Draft Genome Sequence of *Candida auris* Strain LOM, a Human Clinical Isolate from Greater Metropolitan Houston, Texas

Running Head: *Candida auris* LOM

S. Wesley Long[1,2,#], Randall J. Olsen[1,2,#], Hoang A. T. Nguyen[1], Matthew Ojeda Saavedra[1], and James M. Musser[1,2]

[1]Center for Molecular and Translational Human Infectious Diseases Research and Houston Methodist Diagnostic Laboratory, Department of Pathology and Genomic Medicine, Houston Methodist Research Institute and Houston Methodist Hospital, Houston, Texas, USA

[2]Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, New York USA

[#]Address correspondence to: S. Wesley Long, M.D., Ph.D., or Randall J. Olsen, M.D., Ph.D., Department of Pathology and Genomic Medicine, Houston Methodist Hospital, 6565 Fannin St., B250, Houston, Texas 77030.

Abstract word count: 58

Text word count: 590

ABSTRACT

Candida auris is an emerging pathogen with considerable public health importance. We present the draft genome sequence of a strain recently cultured from the urine of a patient hospitalized in the greater Houston metropolitan region. Two combined Oxford Nanopore sequencing runs provided sufficient data to rapidly generate a draft genome with 9 contigs covering the expected 7 chromosomes.

ANNOUNCEMENT

Candida auris recently has emerged worldwide and caused outbreaks in healthcare facilities (1, 2). An isolate of *C. auris* cultured from the urine of a patient from the Houston metropolitan region was identified by Brucker MALDI-TOF analysis using the RUO database. Given the potential infection control and public health importance of this isolate, we rapidly characterized its genome using two Oxford Nanopore GridION runs. Initial DNA extraction was performed using ballistic lysis with FastPrep matrix B beads and the Qiagen DNA Blood/Tissue Kit. The first GridION run was performed with the Rapid Barcoding Kit from Oxford Nanopore (SQK-RBK004). The second run was performed using DNA extracted using FastPrep matrix Y beads, the Qiagen DNA Blood/Tissue Kit. The Ligation Sequencing Kit from Oxford Nanopore (SQK-LSK109) with a long read wash step to select for longer reads as suggested in the manufacturer's protocol. Initial genome assembly of GridION data from the first 16 hours of the first sequencing run using Unicycler v0.4.3 yielded 11 contigs in the aggregate corresponding to 12,237,085

bps (3). After completion of the second GridION run, the long reads from both runs were combined and filtered using FiltLong v0.1.1 (https://github.com/rrwick/Filtlong) with a minimum read length cutoff at 5 kb and a target of 100x-fold coverage. Unicycler v0.4.3 was used to assemble the genome from the long reads. The final assembly had 9 contigs with a total length of 12,293,266 bps. The largest contig is 4,297,164 bps, and the N50 is 2,304,466 bps. The average GC content is 45.19%. This work was approved by our Institutional Review Board (IRB1010-0199).

We compared our assemblies with reference genome assemblies deposited in GenBank using progressiveMauve v2.4.0 and discovered that our 7 longest contigs correspond to the expected 7 chromosomes present in other *C. auris* strains (4). Phylogenetic analysis using mummer v4.0 to compare against the available reference genomes shows that *C. auris* strain LOM is most closely related to strain B11221, a representative strain from clade III, a lineage with many South African strains (5, 6). This relationship was confirmed using prephix and phrecon to generate an alignment for FastTree to create a Neighbor Joining phylogenetic tree (https://github.com/codinghedgehog/). The *C. auris* strain LOM genome has the *ERG11* gene with a F126T amino acid replacement that is common to clade III strains. (6).

Although an in-depth epidemiological investigation had not been completed at the time of manuscript submission, the patient had not recently traveled outside the United States, had a history of long-term care at outside facilities, and had very recently been transferred to hospital A from another facility. The organism was

classified as present on admission. This work further demonstrates the usefulness of real-time long read whole genome sequencing to rapidly provide relevant information concerning emerging pathogens of significant concern (7, 8). The availability of this high-quality draft genome sequence serves as a useful resource if additional isolates are recovered in greater Houston, a metropolitan area with an international, ethnically diverse population of approximately 6 million. The sequence data also will assist translational research efforts designed to more fully understand the molecular mechanisms underlying host interactions in an emerging pathogen that has substantial detrimental public health potential.

Data availability and accession numbers

The BioProject accession for the draft genome sequence data of *C. auris* strain LOM is PRJNA540998. The GridION runs are present in the SRA (SRR9017243 and SRR9017244). The assembled draft genome can be found at (Genbank # pending).

ACKNOWLEDGMENTS

We thank the technical staff in the Houston Methodist Hospital Clinical Microbiology Laboratory for assistance. We thank Andrew Pann for his development and continued support of the prephix and phrecon tools. Supported by funds from the L. E. Simmons Family Foundation and Fondren Foundation to J.M.M.

1. Spivak ES, Hanson KE. 2018. *Candida auris*: an Emerging Fungal Pathogen. J Clin Microbiol 56:e01588-17.
2. Jeffery-Smith A, Taori SK, Schelenz S, Jeffery K, Johnson EM, Borman A, Manuel R, Brown CS. 2018. Candida auris: a Review of the Literature. Clin Microbiol Rev 31.
3. 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.
4. Darling AE, Tritt A, Eisen JA, Facciotti MT. 2011. Mauve assembly metrics. Bioinformatics (Oxford, England) 27:2756-2757.
5. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biol 5:R12.
6. Muñoz JF, Gade L, Chow NA, Loparev VN, Juieng P, Berkow EL, Farrer RA, Litvintseva AP, Cuomo CA. 2018. Genomic insights into multidrug-resistance, mating and virulence in Candida auris and related emerging species. Nature communications 9:5346-5346.
7. Long SW, Kachroo P, Musser JM, Olsen RJ. 2017. Whole-genome sequencing of a human clinical isolate of *emm28 Streptococcus pyogenes* causing necrotizing fasciitis acquired contemporaneously with Hurricane Harvey. Genome Announc 5.
8. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouedraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Camino-Sanchez A, Carter LL, Doerrbecker J, Enkirch T, Dorival IGG, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallash E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Trina R, Bello A, Sall AA, Faye O, et al. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530:228-232.