Draft Genome Sequence of a Clinical Isolate of *Mycobacterium tuberculosis* Strain PR05

Azwin Ismail,ᵃ L. K. Teh,ᵃ Y. F. Ngeow,ᵇ M. N. Norazmi,ᶜ·ᵈ Zainuddin F. Zainul,ᶜ T. H. Tang,ᵉ N. Najimudin,ᶠ M. Z. Sallehᵃ

Pharmacogenomics Centre (PROMISE), Faculty of Pharmacy, Universiti Teknologi MARA Malaysia, Puncak Alam Campus, Selangor, Malaysiaᵃ; Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysiaᵇ; School of Health Sciences, Health Campus, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysiaᶜ; Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, Pulau Pinang, Malaysiaᵈ; Advanced Medical and Dental Institute, Universiti Sains Malaysia, Kepala Batas, Pulau Pinang, Malaysiaᵉ; School of Biological Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysiaᶠ

We report the annotated genome sequence of a clinical isolate, *Mycobacterium tuberculosis* strain PR05, which was isolated from the human cerebrospinal fluid of a patient diagnosed with tuberculosis.

Address correspondence to M. Z. Salleh, zakisalleh@puncakalam.uitm.edu.my.

Tuberculosis meningitis (TM) is a severe manifestation of extrapulmonary tuberculosis in humans caused by *Mycobacterium tuberculosis*. It is a slow progressive granulomatous inflammation of the basal meninges that can lead to various complications such as cerebral vascular infarction, hydrocephalus, and cranial nerve palsy (1). TM is also known to cause fatal tuberculosis disease associated with a high frequency of neurologic sequelae and mortality (2–6).

Here we report a draft genome sequence of a clinical isolate of *Mycobacterium tuberculosis*, strain PR05, isolated from the human cerebrospinal fluid of a patient diagnosed with tuberculosis. The whole-genome sequencing of *Mycobacterium tuberculosis* strain PR05 was performed using an Illumina Genome Analyzer IIx sequencer with sequencing by synthesis (SBS) technology (7). A total of 25,189,322 nucleotides were generated and 62,307 reads were trimmed off by 99.75%, providing 205.13-fold genome coverage and four megabases of genome size. These reads were assembled using CLCbio (CLC Genomics Workbench version 5.5.1). OSlay was used for ordering and sorting out the assemblies of *M. tuberculosis* strain PR05 (8). The draft genome was annotated using the automated annotation server RAST (Rapid Annotation using Subsystem Technology) (9) and BASys (Bacterial Annotation System) (10).

In *de novo* assembly, the generated 25,065,512 reads covered a total of 887,991,375 bp. The genome of PR05 has a G+C content of 65.6% with approximately 1,412,995 read counts. The number of contigs derived from the *de novo* assembly was 225, ranging from 200 to 181,900 bp with an average length of 19,243 bp and an N_{50} of 64,609 bp. The number of contigs greater than 500 bp was 180. Strain PR05 was mapped to H37Rv (11) and the CCDC5180 reference strain (12). The PR05 genome is a circular DNA of 4,410,213 bp, which is 1,319 bp shorter than the virulent H37Rv strain and 4,232 bp longer than the Beijing CCDC5180 strain. The genome was most similar to strain H37Rv, with 99% similarity in the nucleotide sequence over the alignable portions by using fast sequence alignment tools BLAST (13).

The annotation for the PR05 genome sequence comprises 4,184 coding sequences and 44 RNA genes. We found 394 subsystems, compared to 368 subsystems found in the H37Rv strain. We detected 109 protein-encoding genes, which may play a role in virulence and invasion. Another 86 protein-encoding genes that are related to invasion and intracellular resistance and 22 protein-encoding genes that are involved in antibiotic resistance were detected. A total of 4,771 genes were identified and 4,768 genes were annotated through the BASys.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AOMG00000000. The version described in this paper is the second version, AOMG02000000.

ACKNOWLEDGMENTS

REFERENCES

1. **Christensen ASH, Andersen AB, Thomsen VO, Andersen PH, Johansen IS.** 2011. Tuberculous meningitis in Denmark: a review of 50 cases. BMC Infect. Dis. **11**:47. doi:10.1186/1471-2334-11-47.
2. **Bidstrup C, Andersen PH, Skinhøj P, Andersen ÅB.** 2002. Tuberculous meningitis in a country with a low incidence of tuberculosis: still a serious disease and a diagnostic challenge. Scand. J. Infect. Dis. **34**:811–814.
3. **Girgis NI, Sultan Y, Farid Z, Mansour MM, Erian MW, Hanna LS, Mateczun AJ.** 1998. Tuberculous meningitis, Abbassia Fever Hospital–Naval Research Unit No. 3–Cairo, Egypt, from 1976 to 1996. Am. J. Trop. Med. Hyg. **58**:28–34.
4. **Kent SJ, Crowe SM, Yung A, Lucas CR, Mijch AM.** 1993. Tuberculous meningitis: a 30-year review. Clin. Infect. Dis. **17**:987–994.
5. **Rock RB, Hu S, Gekker G, Sheng WS, May B, Kapur V, Peterson PK.** 2005. Mycobacterium tuberculosis-induced cytokine and chemokine expression by human microglia and astrocytes: effects of dexamethasone. J. Infect. Dis. **192**:2054–2058.
6. **Verdon R, Chevret S, Laissy JP, Wolff M.** 1996. Tuberculous meningitis in adults: review of 48 cases. Clin. Infect. Dis. **22**:982–988.
7. **Li H, Ruan J, Durbin R.** 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. **18**:1851–1858. doi:10.1101/gr.078212.108.

8. **Richter DC, Schuster SC, Huson DH.** 2007. OSLay: optimal syntenic layout of unfinished assemblies. Bioinformatics **23:**1573–1579.

9. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75.

10. **Van Domselaar GH, Stothard P, Shrivastava S, Cruz JA, Guo A, Dong X, Lu P, Szafron D, Greiner R, Wishart DS.** 2005. BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res. **33:** W455–W459.

11. **Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG.** 1998. Deciphering the biology of *Mycobacterium* tuberculosis from the complete genome sequence. Nature **393:**537–544.

12. **Zhang Y, Chen C, Liu J, Deng H, Pan A, Zhang L, Zhao X, Huang M, Lu B, Dong H, Du P, Chen W, Wan K.** 2011. Complete genome sequences of Mycobacterium tuberculosis strains CCDC5079 and CCDC5080, which belong to the Beijing family. J. Bacteriol. **193:** 5591–5592.

13. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.