Whole-Genome Sequence of *Haloimpatiens lingqiaonensis* Strain P8956

Hussein Anani,[a,b] Didier Raoult,[b,c,d] Pierre-Edouard Fournier[a,b]

aAix Marseille Université, Institut de Recherche pour le Développement (IRD), Service de Santé des Armées, AP-HM, UMR Vecteurs Infections Tropicales et Méditerranéennes (VITROME), Institut Hospitalo-Universitaire Méditerranée Infection, Marseille, France

bInstitut Méditerranée-Infection, Marseille, France

cAix-Marseille Université, Institut de Recherche pour le Développement (IRD), UMR Microbes Evolution Phylogeny and Infections (MEPHI), Institut Hospitalo-Universitaire Méditerranée-Infection, Marseille, France

dSpecial Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT In 2016, *Haloimpatiens lingqiaonensis* was described as a bacterial isolate from paper mill wastewater. Previously, no whole-genome sequence was available for this microorganism. Whole-genome sequencing of strain P8956 yielded a 3,295,388-bp genome with a 30.7% G+C content, 2,917 protein-coding genes, and 95 predicted RNA genes.

In 2016, Wu et al. proposed the creation of the bacterial genus *Haloimpatiens* within the family *Clostridiaceae* and the phylum *Firmicutes* (1). Strain ZC-CMC3[T] was described as the type strain of the species *Haloimpatiens lingqiaonensis* (2). Strain ZC-CMC3[T] had been isolated from wastewater samples collected from a paper mill in Lingqiao City, Zhejiang, China (2). The organism was isolated as described by Wu et al. in 2016 (1). In our laboratory, growth of *Haloimpatiens lingqiaonensis* strain P8956 was obtained after 24 h of culture in 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) in an anaerobic atmosphere at 37°C (3–5). This bacterium exhibited a 94.04% 16S rRNA sequence similarity with *Hathewaya histolytica* strain JCM 1403[T] (GenBank accession number NR_113187), its closest phylogenetic neighbor.

In the present study, DNA from *Haloimpatiens lingqiaonensis* strain P8956 was extracted using the EZ1 biorobot and the EZ1 DNA tissue kit (Qiagen, Hilden, Germany). Extracted DNA was quantified at 0.2 μg/μl using a Qubit assay with a high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) and sequenced with the MiSeq sequencer (Illumina, Inc., San Diego, CA, USA). To prepare the paired-end library, dilution was performed to require 1 ng of each genome as input to prepare the paired-end library. The "tagmentation" step fragmented and tagged the DNA. The DNA was fragmented and amplified by limited PCR (12 cycles), introducing dual-index barcodes and sequencing adapters. After purification on AMPure XP beads (Beckman Coulter, Inc., Fullerton, CA, USA), libraries were normalized and pooled for sequencing following the MiSeq System Denature and Dilute Libraries Guide 15039740-10 (Illumina kit). Paired-end sequencing and automated cluster generation with dual-indexed 2 × 250-bp reads were performed during a 39-hour run. Total information of 8.2 Gb was obtained from a 1,207,000/mm² cluster density with a cluster passing quality control filters of 89.3% (10,507.2 passed filtered reads). A total of 5,876,657 reads were quality checked using FastQC and trimmed using Trimmomatic version 0.36.6 (6). MiSeq reads were assembled using SPAdes version 3.5.0 software (7). The "careful" option was used in order to reduce the number of mismatches and short indels. Default parameters were applied here and for all software (for k values, i.e., k-mer values of 127, 99, 77, 55, 33, and 21). SSPACE (8) and GapFiller (9) were used to combine contigs with default parameters. The draft genome sequence of *Haloimpatiens lingqiaonensis* strain P8956 is composed

of 185 contigs (N_{50}, 36,473 contigs; L_{50}, 30 contigs; coverage, 35×) for a total of 3,295,388-bp with a 30.7% G+C content. Annotation using Prokka version 1.13 (10) predicted 3,012 genes and 2,917 protein-coding genes, 1,360 (45.15%) of which were assigned to clusters of orthologous group categories. In addition, 95 RNA genes were detected (10 rRNAs and 85 tRNAs). Using BLAST methods against resistance and toxin-antitoxin databases (11–13) with default parameters, genes with identity above 80% were taken into account. Six antibiotic resistance-associated genes were predicted, coding resistance to lincosamides (3 genes), oxacillin (1), beta-lactams (1), and mupirocin (1). No toxin-antitoxin module or bacteriocin-associated gene could be found.

Data availability. The draft genome and read sequences of *H. lingqiaonensis* strain P8956 (BioProject number PRJEB32392 and BioSample number SAMEA5587607) have been deposited at EBI/GenBank under the accession numbers CABDWS010000001 to CABDWS010000185 and ERR3393035.

REFERENCES

1. Wu D, Zhang N-F, Sun C, Zhang W-W, Han S-B, Pan J, Wu M, Th D, Zhu X-F. 2016. *Haloimpatiens lingqiaonensis* gen. nov., sp. nov., an anaerobic bacterium isolated from paper-mill wastewater. Int J Syst Evol Microbiol 66:628–632. https://doi.org/10.1099/ijsem.0.000765.

2. Parte AC. 2018. LPSN—List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. Int J Syst Evol Microbiol 68:1825–1829. https://doi.org/10.1099/ijsem.0.002786.

3. Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, Trape J-F, Koonin EV, La Scola B, Raoult D. 2012. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193. https://doi.org/10.1111/1469-0691.12023.

4. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, Raoult D. 2015. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 28:237–264. https://doi.org/10.1128/CMR.00014-14.

5. Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, Caputo A, Cadoret F, Traore SI, Seck EH, Dubourg G, Durand G, Mourembou G, Guilhot E, Togo A, Bellali S, Bachar D, Cassir N, Bittar F, Delerce J, Mailhe M, Ricaboni D, Bilen M, Dangui Nieko NPM, Dia Badiane NM, Valles C, Mouelhi D, Diop K, Million M, Musso D, Abrahão J, Azhar EI, Bibi F, Yasir M, Diallo A, Sokhna C, Djossou F, Vitton V, Robert C, Rolain JM, La Scola B, Fournier P-E, Levasseur A, Raoult D. 2016. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 1:16203. https://doi.org/10.1038/nmicrobiol.2016.203.

6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

7. 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.

8. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. https://doi.org/10.1093/bioinformatics/btq683.

9. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13:R56. https://doi.org/10.1186/gb-2012-13-6-r56.

10. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

11. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain J-M. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 58:212–220. https://doi.org/10.1128/AAC.01310-13.

12. Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive Web interface. Nucleic Acids Res 47:D687–D692. https://doi.org/10.1093/nar/gky1080.

13. Anani H, Abou Abdallah R, Chelkha N, Fontanini A, Ricaboni D, Mailhe M, Raoult D, Fournier P-E. 2019. Draft genome and description of *Merdibacter massiliensis* gen. nov., sp. nov., a new bacterium genus isolated from the human ileum. Sci Rep 9:7931. https://doi.org/10.1038/s41598-019-44343-8.