Complete Genome Sequences of *Mycobacterium tuberculosis* Isolates Subjected to 200 Continuous Passages

Lucio Vera-Cabrera, Carmen A. Molina-Torres, Jose Marco Lopez-Ortiz, Jorge Castro-Garza, Carlos Cordova-Fletes, Jorge Ocampo-Candiani

Laboratorio Interdisciplinario de Investigación Dermatológica, Servicio de Dermatología, Hospital Universitario “Dr. José Eleuterio González,” Universidad Autónoma de Nuevo Leon, Monterrey, NL, México
Centro de Investigación Biomédica del Noreste, Instituto Mexicano del Seguro Social, Monterrey, NL, México
Departamento de Bioquímica, Facultad de Medicina, Universidad Autónoma de Nuevo Leon, Monterrey, NL, México

**ABSTRACT** We report 6 draft genome sequences corresponding to *Mycobacterium tuberculosis* H37Rv and *M. tuberculosis* DR689, a Beijing isolate, plus their counterparts subjected to 200 continuous passages in Middlebrook 7H9 broth, either alone or with ox bile.

*M. tuberculosis* is still a public health problem worldwide (1). In Mexico, the number of new pulmonary and meningeal tuberculosis cases has increased in the last several years (2). New drugs and immunoprophylactic media are needed to control this disease. Considering that continuous passaging produces important biological changes in microorganisms (3, 4), we decided to subculture *Mycobacterium tuberculosis* isolates. Although *M. tuberculosis* and *M. bovis* belong to the same bacterial complex, they have important biological differences (5).

In our project, we subcultured isolates of different lineages of *M. tuberculosis* (6), including MTB H37Rv and a Beijing isolate, DR689. The only isolate that has already been sequenced is strain H37Rv, which constitutes the reference genome sequence used worldwide (GenBank accession number NC_000962.3). This strain was donated to J.C.-G. in 1998 by Fred Quinn, when he was a research microbiologist at the Tuberculosis and Mycobacteriology Branch, Division of AIDS, STD, and TB Laboratory Research, Centers for Disease Control and Prevention (Atlanta, GA). It was cryopreserved at −70°C until we started the subculture. Both strains MTB H37Rv and MTBDR689 were subjected to 200 continuous passages on M7H9-Tween 80-oleic acid-albumin-dextrose-catalase (OADC); because Calmette and Guerin used ox bile to better suspend the bacteria, we processed some samples with and some without bile. We use the following nomenclature to refer to our cultures: *M. tuberculosis* H37Rv parental strain (MTB H37RvP), without bile (MTB H37Rv-P200) and with bile (MTB H37Rv-P200B), and *M. tuberculosis* DR689 parental strain (MTBDR689P) without bile (MTBDR689-P200) and with bile (MTBDR689P-200B).

DNA was extracted from the colonies grown on Lowenstein-Jensen agar using the cetyltrimethylammonium bromide (CTAB)-NaCl technique with modifications described previously (6). Subsequently, high-throughput sequencing was carried out using the Illumina MiSeq genome analyzer (San Diego, CA, USA). Sequencing libraries were prepared using the NuGEN Ultralow v2 kit or the Nextera XT v2 kit following the manufacturer’s recommended protocol. The libraries were sequenced on a NextSeq 500 instrument using midoutput v2 chemistry (2 × 150 bp) or on a MiSeq instrument using V2 chemistry (2 × 250 bp); the fastq reads obtained were mapped against the *M. tuberculosis* H37Rv reference genome sequence reported in GenBank (accession number NC_000962.3) using the BWA-MEM aligner (7) from the Sequencher v5.4.6 suite.
### TABLE 1 Draft genome characteristics and accession numbers of M. tuberculosis H37RvP, M. tuberculosis DR689P, and their counterparts, subjected to 200 continuous passages with or without bile

<table>
<thead>
<tr>
<th>Strain/isolate</th>
<th>GenBank accession no.</th>
<th>SRA accession no.</th>
<th>Genome size (Mbp)</th>
<th>G+C content (%)</th>
<th>No. of contigs</th>
<th>N50 value (bp)</th>
<th>No. of reads</th>
<th>Coverage (X)</th>
<th>No. of CDSs</th>
<th>% coding proteins</th>
<th>No. of tRNAs</th>
<th>No. of rRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. tuberculosis H37RvP</td>
<td>VRNC000000000</td>
<td>SRA8374041</td>
<td>4.3</td>
<td>65.55</td>
<td>103</td>
<td>142,434</td>
<td>1,690,242</td>
<td>162</td>
<td>5,489</td>
<td>98.2</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>M. tuberculosis H37Rv-P200</td>
<td>VRND000000000</td>
<td>SRA8374042</td>
<td>4.3</td>
<td>65.55</td>
<td>155</td>
<td>64,842</td>
<td>7,851,072</td>
<td>162</td>
<td>4,378</td>
<td>97.1</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>M. tuberculosis H37Rv-P200B</td>
<td>VRNE000000000</td>
<td>SRA8374043</td>
<td>4.3</td>
<td>65.43</td>
<td>205</td>
<td>57,433</td>
<td>3,899,892</td>
<td>109.5</td>
<td>4,622</td>
<td>93.5</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td>M. tuberculosis DR689P</td>
<td>VRNF000000000</td>
<td>SRA8374038</td>
<td>4.3</td>
<td>65.59</td>
<td>115</td>
<td>105,204</td>
<td>1,373,090</td>
<td>34</td>
<td>4,464</td>
<td>95.8</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>M. tuberculosis DR689-P200</td>
<td>VRNG000000000</td>
<td>SRA8374039</td>
<td>4.3</td>
<td>65.44</td>
<td>230</td>
<td>45,562</td>
<td>2,040,850</td>
<td>58</td>
<td>4,523</td>
<td>94.2</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>M. tuberculosis DR689-P200B</td>
<td>VRNH000000000</td>
<td>SRA8374040</td>
<td>4.3</td>
<td>65.56</td>
<td>160</td>
<td>64,269</td>
<td>5,448,990</td>
<td>81</td>
<td>4,359</td>
<td>96.4</td>
<td>47</td>
<td>1</td>
</tr>
</tbody>
</table>

*CDs, coding DNA sequences.*
The genomes were assembled into contigs using SPAdes Genome Assembler v3.9.0 and deposited in GenBank. Annotations were made using the single Prokaryotic Genome Annotation Pipeline (PGAP) of the NCBI (8).

Data availability. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the following accession numbers: VRNC00000000 (Mycobacterium tuberculosis H37RvP), VRND00000000 (Mycobacterium tuberculosis H37Rv-P200), VRNE00000000 (Mycobacterium tuberculosis H37Rv-P2008), VRNF00000000 (Mycobacterium tuberculosis DR689P), VRNG00000000 (Mycobacterium tuberculosis DR689-P200), and VRNH00000000 (Mycobacterium tuberculosis DR689-P2008). All these sequences belong to BioProject accession number PRJNA559369. The characteristics of the draft genomes are summarized in Table 1.

REFERENCES