genomeA
Journals.ASM.org

Draft Genome Sequence of *Mycobacterium cosmeticum* DSM 44829

Olivier Croce, Catherine Robert, Didier Raoult, Michel Drancourt

Aix Marseille Université, URMITE, UM 63, UMR_S 1095, UMR 7278, Marseille, France

We announce the draft genome sequence of *Mycobacterium cosmeticum* strain DSM 44829, a nontuberculous species responsible for opportunistic infection. The genome described here is composed of 6,462,090 bp, with a G+C content of 68.24%. It contains 6,281 protein-coding genes and 75 predicted RNA genes.

Received 24 March 2014 **Accepted** 25 March 2014 **Published** 10 April 2014

Citation Croce O, Robert C, Raoult D, Drancourt M. 2014. Draft genome sequence of *Mycobacterium cosmeticum* DSM 44829. Genome Announc. 2(2):e00315-14. doi:10.1128/genomeA.00315-14.

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

Mycobacterium cosmeticum belongs to a poorly defined group of rapidly growing nontuberculous mycobacteria, which are most closely related to *Mycobacterium frederiksbergense*, *Mycobacterium hodleri*, *Mycobacterium diernhoferi*, and *Mycobacterium neoaurum* (1). It is an environmental organism recovered from water, including household potable water (2) and water collected at a nail salon (1), activated sludge from a wastewater treatment plant (3), and monument sandstones (4). Accordingly, *M. cosmeticum* can use the benzene series as the unique source of carbon (3). In medicine, *M. cosmeticum* is an opportunistic pathogen most frequently implicated in cutaneous granulomatous lesions following mesotherapy (1, 5). Further isolates have been obtained from blood samples collected from patients with indwelling catheters and from sputum specimens (6). *M. cosmeticum* has also been implicated as a gastrointestinal tract pathogen (7, 8). The optimal treatment of *M. cosmeticum* infection is not known, but one isolate has recently been shown *in vitro* to be susceptible to amikacin combined with clofazimine (9).

In this study, we sequenced the whole genome of the *M. cosmeticum* DSM 44829 strain in order to help depict its phylogenetic relationship with closely related mycobacteria and unique metabolic capabilities.

Genomic DNA was isolated from the *M. cosmeticum* DSM 44829 strain grown in MGIT Middlebrook broth at 37°C (Becton, Dickinson, Sparks, MD). It was then sequenced using three high-throughput next-generation sequencing (NGS) technologies: Roche 454 (Roche Diagnostics Corporation, Indianapolis, IN) (10), SOLiD version 4 (Life Technologies, Carlsbad, CA), and MiSeq Illumina (Illumina Inc., San Diego, CA). A 3.3-kb paired-end library was loaded on a picotiter plate and sequenced with the Roche-GS FLX Titanium sequencing kit XLR70. The run yielded 102 Mb with 256,437 passed filters and an average length of 397 bp. The barcoded paired-end SOLiD library generated 1,144,665 reads of 50 × 35-bp length. Finally, a paired-end Nextera library sequenced on MiSeq at 2 × 250 bp yielded 2,618,618 reads with an indexing of 22.89% on the flow cell.

The reads from the various sequencing technologies were first assembled separately. The 454 reads were assembled into contigs and scaffolds using Newbler version 2.8 (Roche). The Illumina reads were trimmed using Trimmomatic (11) and then assembled using the SPAdes software (12, 13). Contigs obtained were combined by using SSPACE (14) and Opera software version 1.2 (15), helped by GapFiller software version 1.10 (16). Some manual refinements using CLC Genomics version 6 software (CLC bio, Aarhus, Denmark) and homemade tools improved the genome. It was found that the *M. cosmeticum* draft genome consists of five contigs without gaps, containing 6,462,090 bp and a G+C content of 68.24%.

Noncoding genes and miscellaneous features were predicted using RNAmmer (17), Aragorn (18), Rfam (19), and Pfam (20). Open reading frames were predicted using Prodigal (21), and functional annotation was achieved using BLASTp against the GenBank database (22) and the Clusters of Orthologous Groups (COG) database (23, 24). Using these tools, it was found that the *M. cosmeticum* genome contains ≥75 predicted RNAs, including six rRNAs, 53 tRNAs, one transfer-messenger RNA, and 15 miscellaneous RNAs. A total of 6,281 genes were also identified, representing a coding capacity of 5,995,551 bp (coding percentage, 92.7%). Among these genes, 926 (14.74%) were found to encode putative proteins and 1,033 (16.44%) were assigned as genes for hypothetical proteins. Moreover, 6,211 genes matched at least one sequence in the COG database using BLASTp default parameters.

Nucleotide sequence accession numbers. The *M. cosmeticum* strain DSM 44829 genome sequence has been deposited at DDBJ/EMBL/GenBank under accession no. CCBB010000001 to CCBB010000005.

ACKNOWLEDGMENT

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

REFERENCES

1. **Cooksey RC, de Waard JH, Yakrus MA, Rivera I, Chopite M, Toney SR, Morlock GP, Butler WR.** 2004. *Mycobacterium cosmeticum* sp. nov., a novel rapidly growing species isolated from a cosmetic infection and from a nail salon. Int. J. Syst. Evol. Microbiol. **54**:2385–2391. http://dx.doi.org/10.1099/ijs.0.63238-0.
2. **Perez-Martinez I, Aguilar-Ayala DA, Fernandez-Rendon E, Carrillo-Sanchez AK, Helguera-Repetto AC, Rivera-Gutierrez S, Estrada-Garcia T, Cerna-Cortes JF, Gonzalez-Y-Merchand JA.** 2013. Occurrence of potentially pathogenic nontuberculous mycobacteria in Mexican house-

hold potable water: a pilot study. BMC Res. Notes **6:**531. http://dx.doi.org/10.1186/1756-0500-6-531.

3. **Zhang L, Zhang C, Cheng Z, Yao Y, Chen J.** 2013. Biodegradation of benzene, toluene, ethylbenzene, and *o*-xylene by the bacterium *Mycobacterium cosmeticum* byf-4. Chemosphere **90:**1340–1347. http://dx.doi.org/10.1016/j.chemosphere.2012.06.043.
4. **Kusumi A, Li XS, Katayama Y.** 2011. Mycobacteria isolated from Angkor monument sandstones grow chemolithoautotrophically by oxidizing elemental sulfur. Front Microbiol. **2:**104. http://dx.doi.org/10.3389/fmicb.2011.00104.
5. **Beer K, Waibel J.** 2009. Disfiguring scarring following mesotherapy-associated *Mycobacterium cosmeticum infection*. J. Drugs Dermatol. **8:**391–393.
6. **Cooksey RC, de Waard JH, Yakrus MA, Toney SR, Da Mata O, Nowicki S, Sohner K, Koch E, Petti CA, Morey RE, Srinivasan A.** 2007. *Mycobacterium cosmeticum*, Ohio and Venezuela. Emerg. Infect. Dis. **13:**1267–1269. http://dx.doi.org/10.3201/eid1308.061061.
7. **Addley J, McKeagney P, Turner G, Kelly M.** 2010. *Mycobacterium cosmeticum* as an unusual cause of ascites. BMJ Case Rep. **2010:**bcr0420091733. http://dx.doi.org/10.1136/bcr.04.2009.1733.
8. **Boschetti G, Cotte E, Moussata D, Chauvenet M, Breysse F, Chomarat M, Isaac S, Berger F, Kaiserlian D, Nancey S, Flourie B.** 2011. Identification of *Mycobacterium cosmeticum* sp. as a novel colitogenic infectious agent in a nonimmunocompromised patient. Inflamm. Bowel Dis. **17:**E128–E130. http://dx.doi.org/10.1002/ibd.21804.
9. **Van Ingen J, Totten SE, Helstrom NK, Heifets LB, Boeree MJ, Daley CL.** 2012. *In vitro* synergy between clofazimine and amikacin in treatment of nontuberculous mycobacterial disease. Antimicrob. Agents Chemother. **56:**6324–6327. http://dx.doi.org/10.1128/AAC.01505-12.
10. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, 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, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380. http://dx.doi.org/10.1038/nature03959.
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 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.
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, 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.
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 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.
21. **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.
22. **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.
23. **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.
24. **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.