




Genome Sequences of a *Staphylococcus aureus* Clinical Isolate, Strain SMA0034-04 (UGA22), from Siaya County Referral Hospital in Siaya, Kenya

Gary Xie,^a Qiuying Cheng,^b Hajnalka Daligault,^a Karen Davenport,^a Cheryl Gleasner,^a Lindsey Jacobs,^c Jessica Kubicek-Sutherland,^d  Tessa LeCuyer,^b Vincent Otieno,^e Evans Raballah,^f Harshini Mukundan,^d Douglas J. Perkins,^b Benjamin McMahon,^c Norman Doggett^a

^aBiosecurity and Public Health, Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

^bCenter for Global Health, Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, New Mexico, USA

^cTheoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

^dPhysical Chemistry and Applied Spectroscopy, Chemistry Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

^eUniversity of New Mexico Laboratories of Parasitic and Viral Diseases, Kisumu, Kenya

^fDepartment of Medical Laboratory Sciences, School of Public Health, Biomedical Sciences and Technology, Masinde Muliro University of Science and Technology, Kakamega, Kenya

ABSTRACT Here, we report the genome sequences of a *Staphylococcus aureus* clinical isolate, strain SMA0034-04 (UGA22), which contains one chromosome and one plasmid. We also reveal that isolate SMA0034-04 (UGA22) contains loci in the genome that encode multiple exotoxins.

Staphylococcus aureus, a Gram-positive, round-shaped *Firmicutes* bacterium, is a common cause of skin infections, including abscesses, respiratory infections (such as sinusitis), and food poisoning. Pathogenic strains often promote infections by producing virulence factors, such as potent protein toxins, and by the expression of a cell surface protein that binds and inactivates antibodies. We report here the draft genome sequences of *S. aureus* strain SMA0034-04 (UGA22), isolated from the venous blood of a febrile female pediatric patient (10.8 months old) at the Siaya County Referral Hospital in western Kenya in 2004. Diagnostic results revealed that the bacteremic patient was HIV negative but had *Plasmodium falciparum* malaria with thrombocytopenia.

Although *S. aureus* is a common nosocomial pathogen, this was not the case here. Prior to any treatment interventions, blood was collected upon admission into a pediatric 1.5-ml microbial tube (Isolator, Wampole Laboratories, Cranbury, NJ) and was cultivated at 35°C for 18 to 24 h in 5% CO₂ on 5% sheep blood agar. Bacterial DNA was extracted from a pure culture using the UltraClean microbial DNA isolation kit (Qiagen, Germantown, MD) according to the manufacturer's instructions with minimal modifications. The library was prepared from 100 ng of bacterial DNA using a NEBNext Ultra DNA library prep kit for Illumina (New England Biolabs, Ipswich, MA). *S. aureus* strain SMA0034-04 DNA was draft sequenced to 122-fold coverage using a MiSeq v2 500-cycle sequencing kit (Illumina, San Diego, CA), resulting in 3,480,522 paired-end 251-bp reads. The reference genome and plasmid were determined by the top megablast search hit against the NCBI NR database. The Burrows-Wheeler Aligner (BWA) version 0.7.2 (1) was used to map SMA0034-04 reads covering 94.78% of the *S. aureus* subsp. *aureus* MW2 chromosome (2) and 100% of *S. aureus* pWBG745, a conjugative plasmid that has been known to mobilize unrelated antimicrobial resistance/virulence genes (3). Data quality was assessed and the data files were filtered and trimmed with FaQCs version 1.3 (4) and then assembled with Velvet version

Citation Xie G, Cheng Q, Daligault H, Davenport K, Gleasner C, Jacobs L, Kubicek-Sutherland J, LeCuyer T, Otieno V, Raballah E, Mukundan H, Perkins DJ, McMahon B, Doggett N. 2019. Genome sequences of a *Staphylococcus aureus* clinical isolate, strain SMA0034-04 (UGA22), from Siaya County Referral Hospital in Siaya, Kenya. *Microbiol Resour Announc* 8:e01627-18. <https://doi.org/10.1128/MRA.01627-18>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2019 Xie et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Norman Doggett, doggett@lanl.gov.

Received 21 December 2018

Accepted 27 February 2019

Published 25 April 2019

1.2.08 (5, 6) and IDBA version 1.1.0 (7). The consensus sequences were computationally shredded and reassembled with Phrap version SPS-4.24 (8, 9) to allow some manual editing with Consed (10), resulting in a final 80 contigs (81.43% of the reads) of >200 bp, with an N_{50} value of 621,451 bp, for SMA0034-04. Among these, contig 71 (pUGA22) encodes the intact *rep24*-type plasmid of an *S. aureus* pWBG745 homolog. The draft genome sequence of SMA0034-04 consists of 2,608,371 bp, with an average G+C content of 32.9%. Annotations were completed at Los Alamos National Laboratory (LANL) using an automated system with the Ergatis workflow manager (11) and in-house scripts.

There are 2,512 predicted protein-coding genes, 26 tRNA genes, and 12 rRNA genes within the genome of SMA0034-04. Of these, 37% of the protein-coding genes were annotated in a SEED subsystem (12), whereas 63% were not associated with a SEED subsystem. A total of 740 genes were annotated as hypothetical proteins. Of all the predicted genes, 2,295 are shared between SMA0034-04 and the *S. aureus* subsp. *aureus* MW2 genome, with 217 and 273 genes being unique to SMA0034-04 and MW2, respectively. The plasmid pUGA22 shares 99.93% identity and 100% query coverage with its counterpart, pWBG745. The main differences are 4 hypothetical protein genes that are deleted in pUGA22 and that antibacterial resistance genes are absent in pUGA22. The estimated antibacterial resistance profile (13) is similar between SMA0034-04 and MW2, except for methicillin resistance genes in the MW2 chromosome. A genome survey of toxin and virulence factor genes revealed the presence of protein A and superantigen-like exotoxins (SSLs) on SMA0034-04 contig 78; the presence of exfoliative toxin and staphylococcal alpha and delta toxins on contig 80; and the presence of a gamma-hemolysin hlgACB cluster, Pantone-Valentine leukocidin (PVL) cytotoxin, and extracellular metalloprotease aureolysin (AUR) on contig 79. The enterotoxin type B gene (*entB*) is truncated by a missense mutation on SMA0034-04 contig 79, and the beta-hemolysin (beta toxin) gene is interrupted by a phage element on contigs 74 and 80. This phage integrated into beta-hemolysin has been recorded in many *S. aureus* strains of human origin (14).

Data availability. The GenBank accession number for the *Staphylococcus aureus* SMA0034-04 (UGA22) genome sequences is [NWTZ00000000](https://ncbi.nlm.nih.gov/nucl/NWTZ00000000). The raw reads have been deposited in the NCBI Sequence Read Archive under accession number [SRR8648187](https://ncbi.nlm.nih.gov/sra/SRR8648187).

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Research grants R01AI51305, R01AI130473-01A1, and D43TW05884 (to D.J.P.) and by Los Alamos National Laboratory LDRD grant 20150090DR (to B.M.).

LANL is operated by Los Alamos National Security, LLC, for the Department of Energy, under contract DE-AC52-06NA25396.

REFERENCES

- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 359:1819–1827. [https://doi.org/10.1016/S0140-6736\(02\)08713-5](https://doi.org/10.1016/S0140-6736(02)08713-5).
- O'Brien FG, Ramsay JP, Monecke S, Coombs GW, Robinson OJ, Htet Z, Alshaiikh FAM, Grubb WB. 2015. *Staphylococcus aureus* plasmids without mobilization genes are mobilized by a novel conjugative plasmid from community isolates. *J Antimicrob Chemother* 70:649–652. <https://doi.org/10.1093/jac/dku454>.
- Lo C-C, Chain PSG. 2014. Rapid evaluation and quality control of next generation sequencing data with FaQCs. *BMC Bioinformatics* 15:366. <https://doi.org/10.1186/s12859-014-0366-2>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Zerbino DR, McEwen GK, Margulies EH, Birney E. 2009. Pebble and Rock Band: heuristic resolution of repeats and scaffolding in the Velvet short-read *de novo* assembler. *PLoS One* 4:e8407. <https://doi.org/10.1371/journal.pone.0008407>.
- Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
- Ewing B, Green P. 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* 8:186–194. <https://doi.org/10.1101/gr.8.3.186>.
- Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* 8:175–185. <https://doi.org/10.1101/gr.8.3.175>.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for

- sequence finishing. *Genome Res* 8:195–202. <https://doi.org/10.1101/gr.8.3.195>.
11. Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. 2010. An Ergatis-based prokaryotic genome annotation Web server. *Bioinformatics* 26:1122–1124. <https://doi.org/10.1093/bioinformatics/btq090>.
 12. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
 13. Clinical and Laboratory Standards Institute. 2018. M100 performance standards for antimicrobial susceptibility testing, 28th ed. Clinical and Laboratory Standards Institute, Wayne, PA.
 14. Goerke C, Koller J, Wolz C. 2006. Ciprofloxacin and trimethoprim cause phage induction and virulence modulation in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 50:171–177. <https://doi.org/10.1128/AAC.50.1.171-177.2006>.