



# Whole-Genome Sequencing of *Mycobacterium tilburgii* Strain MEPHI

Jamal Saad,<sup>a,b</sup> Michel Drancourt,<sup>b</sup> Margaret M. Hannan,<sup>c</sup> Patrick J. Stapleton,<sup>c</sup> Simon Grandjean Lapierre<sup>d</sup>

<sup>a</sup>IHU Méditerranée Infection, Marseille, France

<sup>b</sup>IRD, MEPHI, IHU Méditerranée Infection, Aix-Marseille Université, Marseille, France

<sup>c</sup>Mater Misericordiae University Hospital, University College Dublin, Dublin, Ireland

<sup>d</sup>Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal, Quebec, Canada

**ABSTRACT** *Mycobacterium tilburgii* is a fastidious mycobacterium which has previously been reported to cause severe disseminated infections. Genome sequencing of the *M. tilburgii* MEPHI clinical isolate yielded 3.14 Mb, with 66.3% GC content, and confirmed phylogenetic placement within the *Mycobacterium simiae* complex.

Only 13 cases of *Mycobacterium tilburgii* infection have been reported, highlighting host immunosuppression as a risk factor for *M. tilburgii* infection (1–3). *M. tilburgii* was never previously isolated in culture; its detection relied on the direct sequencing of internal transcribed spacer 1 (ITS1), *hsp65*, *rpoB*, and 16S rRNA genes from clinical material (2). We isolated *M. tilburgii* from a bone marrow sample by using a shell vial assay, as previously described (4, 5) (Fig. 1A). This organism had been previously detected in this immunocompromised patient in a biopsy specimen taken from a duodenal plaque by direct 16S rRNA gene sequencing. DNA (0.2  $\mu\text{g}/\mu\text{l}$ ) extracted using InstaGene matrix (Bio-Rad, Marnes-la-Coquette, France) was sequenced using MiSeq platform (Illumina, Inc., San Diego, CA, USA) paired-end sequencing and automated cluster generation with dual-indexed 2  $\times$  250-bp reads. A 535,000/mm<sup>2</sup> cluster density filter and a 96.3% cluster passing quality control filter were applied. A total of 8,690,521 reads were assembled using SPAdes version 3.12.0 with the option “–careful” in order to reduce the number of mismatches and short indels (6). Default parameters for k values, i.e., k-mer values of 127, 99, 77, 55, 33, and 21, were applied. A total of 977 human DNA contigs were removed via a BLASTn script that connected remotely to the NCBI database (nr/nt) (BLAST+ 2.3.0). Annotation using Prokka version 1.13 (7) of the *M. tilburgii* strain MEPHI genome sequence (mean coverage depth, 6.5 $\times$ ) comprises 102 contigs, with a total assembly size of 3,238,527 bp and 66.3% GC content ( $N_{50}$  value, 56,170 bp; coverage, 5.6 $\times$ ). One 19,595-bp contig (62.9 GC% content), exhibited 77% nucleotide similarity with *Mycobacterium* sp. strain KMS plasmid pMKMS02 (GenBank accession no. [CP000520](https://doi.org/10.1128/MRA.00933-19)) using BLASTn. The *M. tilburgii* strain MEPHI genome was predicted to contain 4,207 genes, including 4,154 protein-coding genes, and 53 RNAs (49 tRNAs, 3 rRNAs, and one transfer-messenger RNA [tmRNA]). Phylogenetic trees based on the 3,504-bp *rpoB* and the 1,523-bp 16S RNA genes confirmed phylogenetic placement into the *Mycobacterium simiae* complex (Fig. 1B and C). The genomic similarities estimated using OrthoANI version 0.93.1 (8) and *in silico* DNA-DNA hybridization estimated using the GGDC version 2.0 online tool (9) were 87.17% and 33.8% with *M. simiae* (GenBank assembly accession no. [GCA\\_000455305](https://doi.org/10.1128/MRA.00933-19)), 87.15% and 33.6% with *Mycobacterium sherrisii* (GenBank assembly accession no. [GCA\\_001722325](https://doi.org/10.1128/MRA.00933-19)), 82.88% and 26.8% with *Mycobacterium triplex* (GenBank assembly accession no. [GCA\\_000689255](https://doi.org/10.1128/MRA.00933-19)), 82.76% and 26.6% with *Mycobacterium florentinum* (NCBI RefSeq accession no. [NZ\\_LQOV00000000](https://doi.org/10.1128/MRA.00933-19)), 82.62% and 26.1% with *Mycobacterium genavense*

**Citation** Saad J, Drancourt M, Hannan MM, Stapleton PJ, Grandjean Lapierre S. 2019. Whole-genome sequencing of *Mycobacterium tilburgii* strain MEPHI. *Microbiol Resour Announc* 8:e00933-19. <https://doi.org/10.1128/MRA.00933-19>.

**Editor** Catherine Putonti, Loyola University Chicago

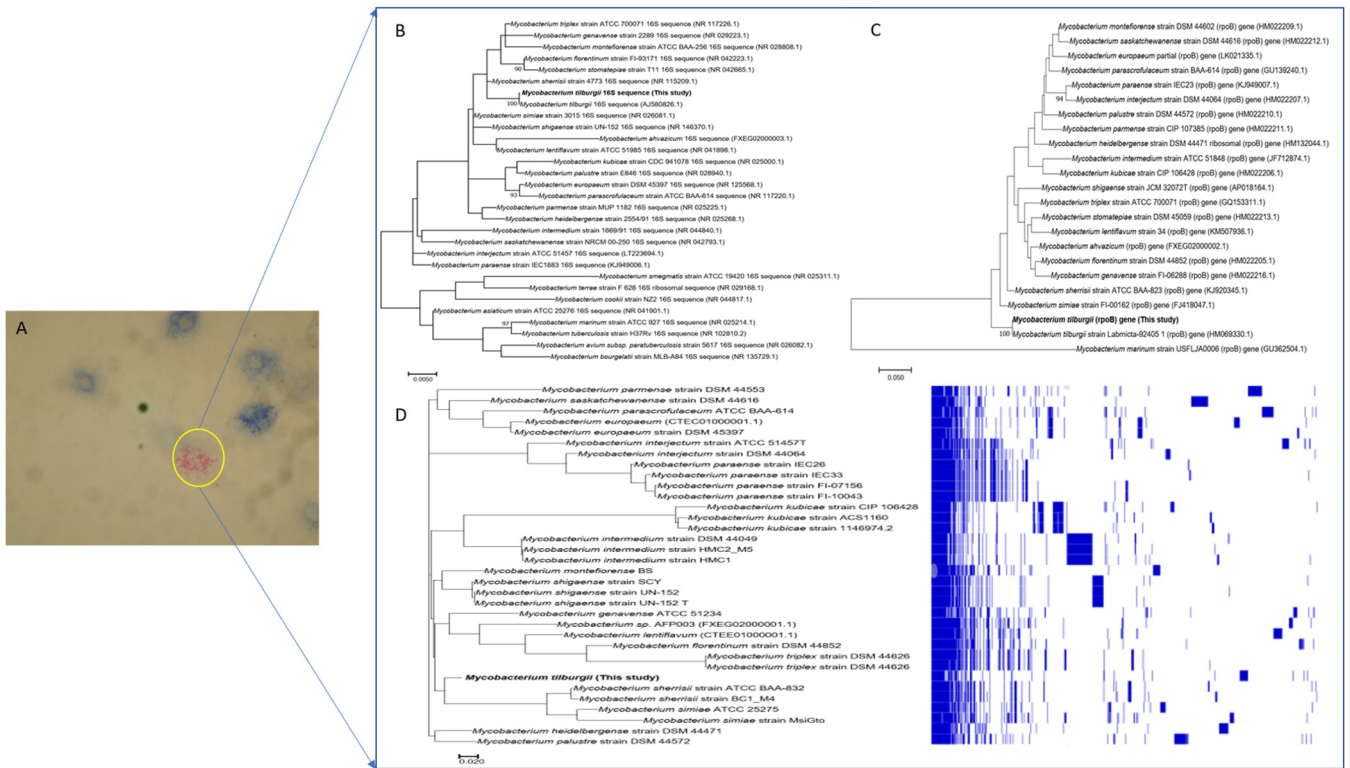
**Copyright** © 2019 Saad et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Michel Drancourt, [michel.drancourt@univ-amu.fr](mailto:michel.drancourt@univ-amu.fr).

**Received** 1 August 2019

**Accepted** 30 August 2019

**Published** 3 October 2019



**FIG 1** (A) Ziehl-Neelsen stain visualization of *Mycobacterium tuberculosis* MEPHI strain in cell culture. (B and C) Phylogenetic tree based on the 16S rRNA gene sequence (B) and *rpoB* gene sequence (C). Sequences were aligned using Muscle v3.8.31 implemented in the MEGA7 software (13, 14). Phylogenetic inferences were obtained using the maximum likelihood method based on the Tamura and Nei model (1,000 bootstrap replicates); bootstrap values of  $\geq 90\%$  are given at the nodes. (D, Left) Newick tree of *Mycobacterium tuberculosis* strain MEPHI and other members of the *Mycobacterium simiae* complex, generated by Roary from binary gene presence/absence in the accessory genome. (D, Right) Plot of Roary gene presence/absence analysis results (blue, gene presence; white, gene absence). The continuous blue block at the beginning shows the conserved genes of all members of the *M. simiae* complex, while regions of discontinuous blue denote accessory genome content, which varies between members.

(NCBI RefSeq accession no. [NZ\\_JAGZ00000000](#)), 82.42% and 26.3% with *Mycobacterium ahvazicum* (NCBI RefSeq accession no. [NZ\\_FXEG00000000](#)), and 82.35% and 26% with *Mycobacterium lentiflavum* (NCBI RefSeq accession no. [NZ\\_CTEE00000000](#)), respectively. Roary pangenome analysis (10) of the *M. simiae* complex including *M. tuberculosis* strain MEPHI yielded a total of 37,895 genes distributed as 1,077 core genome genes, 403 soft-core genes, and 5,547 shell genes (Fig. 1D). PathogenFinder 1.1 (11) predicted *M. tuberculosis* strain MEPHI to be a human pathogen (probability, 0.736) with detection of 14 pathogenicity-associated proteins. Analysis using the Resistance Gene Identifier version 5.0.5 database (12) indicated that the *M. tuberculosis* strain MEPHI genome harbored *rpBA*, *murA* (C117D), *mtrA*, *katG* (I335T), and *pncA* (H82R) mutations conferring resistance to rifampin, fosfomicin, macrolides, penam, isoniazid, and pyrazinamide, respectively. The whole-genome sequence confirms that *M. tuberculosis* belongs to the *M. simiae* complex. Targeted molecular assays could be designed for the detection of *M. tuberculosis* in clinical specimens.

**Data availability.** Illumina raw sequences and the assembled whole-genome sequence are available at the EMBL/GenBank under accession no. [ERR3436035](#) and [ERP116312](#), respectively.

**ACKNOWLEDGMENTS**

This work was supported by the French Government under the “Investissements d’avenir” (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, National Agency for Research; reference Méditerranée Infection 10-IAHU-03). This work was supported by Région Sud (Provence Alpes Côte d’Azur) and

European funding from FEDER PRIMI. S.G.L. is supported by the Canadian Association for Microbiology and Infectious Diseases.

The funders had no role to play in the study design, data collection and interpretation, or the decision to submit the work for publication.

## REFERENCES

1. Buiting AVM, Bergmans A, Schouls L. 1995. A new mycobacterial species causing disseminated infection. Abstr 35th Interscience Conference on Antimicrobial Agents and Chemotherapy, 17 to 20 September 1995, San Francisco, CA.
2. Hartwig NG, Warris A, van de Vosse E, van der Zanden AGM, Schülin-Casonato T, van Ingen J, van Hest R. 2011. "Mycobacterium tuberculosis" infection in two immunocompromised children: importance of molecular tools in culture-negative mycobacterial disease diagnosis. *J Clin Microbiol* 49:4409–4411. <https://doi.org/10.1128/JCM.05460-11>.
3. Heyckendorf J, Aries SP, Greinert U, Richter E, Schultz H, Lange C. 2015. Functional immune reconstitution by interleukin-2 adjunctive therapy for HIV/mycobacterial co-infection. *Emerg Infect Dis* 21:1685–1687. <https://doi.org/10.3201/eid2109.150461>.
4. Gouriet F, Fenollar F, Patrice JY, Drancourt M, Raoult D. 2005. Use of shell-vial cell culture assay for isolation of bacteria from clinical specimens: 13 years of experience. *J Clin Microbiol* 43:4993–5002. <https://doi.org/10.1128/JCM.43.10.4993-5002.2005>.
5. Doughty EL, Sergeant MJ, Adetifa I, Antonio M, Pallen MJ. 2014. Culture-independent detection and characterisation of *Mycobacterium tuberculosis* and *M. africanum* in sputum samples using shotgun metagenomics on a benchtop sequencer. *PeerJ* 2:e585. <https://doi.org/10.7717/peerj.585>.
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. <https://doi.org/10.1089/cmb.2012.0021>.
7. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
8. Lee I, Ouk Kim Y, Park SC, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>.
9. Auch AF, von Jan M, Klenk H-P, 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. <https://doi.org/10.4056/sigs.531120>.
10. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
11. Cosentino S, Voldby Larsen M, Moller Aarestrup F, Lund O. 2013. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. *PLoS One* 8:e77302. <https://doi.org/10.1371/journal.pone.0077302>.
12. Jia B, Raphenya AR, Alcock B, Wagglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 45:D566–D573. <https://doi.org/10.1093/nar/gkw1004>.
13. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
14. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.