Complete Genome Sequences of Two USA300-Related Community-Associated Methicillin-Resistant *Staphylococcus aureus* Clinical Isolates

Jo-Ann McClure,a Kunyan Zhanga,b,c,d,e

aCentre for Antimicrobial Resistance, Alberta Health Services/Alberta Public Laboratories/University of Calgary, Calgary, Alberta, Canada
bDepartment of Pathology and Laboratory Medicine, University of Calgary, Calgary, Alberta, Canada
cDepartment of Microbiology, Immunology, and Infectious Diseases, University of Calgary, Calgary, Alberta, Canada
dDepartment of Medicine, University of Calgary, Calgary, Alberta, Canada
eThe Calvin, Phoebe, and Joan Snyder Institute for Chronic Diseases, University of Calgary, Calgary, Alberta, Canada

ABSTRACT USA300 is a predominant community-associated methicillin-resistant *Staphylococcus aureus* strain causing significant morbidity and mortality in North America. We present the full annotated genome sequences of two methicillin-resistant *Staphylococcus aureus* isolates related to the USA300 pulsotype with the goal of studying the evolutionary relationships of this highly successful strain type.

Community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) pulsed-field gel electrophoresis (PFGE) strain type USA300 has rapidly become the predominant MRSA strain in North America (1–7). The group comprises a family of strains with related PFGE banding patterns belonging to clonal complex 8 (CC8) and multilocus sequence type 8 (ST8) and carrying staphylococcal cassette chromosome *mec* type IV. Although it is the leading cause of MRSA skin and soft tissue infections in North America, USA300 can also cause life-threatening diseases, such as sepsis and necrotizing pneumonia, and is appearing in outbreaks worldwide. Because of the unique success of USA300, there is significant interest in investigating the genetic diversity of this group, as well as in understanding its evolutionary history. We selected two MRSA isolates of human origin that are very closely related to USA300 based on PFGE and molecular analyses, with both being ST8, *spa* type t008, and staphylococcal cassette chromosome *mec* (SCC*mec*) type IV, for whole-genome sequencing. Strain C8879 was collected in 2002, prior to a USA300 outbreak in the Calgary Health Region in 2004 (8), while strain JK3137 was collected in 2011, 7 years after the outbreak. Both strains were from hospital patients. These sequences, spanning a 9-year window, will help increase our understanding of the evolution of this successful lineage.

Overnight bacterial cultures were started from single colonies and grown at 37°C, and then genomic DNA was obtained from the MRSA isolates using a phenol-chloroform extraction. Library preparation, DNA sequencing, contig assembly, and genome circularization were performed at the Génome Québec Innovation Centre in Montreal, Quebec, Canada. Sheared large-insert libraries were generated with Covaris g-TUBEs and the SMRTbell template prep kit version 1.0 and then sequenced with Pacific Biosciences (PacBio) RSII sequencing technology using one single-molecule real-time (SMRT) cell. Contig assembly was done using the RS Hierarchical Genome Assembly Process (HGAP) protocol version 2.3.0.140936.p5 using default settings (9–11). The read quality was ensured by aligning shorter reads on longer reads using Basic Local Alignment with Successive Refinement (BLASR) (9). The finalized genome was generated by circularizing with Circlator version 1.4.1 (12) and adjusted to the origin of

replication. Gene annotation was done using the NCBI's Prokaryotic Genome Annotation Pipeline version 4.1 (13).

Three contigs resulted from the assembly of the MRSA strain C8879 reads. There were 91,261 raw reads covering 1,077,017,883 sequenced bases with an average read length of 11,801 bp, an estimated genome coverage of 348×, and a GC content of 32.81%. The resulting chromosome was 2,832,897 bp long with 2,972 genes identified, of which 2,890 were coding DNA sequences (CDS), 82 were RNA genes, and 74 were pseudogenes. Four contigs were produced from the assembly of the JK3137 reads. There were 105,404 raw reads covering 1,138,170,110 sequenced bases with an average read length of 10,798 bp, an estimated genome coverage of 366×, and a GC content of 32.82%. The resulting chromosome was 2,955,000 bp long with 3,150 genes identified, of which 3,068 were CDS, 82 were RNA genes, and 106 were pseudogenes.

A complete analysis is under way looking at the chromosomal components in each of these strains and how they compare to those in the highly successful USA300 outbreak isolate C2406 (14).

Data availability. The chromosomal genome sequences have been deposited at GenBank under the accession numbers CP020956 (C8879) and CP020960 (JK3137) and SRA accession numbers SRX5621936 (C8879) and SRX5627280 (JK3137).

REFERENCES

1. Carrel M, Perencevich EN, David MZ. 2015. USA300 methicillin-resistant *Staphylococcus aureus*, United States, 2000–2013. Emerg Infect Dis 21: 1973–1980. https://doi.org/10.3201/eid2111.150452.

2. Diekema DJ, Richter SS, Heilmann KP, Dohrn CL, Riahi F, Tendolkar S, McDanel JS, Doern GV. 2014. Continued emergence of USA300 methicillin-resistant *Staphylococcus aureus* in the United States: results from a nationwide surveillance study. Infect Control Hosp Epidemiol 35:285–292. https://doi.org/10.1086/675283.

3. Jenkins TC, McCollister BD, Sharma R, McFann KK, Madinger NE, Barron M, Bessesen M, Price CS, Burman WJ. 2009. Epidemiology of healthcare-associated bloodstream infection caused by USA300 strains of methicillin-resistant *Staphylococcus aureus* in 3 affiliated hospitals. Infect Control Hosp Epidemiol 30:233–241. https://doi.org/10.1086/595963.

4. Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Zell ER, Fosheim GE, McDougal LK, Carey RB, Fridkin SK, Active Bacterial Core Surveillance (ABCs) MRSA Investigators. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. JAMA 298:1763–1771. https://doi.org/10.1001/jama.298.15.1763.

5. Kreisel KM, Stine OC, Johnson JK, Perencevich EN, Shardell MD, Lesse AJ, Gordin FM, Climo MW, Roghmann MC. 2011. USA300 methicillin-resistant *Staphylococcus aureus* bacteremia and the risk of severe sepsis: is USA300 methicillin-resistant *Staphylococcus aureus* associated with more severe infections? Diagn Microbiol Infect Dis 70:285–290. https://doi.org/10.1016/j.diagmicrobio.2011.03.010.

6. McCaskill ML, Mason EO, Jr, Kaplan SL, Hammerman W, Lamberth LB, Hulten KG. 2007. Increase of the USA300 clone among community-acquired methicillin-susceptible *Staphylococcus aureus* causing invasive infections. Pediatr Infect Dis J 26:1122–1127. https://doi.org/10.1097/INF.0b013e31814536e0.

7. Simor AE, Gilbert NL, Gravel D, Mulvey MR, Bryce E, Loeb M, Matlow A, McGeer A, Louie L, Campbell J, Canadian Nosocomial Infection Surveillance Program. 2010. Methicillin-resistant *Staphylococcus aureus* colonization or infection in Canada: national surveillance and changing epidemiology, 1995–2007. Infect Control Hosp Epidemiol 31:348–356. https://doi.org/10.1086/651313.

8. Gilbert M, MacDonald J, Gregson D, Siushansian J, Zhang K, Elsayed S, Laupland K, Louie T, Hope K, Mulvey M, Gillespie J, Nielsen D, Wheeler V, Louie M, Honish A, Keays G, Conly J. 2006. Outbreak in Alberta of community-acquired (USA300) methicillin-resistant *Staphylococcus aureus* in people with a history of drug use, homelessness or incarceration. CMAJ 175:149–154. https://doi.org/10.1503/cmaj.051565.

9. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using Basic Local Alignment with Successive Refinement (BLASR): application and theory. BMC Bioinformatics 13:238. https://doi.org/10.1186/1471-2105-13-238.

10. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.

11. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. 2000. A whole-genome assembly of *Drosophila*. Science 287:2196–2204. https://doi.org/10.1126/science.287.5461.2196.

12. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.

13. 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.

14. McClure J-A, Zhang K. 2017. Complete genome sequence of a community-associated methicillin-resistant *Staphylococcus aureus* hypervirulent strain, USA300-C2406, isolated from a patient with a lethal case of necrotizing pneumonia. Genome Announc 5:e00461-17. https://doi.org/10.1128/genomeA.00461-17.