Whole genome sequencing data and analysis of a rifampicin-resistant Mycobacterium tuberculosis strain SBH162 from Sabah, Malaysia

ABSTRACT

A Mycobacterium tuberculosis strain SBH162 was isolated from a 49-year-old male with pulmonary tuberculosis. GeneXpert MDR/ RIF identified the strain as rifampicin-resistant M. tuberculosis. The whole genome sequencing was performed using Illumina HiSeq 4000 system to further investigate and verify the mutation sites of the strain through genetic analyses namely variant calling using bioinformatics tools. The de novo assembly of genome generated 100 contigs with N50 of 156,381bp. The whole genome size was 4,343,911 bp with G b C content of 65.58% and consisted of 4,306 predicted genes. The mutation site, S450L, for rifampicin resistance was detected in the rpoB gene. Based on the phylogenetic analysis using the Maximum Likelihood method, the strain was identified as belonging to the Europe America Africa lineage (Lineage 4). The genome dataset has been deposited at DDBJ/ENA/ GenBank under the accession number SMOE00000000.