UNIVERSITI TEKNOLOGI MARA

GENOMIC EPIDEMIOLOGY OF SPORADIC HUMAN *Brucella melitensis* CASES IN MALAYSIA THROUGH WHOLE GENOME ANALYSIS

JAMA'AYAH BINTI MOHAMED ZAHIDI @ AHMAD BAKRI

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science
(Molecular Biology)

Faculty of Applied Science

February 2023

ABSTRACT

Brucella melitensis is the most pathogenic and genetically diverged species within the genus *Brucella*. In Malaysia, human infections by *B. melitensis* are reported yearly and the cases are primarily sporadic. Whole genome sequencing (WGS) analysis of *B. melitensis* strains can provide high-quality information for genotyping, monitoring and trace-back purposes in epidemiological studies. This research aim is to determine the genetic relatedness of local *B. melitensis* strains compared to strains from other geographical locations and to study the genetic variation of these local *B. melitensis* strains isolated from sporadic human infections. In this study, whole genomic analysis of 25 sporadic human *B. melitensis* cases in Malaysia and 75 genome sequences of *B. melitensis* retrieved from the NCBI database. From a total of 100 genome sequences, thirty-four Malaysian strains (25 sporadic isolates in this study, six human isolates from other studies and three animal isolates) can be grouped into genotype II and segregated into a distinct lineage clade of South-East Asian (SEA). Phylogenetics analysis showed that 24 of these 25 strains clustered within a single clade supported by 95% bootstrap. One of the isolate was clustered closely with the Indian isolates strain Bm IND1. The strain also had different SNPs number as compared to 24 sporadic *B. melitensis* strain. MLST typing indicated that 24 out of 25 isolates (96%) are grouped as type ST39, except for strain BMM36/11 which is ST8. WGS analysis indicated that the genome of all 25 strains appears to be highly conserved. However, SNP analysis indicates that the Malaysian strains are diverged from the reference *B. melitensis* 16M strain by about 2234 SNPs. WGS analysis revealed 44 virulence factors genes of which eleven are from the type IV secretion system (*virB1-virB11*). AMR genes *adeF* and *mprF* were detected in all 25 strains. In conclusion, the WGS analysis provides useful data with high-quality information and allows us to help identify the source of origin in *B. melitensis* strains, which beneficial to clinicians, public health, researcher and scientist to overcome the trace-back epidemiology investigation.

ACKNOWLEDGEMENT

TABLE OF CONTENTS