JB Journals.ASM.org

GENOME ANNOUNCEMENT

Complete Annotated Genome Sequence of *Mycobacterium tuberculosis* Erdman

Tohru Miyoshi-Akiyama,[a] Kazunori Matsumura,[a] Hiroki Iwai,[a] Keiji Funatogawa,[b] and Teruo Kirikae[a]

Department of Infectious Diseases, Research Institute, National Center for Global Health and Medicine, Shinjuku, Tokyo, Japan,[a] and Tochigi Prefectural Institute of Public Health and Environmental Science, Kawachi-machi, Tochigi, Japan[b]

We report the completely annotated genome sequence of *Mycobacterium tuberculosis* Erdman (TMC 107; ATCC 35801), which is a well-known laboratory strain of *M. tuberculosis*.

Mycobacterium tuberculosis strain Erdman was isolated from human sputum by William H. Feldman in 1945, at Mayo Clinic, Rochester, MN, and deposited with the Trudeau Mycobacterium Culture Collection in 1946 (7). There is no description of the naming of "Erdman" in reference 7. Due to its consistently high virulence (7), it has been widely used as a standard virulent laboratory strain for virulence and immunization studies. *M. tuberculosis* Erdman has a faster *in vivo* doubling time than two attenuated strains, *M. tuberculosis* H37Ra and *Mycobacterium bovis* Bacillus Calmette-Guerin (BCG), and a slightly faster *in vivo* doubling time than the virulent H37Rv strain in mice (5). The *M. tuberculosis* Erdman strain showed the same level of virulence as *M. tuberculosis* CDC1551 in C57BL/6 mice (2). We recently found that strain Erdman was quite virulent compared with a clinical isolate, *M. tuberculosis* NCGM2209 (4), in BALB/c mice (strain Erdman 50% lethal dose [LD_{50}], 1×10^3/mouse; strain NCGM2209 LD_{50}, $>2 \times 10^7$/mouse) (data not shown; estimated from 180-day-mortality data by the method of Reed and Muench [6]).

First, the *M. tuberculosis* Erdman genome was analyzed using a GS FLX Titanium sequencer (Roche) with an 8-kb pair-end library prepared from the genome. This generated 976,847 reads, covering 358,569,827 bp, which were assembled into scaffolds and contigs by GS De Novo Assembler 2.6 (Newbler; Roche). Then, gap filling was performed using conventional Sanger sequencing of the PCR fragments, brute-force PCR for the contigs and scaffolds, and an ABI 3730xl DNA sequencer. Finally, 1,274,470-kb single-end reads obtained with Genome Analyzer IIx (Illumina) were used to add to the draft genome sequence by the use of Maq software (3). Primary coding sequence extractions and initial function assignments were performed using the automated annotation server RAST (Rapid Annotation using Subsystem Technology) (1). The results were compared to verify the annotation and were corrected manually by *in silico* molecular cloning (In Silico Biology Inc., Kanagawa, Japan). The *M. tuberculosis* Erdman genome consists of a single circular chromosome of 4,392,353 bp, with an average GC content of 65.6%. The chromosome was shown to contain a total of 4,246 protein-coding genes, 52 tRNA genes, one transfer mRNA for all amino acids, and 1 *rrn* operon. In addition, the chromosome harbors 11 IS*6110* sequences.

Nucleotide sequence accession number. Nucleotide sequences of the chromosome of *M. tuberculosis* Erdman have been deposited in the DNA Database of Japan under accession no. AP012340.

ACKNOWLEDGMENTS

This study was supported by Health Sciences Research grants (H21-SHINKO-IPPAN-016). T.K. and T.M.-A. were supported by Grants for International Health Research from the Ministry of Health, Labor, and Welfare of Japan (21A-6 and 23A-301, respectively). We thank N. Saito, S. Suzuki, and Y. Sakurai for their excellent work in the genome analysis.

REFERENCES

1. **Aziz RK, et al.** 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
2. **Kelley CL, Collins FM.** 1999. Growth of a highly virulent strain of Mycobacterium tuberculosis in mice of differing susceptibility to tuberculous challenge. Tuber. Lung Dis. **79**:367–370.
3. **Li H, Ruan J, Durbin R.** 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. **18**:1851–1858.
4. **Miyoshi-Akiyama T, Matsumura K, Kobayashi N, Maeda S, Kirikae T.** 2011. Genome sequence of clinical isolate Mycobacterium tuberculosis NCGM2209. J. Bacteriol. **193**:6792.
5. **North RJ, Izzo AA.** 1993. Mycobacterial virulence. Virulent strains of Mycobacteria tuberculosis have faster in vivo doubling times and are better equipped to resist growth-inhibiting functions of macrophages in the presence and absence of specific immunity. J. Exp. Med. **177**:1723–1733.
6. **Reed LJ, Muench GP.** 1938. A simple method of estimating fifty per cent end points. Am. J. Epidemiol. (London) **3**:493–497.
7. **Trudeau Institute.** 1972. TMC #107, p 13. *In* Trudeau mycobacterial culture collection, U.S.-Japan Cooperative Medical Science Program, Geographic Medical Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland. Trudeau Institute, Saranac Lake, NY.

Received 4 March 2012 Accepted 7 March 2012

Address correspondence to Tohru Miyoshi-Akiyama, takiyam@ri.ncgm.go.jp.

doi:10.1128/JB.00353-12