Draft Genome Sequence of Mycobacterium avium 11

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1128/genomeA.00766-17</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:34375264">http://nrs.harvard.edu/urn-3:HUL.InstRepos:34375264</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Draft Genome Sequence of
Mycobacterium avium 11

Michelle Yee,a David Klinzing,b Jun-Rong Wei,c,d Martin Gengenbacher,a,e Eric J. Rubin,c,d Jung-Yien Chien,f Po-Ren Hsueh,f,g Thomas Dick,a,e

Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore; AITbiotech Singapore, Singapore, Singapore; Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA; Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA; Public Health Research Institute, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, New Jersey, USA; Department of Internal Medicine, National Taiwan University Hospital, National Taiwan University College of Medicine, Taipei, Taiwan; Department of Laboratory Medicine, National Taiwan University Hospital, National Taiwan University College of Medicine, Taipei, Taiwan

ABSTRACT

Mycobacterium avium accounts for most lung disease caused by nontuberculous mycobacteria (NTM). The lack of effective chemotherapy calls for the discovery of new drugs. Here, we report the draft genome sequence of M. avium 11, a clinical isolate used as a screening strain for NTM-focused drug discovery.

The slow-growing nontuberculous mycobacterium (NTM) Mycobacterium avium represents a complex (M. avium complex [MAC]) of at least four subspecies. Members of the MAC are opportunistic pathogens ubiquitous in nature. The organisms have been isolated from soil and water samples and infect a diverse range of hosts, including birds, swine, ruminants, and humans (1). MAC pathogens cause the majority of all NTM infections in the United States and other countries, with M. avium subsp. hominisuis being of high clinical significance (1, 2). M. avium subsp. hominisuis can cause lymphadenitis, pulmonary and soft tissue infections, as well as disseminated disease (2, 3). The absence of effective multidrug regimens renders treatment of MAC infections difficult (2). Disseminated infections are commonly observed in patients with AIDS and are very challenging to cure due to adverse drug effects and drug–drug interaction with background HIV therapy (2, 4). Hence, there is an urgent need for new effective drugs with minimal side effects and no drug–drug interaction. Here, we report the draft genome sequence of M. avium 11, which is used as a screening strain in our ongoing NTM-focused drug discovery program (5–7).

M. avium 11 was isolated from the bone marrow of an AIDS patient with disseminated MAC infection, including pulmonary infection (National Taiwan University Hospital, Taipei, Taiwan). This clinical isolate forms smooth, opaque, and domed colonies (3) on Middlebrook 7H11 agar after 3 weeks of incubation at 37°C. M. avium 11 is classified as M. avium subsp. hominisuis based on the 3’ region of the hsp65 gene sequence (8) that shows 100% identity when aligned with the hsp65 3’ region of M. avium subsp. hominisuis 104 (GenBank accession no. NC_008595).

Genomic DNA of M. avium 11 was extracted from a log-phase liquid culture using the phenol-chloroform method as described previously (9). The DNA libraries were prepared using Covaris shearing and Illumina TruSeq Nano DNA library preparation kits. Two sequencing runs of 2 × 300-bp read lengths were performed using an Illumina MiSeq platform, generating a total of 8,879,671 paired-end reads (AITbiotech, Singapore). These paired-end reads were quality checked using FastQC and subsequently quality trimmed using fqtrim (https://ccb.jhu.edu/software/fqtrim) with a window size of 7, a minimum average Q score of 28, and a minimum posttrim length of 35. The resulting 8,386,266 paired-end reads after trimming were de novo assembled using
SPAdes version 3.6.2 (10) (k-mer sizes of 33, 55, 77, 99, and 127), producing 70 contigs with an N_{50} contig size of 216,071 bp. This assembled draft genome is 5,448,889 bp in size with a GC content of 68.99%. Annotations were performed using the Rapid Annotations using Subsystems Technology (RAST) server version 2 (11) predicted a total of 5,141 coding sequences and 49 RNAs, which includes 46 tRNAs and 3 rRNAs. Interestingly, PHASTER (12) predicted two prophage regions in the genome, including one partial prophage containing 68 coding sequences (65.56% GC content) and one intact prophage containing 56 coding sequences (66.74% GC content).

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number NISH00000000. The version described in this paper is the first version, NISH01000000.

**ACKNOWLEDGMENTS**

This work was supported by the Singapore Ministry of Health’s National Medical Research Council under its Translational Clinical Research flagship grant NMRC/TCR/011-NUHS/2014 and is part of the Singapore Programme of Research Investigating New Approaches to Treatment of Tuberculosis (SPRINT-TB; http://www.sprinttb.org) led by Nick Paton.

**REFERENCES**