genomeA
Journals.ASM.org

Draft Genome Sequence of *Mycobacterium mucogenicum* Strain CSUR P2099

Shady Asmar, Nicolás Rascovan, Catherine Robert, Michel Drancourt

Aix Marseille Université, URMITE, UMR 63, CNRS 7278, IRD 198, Inserm 1095, Faculté de Médecine, Marseille, France

***Mycobacterium mucogenicum* is a rapid-growing, nontuberculosis *Mycobacterium* species. The draft genome of *M. mucogenicum* CSUR P2099 comprises 6,210,127 bp exhibiting a 67.2% G+C content, 6,003 protein-coding genes, and 91 predicted RNA genes.**

Received 1 October 2015 **Accepted** 5 October 2015 **Published** 19 November 2015

Citation Asmar S, Rascovan N, Robert C, Drancourt M. 2015. Draft genome sequence of *Mycobacterium mucogenicum* strain CSUR P2099. Genome Announc 3(6):e01369-15. doi:10.1128/genomeA.01369-15.

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

Mycobacterium mucogenicum is a rapid-growing mycobacterium formerly known as a *Mycobacterium chelonae*–like organism (1, 2). *M. mucogenicum* was commonly implicated in outbreaks of opportunistic infections traced to contaminated water systems (3, 4) and contaminated hospital equipment (5–7). Catheter-related infections are the major form of *M. mucogenicum* opportunistic infection (8, 9), and *M. mucogenicum* is the leading mycobacterium to cause catheter-related bloodstream infections (9). However, respiratory tract, skin, soft tissue, and central nervous system infections due to *M. mucogenicum* have also been reported (10). We performed the whole-genome sequencing of *M. mucogenicum* CSUR P2099 in order to promote the development of advanced molecular tools for the detection and identification of this species.

M. mucogenicum CSUR P2099 was cultured in MGIT Middlebrook liquid culture (Becton, Dickinson, Le Pont-de-Claix, France) at 37°C in a 5% CO_2 atmosphere. Genomic DNA was then sequenced by Illumina MiSeq runs (Illumina, San Diego, CA, USA) using a 5-kb mate-paired library. Reads were trimmed using Trimmomatic (11) and assembled using Spades version 3.5 (12, 13). Contigs were combined together by SSPACE version 2 (14) and Opera version 2 (15) and helped by GapFiller version 1.10 (16). This resulted in a draft genome consisting of seven scaffolds and seven contigs for a total of 6,210,127 bp and a G+C content of 67.2%. Noncoding genes and miscellaneous features were predicted using RNAmmer (17), ARAGORN (18), Rfam (19), PFAM (20), and Infernal (21). Coding DNA sequences were predicted using Prodigal (22), and functional annotation was achieved using BLASTp against the GenBank database (23) and the Clusters of Orthologous Groups (COGs) database (24, 25). The genome was shown to encode at least 91 predicted RNAs, including 6 rRNAs, 57 tRNAs, 1 tmRNA, and 27 miscellaneous RNAs. A total of 6,003 identified genes, including 3,678 (61.27%) encoding putative proteins and 1,884 (31.38%) assigned as hypothetical proteins, yielded a coding capacity of 6,145,630 bp (coding percentage, 98.96%). A total of 4,199 (69.9%) genes matched at least one sequence in the COGs database with BLASTp default parameters.

Nucleotide sequence accession numbers. The *M. mucogenicum* genome sequence has been deposited in EMBL under the accession numbers CYSI01000001 to CYSI01000007.

ACKNOWLEDGMENT

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

REFERENCES

1. **Adékambi T, Drancourt M.** 2004. Dissection of phylogenetic relationships among 19 rapidly growing *Mycobacterium* species by 16S rRNA, *hsp*65, *sodA*, *recA* and *rpoB* gene sequencing. Int J Syst Evol Microbiol **54:**2095–2105. http://dx.doi.org/10.1099/ijs.0.63094-0.
2. **Runyon EH.** 1970. Identification of mycobacterial pathogens using colony characteristics. Am J Clin Pathol **54:**578–586.
3. **Kline S, Cameron S, Streifel A, Yakrus MA, Kairis F, Peacock K, Besser J, Cooksey RC.** 2004. An outbreak of bacteremias associated with *Mycobacterium mucogenicum* in a hospital water supply. Infect Control Hosp Epidemiol **25:**1042–1049. http://dx.doi.org/10.1086/502341.
4. **Livni G, Yaniv I, Samra Z, Kaufman L, Solter E, Ashkenazi S, Levy I.** 2008. Outbreak of *Mycobacterium mucogenicum* bacteraemia due to contaminated water supply in a paediatric haematology-oncology department. J Hosp Infect **70:**253–258. http://dx.doi.org/10.1016/j.jhin.2008.07.016.
5. **Tagashira Y, Kozai Y, Yamasa H, Sakurada M, Kashiyama T, Honda H.** 2015. A cluster of central line–associated bloodstream infections due to rapidly growing nontuberculous mycobacteria in patients with hematologic disorders at a Japanese tertiary care center: an outbreak investigation and review of the literature. Infect Control Hosp Epidemiol **36:**76–80. http://dx.doi.org/10.1017/ice.2014.14.
6. **Maybrook RJ, Campsen J, Wachs ME, Levi ME.** 2013. A case of *Mycobacterium mucogenicum* infection in a liver transplant recipient and a review of the literature. Transpl Infect Dis **15:**E260–E263. http://dx.doi.org/10.1111/tid.12154.
7. **Shachor-Meyouhas Y, Sprecher H, Eluk O, Ben-Barak A, Kassis I.** 2011. An outbreak of *Mycobacterium mucogenicum* bacteremia in pediatric hematology-oncology patients. Pediatr Infect Dis J **30:**30–32. http://dx.doi.org/10.1097/INF.0b013e3181ee31d7.
8. **Band JD, Ward JI, Fraser DW, Peterson NJ, Silcox VA, Good RC, Ostroy PR, Kennedy J.** 1982. Peritonitis due to a mycobacterium chelonei-like organism associated with intermittent chronic peritoneal dialysis. J Infect Dis **145:**9–17. http://dx.doi.org/10.1093/infdis/145.1.9.
9. **Han X, Dé I, Jacobson K.** 2007. Rapidly growing mycobacteria: clinical and microbiologic studies of 115 cases. Am J Clin Pathol **128:**612–621. http://dx.doi.org/10.1309/1KB2GKYT1BUEYLB5.

10. **Adékambi T**. 2009. *Mycobacterium mucogenicum* group infections: a review. Clin Microbiol Infect **15:**911–918. http://dx.doi.org/10.1111/j.1469-0691.2009.03028.x.
11. **Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B**. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res **40:**W622–W627. http://dx.doi.org/10.1093/nar/gks540.
12. **Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20:**714–737. http://dx.doi.org/10.1089/cmb.2013.0084.
13. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA**. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19:**455–477. http://dx.doi.org/10.1089/cmb.2012.0021.
14. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W**. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27:**578–579. http://dx.doi.org/10.1093/bioinformatics/btq683.
15. **Gao S, Sung W, 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.
16. **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.
17. **Lagesen K, Hallin P, Rodland EA, Staerfeldt H-, 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.
18. **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.
19. **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.
20. **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 ELL, 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.
21. **Nawrocki EP, Kolbe DL, Eddy SR**. 2009. Infernal 1.0: inference of RNA alignments. BioInformatics **25:**1335–1337. http://dx.doi.org/10.1093/bioinformatics/btp157.
22. **Hyatt D, Chen G, 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.
23. **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.
24. **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.
25. **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.