Resour Announc 8:e00379-19. https://doi.org/10.1128/MRA.00379-19.

6. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

7. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.

8. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

9. Lomsadze A, Gemayel K, Tang S, Borodovsky M. 2018. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. Genome Res 28:1079–1089. https://doi.org/10.1101/gr.230615.117.

10. Kleinheinz KA, Joensen KG, Larsen MV. 2014. Applying the ResFinder and VirulenceFinder. Bacteriophage 4:e27943-7. https://doi.org/10.4161/bact.27943.

11. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agersø Y, Lund O, Larsen MV, Aarestrup FM. 2013. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother 68:771–777. https://doi.org/10.1093/jac/dks496.