AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of *Mycobacterium acapulcensis* Strain CSURP1424

Shady Asmar, Nicolás Rascovan, Catherine Robert, Michel Drancourt

Unité de Recherche sur les Maladies Infectieuses et Tropicales Émergentes, CNRS, UMR 7278, IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France

***Mycobacterium acapulcensis* is a rapidly growing scotochromogenic acid-fast bacillus. The draft genome of *M. acapulcensis* CSURP1424 comprises 5,290,974 bp, exhibiting a 66.67% G+C content, 4,870 protein-coding genes, and 71 predicted RNA genes.**

Received 20 June 2016 **Accepted** 21 June 2016 **Published** 11 August 2016

Citation Asmar S, Rascovan N, Robert C, Drancourt M. 2016. Draft genome sequence of *Mycobacterium acapulcensis* strain CSURP1424. Genome Announc 4(4):e00836-16. doi:10.1128/genomeA.00836-16.

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

Mycobacterium acapulcensis is a rapidly growing scotochromogenic acid-fast bacillus that was first isolated from sputum in Acapulco, a town on the Pacific coast of Mexico, during a campaign against tuberculosis (1). *M. acapulcensis* was regarded for a long time as a synonym of *Mycobacterium flavescens*, a closely related species (2, 3).

We performed whole-genome sequencing of *M. acapulcensis* CSURP1424 in order to facilitate the development of advanced molecular tools for the detection and identification of this species.

Genomic DNA was isolated from *M. acapulcensis* CSURP1424 cultured in MGIT Middlebrook liquid culture (Becton, Dickinson, Le Pont-de-Claix, France) at 37°C in a 5% CO_2 atmosphere. *M. acapulcensis* genomic DNA was sequenced by Illumina MiSeq runs (Illumina Inc, San Diego, CA, USA). A library with a 5.48-kb insert was loaded twice on a flow cell. Within these runs, the index representation for *M. acapulcensis* was determined to be 7.33% and 5.99%. The 1,841,967 paired reads were trimmed using Trimmomatic (4) and then assembled into scaffolds using Spades version 3.5 (5, 6) before finishing. Contigs were combined together by SSPACE version 2 (7) and Opera version 2 (8) helped by GapFiller version 1.10 (9). This resulted in a draft genome consisting of 29 scaffolds and 97 contigs, for a total of 5,290,974 bp and a G+C content of 66.67%. Noncoding genes and miscellaneous features were predicted using RNAmmer (10), ARAGORN (11), Rfam (12), PFAM (13), and Infernal (14). Coding DNA sequences (CDSs) were predicted using Prodigal (15), and functional annotation was achieved using BLASTp against the GenBank database (16) and the Clusters of Orthologous Groups (COGs) database (17, 18). The genome was shown to encode at least 71 predicted RNAs, including four rRNAs, 46 tRNAs, one tmRNAs, and 20 miscellaneous RNAs. A total of 4,920 identified genes yielded a coding capacity of 4,563,841 bp (coding percentage, 86.25%). Among these genes, 3,915 (80.39%) were found to encode for putative proteins and 723 (14.85%) were assigned as hypothetical proteins. Moreover, 2,929 (60.1%) genes matched at least one sequence in the Clusters of Orthologous Groups database with BLASTp default parameters. Further, the *M. acapulcensis* CSURP1424 genome was incorporated into *in silico* DNA-DNA hybridization (DDH) (19) with reference genomes selected on the basis of their 16S rRNA gene proximity, and DDH values were estimated using the GGDC version 2.0 online tool (20). This analysis yielded 22.8% ± 2.37 similarity with *Mycobacterium pyrenivorans* (21); 21.3% ± 2.34 with *Mycobacterium rhodesiae* NBB3 (22) and *Mycobacterium tusciae* (23); 21.1% ± 2.33 with *Mycobacterium austroafricanum* (24) and *Mycobacterium vanbaalenii* (25); 20.8% ± 2.33 with *Mycobacterium gilvum* Spyr1 (26); 20.5% ± 2.32 with *Mycobacterium aurum* (27); 20.3% ± 2.31 with *Mycobacterium fortuitum* ATCC 6841 (28); 20% ± 2.31 with *Mycobacterium neoaurum* DSM 44074 (29); and 19.9% ± 2.3 with *Mycobacterium marinum* E11 (30).

Accession number(s). The *M. acapulcensis* genome sequence has been deposited at EMBL under the accession numbers LT592221 to LT592249.

ACKNOWLEDGMENTS

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

REFERENCES

1. **Bojalil LF, Cerbon J, Trujillo A.** 1962. Adansonian classification of mycobacteria. J Gen Microbiol **28:**333–346. http://dx.doi.org/10.1099/00221287-28-2-333.
2. **Turenne CY, Tschetter L, Wolfe J, Kabani A.** 2001. Necessity of quality-controlled 16S rRNA gene sequence databases: identifying nontuberculous *Mycobacterium* species. J Clin Microbiol **39:**3637–3648. http://dx.doi.org/10.1128/JCM.39.10.3638-3648.2001.
3. **Tsukamura M, Mizuno S.** 1977. Numerical analysis of relationships among rapidly growing, scotochromogenic mycobacteria. J Gen Microbiol **98:**511–517. http://dx.doi.org/10.1099/00221287-98-2-511.
4. **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.
5. **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.
6. **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.

7. **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.
8. **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.
9. **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.
10. **Lagesen K, Hallin P, Rødland EA, Staerfeldt H-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.
11. **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.
12. **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.
13. **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.
14. **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.
15. **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.
16. **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.
17. **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.
18. **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.
19. **Richter M, Rosselló-Móra R**. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A **106:**19126–19131. http://dx.doi.org/10.1073/pnas.0906412106.
20. **Auch AF, von Jan M, Klenk HP, Göker M**. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci **2:**117–134. http://dx.doi.org/10.4056/sigs.531120.
21. **Derz K, Klinner U, Schuphan I, Stackebrandt E, Kroppenstedt RM**. 2004. *Mycobacterium pyrenivorans* sp. nov., a novel polycyclic-aromatic-hydrocarbon-degrading species. Int J Syst Evol Microbiol **54:**2313–2317. http://dx.doi.org/10.1099/ijs.0.03003-0.
22. **Tsukamura M, Mizuno S, Gane NFF, Mills A, King L**. 1971. *Mycobacterium rhodesiae* sp. nov.: a new species of rapid-growing scotochromogenic mycobacteria. Jpn J Microbiol **15:**407–416. http://dx.doi.org/10.1111/j.1348-0421.1971.tb00598.x.
23. **Tortoli E, Kroppenstedt RM, Bartoloni A, Caroli G, Jan I, Pawlowski J, Emler S**. 1999. *Mycobacterium tusciae* sp. nov. Int J Syst Bacteriol **49:**1839–1844. http://dx.doi.org/10.1099/00207713-49-4-1839.
24. **Croce O, Robert C, Raoult D, Drancourt M**. 2014. Draft genome sequence of *Mycobacterium austroafricanum* DSM 44191. Genome Announc **2**(2):e00317-14. http://dx.doi.org/10.1128/genomeA.00317-14.
25. **Khan AA, Kim SJ, Paine DD, Cerniglia CE**. 2002. Classification of a polycyclic aromatic hydrocarbon-metabolizing bacterium, *Mycobacterium* sp. strain PYR-1, as *Mycobacterium vanbaalenii* sp. nov. Int J Syst Evol Microbiol **52:**1997–2002. http://dx.doi.org/10.1099/00207713-52-6-1997.
26. **Kallimanis A, Karabika E, Mavromatis K, Lapidus A, Labutti KM, Liolios K, Ivanova N, Goodwin L, Woyke T, Velentzas AD, Perisynakis A, Ouzounis CC, Kyrpides NC, Koukkou AI, Drainas C**. 2011. Complete genome sequence of *Mycobacterium* sp. strain (Spyr1) and reclassification to *Mycobacterium gilvum* Spyr1. Stand Genomic Sci **5:**144–153. http://dx.doi.org/10.4056/sigs.2265047.
27. **Honarvar B, Movahedan H, Mahmoodi M, Sheikholeslami FM, Farnia P**. 2012. *Mycobacterium aurum* keratitis: an unusual etiology of a sight-threatening infection. Braz J Infect Dis **16:**204–208.
28. **Wells AQ, Agius E, Smith N**. 1955. *Mycobacterium fortuitum*. Am Rev Tuberc **72:**53–63.
29. **Phelippeau M, Robert C, Croce O, Raoult D, Drancourt M**. 2014. Draft genome sequence of *Mycobacterium neoaurum* strain DSM 44074T. Genome Announc 2(4):e00699-14. http://dx.doi.org/10.1128/genomeA.00699-14.
30. **Sette CS, Wachholz PA, Masuda PY, da Costa Figueira RB, de Oliveira Mattar FR, Ura DG**. 2015. *Mycobacterium marinum* infection: a case report. J Venom Anim Toxins Incl Trop Dis **21:**7. http://dx.doi.org/10.1186/s40409-015-0008-9.