AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Reference Genome Sequence of *Corynebacterium mastitidis* RC, an Ocular Commensal, Isolated from Mouse Conjunctiva

Vijayaraj Nagarajan,[a] Anthony J. St. Leger,[b,c] Amy Zhang,[a] Phyllis Silver,[a] Rachel R. Caspi[a]

[a]Laboratory of Immunology, National Eye Institute, NIH, Bethesda, Maryland, USA
[b]Department of Ophthalmology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
[c]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

ABSTRACT Here, we report the genome sequence of a protective commensal, *Corynebacterium mastitidis* RC, isolated from mouse conjunctiva. The *C. mastitidis* RC genome sequence is 2,153,054 bp in size and 96.95% complete, and we believe that it can contribute to the understanding of the functional immune attributes of the ocular commensal microbiome.

Corynebacterium is a Gram-positive, pleomorphic bacterium that exists frequently as a normal flora inhabiting skin, mucous membrane, and conjunctiva. Several *Corynebacterium* species have also been identified as either pathogens or opportunistic pathogens causing diseases like diphtheria, endocarditis, and mastitis (1). Genome sequences of two *Corynebacterium mastitidis* species, namely, one isolated from the milk of a sheep with subclinical mastitis (2) and the other isolated from a mouse with keratitis (3), are currently available in NCBI genome database. Here, we report the nearly complete draft genome sequence of *C. mastitidis* strain RC, identified as a protective commensal on the ocular surface in our previous study (4). The *C. mastitidis* strain RC was isolated from the conjunctiva of a wild-type, 8-week-old, female NIH mouse (C57BL/6). The single-colony isolates were identified as *C. mastitidis* using 16S rRNA gene sequencing.

Single-colony isolates of *C. mastitidis* were obtained using methods described in our previous study (4). Briefly, mice conjunctiva homogenate was spread on tryptic soy agar (TSA) plates and incubated for 7 days. The isolated *C. mastitidis* colonies were confirmed as axenic using 16S rRNA gene sequencing. For sequencing, single-colony isolates of *C. mastitidis* RC were cultured using tryptic soy agar + 5% blood agar plates, at 37°C under aerobic conditions. DNA was extracted using the bacterial DNA prep kit (Zymo Research, Irvine, CA) as per the manufacturer's instructions. DNA concentration was quantified with Qubit (ThermoFisher, Waltham, MA). Sequencing libraries were prepared with the Nextera XT DNA library prep kit (Illumina, San Diego, CA). Library quality was accessed using TapeStation system (TapeStation Analysis Software A.02.01 SR1; high-sensitivity D5000 ScreenTape kit; Agilent Technologies, Inc.). Paired-end sequencing was done in the Illumina MiSeq platform, with the 600 cycle V3 kit box 1, using one flow cell. FastQC (v0.11.9) (5) was used to analyze the quality of the reads before and after processing. BBTools (v38.87) (6) was used to remove adapters and preprocess the reads. A *de novo* genome assembly was done using the SPAdes optimizer Unicycler (v0.4.8) (7), comprising of SPAdes (v3.15.2) (8) and Pilon (v1.23) (9), with default parameters for all three available modes. Assembly quality was assessed using CheckM (v1.1.3) (10) and QUAST (v5.0.2) (11), using default parameters. The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (v6.0) was used for annotation (12). The complete code, with parameters and custom scripts used for each of the analysis components of the workflow,

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Address correspondence to Vijayaraj Nagarajan, nagarajanv@nih.gov.

The authors declare no conflict of interest.

Received 10 March 2022
Accepted 26 April 2022
Published 10 May 2022

is provided at the GitHub repository online at https://github.com/NIH-NEI/cmast-genome-assembly.

A total of 3.8 million reads with a 250-bp read length were generated using the paired-end sequencing technique. The assembled *C. mastitidis* RC genome is 2,153,054 bp long, with 42 contigs at the *N*50 of 130,636 bp and NG50 of 130,636 bp, 80× coverage, and the GC content of 69.06%, after removing 5 contigs with a size less than 200 bp. Compared with the 585 gene markers in the *Corynebacterium* lineage, *C. mastitidis* RC shows a 96.95% completeness. The PGAP annotations identified 2,132 total genes in *C. mastitidis* RC, comprising of 2,050 protein-coding genes, 52 tRNAs, 3 rRNAs, 2 noncoding RNAs (ncRNAs), and 2 CRISPR arrays.

Data availability. The raw sequencing data are deposited and available at NCBI SRA under the accession number SRR15665287. The complete project information is available under the NCBI BioProject identifier (ID) PRJNA758739. The sample information is available under the NCBI BioSample ID SAMN21033600. The annotated genome sequence is deposited in NCBI under the accession JAKRKB000000000. The code and the workflow documentation describing the genome assembly are deposited and available at the github repository online at https://github.com/NIH-NEI/cmast-genome-assembly. The permanently linked code release is also published in Zenodo, online at https://doi.org/10.5281/zenodo.6282054.

ACKNOWLEDGMENTS

This work utilized the computational resources of the NIH HPC Biowulf cluster. The library preparation and sequencing were carried out at the NCI Microbiome and Genetics Core.

This research was supported in part by the Intramural Research Program of the NIH, National Eye Institute, project number EY000184 and R01 EY032482.

REFERENCES

1. Aoki T, Kitazawa K, Deguchi H, Sotozono C. 2021. Current evidence for Corynebacterium on the ocular surface. Microorganisms 9:254. https://doi.org/10.3390/microorganisms9020254.
2. Fernandez-Garayzabal JF, Collins MD, Hutson RA, Fernandez E, Monasterio R, Marco J, Dominguez L. 1997. Corynebacterium mastitidis sp. nov., isolated from milk of sheep with subclinical mastitis. Int J Syst Bacteriol 47:1082–1085. https://doi.org/10.1099/00207713-47-4-1082.
3. Cheleuitte-Nieves C, Gulvik CA, Humrighouse BW, Bell ME, Villarma A, Westblade LF, Lipman NS, Fischetti VA, McQuiston JR. 2018. Draft reference genome sequence of Corynebacterium mastitidis 16–1433, osolated from a mouse. Genome Announc 6:e00050-18. https://doi.org/10.1128/genomeA.00050-18.
4. St Leger AJ, Desai JV, Drummond RA, Kugadas A, Almaghrabi F, Silver P, Raychaudhuri K, Gadjeva M, Iwakura Y, Lionakis MS, Caspi RR. 2017. An ocular commensal protects against corneal infection by driving an interleukin-17 response from mucosal gammadelta T cells. Immunity 47:148–158.e5. https://doi.org/10.1016/j.immuni.2017.06.014.
5. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data [online]. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
6. Bushnell B. 2022. BBMap short read aligner, and other bioinformatic tools. https://sourceforge.net/projects/bbmap/.
7. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
8. 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.
9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
11. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
12. 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.