genomeA
Journals.ASM.org

Draft Genome Sequence of *Mycobacterium asiaticum* Strain DSM 44297

Olivier Croce, Catherine Robert, Didier Raoult, Michel Drancourt

Aix Marseille Université, URMITE, Marseille, France

We report the draft genome sequence of *Mycobacterium asiaticum* strain DSM 44297, a tropical mycobacterium seldom responsible for human infection. The genome of *M. asiaticum* has a size of 5,935,986 bp, with a 66.03% G+C content, encoding 5,591 proteins and 81 RNAs.

Received 25 March 2014 **Accepted** 3 April 2014 **Published** 17 April 2014

Citation Croce O, Robert C, Raoult D, Drancourt M. 2014. Draft genome sequence of *Mycobacterium asiaticum* strain DSM 44297. Genome Announc. 2(2):e00320-14. doi: 10.1128/genomeA.00320-14.

Address correspondence to Michel Drancourt, michel.drancourt@univ-amu.fr.

Mycobacterium asiaticum is a nontuberculous mycobacterium initially described in 1971 (1) and recognized as an organism of medical interest in 1982 (2). *M. asiaticum* has been isolated in monkeys (3) and pigs (4). In humans, *M. asiaticum* isolates have been obtained from respiratory tract specimens (5, 6), lyphadenitis, bursitis, and wound specimens (7), and keratitis samples (8). In respiratory tract specimens, it is mainly recognized as a harmless contaminant but a few patients have been documented with *M. asiaticum* pneumonia (2, 9). Interestingly, this mycobacterium has been mainly documented in subtropical and tropical countries and areas, including Australia (7), California (9), Florida (8), Brazil (5), and Uganda (4).

In order to gain further insights into this organism, we performed whole-genome sequencing of the *M. asiaticum* DSM 44297 strain. Genomic DNA was extracted from mycobacteria grown on Middlebrook 7H10 agar medium at 37°C under a 5% CO_2 atmosphere. A shotgun XL+ library and 3-kb paired-end library were pyrosequenced on a 454_Roche_Titanium platform (Roche-454 Life Sciences, Boulogne-Billancourt, France) (10). This project was loaded on a 1/4 region for each application on picotiter plates. The 454 sequencing generated 362,293 reads assembled into contigs and scaffolds using Newbler version 2.8 (Roche, 454 Life Sciences). Contigs obtained were combined together using Opera software v1.2 (11) and GapFiller v1.10 (12) to reduce the set. Finally, some manual refinements using CLC Genomics v5 software (CLC bio, Aarhus, Denmark) improved the genome. The draft genome sequence of *M. asiaticum* consists of 10 scaffolds of 106 contigs containing 5,910,460 bp, with an estimated genome size including gaps of 5,935,986 bp. Its G+C content is 66.03%.

Noncoding genes and miscellaneous features were predicted using RNAmmer (13), ARAGORN (14), Rfam (15), and PFAM (16). Open reading frames were predicted using Prodigal (17), and functional annotation was achieved using BLASTp against the GenBank database (18) and the Clusters of Orthologous Groups (COG) database (19, 20). The genome was shown to encode at least 81 predicted RNAs, including 3 rRNAs in a single operon, 51 tRNAs, 1 transfer-messenger RNA, and 26 miscellaneous RNAs. A total of 5,591 genes were also identified, representing a coding capacity of 5,458,539 bp (coding percentage: 91.9%). Among these genes, 742 (13.27%) were found to encode putative proteins and 881 (15.75%) were assigned as encoding hypothetical proteins. Moreover, 5,537 genes matched a least one sequence in the COG database with BLASTp default parameters.

Nucleotide sequence accession numbers. The *M. asiaticum* DSM 44297 strain genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession no. HG964936 to HG964945. The whole-genome shotgun master numbers are CCBD010000001 to CCBD010000103.

ACKNOWLEDGMENT

This study was financially supported by URMITE, IHU Méditerranée Infection, Marseille, France.

REFERENCES

1. **Weiszfeiler G, Karasseva V, Karczag E.** 1971. A new mycobacterium species: *Mycobacterium asiaticum* n. sp. Acta Microbiol. Acad. Sci. Hung. **18:**247–252.
2. **Blacklock ZM, Dawson DJ, Kane DW, McEvoy D.** 1983. *Mycobacterium asiaticum* as a potential pulmonary pathogen for humans. A clinical and bacteriologic review of five cases. Am. Rev. Respir. Dis. **127:**241–244.
3. **Weiszfeiler JG, Karasseva V, Karczag E.** 1981. *Mycobacterium simiae* and related mycobacteria. Rev. Infect. Dis. **3:**1040–1045. http://dx.doi.org/10.1093/clinids/3.5.1040.
4. **Muwonge A, Kankya C, Johansen TB, Djønne B, Godfroid J, Biffa D, Edvardsen V, Skjerve E.** 2012. Non-tuberculous mycobacteria isolated from slaughter pigs in Mubende district, Uganda. BMC Vet. Res. **8:**52. http://dx.doi.org/10.1186/1746-6148-8-52.
5. **De Lima CA, Gomes HM, Oelemann MA, Ramos JP, Caldas PC, Campos CE, Pereira MA, Montes FF, de Oliveira MS, Suffys PN, Moura MM.** 2013. Nontuberculous mycobacteria in respiratory samples from patients with pulmonary tuberculosis in the state of Rondônia, Brazil. Mem. Inst. Oswaldo Cruz **108:**457–462. http://dx.doi.org/10.1590/S0074-027610804201301O.
6. **Varghese B, Memish Z, Abuljadayel N, Al-Hakeem R, Alrabiah F, Al-Hajoj SA.** 2013. Emergence of clinically relevant non-tuberculous mycobacterial infections in Saudi Arabia. PLoS Negl. Trop. Dis. **7:**e2234. http://dx.doi.org/10.1371/journal.pntd.0002234.
7. **Grech M, Carter R, Thomson R.** 2010. Clinical significance of *Mycobacterium asiaticum* isolates in Queensland, Australia. J. Clin. Microbiol. **48:**162–167. http://dx.doi.org/10.1128/JCM.01602-09.

8. **Ford JG, Huang AJ, Pflugfelder SC, Alfonso EC, Forster RK, Miller D.** 1998. Nontuberculous mycobacterial keratitis in south Florida. Ophthalmology **105:**1652–1658. http://dx.doi.org/10.1016/S0161-6420(98)99034-0.
9. **Taylor LQ, Williams AJ, Santiago S.** 1990. Pulmonary disease caused by *Mycobacterium asiaticum*. Tubercle **71:**303–305. http://dx.doi.org/10.1016/0041-3879(90)90047-C.
10. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380. http://dx.doi.org/10.1038/nature03959.
11. **Gao S, Sung WK, Nagarajan N.** 2011. Opera: reconstructing optimal genomic scaffolds with high-throughput paired-end sequences. J. Comput. Biol. **18:**1681–1691. http://dx.doi.org/10.1089/cmb.2011.0170.
12. **Boetzer M, Pirovano W.** 2012. Toward almost closed genomes with GapFiller. Genome Biol. **13:**R56. http://dx.doi.org/10.1186/gb-2012-13-6-r56.
13. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.
14. **Laslett D, Canback B.** 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. **32:**11–16. http://dx.doi.org/10.1093/nar/gkh152.
15. **Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR.** 2003. Rfam: an RNA family database. Nucleic Acids Res. **31:**439–441. http://dx.doi.org/10.1093/nar/gkg006.
16. **Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD.** 2012. The Pfam protein families database. Nucleic Acids Res. **40:**D290–D301. http://dx.doi.org/10.1093/nar/gkr1065.
17. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.
18. **Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW.** 2012. GenBank. Nucleic Acids Res. **40:**D48–D53. http://dx.doi.org/10.1093/nar/gkr1202.
19. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36. http://dx.doi.org/10.1093/nar/28.1.33.
20. **Tatusov RL, Koonin EV, Lipman DJ.** 1997. A genomic perspective on protein families. Science **278:**631–637. http://dx.doi.org/10.1126/science.278.5338.631.