AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of 27 Hospital-Associated Methicillin-Resistant *Staphylococcus aureus* Strains Isolated in Minnesota

Saeed A. Khan,[a] Dereje D. Gudeta,[a] Nesreen Aljahdali,[a,d] Jungwhan Chon,[b] Paula Snippes Vagnone,[c] Mohamed S. Nawaz,[a] Steven L. Foley,[a] Kidon Sung[a]

[a]Division of Microbiology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas, USA
[b]Department of Pet Total Care, Division of Nursing and Welfare, Kyung-in Women's University, Incheon, South Korea
[c]Public Health Laboratory Division, Minnesota Department of Health, St. Paul, Minnesota, USA
[d]Biological Science Department, College of Science, King Abdul-Aziz University, Jeddah, Saudi Arabia

ABSTRACT Infections caused by hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) strains have higher morbidity and mortality rates and require longer hospital stays than do those caused by hospital-associated methicillin-sensitive *Staphylococcus aureus* strains. To gain insight into their genomic makeup, antimicrobial resistance, biofilm formation, and virulence potentials, here we present the draft whole-genome sequences of 27 HA-MRSA strains isolated in Minnesota.

Hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) accounts for the majority of invasive MRSA infections and is usually resistant to non-β-lactam as well as β-lactam antibiotics (1, 2). Most HA-MRSA strains harbor large staphylococcal cassette chromosome *mec* (SCC*mec*) elements I, II, or III, but community-associated MRSA strains harbor small SCC*mec* elements IV or V (3). Among 27 HA-MRSA isolates, 18 of the 27 (~66.67%) were SCC*mec* type II and 9 (~33.33%) were SCC*mec* type IV.

The *S. aureus* isolates (1 from pleural fluid, 1 from hip fluid, 1 from bone, 1 from unnamed body fluid, and 23 from blood) in this study were collected in 2016 (from multiple clinical sites in Hennepin and Ramsey counties in Minnesota) as part of the normal surveillance by the Minnesota Department of Health (MDH) and were approved for use in research by the institutional review board of the National Center for Toxicological Research. The isolates were obtained as cryostocks from the MDH, streaked on standard blood agar plates, and incubated at 37°C to obtain single colonies, which were inoculated and grown overnight at 37°C in tryptic soy broth. Genomic DNA was isolated using the Qiagen QIAamp minikit, supplementing the AL buffer with lysostaphin. The DNA concentration and quality were determined using a NanoDrop 2000 spectrophotometer and a Qubit Flex fluorometer (Thermo Fisher Scientific). Sequencing libraries were made using a Nextera XT DNA library preparation kit (Illumina, San Diego, CA) and multiplexed using a Nextera XT index kit (Illumina). Whole-genome sequencing (WGS) was performed on a MiSeq platform (Illumina) using a 2 × 300-bp paired-end run (4).

Removal of the adapter sequences, filtering of the low-quality sequence reads, quality control, and *de novo* assembly of high-quality sequences were performed using the Trim Reads, QC for Sequencing Reads, and De Novo Assembly tools of CLC Genomics Workbench v21.0.4 (Qiagen) (5, 6). QUAST v5.1 (http://quast.sourceforge.net) was employed to check the assembly quality and the completeness of the draft genomes and to compare contig and scaffold statistics (7). The draft genome assembly was first annotated using the Rapid Annotations using Subsystem Technology (RAST) tool kit (RASTtk) within Pathosystems Resource Integration Centre (PATRIC) v3.6.9 (https://www.patricbrc.org) (8). The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.2 was then employed to annotate the publicly available genome sequences (9). Multilocus sequence typing (MLST), *spa* typing, and SCC*mec* typing

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

Address correspondence to Kidon Sung, kidon.sung@fda.hhs.gov.

The authors declare no conflict of interest.

Received 7 December 2021
Accepted 9 January 2022
Published 27 January 2022

TABLE 1 WGS analyses of 27 HA-MRSA isolates

Isolate	MLST sequence type[a]	*spa* type	SCC*mec* type	Assembly length (bp)	Total no. of reads	Genome coverage (×)	N_{50} (bp)	No. of contigs[b]	No. of CDSs[b]	G+C content (%)[b]	No. of tRNAs[b]	No. of rRNAs[b]	No. of antibiotic resistance genes[c]	No. of virulence genes[d]	GenBank accession no.	SRA accession no.
HAR1	231	t002	II(2A)	2,815,548	3,450,998	253	51,999	124	2,708	32.75	57	4	40	68	JAHNFW000000000	SRX11183815
HAR2	5	t002	II(2A)	2,876,411	1,207,112	258	84,462	109	2,839	32.72	50	4	41	68	JAHNFX000000000	SRX11183816
HAR3	8	t008	IVa(2B)	2,905,856	1,402,652	261	71,659	125	2,856	32.63	55	3	42	74	JAHNFY000000000	SRX11183827
HAR4	105	t045	II(2A)	2,878,388	1,515,458	258	168,260	79	2,798	32.73	50	2	41	68	JAHNFZ000000000	SRX11183835
HAR5	105	t586	II(2A)	2,855,492	1,215,200	256	169,609	60	2,756	32.71	57	4	44	64	JAHNGA000000000	SRX11183836
HAR6	8	NA	IVa(2B)	2,838,926	1,489,186	255	46,508	173	2,751	32.58	53	4	44	74	JAHNGB000000000	SRX11183837
HAR7	231	NA	II(2A)	2,719,526	1,630,512	243	35,658	199	2,624	32.73	52	4	42	68	JAHNGC000000000	SRX11183838
HAR8	NA	t002	II(2A)	2,803,227	3,196,628	251	5,926	912	3,073	32.91	50	4	49	77	JAHNGD000000000	SRX11183839
HAR9	231	t002	II(2A)	2,844,465	2,315,796	255	144,785	76	2,747	32.70	57	4	42	64	JAHNGE000000000	SRX11183840
HAR10	NA	t1154	II(2A)	2,834,834	2,011,880	254	108,757	75	2,729	32.72	65	7	41	64	JAHNGF000000000	SRX11183841
HAR11	8	NA	IVa(2B)	2,909,277	2,190,598	261	120,590	98	2,843	32.66	57	5	46	73	JAHNGG000000000	SRX11183817
HAR12	2253	t008	IVg(2B)	2,849,975	2,224,754	256	108,887	80	2,755	32.64	53	4	40	69	JAHNGH000000000	SRX11183818
HAR13	8	t008	IVa(2B)	2,852,089	2,302,916	256	144,828	64	2,731	32.63	57	4	43	67	JAHNGI000000000	SRX11183819
HAR14	105	t002	II(2A)	2,870,713	4,028,550	258	102,669	103	2,799	32.77	57	4	43	67	JAHNGJ000000000	SRX11183820
HAR15	8	t008	IVa(2B)	2,909,441	2,292,566	261	101,907	84	2,835	32.65	57	5	42	76	JAHNGK000000000	SRX11183821
HAR16	5	t002	IVa(2B)	2,731,741	4,383,744	245	62,045	119	2,610	32.70	54	4	43	67	JAHNGL000000000	SRX11183822
HAR17	5	t002	IVc(2B)	2,734,797	5,360,542	245	139,237	61	2,611	32.68	57	5	40	67	JAHNGM000000000	SRX11183823
HAR18	105	NA	II(2A)	2,864,397	1,154,926	257	50,895	136	2,797	32.76	61	4	43	66	JAHNGN000000000	SRX11183824
HAR19	105	t002	II(2A)	2,842,311	2,170,636	255	199,380	52	2,735	32.75	57	4	45	63	JAHNGO000000000	SRX11183825
HAR20	5	t002	IVc(2B)	2,763,279	1,938,616	248	101,535	61	2,651	32.74	56	3	39	62	JAHNGP000000000	SRX11183826
HAR21	5	t003	II(2A)	2,803,767	1,397,856	251	50,556	99	2,693	32.75	48	4	40	65	JAHNGQ000000000	SRX11183828
HAR22	496	NA	II(2A)	2,793,012	1,657,336	250	68,218	84	2,676	32.70	51	4	41	69	JAHNGR000000000	SRX11183829
HAR23	NA	t458	II(2A)	2,832,881	3,645,406	254	68,354	76	2,733	32.71	49	4	41	61	JAHNGS000000000	SRX11183830
HAR24	5	t003	II(2A)	2,806,266	1,473,668	252	68,160	94	2,696	32.77	51	4	40	67	JAHNGT000000000	SRX11183831
HAR25	105	t002	II(2A)	2,801,912	426,454	251	8,649	680	2,937	32.87	42	4	48	81	JAHNGU000000000	SRX11183832
HAR26	231	t586	II(2A)	2,846,528	1,092,370	255	32,168	230	2,794	32.73	55	4	43	68	JAHNGV000000000	SRX11183833
HAR27	105	t002	II(2A)	2,845,911	1,744,514	255	38,233	211	2,806	32.76	49	4	45	73	JAHNGW000000000	SRX11183834

[a] NA, not available.

[b] The numbers of contigs, CDSs, tRNAs, and rRNAs and the G+C content were determined by PATRIC.

[c] Searched by PATRIC.

[d] Searched in the Virulence Factor Database (VFDB).

were conducted with MLST v2.0, spaTyper v1.0, and SCC*mec*Finder v1.2, respectively (10, 11). Default parameters were used for all software unless otherwise specified. The MLST, *spa*, and SCC*mec* types, assembly length, total numbers of reads, contigs, coding sequences (CDSs), tRNA genes, rRNA genes, antibiotic resistance genes, and virulence genes, genome coverage, N_{50} values, G+C contents, and GenBank and Sequence Read Archive (SRA) accession numbers are listed in Table 1. The genome coverages of the draft whole-genome sequences ranged between 244× and 261×, with the genome sizes ranging from 2,719,526 bp to 2,909,441 bp; the overall average number of CDSs and G+C content were 2,762 and 32.72%, respectively. The numbers of contigs ranged from 52 to 912, while the N_{50} values ranged from 5,926 bp to 199,380 bp.

Data availability. The draft genome sequences described in this study were deposited in DDBJ/ENA/GenBank with the accession numbers listed in Table 1. The raw sequence reads were deposited in the SRA under BioProject accession number PRJNA737146.

ACKNOWLEDGMENTS

We recognize Jing Han and Jinshan Jin for their critical review of the manuscript.

This project was supported by the National Center for Toxicological Research and the U.S. Food and Drug Administration (grant E0770101).

The manuscript reflects the views of the authors and does not necessarily reflect those of the U.S. Food and Drug Administration. Any mention of commercial products is for clarification only and is not intended as approval, endorsement, or recommendation.

REFERENCES

1. Stefani S, Chung DR, Lindsay JA, Friedrich AW, Kearns AM, Westh H, Mackenzie FM. 2012. Meticillin-resistant *Staphylococcus aureus* (MRSA): global epidemiology and harmonisation of typing methods. Int J Antimicrob Agents 39:273–282. https://doi.org/10.1016/j.ijantimicag.2011.09.030.
2. Henderson A, Nimmo GR. 2018. Control of healthcare- and community-associated MRSA: recent progress and persisting challenges. Br Med Bull 125:25–41. https://doi.org/10.1093/bmb/ldx046.
3. Peng H, Liu D, Ma Y, Gao W. 2018. Comparison of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus* isolates at a Chinese tertiary hospital, 2012–2017. Sci Rep 8:17916. https://doi.org/10.1038/s41598-018-36206-5.
4. Sung K, Khajanchi BK, Hiett KL, Line JE, Khan SA. 2020. Draft genome sequences of two *Campylobacter jejuni* strains that show significantly different colonization potentials in chickens. Microbiol Resour Announc 9:e00687-20. https://doi.org/10.1128/MRA.00687-20.
5. Kolle M, Horta MAC, Nowrousian M, Ohm RA, Benz JP, Pilgard A. 2020. Degradative capacity of two strains of *Rhodonia placenta*: from phenotype to genotype. Front Microbiol 11:1338. https://doi.org/10.3389/fmicb.2020.01338.
6. Utturkar SM, Klingeman DM, Land ML, Schadt CW, Doktycz MJ, Pelletier DA, Brown SD. 2014. Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences. Bioinformatics 30:2709–2716. https://doi.org/10.1093/bioinformatics/btu391.
7. 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.
8. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJC, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res 42:D581–D591. https://doi.org/10.1093/nar/gkt1099.
9. 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.
10. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol 50:1355–1361. https://doi.org/10.1128/JCM.06094-11.
11. Bartels MD, Petersen A, Worning P, Nielsen JB, Larner-Svensson H, Johansen HK, Andersen LP, Jarløv JO, Boye K, Larsen AR, Westh H. 2014. Comparing whole-genome sequencing with Sanger sequencing for *spa* typing of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 52:4305–4308. https://doi.org/10.1128/JCM.01979-14.