

# Draft Genome Sequence of *Mycobacterium neworleansense* Strain ATCC 49404<sup>T</sup>

Shady Asmar, Catherine Robert, Olivier Croce, Aurelia Caputo, Michel Drancourt

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

***Mycobacterium neworleansense* is a rapid growing nontuberculosis species belonging to the *Mycobacterium fortuitum* complex. The draft genome of *M. neworleansense* ATCC 49404<sup>T</sup> comprises 6,287,317 bp exhibiting a 66.85% G+C content, 5,997 protein-coding genes, and 89 predicted RNA genes.**

Received 22 September 2015 Accepted 25 September 2015 Published 5 November 2015

**Citation** Asmar S, Robert C, Croce O, Caputo A, Drancourt M. 2015. Draft genome sequence of *Mycobacterium neworleansense* strain ATCC 49404<sup>T</sup>. *Genome Announc* 3(6): e01314-15. doi:10.1128/genomeA.01314-15.

**Copyright** © 2015 Asmar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

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

*Mycobacterium neworleansense* is a rapidly growing nontuberculous mycobacterium formerly classified as *Mycobacterium fortuitum* sorbitol-negative third biovariant (1). Within the *M. fortuitum* complex, *M. neworleansense* clusterizes with *Mycobacterium peregrinum*, *Mycobacterium septicum*, and *Mycobacterium porcicum* (2, 3). The name *M. neworleansense* derives from New Orleans, USA, where this mycobacterium was first isolated from a scalp wound (1).

We performed the whole-genome sequencing of *M. neworleansense* ATCC 49404<sup>T</sup> in order to refine understanding of its taxonomic relationships with closely related species of the *M. fortuitum* complex.

From this perspective, *M. neworleansense* ATCC 49404<sup>T</sup> was cultured in MGIT Middlebrook liquid culture (Becton Dickinson, Le Pont-de-Claix, France) at 37°C in a 5% CO<sub>2</sub> atmosphere. *M. neworleansense* genomic DNA was then sequenced by Illumina MiSeq runs (Illumina Inc., San Diego, USA) using a 5-kb mate-paired library. Reads were trimmed using Trimmomatic (4) and assembled using Spades v3.5 (5, 6). Contigs were combined together by SSPACE v2 (7), Opera v2 (8) helped by GapFiller v1.10 (9), and homemade tools in Python to refine the set. This resulted in a draft genome consisting of nine contigs without gap for a total of 6,287,317 bp and a G+C content of 66.85%. Noncoding genes and miscellaneous features were predicted using RNAmmer (10), ARAGORN (11), Rfam (12), PFAM (13), and Infernal (14). Coding DNA sequences were predicted using Prodigal (15), and functional annotation was achieved using BLAST+ (16) and HMMER3 (17) against the UniProtKB database (18). The genome was shown to encode at least 89 predicted RNAs, including 9 rRNAs, 59 tRNAs, 1 transfer-messenger RNA (tmRNA), and 20 miscellaneous RNAs. A total of 5,997 identified genes yielded a coding capacity of 5,837,343 bp (coding percentage, 92.84%). Among these genes, 246 (4.1%) were found to be putative proteins and 1,011 (16.86%) were assigned as hypothetical proteins. Moreover, 4,296 genes matched a least one sequence in the Clusters of Orthologous Groups (19, 20) with BLASTP default parameters.

*In silico* DNA-DNA hybridation (DDH) (21) was performed with the *M. fortuitum* complex species and other reference ge-

nomes selected on the basis of their 16S rRNA gene sequence proximity with *M. neworleansense*. The *M. neworleansense* genome was locally aligned 2 by 2 using a BLAT algorithm (22, 23) against each the selected genomes, and DDH values were estimated from a generalized linear model (24). The DDH was 35.9% ± 2.48% with *Mycobacterium septicum* DSM 44393 (25), 32.8% ± 2.46% with *Mycobacterium senegalense* CK2 M4421 (1), 32.8% ± 2.46% with *Mycobacterium conceptionense* (26), 32.6% ± 2.46% with *Mycobacterium farcinogenes* (27), 32.3% ± 2.46% with *Mycobacterium peregrinum* (28), 31.7% ± 2.46% with *Mycobacterium fortuitum* ATCC 6841 (29), 20.5% ± 2.32% with *Mycobacterium chubuense* NBB4 (30) and *Mycobacterium gilvum* Spyr1 (31), and 20.1% ± 2.31% with *Mycobacterium aromaticivorans* JS19b1 (32).

These data confirm *M. neworleansense* as a unique species more closely related to *M. septicum* in the *M. fortuitum* complex.

**Nucleotide sequence accession numbers.** The *M. neworleansense* ATCC 49404<sup>T</sup> genome sequence has been deposited at EMBL under the accession numbers [CWKH01000001](https://www.ebi.ac.uk/ena/record/CWKH01000001) to [CWKH01000009](https://www.ebi.ac.uk/ena/record/CWKH01000009).

## ACKNOWLEDGMENT

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

## REFERENCES

- Schinsky MF, Morey RE, Steigerwalt AG, Douglas MP, Wilson RW, Floyd MM, Butler WR, Daneshvar MI, Brown-Elliott BA, Wallace RJ, Jr, McNeil MM, Brenner DJ, Brown JM. 2004. Taxonomic variation in the *Mycobacterium fortuitum* third biovariant complex: description of *Mycobacterium boenickei* sp. nov., *Mycobacterium houstonense* sp. nov., *Mycobacterium neworleansense* sp. nov. and *Mycobacterium brisbanense* sp. nov. and recognition of *Mycobacterium porcinum* from human clinical isolates. *Int J Syst Evol Microbiol* 54:1653–1667. <http://dx.doi.org/10.1099/ijs.0.02743-0>.
- Adékambi T, Drancourt M. 2004. Dissection of phylogenetic relationships among 19 rapidly growing mycobacterium species by 16S rRNA, hsp65, sodA, recA and rpoB gene sequencing. *Int J Syst Evol Microbiol* 54:2095–2105. <http://dx.doi.org/10.1099/ijs.0.63094-0>.
- Hamid ME. 2014. Current perspectives on *Mycobacterium farcinogenes*

- and *Mycobacterium senegalense*, the causal agents of bovine Farcy. *Vet Med Int* 2014;247906. <http://dx.doi.org/10.1155/2014/247906>.
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 preassembled contigs using SSPACE. *Bioinformatics* 27:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
  8. 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>.
  9. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
  10. 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>.
  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, Coghill 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>.
  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 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>.
  16. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <http://dx.doi.org/10.1186/1471-2105-10-421>.
  17. Eddy SR. 2011. Accelerated profile HMM searches. *PLOS Comput Biol* 7:e1002195. <http://dx.doi.org/10.1371/journal.pcbi.1002195>.
  18. The UniProt Consortium. 2011. Ongoing and future developments at the universal protein resource. *Nucleic Acids Res* 39:D214–D219. <http://dx.doi.org/10.1093/nar/gkq1020>.
  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>.
  21. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106: 19126–19131. <http://dx.doi.org/10.1073/pnas.0906412106>.
  22. Kent WJ. 2002. BLAT—the blast-like alignment tool. *Genome Res* 12: 656–664. <http://dx.doi.org/10.1101/gr.229202>.
  23. Auch AF, Von Jan M, Klenk H, 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>.
  24. Meier-Kolthoff JP, Auch AF, Klenk H, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:14–60. <http://dx.doi.org/10.1186/1471-2105-14-60>.
  25. Sassi M, Robert C, Raoult D, Drancourt M. 2013. Noncontiguous genome sequence of *Mycobacterium septicum* strain dsm 44393T. *Genome Announc* 1(4):e00574–13. <http://dx.doi.org/10.1128/genomeA.00574-13>.
  26. Adékambi T, Stein A, Carvajal J, Raoult D, Drancourt M. 2006. Description of *Mycobacterium conceptionense* sp. nov., a *Mycobacterium fortuitum* group organism isolated from a posttraumatic osteitis inflammation. *J Clin Microbiol* 44:1268–1273. <http://dx.doi.org/10.1128/JCM.44.4.1268-1273.2006>.
  27. Croce O, Robert C, Raoult D, Drancourt M. 2014. Draft genome sequence of *Mycobacterium farcinogenes* NCTC 10955. *Genome Announc* 2(3):e00523–14. <http://dx.doi.org/10.1128/genomeA.00523-14>.
  28. Asmar S, Rascovan N, Robert C, Drancourt M. 2015. Draft Genome Sequence of *Mycobacterium peregrinum* Strain CSUR P2098. *Genome Announc* 3(6):e01274–15. <http://dx.doi.org/10.1128/genomeA.01274-15>.
  29. Wells AQ, Agius E, Smith N. 1955. *Mycobacterium fortuitum*. *Am Rev Tuberc* 72:53–63.
  30. 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>.
  31. 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, Drinas 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>.
  32. Kwak Y, Park GS, Lee SE, Li QX, Shin JH. 2014. Genome sequence of *Mycobacterium aromaticivorans* JS19b1(T), a novel isolate from Hawaiian soil. *J Biotechnol* 186:137–138. <http://dx.doi.org/10.1016/j.jbiotec.2014.06.034>.