Draft Genome Sequence of *Mycobacteroides* sp. Strain LB1, Isolated from the Sputum of a Cystic Fibrosis Patient

Peter Menzel,[a] Rolf Schwarzer,[a] Johannes Helmuth,[a] Inna Friesen[a]

aLabor Berlin-Charité Vivantes GmbH, Berlin, Germany

ABSTRACT   We report the draft genome sequence of *Mycobacteroides* sp. strain LB1, isolated from the sputum of a cystic fibrosis patient in Berlin, Germany. The genome size is 4.9 Mbp with a GC content of 63.8%. The genome is only distantly related to other *Mycobacteroides* species, suggesting that it may represent a novel species.

Mycobacteroides is a recently proposed bacterial genus in the family *Mycobacteriaceae*, comprising rapid-growing species in the former *Mycobacterium abscessus*/M. *chelonae* cluster (1). While *Mycobacteroides* species typically inhabit natural environments, such as soil or water, some species are opportunistic human pathogens that can cause infections in a variety of organs (2). The risk of infection is typically increased by underlying conditions, such as depression of the immune system or lung diseases, e.g., cystic fibrosis (3). Here, we report the genome of a possibly novel species, *Mycobacteroides* sp. strain LB1, which was isolated from the sputum of a 22-year-old patient from Berlin (Germany) with the underlying conditions of cystic fibrosis and severe bronchiectasis.

Cultivation of the isolate from an undiluted and decontaminated sample on Middlebrook 7H9 broth under aerobic conditions was successful after 21 days in a Bactec MGIT 960 culture system (Becton Dickinson) at 30°C. No growth on solid media (Löwenstein-Jensen and Stonebrink agars, Artelt-Enclit) could be obtained after subcultivation with incubation for 16 weeks. Due to the failed growth during subcultivation, antibiotic susceptibility testing could not be performed. The tuberculosis antigen MPT64 rapid test (Becton Dickinson) for detection of the presence of the *Mycobacterium tuberculosis* complex was negative. Thereafter, species identification was performed using the GenoType *Mycobacterium* common mycobacteria (CM) assay v2.0 (Hain Lifescience GmbH), which identified the isolate as an unspecified *Mycobacterium* species. Further, sequencing of the V3-V4 region of the 16S rRNA gene, using the Illumina 16S metagenomic sequencing protocol (https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf) and 2 × 250-bp reads on the Illumina MiSeq platform, also could not resolve the species, and we therefore opted for whole-genome sequencing.

Genomic DNA was extracted from the liquid culture (see above) and purified using the DNeasy PowerSoil pro kit (Qiagen), and an Illumina sequencing library was prepared using the QiaSeq FX kit (Qiagen) following the manufacturer's protocol. Sequencing was performed on an Illumina MiSeq platform using 2 × 250-bp paired-end reads, resulting in 1.5 million read pairs. After adapter trimming by fastp v0.20.0 (4), reads were assembled *de novo* with the assembly pipeline shovill v1.0.9 (https://github.com/tseemann/shovill), which uses SPAdes v3.14.0 (5). The assembly comprises 16 contigs longer than 500 bp, with a combined length of 4.9 Mbp and an N_{50} value of 596 kbp. The GC content is 63.8%. Default parameters were used for all programs.

The assembly was screened against all complete and draft bacterial genomes available in the NCBI RefSeq database (6) on 30 June 2020 using Mash Screen v2.0 (7). Genome-wide average nucleotide identities (ANI) between the identified most similar

genomes and our assembly were calculated using FastANI v1.3 (8). The genomes most similar to that of *Mycobacteroides* sp. strain LB1 belong to several strains of *Mycobacteroides salmoniphilum*, with a maximum ANI of 85.9%. Other *Mycobacteroides* species show lower similarities, such as *M. chelonae* (84.5% ANI), *M. saopaulense* (84.3% ANI), and *M. franklinii* (84.2% ANI), and the most similar species from a different genus is *Mycolicibacterium fortuitum* (78.0% ANI). The *Mycobacteroides* sp. strain LB1 genome was annotated via the NCBI Prokaryotic Genome Annotation Pipeline (9), which identified 4,659 protein coding sequences.

Given the atypical slow growth and low genome similarity to other *Mycobacteroides* species, *Mycobacteroides* sp. strain LB1 possibly represents a novel species within this genus.

Data availability. Sequencing data are available at NCBI BioProject PRJNA644770, and the genome assembly is available at DDBJ/ENA/GenBank under the accession number JACCFF000000000.

REFERENCES

1. Gupta RS, Lo B, Son J. 2018. Phylogenomics and comparative genomic studies robustly support division of the genus Mycobacterium into an emended genus Mycobacterium and four novel genera. Front Microbiol 9:67. https://doi.org/10.3389/fmicb.2018.00067.

2. Tortoli E. 2014. Microbiological features and clinical relevance of new species of the genus Mycobacterium. Clin Microbiol Rev 27:727–752. https://doi.org/10.1128/CMR.00035-14.

3. Simmon KE, Brown-Elliott BA, Ridge PG, Durtschi JD, Mann LB, Slechta ES, Steigerwalt AG, Moser BD, Whitney AM, Brown JM, Voelkerding KV, McGowan KL, Reilly AF, Kirn TJ, Butler WR, Edelstein PH, Wallace RJ, Jr, Petti CA. 2011. Mycobacterium chelonae-abscessus complex associated with sinopulmonary disease, Northeastern USA. Emerg Infect Dis 17:1692–1700. https://doi.org/10.3201/eid1709.101667.

4. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

5. 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.

6. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44:D733–D745. https://doi.org/10.1093/nar/gkv1189.

7. Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. 2019. Mash Screen: high-throughput sequence containment estimation for genome discovery. Genome Biol 20:232. https://doi.org/10.1186/s13059-019-1841-x.

8. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9.

9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.