



Whole-Genome Sequences of *Mycobacterium tuberculosis* TB282 and TB284, a Widespread and a Unique Strain, Respectively, Identified in a Previous Study of Tuberculosis Transmission in Central Los Angeles, California, USA

Lixin Zhang,^{a,b} Zhenhua Yang^c

Department of Epidemiology & Biostatistics, Michigan State University, East Lansing, Michigan, USA^a;
Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA^b; Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA^c

ABSTRACT We report here the genome sequences of two *Mycobacterium tuberculosis* clinical isolates previously identified in central Los Angeles, CA, in the 1990s using a PacBio platform. Isolate TB282 represents a large-cluster strain that caused 27% of the tuberculosis cases, while TB284 represents a strain that caused disease in only one patient.

Tuberculosis transmission has been associated with both patient- and environmental-related risk factors, such as the ability to generate infectious aerosols by the source patient, environmental conditions that affect the dispersal of aerosols, and susceptibility of exposed individuals (1). Molecular epidemiological studies have also shown that some *Mycobacterium tuberculosis* strains cause disease in many persons, whereas others only cause disease in a single or a limited number of patients (2–6). Why *M. tuberculosis* strains vary widely in the extent to which they cause disease in the population cannot always be explained by these epidemiological factors. In a previous study of tuberculosis transmission in central Los Angeles, CA, it was found that one *M. tuberculosis* strain, designated strain 210, was responsible for more than one-quarter of the tuberculosis cases (5). This large-cluster strain was found to grow more rapidly in human macrophages than strains that caused disease in a small number of persons or in a single patient who frequented the identified tuberculosis transmission sites when the patient was highly infectious (7). These findings suggest the potential importance of bacterial factors in tuberculosis transmission and pathogenesis. Here, we sequenced isolates TB282 and TB284 that represent strain 210 and the single-case-causing strain, respectively.

Genomic DNA was isolated using the cetyltrimethylammonium bromide (CTAB) method (8). Each genome was sequenced on a PacBio RSII platform with P5-C3 chemistry using one single-molecule real-time (SMRT) cell and a 10-kb insert library. *De novo* genome assemblies were created using PacBio's SMRT Portal (version 2.2.0) and the Hierarchical Genome Assembly Process (HGAP) (9), with default settings and a seed read cutoff length of 5,000 bp. Annotation of the genomes was done using the RAST annotation server (10).

The single SMRT cell sequencing resulted in a total of 61,505 reads, with a mean read length of 7,292 bp (N_{50} , 11,370 bp) and an average coverage of $82.95\times$ for TB282, and

Received 1 November 2016 Accepted 8 November 2016 Published 12 January 2017

Citation Zhang L, Yang Z. 2017. Whole-genome sequences of *Mycobacterium tuberculosis* TB282 and TB284, a widespread and a unique strain, respectively, identified in a previous study of tuberculosis transmission in central Los Angeles, California, USA. *Genome Announc* 5:e01418-16. <https://doi.org/10.1128/genomeA.01418-16>.

Copyright © 2017 Zhang and Yang. 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 Lixin Zhang, lxzhang@msu.edu.

a total 38,740 reads with a mean read length of 7,239 bp (N_{50} , 11,083 bp) and an average coverage of $52.73\times$ for TB284. Using filtered reads, a genome assembly with a single contig of 4,424,438 bp was obtained for TB282, and for TB284, an assembly with three contigs of 4,394,243 bp, 4,245 bp, and 1,737 bp, respectively, was obtained.

Both genome sequences have a mean G+C content of 65.6% and contain 45 tRNA genes and three rRNA genes. TB282 has 4,285 protein-coding regions, while TB284 has 4,230 protein-coding regions. The two genomes were compared by performing whole-genome alignment using Mauve (11) to identify sequence differences. A total of 2,147 single nucleotide polymorphisms (SNPs) were identified. In addition, there are 78 insertion or deletion differences greater than 100 bp that exist between the two genomes.

The genome sequences of this pair of isolates, one representing a widespread strain and the other a unique strain, will provide an additional resource for comparative genomic analyses in studying the bacterial factors important for tuberculosis transmission.

Accession number(s). The genome sequences of the *M. tuberculosis* strains TB282 and TB284 have been deposited at DDBJ/ENA/GenBank under GenBank accession numbers [CP017920](#) and [MNBY00000000](#), respectively.

ACKNOWLEDGMENTS

This research was supported by the University of Michigan internal fund. The PacBio sequencing was performed at Sequencing Core at the University of Michigan Biomedical Research Core Facilities. We thank Dr. M. Donald Cave for assistance in reviewing background information about the two study strains.

REFERENCES

1. CDC. 2013. Core curriculum on tuberculosis: what the clinician should know, 6th ed. Centers for Disease Control and Prevention, Atlanta, GA. <http://www.cdc.gov/tb/education/corecurr/default.htm>.
2. Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, Drucker E, Bloom BR. 1994. Transmission of tuberculosis in New York City: an analysis by DNA fingerprinting and conventional epidemiologic methods. *N Engl J Med* 330:1710–1716. <https://doi.org/10.1056/NEJM199406163302403>.
3. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, Schechter GF, Daley CL, Schoolnik GK. 1994. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 330:1703–1709. <https://doi.org/10.1056/NEJM199406163302402>.
4. Burman WJ, Reeves RR, Hawkes AP, Rietmeijer CA, Yang Z, el-Hajj H, Bates JH, Cave MD. 1997. DNA fingerprinting with two probes decreases clustering of *Mycobacterium tuberculosis*. *Am J Respir Crit Care Med* 155:1140–1146. <https://doi.org/10.1164/ajrccm.155.3.9117000>.
5. Barnes PF, Yang Z, Preston-Martin S, Pogoda JM, Jones BE, Otaya M, Eisenach KD, Knowles L, Harvey S, Cave MD. 1997. Patterns of tuberculosis transmission in central Los Angeles. *JAMA* 278:1159–1163.
6. Verhagen LM, van den Hof S, van Deutekom H, Hermans PW, Kremer K, Borgdorff MW, van Soolingen D. 2011. Mycobacterial factors relevant for transmission of tuberculosis. *J Infect Dis* 203:1249–1255. <https://doi.org/10.1093/infdis/jir013>.
7. Zhang M, Gong J, Yang Z, Samten B, Cave MD, Barnes PF. 1999. Enhanced capacity of a widespread strain of *Mycobacterium tuberculosis* to grow in human macrophages. *J Infect Dis* 179:1213–1217. <https://doi.org/10.1086/314738>.
8. Belisle JT, Sonnenberg MG. 1998. Isolation of genomic DNA from mycobacteria, p 31–44. *In* Parish T, Stoker NG (ed.), *Methods in molecular biology*, vol 101. Mycobacteria protocols. Humana Press, Totowa, NJ.
9. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
10. 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. <https://doi.org/10.1186/1471-2164-9-75>.
11. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.