Draft Genome Sequence of *Mycobacterium peregrinum* Strain CSUR P2098

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 peregrinum is a nonpigmented rapid growing nontuberculosis species belonging to the *Mycobacterium fortuitum* group. The draft genome of *M. peregrinum* type I CSUR P2098 comprises 7,109,836 bp exhibiting a 66.23% G+C content, 6,894 protein-coding genes, and 100 predicted RNA genes. Its genome analysis suggests this species differs from *Mycobacterium senegalense*.

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

Initially reported in 1962, *Mycobacterium peregrinum* (1) was officially described 30 years later as a rapidly growing nontuberculous mycobacterium in the *Mycobacterium fortuitum* complex (2, 3). It was further divided into pipemidic acid-susceptible *M. peregrinum* type I and pipemidic acid-resistant *M. peregrinum* type II (4), regarded as a human strain of *Mycobacterium senegalense* (5). *M. peregrinum* has been isolated from various water-related sources (6–8). *M. peregrinum* accounts for only 1% to 2% of infections due to rapid growing mycobacteria (9–13).

We performed whole-genome sequencing of *M. peregrinum* CSUR P2098 strain, a type I, in order to firmly ensure its taxonomic relationships within the *M. fortuitum* complex.

Genomic DNA was isolated from *M. peregrinum* CSUR P2098 cultured in MGIT Middlebrook liquid culture (Becton, Dickinson, Le Pont-de-Claix, France) at 37°C in a 5% CO_2 atmosphere. *M. peregrinum* genomic DNA was sequenced by Illumina MiSeq runs (Illumina Inc., San Diego) using a 5-kb mate-paired library. Reads were trimmed using Trimmomatic (14), and assembled using Spades v3.5 (15, 16). Contigs were combined together by SSPACE v2 (17) and Opera v2 (18) helped by GapFiller v1.10 (19). This resulted in a draft genome consisting in six scaffolds and seven contigs for a total of 7,109,836 bp and a G+C content of 66.23%. Noncoding genes and miscellaneous features were predicted using RNAmmer (20), ARAGORN (21), Rfam (22), PFAM (23), and Infernal (24). Coding DNA sequences (CDSs) were predicted using Prodigal (25) and functional annotation was achieved using BLASTp against the GenBank database (26) and the Clusters of Orthologous Groups (COGs) database (27, 28). The genome was shown to encode at least 100 predicted RNA including six rRNA, 64 tRNA, one tmRNA, and 29 miscellaneous RNA. A total of 6,894 identified genes yielded a coding capacity of 7,054,578 bp (coding percentage, 99.22%). Among these genes, 5,101 (74%) were found as putative proteins and 1,282 (18.6%) were assigned as hypothetical proteins. Moreover, 4,020 (58.3%) genes matched a least one sequence in the COGs database with BLASTp default parameters. Further, *M. peregrinum* CSUR P2098 genome was incorporated into *in silico* DNA-DNA hybridation

(DDH) (29) with the *M. fortuitum* complex species and other reference genomes selected on the basis of their 16S rRNA gene sequence proximity with *M. peregrinum* and DDH values were estimated using the GGDC v2.0 online tool (30). This analysis yielded 33.8% ± 2.47 similarity with *Mycobacterium septicum* DSM 44393 (31), 33.6% ± 2.47 *Mycobacterium fortuitum* ATCC 6841 (32), 33.5% ± 2.47 with *Mycobacterium* sp. VKM Ac-1817D (33), 32.3% ± 2.46 with *Mycobacterium senegalense* CK2 and *Mycobacterium neworleansense* ATCC 49404 (34), 31.8% ± 2.46 with *Mycobacterium conceptionense* D16 (unpublished data) (35), 31.6% ± 2.46 with *Mycobacterium farcinogenes* DSM43637 (36), and 20.7% ± 2.32 to *Mycobacterium gilvum* Spyr1 (37).

These data confirm *M. peregrinum* as a unique species within the *M. fortuitum* complex. In particular, these data suggest that *M. peregrinum* differs from *M. senegalense*.

Nucleotide sequence accession numbers. The *M. peregrinum* genome sequence has been deposited at EMBL under the accession numbers CYSH01000001 to CYSH01000007.

ACKNOWLEDGMENT

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. **Kubica GP, Baess I, Gordon RE, Jenkins PA, Kwapinski JBG, McDurmont C, Pattyn SR, Saito H, Silcox V, Stanford JL, Takeya K, Tsukamura M.** 1972. A co-operative numerical analysis of rapidly growing mycobacteria. J Gen Microbiol **73:**55–70. http://dx.doi.org/10.1099/00221287-73-1-55.
3. **Kusunoki S, Ezaki T.** 1992. Proposal of *Mycobacterium peregrinum* sp. nov., nom. rev., and elevation of *Mycobacterium chelonae* subsp. *abscessus* (Kubica et al.) to species status: *Mycobacterium abscessus* comb. nov. Int J Syst Bacteriol **42:**240–245. http://dx.doi.org/10.1099/00207713-42-2-240.
4. **Telenti A, Marchesi F, Balz M, Bally F, Böttger EC, Bodmer T.** 1993. Rapid identification of mycobacteria to the species level by polymerase

chain reaction and restriction enzyme analysis. J Clin Microbiol **31**: 175–178.

5. **Wallace RJ Jr, Brown-Elliott BA, Brown J, Steigerwalt AG, Hall L, Woods G, Cloud J, Mann L, Wilson R, Crist C, Jost KC Jr, Byrer DE, Tang J, Cooper J, Stamenova E, Campbell B, Wolfe J, Turenne C**. 2005. Polyphasic characterization reveals that the human pathogen *Mycobacterium peregrinum* type II belongs to the bovine pathogen species *Mycobacterium senegalense*. J Clin Microbiol **43**:5925–5935.

6. **Vaerewijck MJM, Huys G, Palomino JC, Swings J, Portaels F**. 2005. Mycobacteria in drinking water distribution systems: ecology and significance for human health. FEMS Microbiol Rev **29**:911–934. http://dx.doi.org/10.1016/j.femsre.2005.02.001.

7. **Thomson R, Tolson C, Carter R, Coulter C, Huygens F, Hargreaves M**. 2013. Isolation of nontuberculous mycobacteria (NTM) from household water and shower aerosols in patients with pulmonary disease caused by NTM. J Clin Microbiol **51**:3006–3011. http://dx.doi.org/10.1128/JCM.00899-13.

8. **Torvinen E, Suomalainen S, Paulin L, Kusnetsov J**. 2014. Mycobacteria in Finnish cooling tower waters. APMIS **122**:353–358. http://dx.doi.org/10.1111/apm.12153.

9. **Brown-Elliott BA, Wallace RJ Jr**. 2002. Clinical and taxonomic status of pathogenic nonpigmented or late-pigmenting rapidly growing mycobacteria. Clin Microbiol Rev **15**:716–746. http://dx.doi.org/10.1128/CMR.15.4.716-746.2002.

10. **Nagao M, Sonobe M, Bando T, Saito T, Shirano M, Matsushima A, Fujihara N, Takakura S, Iinuma Y, Ichiyama S**. 2009. Surgical site infection due to Mycobacterium peregrinum: a case report and literature review. Int J Infect Dis **13**:209–211. http://dx.doi.org/10.1016/j.ijid.2008.06.018.

11. **Sawahata M, Hagiwara E, Ogura T, Komatsu S, Sekine A, Tsuchiya N, Takahashi H**. 2010. Pulmonary mycobacteriosis caused by *Mycobacterium peregrinum* in a young, healthy man. Nihon Kokyuki Gakkai Zasshi **48**:866–870.

12. **Torres-Duque CA, Díaz C, Vargas L, Serpa EM, Mosquera W, Garzón MC, Mejía G, García LM, González LA, Castro CM, Ribón W**. 2010. Disseminated mycobacteriosis affecting a prosthetic aortic valve: first case of *Mycobacterium peregrinum* type III reported in Colombia. Biomedica **30**:332–337.

13. **Kamijo F, Uhara H, Kubo H, Nakanaga K, Hoshino Y, Ishii N, Okuyama R**. 2012. A case of mycobacterial skin disease caused by *Mycobacterium peregrinum*, and a review of cutaneous infection. Case Rep Dermatol **4**:76–79.

14. **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.

15. **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.

16. **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.

17. **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.

18. **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.

19. **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.

20. **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.

21. **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.

22. **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.

23. **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.

24. **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.

25. **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.

26. **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.

27. **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.

28. **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.

29. **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.

30. **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.

31. **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.

32. **Wells AQ, Agius E, Smith N**. 1955. Mycobacterium fortuitum. Am Rev Tuberc **72**:53–63.

33. **Shtratnikova VY, Schelkunov MI, Dovbnya DV, Pekov YA, Bragin EY, Ashapkin VV, Donova MV**. 2015. Complete genome sequence of mycobacterium sp. strain VKM Ac-1817D, capable of producing 9α-Hydroxyandrost-4-ene-3,17-dione from phytosterol. Genome Announc **3**(3):e01447-14. http://dx.doi.org/10.1128/genomeA.01447-14.

34. **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.

35. **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.

36. **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.

37. **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* Spyr. Stand Genomic Sci **5**:144–153. http://dx.doi.org/10.4056/sigs.2265047.