Draft Genome Sequences of Five *Staphylococcus saprophyticus* Strains Isolated from African Fermented Nono in Nigeria

Etinosa O. Igbinosa,[a,b] Erik Brinks,[b] Abeni Beshiru,[a] Isoken H. Igbinosa,[a] Maria Stein,[b] Gyu-Sung Cho,[b] Charles M. A. P. Franz[b]

[a]Applied Microbial Processes and Environmental Health Research Group, Department of Microbiology, Faculty of Life Sciences, University of Benin, Benin City, Nigeria
[b]Max Rubner Institute, Federal Research Institute of Nutrition and Food, Department of Microbiology and Biotechnology, Kiel, Germany

ABSTRACT Five *Staphylococcus saprophyticus* strains were isolated from the fermented milk product nono in Nigeria and were sequenced using an Illumina MiSeq platform. The genome sizes ranged from 2.53 to 2.60 Mbp, while the GC contents ranged from 32.99 to 33.07 mol%. The genomes possessed 2,505 to 2,687 protein-coding sequences.

Coagulase-negative staphylococci (CoNS) are found on the skin and membranes of mammals and birds. These strains are also commonly isolated from different natural niches, such as soil, water, and foods, including meat, cheese, and vegetables (1, 2). Catalase-positive CoNS strains have also been reported to be associated with the raw milk microbiota (3), as have pathogenic coagulase-positive *Staphylococcus aureus* strains. In this study, the draft genome sequences of five *Staphylococcus* strains isolated from nono (fermented milk product) samples from the Aduwawa market in Benin City, Nigeria, were determined in order to identify them in an investigation of the microbiological quality and safety of nono. Ten-milliliter nono samples were collected into 90 ml of tryptic soy broth (Oxoid, UK) and incubated at 35°C for 18 to 24 h, and a loopful of culture was streaked onto mannitol salt agar (Neogen Culture Media, Heywood, Lancashire, UK). Five randomly selected colonies were purified by streaking on brain heart infusion (BHI) agar (Roth, Karlsruhe, Germany) at 37°C for 24 h of incubation, and strains (Table 1) were identified as *Staphylococcus saprophyticus* by 16S rRNA gene sequencing.

The total genomic DNA of each strain, cultured in BHI broth at 37°C for 18 h, was extracted using a bacterial DNA kit (Peqlab, Erlangen, Germany) according to the manufacturer's instructions. Genomic DNA libraries were prepared with a Nextera XT library preparation kit (Illumina, USA), and paired-end sequencing was performed on a MiSeq sequencer (Illumina) with 2×300 cycles. The total numbers of paired reads ranged from 553,037 to 783,645 (Table 1), and all paired-end raw data were trimmed using the Trimmomatic (v.0.32) (4) pipeline with the following parameters: sliding window; 4:15, leading; 3, and minlen; 45. The *de novo* assemblies were performed with SPAdes (v.3.15.0) with the following parameters: k-mer 77, careful, and minimum contig length of 500 bp (5); the quality of the draft genome sequences obtained was evaluated using the QUAST (v.5.0.2) tool (6). A precise species identification was performed with the OrthoANI (v.0.93.1) (7) pipeline with closely related staphylococcal type strains as references. Acquired antibiotic resistance genes and plasmid-related sequences were identified using ResFinder (v.4.1) (8) and PlasmidFinder (v.2.1) (9), respectively. Except for the SPAdes and Trimmomatic pipelines, default parameters were used for all programs. The N_{50} values were between 27,155 and 431,885 bp, and the genome coverage ranged from 65- to 93-fold (Table 1). All contigs were annotated using the PATRIC server and the NCBI Prokaryotic Genome Annotation Pipeline (v.4.13) with default parameters (10–12). The five draft genome sizes range from 2.53 Mbp to 2.60 Mbp, with GC contents between 32.99 and 33.07 mol% (Table 1). Three of the five strains contained

TABLE 1 De novo assembly of strains and draft genome features

Strain[a]	No. of contigs	N_{50} (bp)	GC content (mol%)	Total length (bp)	Genome coverage (X)	Total no. of paired-end sequence reads	No. of CDSs	No. of tRNAs	No. of rRNAs	Acquired resistance gene(s)	Plasmid database(s)	GenBank accession no.	SRA accession no.
ST386	32	431,815	33.07	2,536,544	65	550,037	2,508	46	5	cat(pC233), dfrg, tetK	Rep_trans, RepA_N	JAFJNU000000000	SRR13674887
ST388	33	218,917	33.00	2,578,047	69	587,237	2,542	48	4	dfrg	RepA_N	JAFJNV000000000	SRR13674888
ST390	186	27,155	32.99	2,609,417	85	788,480	2,687	42	8	ND	RepA_N, Inc18	JAFJNW000000000	SRR13674889
ST391	30	431,885	33.07	2,539,289	93	783,645	2,505	45	4	cat(pC233), dfrg, tetK	RepA_N, Rep_trans	JAFJNX000000000	SRR13674890
ST392b	39	373,477	33.04	2,569,600	76	653,274	2,540	47	4	cat(pC233), dfrg, tetK	RepA_N, Rep_trans	JAFJNY000000000	SRR13674891

[a] All strains were precisely identified as S. saprophyticus subsp. saprophyticus using the OrthoANI (7) pipeline. CDS, coding sequence; cat, chloramphenicol acetyltransferase; dfrg, trimethoprim-resistant dihydrofolate reductase; tet, tetracycline resistance gene; Rep, replication protein; ND, not detected.

genes encoding chloramphenicol (*cat*) and tetracycline (*tetK*) resistance, while replication proteins (Rep) in all five strains indicated the presence of plasmids (Table 1). The details of genome sequencing and annotation results are shown in Table 1.

Data availability. The draft genome sequences and the raw read data were deposited under the BioProject accession number PRJNA700496, and the accession numbers are listed in Table 1.

REFERENCES

1. Jeong D-W, Lee B, Her J-Y, Lee K-G, Lee J-H. 2016. Safety and technological characterization of coagulase-negative staphylococci isolates from traditional Korean fermented soybean foods for starter development. Int J Food Microbiol 236:9–16. https://doi.org/10.1016/j.ijfoodmicro.2016.07.011.
2. Khusro A, Aarti C, Salem AZM, Barbabosa-Pilego A. 2020. Techno-functional traits and safety aspects of coagulase-negative *Staphylococcus saprophyticus* isolated from traditional fermented food. Food Biotechnol 34:77–99. https://doi.org/10.1080/08905436.2019.1711114.
3. Freitas Ribeiro L, Akira Sato R, de Souza Pollo A, Marques Rossi GA, do Amaral LA. 2020. Occurrence of methicillin-resistant *Staphylococcus* spp. on Brazilian dairy farms that produce unpasteurized cheese. Toxins (Basel) 12:779. https://doi.org/10.3390/toxins12120779.
4. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
5. 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.
6. 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.
7. Lee I, Ouk Kim Y, Park SC, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103. https://doi.org/10.1099/ijsem.0.000760.
8. Kleinheinz KA, Joensen KG, Larsen MV. 2014. Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. Bacteriophage 4:e27943. https://doi.org/10.4161/bact.27943.
9. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L, Moller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.
10. Gillespie JJ, Wattam AR, Cammer SA, Gabbard JL, Shukla MP, Dalay O, Driscoll T, Hix D, Mane SP, Mao C, Nordberg EK, Scott M, Schulman JR, Snyder EE, Sullivan DE, Wang C, Warren A, Williams KP, Xue T, Yoo HS, Zhang C, Zhang Y, Will R, Kenyon RW, Sobral BW. 2011. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. Infect Immun 79:4286–4298. https://doi.org/10.1128/IAI.00207-11.
11. Wattam AR, Gabbard JL, Shukla M, Sobral BW. 2014. Comparative genomic analysis at the PATRIC, a bioinformatic resource center. Methods Mol Biol 1197:287–308. https://doi.org/10.1007/978-1-4939-1261-2_17.
12. 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.