




Sequencing and Annotation of the Genome of *Mycobacterium tuberculosis* MYC004, a Strain Causing Meningitis in Mexico

Ana Laura Guillén-Nepita,^{a,b} Andrea Monserrat Negrete-Paz,^{a,c} Gerardo Vázquez-Marrufo,^c Andrés Cruz-Hernández,^d Pablo Fresia,^e Hugo Naya,^e  Ma Soledad Vázquez-Garcidueñas^a

^aDivisión de Estudios de Posgrado, Facultad de Ciencias Médicas y Biológicas Dr. Ignacio Chávez, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico

^bFacultad de Ingeniería, Universidad Autónoma de Querétaro, Centro Universitario, Santiago de Querétaro, Mexico

^cCentro Multidisciplinario de Estudios en Biotecnología, Facultad de Medicina Veterinaria y Zootecnia, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico

^dEscuela de Agronomía, Universidad De La Salle Bajío, León, Guanajuato, Mexico

^eUnidad de Bioinformática, Institut Pasteur de Montevideo, Montevideo, Uruguay

ABSTRACT *Mycobacterium tuberculosis* strain MYC004 was isolated from a Mexican patient with tuberculous meningitis, the most aggressive form of tuberculosis. The draft genome sequence is the first of a meningeal strain of *M. tuberculosis* reported from Latin America and consists of 4,411,530 bp, including 4,251 protein-encoding genes.

It is estimated that one-third of the world's population is infected with *Mycobacterium tuberculosis*, although the majority will never develop an active disease. There are many types of extrapulmonary tuberculosis affecting every organ system in the body, with tuberculous meningitis being the most fatal form, killing approximately one-third of patients despite antituberculosis treatment (1). The majority of the tuberculosis burden and deaths are concentrated in middle-income countries. There are few studies making attempts to relate extrapulmonary forms of the disease to a specific genetic variant of *M. tuberculosis* on a genome-wide scale (2–4). In order to continue efforts to identify a genetic variant that could impact disease severity and tissue tropism, more genome sequences of extrapulmonary strains from different geographic regions need to be analyzed. However, no meningeal isolates of *M. tuberculosis* from Latin America were represented in databases or bioinformatics analyses (4). *M. tuberculosis* strain MYC004 was isolated by the Laboratorio Estatal de Salud Pública, from the cerebrospinal fluid of a male patient diagnosed in 2012 with tuberculous meningitis in Michoacán, Mexico. The strain was cultured at 37°C on Lowenstein-Jensen agar for 8 weeks. Genomic DNA was extracted in accordance with previously described methods (5); it was sequenced using the ABI SOLiD sequencing platform with a 50-bp-long single-end library, and a total of 1,323,759 reads were generated. Raw reads in color space fasta format were trimmed, and a reference-guided assembly was performed using SOLiD BioScope version 1.3 (Applied Biosystems) for mapping the reads to the reference genome *M. tuberculosis* H37Rv (6). Annotation was performed using Prokka (7) and Rapid Annotations using Subsystems Technology (RAST) (8). The genes were validated using the external gene databases Mycobrowser (<https://mycobrowser.epfl.ch/>) (9) and UniProtKB (<http://www.ebi.ac.uk/uniprot>) (10). The annotations were manually reviewed and curated using Artemis (11). The length of the draft genome of *M. tuberculosis* MYC004 is 4,411,530 bp, with a GC content of 64.7%. In total, 4,251 protein-coding genes were annotated, including 51 RNA-coding genes, 48 tRNAs, and 3 rRNAs. Of the total protein-coding genes, 2,955 (69.51%) were functionally assigned,

Received 9 May 2018 Accepted 13 May 2018 Published 21 June 2018

Citation Guillén-Nepita AL, Negrete-Paz AM, Vázquez-Marrufo G, Cruz-Hernández A, Fresia P, Naya H, Vázquez-Garcidueñas MS. 2018. Sequencing and annotation of the genome of *Mycobacterium tuberculosis* MYC004, a strain causing meningitis in Mexico. *Genome Announc* 6:e00523-18. <https://doi.org/10.1128/genomeA.00523-18>.

Copyright © 2018 Guillén-Nepita et al. 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 Ma Soledad Vázquez-Garcidueñas, marisolvaz@yahoo.com. A.L.G.-N. and A.M.N.-P. contributed equally to this work.

and the remaining genes were annotated as hypothetical proteins. A total of 386 subsystems were annotated in the *M. tuberculosis* MYC004 genome using the RAST server, with the main subsystems being amino acids, amino acid derivatives, carbohydrates, cofactors, vitamins, prosthetic groups, and pigments.

A comparison of genome sequences using RAST revealed that the strains most similar to *M. tuberculosis* MYC004 are *M. tuberculosis* NCGM2209 (score, 483), *M. tuberculosis* F11 (score, 427), and *M. tuberculosis* UM 1072388579 (score, 420).

Accession number(s). The draft genome sequence of *M. tuberculosis* MYC004 was deposited in GenBank under the accession number [CP024614](#).

ACKNOWLEDGMENTS

G.V.-M. and A.C.-H. acknowledge receipt of the 193512/2012-02 grant from the Fondo Regional para el Desarrollo Científico, Tecnológico y de Innovación (CONACYT). A.L.G.-N. acknowledges her postdoctoral fellowship granted by CONACYT.

REFERENCES

1. Heemskerck AD, Nguyen MTH, Dang HTM, Vinh Nguyen CVV, Nguyen LH, Do TDA, Nguyen TTT, Wolbers M, Day J, Le TTP, Nguyen BD, Caws M, Thwaites GE. 2017. Clinical outcomes of patients with drug-resistant tuberculous meningitis treated with an intensified antituberculosis regimen. *Clin Infect Dis* 65:20–28. <https://doi.org/10.1093/cid/cix230>.
2. Saw S, Tan J, Chan X, Chan K, Ngeow Y. 2016. Chromosomal rearrangements and protein globularity changes in *Mycobacterium tuberculosis* isolates from cerebrospinal fluid. *PeerJ* 4:e2484. <https://doi.org/10.7717/peerj.2484>.
3. Faksri K, Xia E, Ong RT-H, Tan JH, Nonghanphithak D, Makhao N, Thamnongdee N, Thanormchat A, Phurattanakornkul A, Rattanarangsee S, Ratanajaraya C, Suriyaphol P, Prammananan T, Teo Y-Y, Chairprasert A. 2018. Comparative whole-genome sequence analysis of *Mycobacterium tuberculosis* isolated from tuberculous meningitis and pulmonary tuberculosis patients. *Sci Rep* 8:4910. <https://doi.org/10.1038/s41598-018-23337-y>.
4. Ruesen C, Chaidir L, van Laarhoven A, Dian S, Ganiem AR, Nebenzahl-Guimaraes H, Huynen MA, Alisjahbana B, Dutilh BE, van Crevel R. 2018. Large-scale genomic analysis shows association between homoplastic genetic variation in *Mycobacterium tuberculosis* genes and meningeal or pulmonary tuberculosis. *BMC Genomics* 19:122. <https://doi.org/10.1186/s12864-018-4498-z>.
5. Guillén-Nepita AL, Vázquez-Marrufo G, Blanco-Guillot FT, Figueroa-Aguilar GA, Vázquez-Garcidueñas MS. 2013. RAPD discloses high molecular diversity of *Mycobacterium tuberculosis* from Michoacan, Mexico. *New Microbiol* 36:413–418.
6. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, III, Tekaiia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537–544. <https://doi.org/10.1038/31159>.
7. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
8. 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>.
9. Kapopoulou A, Lew J, Cole S. 2011. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* 91:8–13. <https://doi.org/10.1016/j.tube.2010.09.006>.
10. The UniProt Consortium. 2017. UniProt: the universal protein knowledge-base. *Nucleic Acids Res* 45:D158–D169. <https://doi.org/10.1093/nar/gkw1099>.
11. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.