




Whole-Genome Sequencing of Streptomycin-Resistant *Mycobacterium tuberculosis* Strain SBH145 from Sabah, Malaysia

Zainal Arifin Mustapha,^a Jaeyres Jani,^b Cheronie Shely Stanis,^c Dg Syahidah Nadiah Abdull Majid,^b Chin Kai Ling,^c Roddy Teo,^d  Kamruddin Ahmed^{b,e}

^aDepartment of Medical Education, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

^bBorneo Medical and Health Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

^cDepartment of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

^dTuberculosis and Leprosy Control Unit, Sabah State Health Department, Ministry of Health Malaysia, Kota Kinabalu, Sabah, Malaysia

^eDepartment of Pathology and Microbiology, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

ABSTRACT This paper reports on the whole-genome sequencing of a streptomycin-resistant *Mycobacterium tuberculosis* strain that was isolated from a patient with pulmonary tuberculosis in Sabah state of Malaysian Borneo. The strain belongs to the EAI2-Manila family of lineage 1 and is clustered with *M. tuberculosis* strains from the Philippines, India, and Taiwan.

The burden of tuberculosis is high in Sabah state of Malaysia (1). To date, streptomycin-resistant *Mycobacterium tuberculosis* strains have not been reported. Here, we report for the first time the whole-genome sequence analysis of a streptomycin-resistant *M. tuberculosis* strain from Sabah. Streptomycin is not included in standard treatment regimens but is used in retreatment and as a first-line agent against drug-resistant tuberculosis (2).

Tuberculosis was diagnosed in an 81-year-old female patient from Kota Kinabalu, Sabah, Malaysia, with the GeneXpert MTB/RIF test. The strain was grown in a Bactec MGIT 320 system. Using a Bactec MGIT 960 SIRE kit with streptomycin, isoniazid, rifampin, and ethambutol, the strain was found to be streptomycin resistant. The genomic DNA was extracted using a Masterpure complete DNA and RNA purification kit. The quality was determined with a NanoDrop 2000c spectrophotometer.

DNA libraries were prepared using a NEBNext Ultra kit and sequenced on an Illumina HiSeq 4000 system. The sequencing output was 9,773,850 paired-end 150-bp reads, with a genome coverage of 415 \times . The quality of the sequence reads was checked using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>), and the reads were preprocessed using BMap v38.43 with a Phred score of Q30. SPAdes v3.11.1 (3) was used for *de novo* assembly, which generated 125 contigs, with an N_{50} value of 97,881 bp. The 99% draft genome size was 4,401,196 bp, with a GC content of 65.57%. NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (4) was utilized to annotate the generated contigs.

For variant-calling analysis, the raw sequence reads were aligned to a reference genome (*M. tuberculosis* H37Rv, GenBank accession number [NC_000962.3](https://www.ncbi.nlm.nih.gov/nuccore/NC_000962.3)) by BWA MEM v0.7.1231 (5), in SAM-BAM format. This format was converted into readable sequences and alignment was done using SAMtools v0.1.1932 (6). Next, the Genome Analysis Toolkit v3.4.033 (7) was used for local realignment of the sequence reads. Finally, an average mapping rate of more than 99% with respect to the reference genome was generated for the reports in variant-calling analysis. The following criteria were used to filter variant sites: the quality of alignment had to be >50 bp or the quality of the base had to be >20 bp with >10 reads covering each site. For the annotation of single-

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2022 Mustapha 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 Kamruddin Ahmed, ahmed@ums.edu.my.

The authors declare no conflict of interest.

Received 21 October 2021

Accepted 11 December 2021

Published 6 January 2022

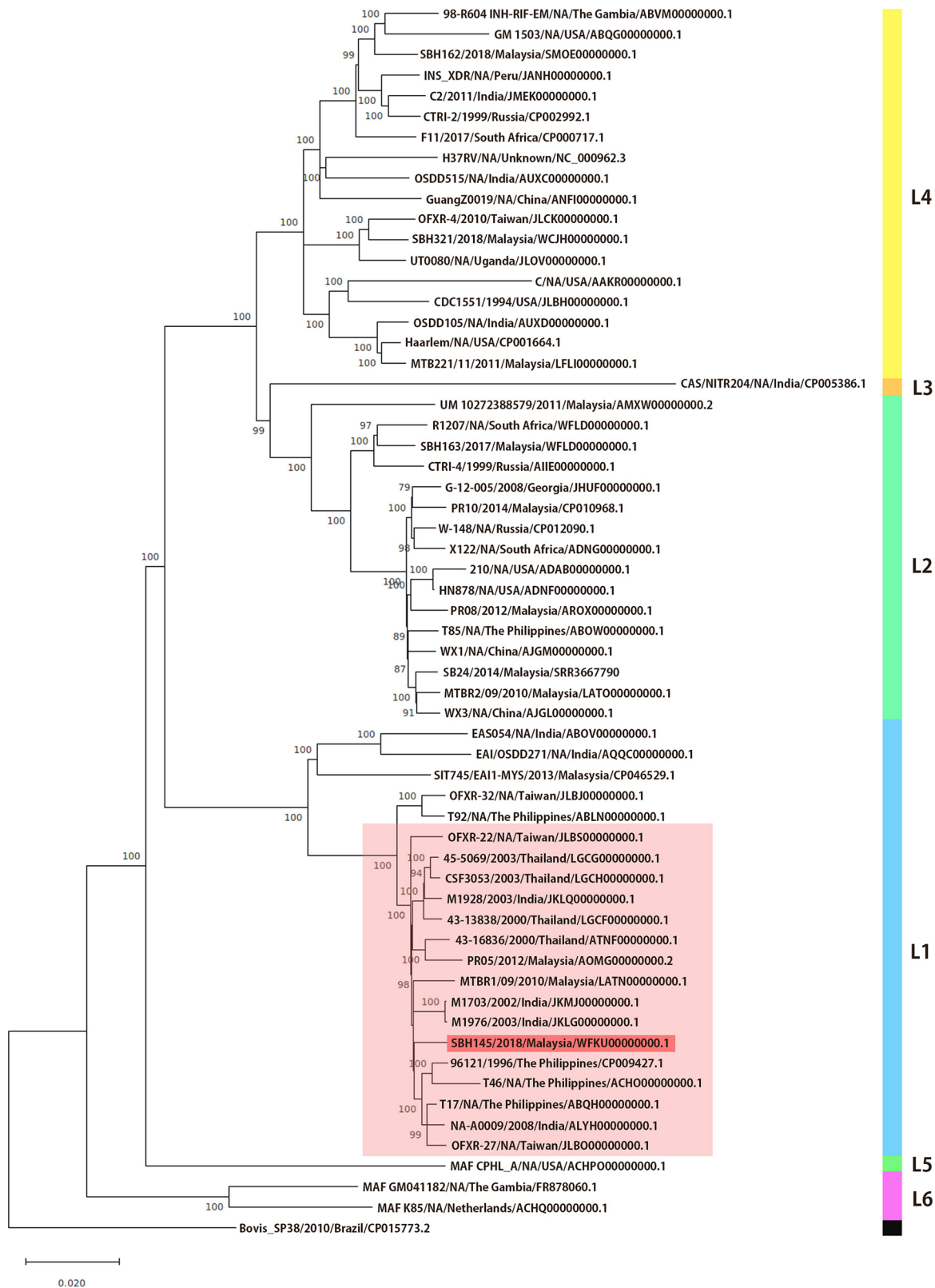


FIG 1 Phylogenetic tree showing that the SBH145 strain belongs to lineage 1 of the EAI2-Manila family and is clustered with strains from Thailand, India, Malaysia, Taiwan, and the Philippines. The phylogenetic tree was constructed using SNP data for the genome sequences of 59 *M. tuberculosis* strains extracted from GenBank; strains are indicated as strain name/country/accession number. *Mycobacterium bovis* strain SP38 was used as an outgroup. The numbers adjacent to nodes represent the bootstrap values; values less than 70% are not shown. The significance of branching was assessed by bootstrap analysis of 1,000 replicates. The scale bar shows the genetic distance, which is expressed as nucleotide substitutions per site.

nucleotide polymorphisms (SNPs), SnpEff v4.134 (8) was used. Moreover, the variant-calling analysis identified a mutation site in the *rpsL* gene (AAG→AGG, locus K43R), which confers high-level resistance to streptomycin (9). Streptomycin interferes with 16S rRNA and interferes with translation proofreading, thereby inhibiting protein synthesis (10). Ribosomal protein S12, encoded by *rpsL*, stabilizes the highly conserved pseudoknot structure formed by 16S rRNA; consequently, amino acid substitutions in RpsL affect the higher-order structure of 16S rRNA and confer streptomycin resistance (10).

For SNP-based phylogenetic analysis, core SNPs were generated by the kSNP3 package (11), nucleotide sequences were aligned, and the maximum likelihood method with the general time-reversible model was used (12). Strain SBH145 clustered with strains from peninsular Malaysia, the Philippines, Thailand, India, and Taiwan (Fig. 1), which are not streptomycin resistant, and it belongs to the Manila clade of lineage 1 of the *M. tuberculosis* complex (13–15). Default parameters were used for all software except where otherwise noted.

Data availability. Raw reads have been deposited in the NCBI SRA under accession number [SRR10204507](https://www.ncbi.nlm.nih.gov/sra/SRR10204507), with BioSample accession number [SAMN12878104](https://www.ncbi.nlm.nih.gov/biosample/SAMN12878104) and BioProject accession number [PRJNA575111](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA575111). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession number [WFKU000000001](https://www.ncbi.nlm.nih.gov/genbank/WFKU000000001).

ACKNOWLEDGMENTS

This work was supported by the Trans-Disciplinary Research Grant Scheme (grant TRG009-2016) from the Ministry of Education Malaysia and a UMGreat Research Grant (grant GUG0288-2/2018) from Universiti Malaysia Sabah.

REFERENCES

- Rashid Ali MR, Parameswaran U, William T, Bird E, Wilkes CS, Lee WK, Yeo TW, Anstey NM, Ralph AP. 2015. A prospective study of tuberculosis drug susceptibility in Sabah, Malaysia, and an algorithm for management of isoniazid resistance. *J Trop Med* 2015:261925. <https://doi.org/10.1155/2015/261925>.
- Hlaing YM, Tongtawe P, Tapchaisri P, Thanongsaksrikul J, Thawornwan U, Archanachan B, Srimanote P. 2017. Mutations in streptomycin resistance genes and their relationship to streptomycin resistance and lineage of *Mycobacterium tuberculosis* Thai isolates. *Tuberc Respir Dis* 80:159–168. <https://doi.org/10.4046/trd.2017.80.2.159>.
- 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>.
- 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>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Susan JL, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms. *Fly (Austin)* 6:80–92. <https://doi.org/10.4161/fly.19695>.
- Villellas C, Aristimuno L, Vitoria MA, Prat C, Blanco S, Garcia de Viedma D, Dominguez J, Samper S, Ainsa JA. 2013. Analysis of mutations in streptomycin-resistant strains reveals a simple and reliable genetic marker for identification of the *Mycobacterium tuberculosis* Beijing genotype. *J Clin Microbiol* 51:2124–2130. <https://doi.org/10.1128/JCM.01944-12>.
- Meier A, Sander P, Schaper KJ, Scholz M, Bottger EC. 1996. Correlation of molecular resistance mechanisms and phenotypic resistance levels in streptomycin-resistant *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 40:2452–2454. <https://doi.org/10.1128/AAC.40.11.2452>.
- Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Majid M, Kumar N, Qureshi A, Yerra P, Kumar A, Kumar MK, Tiruvayipati S, Baddam R, Shaik S, Srikantam A, Ahmed N. 2014. Genomes of two clinical isolates of *Mycobacterium tuberculosis* from Odisha, India. *Genome Announc* 2:e00199-14. <https://doi.org/10.1128/genomeA.00199-14>.
- Phelan JE, Lim DR, Mitarai S, de Sessions PF, Tujan MAA, Reyes LT, Medado IAP, Palparan AG, Mohamed Naim AN, Jie S, Segubre-Mercado E, Simoes B, Campino S, Hafalla JC, Murase Y, Morishige Y, Hibberd ML, Kato S, Ama MCG, Clark TG. 2019. *Mycobacterium tuberculosis* whole genome sequencing provides insights into the Manila strain and drug-resistance mutations in the Philippines. *Nature* 9:9305. <https://doi.org/10.1038/s41598-019-45566-5>.
- Mukherjee S. 2017. Illumina whole genome shotgun sequencing of genomic DNA paired-end library ‘Solexa-189612’ containing sample ‘OFXR-27.’ NCBI BioProject PRJNA221515. <https://www.ncbi.nlm.nih.gov/bioproject/221515>.