Microbiology
Resource Announcements

Complete Genome Sequence of GD1108, a Moderate-Virulence Strain of Human-Associated ST398 Methicillin-Susceptible *Staphylococcus aureus*

Jo-Ann McClure,[a] Steven M. Shideler,[c] Kunyan Zhang[a,b,c,d,e]

[a]Centre for Antimicrobial Resistance, Alberta Health Services/Alberta Public Laboratories/University of Calgary, Calgary, Alberta, Canada
[b]Department of Pathology & Laboratory Medicine, University of Calgary, Calgary, Alberta, Canada
[c]Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Calgary, Alberta, Canada
[d]Department of Medicine, University of Calgary, Calgary, Alberta, Canada
[e]The Calvin, Phoebe and Joan Snyder Institute for Chronic Diseases, University of Calgary, Calgary, Alberta, Canada

ABSTRACT *Staphylococcus aureus* multilocus sequence type 398 (ST398) is responsible for an increasing number of severe infections in humans. There are no reports detailing if all ST398 strains are equally virulent. We present the genome sequence of the moderate-virulence ST398 methicillin-susceptible *Staphylococcus aureus* strain GD1108, determined in a *Caenorhabditis elegans* infection model, to reveal the ST398 sublineage virulence.

*S*taphylococcus aureus multilocus sequence type 398 (ST398) was first reported as an animal pathogen; however, it gained notoriety when it was found associated with human diseases, ranging from minor localized infections to more severe invasive illnesses (1–17). In humans, there appears to be augmented pathogenicity of ST398 infection (2, 8, 18), but no studies existed comparing the relative virulence of the different sublineages. A preliminary analysis from testing a collection of ST398 isolates using the *Caenorhabditis elegans* infection model revealed three lineages with high, moderate, or low virulence and mean nematode killing rates of 90%, 67%, and 44%, respectively. Whole-genome sequencing was done on representatives from each group, with the goal of detecting genetic determinants that could be responsible for the differing levels of toxicity between the strains. In separate reports, we presented the full chromosomal sequences of the high-virulence strain GD487 and the low-virulence strain GD1696. The complete genome sequence of the moderate-virulence strain GD1108 is presented here.

Strain GD1108 was isolated from a school child from a prevalence survey in 2011 in Guangzhou, People's Republic of China. The genome was generated from the hybrid assembly of PacBio and Illumina sequencing reads. Phenol-chloroform extraction was used to isolate genomic DNA from an overnight bacterial culture, started from a single colony of GD1108 in brain heart infusion broth and grown at 37°C. Illumina library preparation and sequencing were done at the Centre for Health Genomics and Informatics at the University of Calgary, Canada, using the recommended conditions. The library was prepared using the NEBNext Ultra II FS DNA fragment library kit, and then 600 cycle MiSeq v3 sequencing was done. A sheared large-insert PacBio library was prepared at the Genome Quebec Innovation Centre in Montreal, Canada, using Covaris g-TUBEs and the SMRTbell template prep kit 1.0. Sequencing of the library with PacBio RS II sequencing technology, using one single-molecule real-time (SMRT) cell, was also performed at Genome Quebec. Raw Illumina reads had adapters trimmed and sequences with quality scores of <20 removed using Cutadapt v1.15, while sequence quality was assessed with FastQC v0.11.5 (19) (http://www.bioinformatics.babraham.ac

.uk/projects/fastqc). Filtered PacBio subreads (prepared by Genome Quebec) and trimmed Illumina reads were used for hybrid genome assembly using the Unicycler v0.4.7 pipeline (SPAdes v3.13.0, minimap, Racon v1.3.2, Pilon v.1.23) (20–24). Once assembled, GC content was determined with QUAST v4.4, and gene annotation was accomplished using NCBI's Prokaryotic Genome Annotation Pipeline using the best-placed reference protein set (GeneMarkS-2+ v4.8) (25, 26). All programs were run using their default settings.

Following assembly, three contigs were generated from the GD1108 reads, including one representing the chromosome and two representing plasmids (3,175 and 18,638 bp). From PacBio sequencing, there were 1,516,326,066 sequenced bases covered by 110,591 raw reads, with an average read length of 13,711 bp. The average read lengths for Illumina sequencing were 261 bp for R1 and 262 bp for R2, with 412,567 reads generated. The estimated genome coverage for Illumina sequencing was $39\times$, while PacBio sequencing had $487\times$ coverage. The assembled GD1108 chromosome was 2,783,012 bp long, with 2,808 genes identified, of which 2,727 were coding DNA sequences (CDS), 81 were RNA genes, and 84 were pseudogenes. The GC content was 32.98%.

Data availability. The chromosomal genome sequence was deposited at GenBank under the accession number CP040230, with SRA accession numbers SRX5915513 (Illumina) and SRX5915514 (PacBio).

REFERENCES

1. Bradley A. 2002. Bovine mastitis: an evolving disease. Vet J 164:116–128. https://doi.org/10.1053/tvjl.2002.0724.
2. Brunel AS, Banuls AL, Marchandin H, Bouzinbi N, Morquin D, Jumas-Bilak E, Corne P. 2014. Methicillin-sensitive *Staphylococcus aureus* CC398 in intensive care unit, France. Emerg Infect Dis 20:1511–1515. https://doi.org/10.3201/eid2009.130225.
3. Cuny C, Friedrich A, Kozytska S, Layer F, Nubel U, Ohlsen K, Strommenger B, Walther B, Wieler L, Witte W. 2010. Emergence of methicillin-resistant *Staphylococcus aureus* (MRSA) in different animal species. Int J Med Microbiol 300:109–117. https://doi.org/10.1016/j.ijmm.2009.11.002.
4. Dorado-Garcia A, Bos ME, Graveland H, Van Cleef BA, Verstappen KM, Kluytmans JA, Wagenaar JA, Heederik DJ. 2013. Risk factors for persistence of livestock-associated MRSA and environmental exposure in veal calf farmers and their family members: an observational longitudinal study. BMJ Open 3:e003272. https://doi.org/10.1136/bmjopen-2013-003272.
5. Graveland H, Duim B, van Duijkeren E, Heederik D, Wagenaar JA. 2011. Livestock-associated methicillin-resistant *Staphylococcus aureus* in animals and humans. Int J Med Microbiol 301:630–634. https://doi.org/10.1016/j.ijmm.2011.09.004.
6. McNamee PT, Smyth JA. 2000. Bacterial chondronecrosis with osteomyelitis ("femoral head necrosis") of broiler chickens: a review. Avian Pathol 29:253–270. https://doi.org/10.1080/03079450050118386.
7. Menzies PI, Ramanoon SZ. 2001. Mastitis of sheep and goats. Vet Clin North Am Food Anim Pract 17:333–358. https://doi.org/10.1016/S0749-0720(15)30032-3.
8. Rasigade JP, Laurent F, Hubert P, Vandenesch F, Etienne J. 2010. Lethal necrotizing pneumonia caused by an ST398 *Staphylococcus aureus* strain. Emerg Infect Dis 16:1330. https://doi.org/10.3201/eid1608.100317.
9. Springer B, Orendi U, Much P, Hoger G, Ruppitsch W, Krziwanek K, Metz-Gercek S, Mittermayer H. 2009. Methicillin-resistant *Staphylococcus aureus*: a new zoonotic agent? Wien Klin Wochenschr 121:86–90. https://doi.org/10.1007/s00508-008-1126-y.
10. Uhlemann AC, Dumortier C, Hafer C, Taylor BS, Sanchez J, Rodriguez-Taveras C, Leon P, Rojas R, Olive C, Lowy FD. 2012. Molecular characterization of *Staphylococcus aureus* from outpatients in the Caribbean reveals the presence of pandemic clones. Eur J Clin Microbiol Infect Dis 31:505–511. https://doi.org/10.1007/s10096-011-1339-2.
11. Uhlemann AC, Hafer C, Miko BA, Sowash MG, Sullivan SB, Shu Q, Lowy FD. 2013. Emergence of sequence type 398 as a community- and healthcare-associated methicillin-susceptible *Staphylococcus aureus* in northern Manhattan. Clin Infect Dis 57:700–703. https://doi.org/10.1093/cid/cit375.
12. van Belkum A, Melles DC, Peeters JK, van Leeuwen WB, van Duijkeren E, Huijsdens XW, Spalburg E, de Neeling AJ, Verbrugh HA, Dutch Working Party on Surveillance Research of MRSA (SOM). 2008. Methicillin-resistant and -susceptible *Staphylococcus aureus* sequence type 398 in pigs and humans. Emerg Infect Dis 14:479–483. https://doi.org/10.3201/eid1403.070760.
13. van Duijkeren E, Moleman M, Sloet van Oldruitenborgh-Oosterbaan MM, Multem J, Troelstra A, Fluit AC, van Wamel WJ, Houwers DJ, de Neeling AJ, Wagenaar JA. 2010. Methicillin-resistant *Staphylococcus aureus* in horses and horse personnel: an investigation of several outbreaks. Vet Microbiol 141:96–102. https://doi.org/10.1016/j.vetmic.2009.08.009.
14. Voss A, Loeffen F, Bakker J, Klaassen C, Wulf M. 2005. Methicillin-resistant *Staphylococcus aureus* in pig farming. Emerg Infect Dis 11:1965–1966. https://doi.org/10.3201/eid1112.050428.
15. Weese JS. 2010. Methicillin-resistant *Staphylococcus aureus* in animals. ILAR J 51:233–244. https://doi.org/10.1093/ilar.51.3.233.
16. Witte W, Strommenger B, Stanek C, Cuny C. 2007. Methicillin-resistant *Staphylococcus aureus* ST398 in humans and animals, Central Europe. Emerg Infect Dis 13:255–258. https://doi.org/10.3201/eid1302.060924.
17. Yu F, Chen Z, Liu C, Zhang X, Lin X, Chi S, Zhou T, Chen Z, Chen X. 2008. Prevalence of *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes among isolates from hospitalised patients in China. Clin Microbiol Infect 14:381–384. https://doi.org/10.1111/j.1469-0691.2007.01927.x.
18. Valentin-Domelier AS, Girard M, Bertrand X, Violette J, Francois P, Donnio PY, Talon D, Quentin R, Schrenzel J, van der Mee-Marquet N, Bloodstream Infection Study Group of the Reseau des Hygienistes du Centre. 2011. Methicillin-susceptible ST398 *Staphylococcus aureus* responsible

for bloodstream infections: an emerging human-adapted subclone? PLoS One 6:e28369. https://doi.org/10.1371/journal.pone.0028369.

19. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.

20. 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.

21. Li H. 2016. minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152.

22. Vaser R, Sovic I, Nagarajan N, Sikic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.

23. 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.

24. 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.

25. 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.

26. 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.