Complete Genome Sequence of Mycobacterium intracellulare Strain ATCC 13950T

Byoung-Jun Kim,a Beom-Soon Choi,b Jong-Sung Lim,b Ik-Young Choi,b Je-Hee Lee,c Jong sik Chun,c,d Yoon-Hoh Kook,a and Bum-Joon Kimb

Department of Microbiology and Immunology, Liver Research Institute, Cancer Research Institute and SNUMRC, College of Medicine, Seoul National University, Seoul, Republic of Korea; National Instrumentation Center for Environmental Management, Seoul National University, Seoul, Republic of Korea; Chunlab, Inc., Seoul National University, Seoul, Republic of Korea; and School of Biological Sciences, Seoul National University, Seoul, Republic of Korea

Here we report the first complete genome sequence of Mycobacterium intracellulare ATCC 13950T, a Mycobacterium avium complex (MAC) strain. This genome sequence will serve as a valuable reference for understanding the epidemiologic, biological, and pathogenic aspects of the disparity between MAC members.

Among the slow-growing mycobacteria responsible for opportunistic infections, members of the Mycobacterium avium complex (MAC) are the nontuberculous mycobacteria most frequently isolated in clinical settings (5–7). Traditionally, the MAC includes two species, M. avium and Mycobacterium intracellulare (1, 3, 6). There are definitely distinct disparities between these two MAC members in epidemiologic, biological, and pathogenic aspects. Currently, of the MAC strains, the complete genome sequences of two MAC strains, M. avium subsp. avium 104 (CP0000479) and M. avium subsp. paratuberculosis (AE016958) (4) and the partial genome sequence of M. colombiense (AFVW00000000) (2), closely related to M. avium, are available. However, the complete genome of M. intracellulare has not yet been determined. To better understand the pathogenic mechanism of M. intracellulare, we report the complete, annotated genome sequence of M. intracellulare ATCC 13950T in the present study.

The M. intracellulare genome was sequenced by a standard shotgun strategy using GS FLX pyrosequencing technology. Sequencing analysis was performed at the National Instrumentation Center for Environmental Management (Genome Analysis Unit) at Seoul National University. A total of 921,179 reads were generated, with an average length of 400, yielding a total sequence of 368,366,484 bp. This represents 68× coverage of the estimated 5.4-Mb genome. The obtained 124 contigs were compared for mapping to the whole-genome sequences of reference strain using the BLASTZ program (http://www.psc.edu/general/software/packages/blastz/). All of the remaining gaps between contigs were completely filled by ~50-fold Solexa reads and PCR amplifications. Genome annotation was performed using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html).

Our data on the M. intracellulare genome show it to have a circular DNA of 5,402,402 bp, which is larger than the genome of M. avium subsp. paratuberculosis (4.8 Mb) and contains more protein coding genes (5,145 versus 4,400) and more tRNA genes (47 versus 45). The genome of M. intracellulare ATCC 13950T has a G+C content of 68.10%, and no plasmid was found. M. intracellulare is known to form a close cluster with M. avium in a phylogenetic analysis based on the 16S rRNA gene sequence. Our phylogenetic analysis based on the complete genome sequences in the NCBI microbial sequence database also supported the close relationships of M. intracellulare with M. avium subsp. avium 104 and M. avium subsp. paratuberculosis. The genome sequence reported here will serve as a valuable reference for understanding the epidemiologic, biological, and pathogenic aspects of the disparity between MAC members.

Nucleotide sequence accession number. The whole-genome sequence of M. intracellulare ATCC 13950T has been deposited in the GenBank database under accession number CP003322.

ACKNOWLEDGMENT

This work was supported by a National Research Foundation of Korea grant funded by the Korean Government Ministry of Education, Science and Technology (2010-0014269).

REFERENCES