Complete Genome Sequence of *Marinobacterium* sp. Strain LSUCC0821, Isolated from the Coastal Gulf of Mexico

Anna M. Lucchesi,a Michael W. Henson,b,* Ben Temperton,c J. Cameron Thrashb

aDepartment of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA
bDepartment of Biological Sciences, University of Southern California, Los Angeles, California, USA
cSchool of Biosciences, University of Exeter, Exeter, United Kingdom

**ABSTRACT** Here, we present the complete genome sequence of *Marinobacterium* sp. strain LSUCC0821, isolated from the coastal Gulf of Mexico with artificial seawater using high-throughput dilution-to-extinction (DTE) cultivation. The 2.36-Mbp circularized genome sequence has 2,231 predicted genes, a 91.5% coding density, and a GC content of 47.8%.

*Marinobacterium* sp. strain LSUCC0821 was cultured from Bay de Pomme D’Or (Louisiana, Gulf of Mexico) in January 2017 using dilution-to-extinction (DTE) cultivation and artificial seawater (1, 2). The genus *Marinobacterium* (order *Oceanospirillales*) has been found in diverse environments, including oil seeps, tidal flats, and coral mucus (3–5). A BLASTn (default settings) search of the 16S rRNA gene sequence (1) against the NCBI nucleotide (nt) database showed the closest cultured representatives of LSUCC0821 (GenBank accession number MK603728) to be *Marinobacterium* sp. IMCC1424 (98% identity; KF146347) and *Marinobacterium marisflavum* IMCC4074 (97% identity; NR_125520) (6). We sequenced the genome of strain LSUCC0821 because of its geographic novelty and ability to enhance future comparative genomic studies of an important marine clade.

Cells for short-read sequencing were revived from four cryostocks of the Thrash lab culture collection in sterile polycarbonate flasks containing 50 ml of MWH3 medium (1). From one revival culture, two additional replicate cultures were grown in 500 ml of MWH3 medium to obtain enough biomass for long-read sequencing. All cultures were grown to peak cell density and filtered onto 0.2-μm polycarbonate filters (Pall, USA). DNA was extracted using lysozyme digestion, phenol-chloroform extraction, and isopropanol precipitation (7, 8). Prior to library preparation, DNA for short- and long-read sequencing was cleaned and concentrated using the genomic DNA clean and concentrator kit (Zymo Research, USA) following the manufacturer’s instructions.

Short-read library preparation and sequencing were completed at the USC Genome Core (University of Southern California, USA). Briefly, the libraries were constructed using the KAPA HyperPlus library preparation kit (Kapa Biosystems, Inc., USA), quantified using the Qubit fluorometer, and analyzed for size and quality using the Agilent BioAnalyzer system. The library DNA was pooled equimolarly and sequenced on an Illumina NextSeq 550 instrument, using a midoutput flow cell in paired-end (PE) 150-bp cycle format, generating 13,776,782 reads. The long-read sequencing libraries were prepared using the SQK-LSK108 1D genomic DNA ligation kit (Oxford Nanopore, UK) with slight modifications (available at https://doi.org/10.17504/protocols.io.bixskfne) and sequenced on a MinION instrument using an R9.4 flow cell (Oxford Nanopore), generating 247,039 raw reads with an N50 value of 7,955 bp. FAST5 sequences were base called with Guppy v.2.3.1 (Oxford Nanopore Technologies) using flipflop mode. Adapters and split reads with
adapters in the middle were removed using Porechop (https://github.com/rrwick/Porechop) with the flags “require_two_barcodes” and “discard_middle.”

Genome assembly was completed using long and short reads with Unicycler v.0.4.8-beta with default settings (9), resulting in a single circular scaffold (rotation and circularity confirmed via Unicycler; assembly graph available at https://doi.org/10.6084/m9.figshare.12857879). The contamination and quality were assessed using CheckM v.1.0.3 with default settings (10). The assembled genome sequence was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (11). The completed genome sequence of LSUCC0821 is 2,360,824 bp (105× coverage), with a GC content of 47.8%, a coding density of 91.5%, and estimated contamination of 0.43%. There are 2,303 predicted genes, 54 tRNA genes, and 3 each of the 5S, 16S, and 23S rRNA genes.

Data availability. Sample information, fastq sequences, and genomic assembly/annotation are accessible under the NCBI BioProject and whole-genome sequence (WGS) accession numbers PRJNA589095 and CP051666, respectively. Cryostocks and/or live cultures of LSUCC0821 are available upon request.

ACKNOWLEDGMENTS

Portions of this research were conducted with high-performance computing resources provided by the University of Southern California.

This work was funded by the Louisiana Board of Regents (Board of Regents) (LEQSF [2014-2017]-RDA-06), the Louisiana State University Department of Biological Sciences, and a National Academy of Science, Engineering, and Medicine Gulf Research Program Early Career Research Fellowship to J.C.T. Nanopore sequencing was supported by an American Natural History Museum Lerner Gray grant to M.W.H.

REFERENCES