AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Streptococcus agalactiae* KALRO-LC1 Strain Isolated from a Mastitis-Infected Camel in Laikipia County, Kenya

Edwin Murungi,[a] Ednah Masila,[b] Irene Ogali,[b] Nathan Langat,[b] Ruth Onywera,[b] Vincent Malonza,[b] Christine Inguyesi,[b] Frank Onyambu,[c] Hezron Wesonga,[b] Monicah Maichomo[b]

[a]Department of Medical Biochemistry, Kisii University, Kisii, Kenya
[b]Veterinary Research Institute, Kenya Agricultural and Livestock Research Organization, Muguga, Kenya
[c]School of Health Sciences, Meru University of Science and Technology, Meru, Kenya

ABSTRACT We report the draft genome sequence of *Streptococcus agalactiae* KALRO-LC1 strain obtained from a mastitis-infected camel in Laikipia County, Kenya. The 2,201,604-bp draft genome is assembled into 3 contigs with a GC content of 35.87% and is predicted to contain 1,192 protein-coding sequences.

Streptococcus agalactiae (Group B *Streptococcus*) causes mastitis in camels resulting in diminished milk production and consequently enormous economic losses that imperil the livelihoods of pastoralists in northern Kenya (1). Moreover, *S. agalactiae* causes serious infections in humans (2, 3). The pathogen is therefore a threat to human and veterinary health.

We report the draft genome of a *S. agalactiae* KALRO-LC1 isolate recovered from a mastitis-infected camel in Laikipia County, Kenya. Milk samples were inoculated on blood agar containing 5% sheep blood and incubated overnight at 37°C. Colonies with zones of beta-hemolysis, and that were catalase negative and Gram positive, were cultured in Edward's selective media to ascertain genus type (4). Lancefield grouping confirmed the isolate as *Streptococcus agalactiae* (5). The isolated bacteria were cultured on nutrient agar plates by incubation under aerobic conditions at 37°C for 16 h followed by enrichment of fresh single colonies in 100 mL nutrient broth and incubation under aerobic conditions at 37°C for 16 h with rotation at 200 rpm (6) to a cell density of 0.5 to 0.7 at OD_{600} (7). Genomic DNA was extracted using the QIAamp DNA kit (Qiagen) according to manufacturer's instructions. An Oxford Nanopore Technologies (ONT) sequencing library was then prepared using the manufacturer's Native Barcoding Expansion 1-12 (EXP-NBD104) in conjunction with the Ligation Sequencing kit (SQK-LSK109), and sequencing was performed on a MinION device using R9.4.1 flow cell to give 33,183 raw reads of N_{50} 103 kb. Basecalling of the raw sequence data (fast5) and subsequent demultiplexing was performed using GUPPY v3.6.1 (https://nanoporetech.com). Adaptors were trimmed using PORECHOP v0.2.4 (8), and NANOFILT v2.5.0 (9) was used for quality filtering by removing reads $<$500 bp or with average quality scores of $<$10. The draft genome was assembled using UNICYCLER v0.5.0 (10) and quality assessed using QUAST v5.2.0 (11). The contigs were annotated using NCBI's PGAP v6.1 (12). Default parameters were used for all software unless otherwise specified.

The assembly contained 3 contigs totaling 2,201,604 bp with a GC content of 35.87%, an N_{50} of 2,139,486 bp, and a coverage of 95×. Annotation revealed a total of 2,269 genes (1,192 coding genes), 23 rRNAs, 85 tRNAs, 3 ncRNAs, and 2 CRISPR arrays.

Using VaxiJen (13), IgPred (14), and AllerTOP v.2 (15), we have delineated several potentially viable antigenic vaccine candidates (Table 1). Our results will foster further investigations into the development of a subunit vaccine for *S. agalactiae* infections.

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

Address correspondence to Edwin Murungi, edwinmurungi@kisiiuniversity.ac.ke.

The authors declare no conflict of interest.

Received 1 September 2022
Accepted 3 September 2022
Published 19 September 2022

TABLE 1 Putative *S. agalactiae* antigenic vaccine candidates

NCBI sequence accession	VaxiJen score	Subcellular localization	Transmembrane helices	IgPred prediction	AllerTOP prediction
MCP9189840.1	1.0828	Cell wall	1	IgG Epitope	Probable nonallergen
MCP9189846.1	0.9236	Cytoplasmic membrane	1	IgG Epitope	Probable nonallergen
MCP9189915.1	0.7253	Unknown	0	IgG Epitope	Probable nonallergen
MCP9189959.1	0.7667	Cytoplasmic	0	IgG Epitope	Probable nonallergen
MCP9189976.1	0.7542	Unknown	0	IgG Epitope	Probable nonallergen
MCP9189841.1	1.5040	Cytoplasmic	0	IgG Epitope	Probable nonallergen
MCP9189875.1	1.1086	Unknown	0	IgG Epitope	Probable nonallergen
MCP9190210.1	1.0150	Unknown	0	IgG Epitope	Probable nonallergen
MCP9190214.1	1.0590	Unknown	1	IgG Epitope	Probable nonallergen
MCP9190551.1	1.0850	Cytoplasmic membrane	0	IgG Epitope	Probable nonallergen
MCP9190938.1	1.0042	Cytoplasmic membrane	1	IgG Epitope	Probable nonallergen
MCP9190954.1	1.0033	Unknown	0	IgG Epitope	Probable nonallergen
MCP9190991.1	1.1004	Extracellular	0	IgG Epitope	Probable nonallergen
MCP9189842	0.9959	Cell wall	0	IgG Epitope	Probable nonallergen
MCP9190131.1	0.9931	Unknown	0	IgG Epitope	Probable nonallergen
MCP9190265.1	0.9462	Cytoplasmic	0	IgG Epitope	Probable nonallergen
MCP9190286.1	0.9441	Cytoplasmic membrane	4	IgG Epitope	Probable nonallergen
MCP9190290.1	0.9461	Cytoplasmic membrane	1	IgG Epitope	Probable nonallergen

Data availability. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under accession JANCLS000000000. Raw sequence reads are available at the Sequence Read Archive (SRA) database under accession SRX16447265.

ACKNOWLEDGMENTS

We acknowledge the director-general of KALRO for support and EU-AgriFI Climate Smart Agricultural Productivity Project for funding.

REFERENCES

1. Ochieng J, Gicheru N, Indieka AS, Maichomo M, Wesonga OH, Murungi E. 2021. Draft genome sequence of *Enterococcus faecalis* 1351, isolated from a mastitis-affected camel in Isiolo County, Kenya. Microbiol Resour Announc 10:e00143-21. https://doi.org/10.1128/MRA.00143-21.
2. Abdelradi A, Murphy A, Ahasic AM. 2020. Invasive *Streptococcus agalactiae* causing meningitis, ventriculitis and endocarditis in a non-pregnant adult. Cureus 12:e11412.
3. Furfaro LL, Chang BJ, Payne MS. 2018. Perinatal *Streptococcus agalactiae* epidemiology and surveillance targets. Clin Microbiol Rev 31:e00049-18. https://doi.org/10.1128/CMR.00049-18.
4. Seligsohn D, Nyman A-K, Younan M, Sake W, Persson Y, Bornstein S, Maichomo M, de Verdier K, Morrell JM, Chenais E. 2020. Subclinical mastitis in pastoralist dairy camel herds in Isiolo, Kenya: prevalence, risk factors, and antimicrobial susceptibility. J Dairy Sci 103:4717–4731. https://doi.org/10.3168/jds.2019-17701.
5. Zorzoli A, Meyer BH, Adair E, Torgov VI, Veselovsky VV, Danilov LL, Uhrin D, Dorfmueller HC. 2019. Group A, B, C, and G *Streptococcus* Lancefield antigen biosynthesis is initiated by a conserved α-d-GlcNAc-β-1,4-l-rhamnosyltransferase. J Biol Chem 294:15237–15256. https://doi.org/10.1074/jbc.RA119.009894.
6. Hoorzook KB, Barnard TG. 2022. Culture independent DNA extraction method for bacterial cells concentrated from water. MethodsX 9:101653. https://doi.org/10.1016/j.mex.2022.101653.
7. Wright MH, Adelskov J, Greene AC. 2017. Bacterial DNA extraction using individual enzymes and phenol/chloroform separation. J Microbiol Biol Educ 18: 18:2–48.
8. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3:e000132. https://doi.org/10.1099/mgen.0.000132.
9. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.
10. 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.
11. 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.
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.
13. Doytchinova IA, Flower DR. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8:4. https://doi.org/10.1186/1471-2105-8-4.
14. Gupta S, Ansari HR, Gautam A, Raghava GPS, Open Source Drug Discovery Consortium. 2013. Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. Biol Direct 8:27. https://doi.org/10.1186/1745-6150-8-27.
15. Dimitrov I, Bangov I, Flower DR, Doytchinova I. 2014. AllerTOP v.2—a server for in silico prediction of allergens. J Mol Model 20:2278. https://doi.org/10.1007/s00894-014-2278-5.