Genome Sequences of *Brucella melitensis*, Isolated from Blood Samples of Brucellosis Patients in Malaysia

Jama'ayah Mohamed Zahidi,ᵃ Norazah Ahmad,ᵃ Bee Yong Tay,ᵃ
Rohaidah Hashim,ᵃ Evie Khoo,ᵇ Nurizan Ahmad,ᵇ Chung Yan Yee,ᵃ
Nur Qhaledaziea Dolhanᵃ

Bacteriology Unit, Infectious Disease Research Centre, Institute for Medical Research, Kuala Lumpur, Malaysiaᵃ;
Veterinary Research Institute, Ipoh, Malaysiaᵇ

ABSTRACT Human brucellosis is a neglected zoonotic disease and has widespread geographical distribution. *Brucella melitensis* has caused outbreaks and sporadic cases in Malaysia. Here, we present the whole-genome sequences of four *B. melitensis* strains isolated from brucellosis patients in Malaysia.

Brucellosis is a highly infectious zoonotic disease affecting humans and diverse groups of wild and domestic animals, with more than 500,000 human cases reported annually worldwide (1, 2). The disease, caused by *Brucella* spp., is mostly transmitted to humans through direct contact or by ingestion of unpasteurized infected dairy products (3, 4). Among the six classically known *Brucella* spp., *B. melitensis* is the most pathogenic species infecting humans (5). It is also the most frequent *Brucella* sp. isolated from blood cultures of humans infected with brucellosis. The genetic characterization of *B. melitensis* isolates is important for epidemiological trace-back investigations of both outbreak and sporadic brucellosis. Multiple-locus variable-number tandem-repeat analysis is generally used for genotyping *Brucella* isolates, but assays are laborious, expensive, and time-consuming.

In order to identify the genetic characteristics of strains from Malaysia, we present here the genome sequences of four *B. melitensis* strains isolated from human patients. The four strains were isolated from blood samples of brucellosis patients who resided in different districts of Johor State (located south of Peninsular Malaysia) in 2015. Strains BMM 13/15 and BMM 26/15 were isolated from Batu Pahat, Johor, while strains BMM 15/15 and BMM 19/15 were isolated from Johor Bahru and Muar, respectively. These strains were inoculated onto brucella agar and incubated at 37°C for 48 h. The genomic DNA was extracted using a MasterPure DNA purification kit (Epicentre, Illumina, USA). Then, DNA libraries were prepared using a Nextera XT DNA library preparation kit (Illumina). Genomic DNA samples of the four clinical strains of *B. melitensis* were sequenced on the Illumina MiSeq platform using a paired-end (2×251-bp) sequencing protocol. The generated sequencing reads were filtered by removing low-quality sequences with the BBDuk version 36 toolkit (http://jgi.doe.gov/data-and-tool/bbtools). The high-quality sequence read data were successfully assembled *de novo* into contigs by using SPAdes version 3.9.0 (6) software.

The draft assemblies of the four clinical strains were annotated using the NCBI Prokaryotic Genome Annotation Pipeline version 4.1. Using Bowtie2 version 2.3, we determined that the sequences shared high nucleotide identity (>99%) with the reference genomes of *B. melitensis* strain 16M (GenBank accession no. NC_003317 and NC_003318) (7). Using these mappings to the reference genomes, we determined the number of single nucleotide polymorphisms (SNPs) by using the SnpEff tool version 4.3i (8). A range of 2,540 to 2,548 SNPs was identified in the four clinical genome sequences

TABLE 1 Summary metrics of whole-genome assemblies of four clinical *B. melitensis* strains isolated from Johor, Malaysia

Strain	No. of contigs	Assembly size (bp)	N_{50} (bp)	Average coverage (×)	No. of genes	No. of tRNAs	No. of SNPs	Accession no.
BMM 13/15	35	3,288,783	189,507	64.55	3,359	49	2,545	NCRL00000000
BMM 15/15	44	3,288,964	176,854	39.36	3,366	49	2,540	NCRJ00000000
BMM 19/15	46	3,288,126	140,693	38.17	3,367	49	2,546	NCRI00000000
BMM 26/15	39	3,288,579	249,479	60.82	3,364	49	2,548	NBBM00000000

(Table 1). This implied that the clinical strains of *B. melitensis* are different from the reference strain. All assemblies contained 3 rRNAs and have a 57.2% G+C content. The other genome assembly metrics (i.e., numbers of contigs, assembly size, N_{50}, average coverage, and number of genes, tRNAs, and SNPs) are provided in Table 1. Further detailed comparative analysis of genetic polymorphisms will help in understanding the expression of certain genes, which can improve epidemiological typing tools and preventive strategies to control brucellosis.

Accession number(s). The whole-genome shotgun sequences reported here have been deposited in GenBank under BioProject PRJNA380526 and are available under the accession numbers listed in Table 1.

REFERENCES

1. Pappas G, Papadimitriou P, Akritidis N, Christou L, Tsianos EV. 2006. The new global map of human brucellosis. Lancet Infect Dis 6:91–99. https://doi.org/10.1016/S1473-3099(06)70382-6.

2. Young EJ. 2009. Brucellosis, p 177–188. *In* Evans AS, Brachman PS, Abrutyn E (ed), Bacterial infections of humans epidemiology and control, 4th ed. Springer, New York, NY.

3. Al-Anazi KA, Al-Jaseer AM. 2013. Brucellosis: a global re-emerging zoonosis; history, epidemiology, microbiology, immunology and genetics, p 66–77. *In* Mascellino MT (ed), Bacterial and mycotic infections in immunocompromised hosts: clinical and microbiological aspects. Omics Group International, Foster City, CA.

4. Wareth G, Melzer F, Elschner MC, Neubauer H, Roesler U. 2014. Detection of *Brucella melitensis* in bovine milk and milk products from apparently healthy animals in Egypt by real-time PCR. J Infect Dev Ctries 8:1339–1343. https://doi.org/10.3855/jidc.4847.

5. Godfroid J, Cloeckaert A, Liautard JP, Kohler S, Fretin D, Walravens K, Garin-Bastuji B, Letesson JJ. 2005. From the discovery of the Malta fever's agent to the discovery of a marine mammal reservoir, brucellosis has continuously been a re-emerging zoonosis. Vet Res 36:313–326. https://doi.org/10.1051/vetres:2005003.

6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin 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. https://doi.org/10.1089/cmb.2012.0021.

7. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

8. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly 6:80–92. https://doi.org/10.4161/fly.19695.