




Complete Genome Sequence of *Streptococcus agalactiae* Strain 01173, Isolated from Kuwaiti Wild Fish

Morena Santi,^a Sharon A. Egan,^a James A. Leigh,^a David W. Verner-Jeffreys,^b  Adam M. Blanchard^a

^aSchool of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington, Leicestershire, United Kingdom

^bCefas Weymouth Laboratory, Centre for Environment, Fisheries, and Aquaculture Science, Weymouth, Dorset, United Kingdom

ABSTRACT Here, we report the complete genome of piscine *Streptococcus agalactiae* 01173 serotype Ia, which was generated using long-read sequencing technology. The bacteria were isolated from wild fish displaying signs of streptococcosis, from a fish kill incident in Kuwait.

Streptococcus agalactiae, a group B streptococcus, is an opportunistic Gram-positive bacterium that is chiefly associated with meningitis and septicemia in infants in the Western world (1) and is increasingly associated with adult infections, especially among the elderly or immunocompromised (2). *S. agalactiae* is a major pathogen of fish that manifests as meningoencephalitis, with high mortality rates being reported for tilapia (*Oreochromis* sp.), and it represents a significant threat to the aquaculture industry (3, 4). Here, we report the whole-genome sequence of *S. agalactiae* strain 01173 serotype Ia, multilocus sequence type 7 (ST7), isolated from infected tissue from a wild fish in Kuwait using sterile swabs and cultured on tryptone soya agar; species confirmation was performed by 16S rRNA sequencing (5). Bacteria were cultured overnight on brain heart infusion (BHI) agar (Oxoid) at 37°C. Single colonies were inoculated into BHI broth, and genomic DNA was extracted as described previously (6). Genomic DNA was sequenced using Illumina MiSeq and Oxford Nanopore Technologies (ONT) GridION X5 Mk1 sequencing (BioProject number [PRJNA627590](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA627590)). ONT libraries were prepared using the SQK-RBK004 rapid barcoding sequencing kit (ONT) and then loaded onto a MinION FLO-MIN106 R9.4.1 flow cell (ONT). Illumina sequence libraries were created using the Illumina TruSeq library preparation kit, following the manufacturer's guidelines, and were sequenced as 250-bp paired-end reads.

Raw data for both Illumina and Nanopore sequences were submitted for rigorous quality filtering. Default parameters were used throughout except where otherwise stated. Illumina data were parsed with Trimmomatic (v0.39) (7) to remove any adaptor contamination using a sliding-window approach to maintain an average Phred score of 30 across the read. Unpaired reads were discarded. Nanopore data were base called using Guppy (v3.2.8) with the high-accuracy model and were checked for adaptors using Porechop (v0.2.4) (<https://github.com/rrwick/Porechop>) with the addition of the discard_middle flag. The data were then parsed with Filtlong (v0.2.0) (<https://github.com/rrwick/Filtlong>) using the Illumina data as a reference and the flags min_length 1000 and keep_percent 90. This resulted in 627,824 Illumina and 6,196 Nanopore high-quality reads (N_{50} , 17.1 kb). Trimmed and filtered reads were used as inputs for Unicycler (v0.4.8) (8) in the normal hybrid assembly mode, circularized, and rotated to *dnaA*. The resulting assembly was polished once using Nanopolish (v0.13.2) (<https://github.com/jts/nanopolish>). The assembled completed genome of *S. agalactiae* strain 01173 (GenBank accession number [CP053027.1](https://www.ncbi.nlm.nih.gov/nuccore/CP053027.1)) was characterized by a single circular chromosome of 2,105,299 bp with a GC content of 35.7%. The assemblies were then annotated using PGAP (v4.11) (9), resulting in an 87.6% coding ratio with 2,024

Citation Santi M, Egan SA, Leigh JA, Verner-Jeffreys DW, Blanchard AM. 2020. Complete genome sequence of *Streptococcus agalactiae* strain 01173, isolated from Kuwaiti wild fish. Microbiol Resour Announc 9:e00674-20. <https://doi.org/10.1128/MRA.00674-20>.

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

Copyright © 2020 Santi 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 Adam M. Blanchard, adam.blanchard@nottingham.ac.uk.

Received 10 June 2020

Accepted 14 August 2020

Published 3 September 2020

protein-coding sequences, 21 rRNAs, 80 tRNAs, and 1 putative clustered regularly interspaced short palindromic repeat (CRISPR) sequence. *S. agalactiae* 01173 contained most virulence factors previously associated with the species, including sortase A (10), C5a peptidase, CAMP factor, complete capsular polysaccharide operon, and serine peptidase (11), and 7 copies of group II intron reverse transcriptase/maturase genes, which are frequently located near potential virulence genes (12). The *S. agalactiae* 01173 whole genome was compared to all publicly available, complete genomes for *Streptococcus agalactiae* (downloaded from the NCBI databases [accessed 24 March 2020]), and individual genomic sequence signatures were created using sourmash (v3.3.0) (13). The resulting hash tables were imported into R to generate heatmaps using a Jaccard similarity index cutoff value of 85%. *S. agalactiae* 01173 clustered with strains isolated from fish, as expected, particularly Mugilidae species, which are commonly found in the Kuwaiti Gulf. Two previously identified fish-associated genes were identified, i.e., genes for an α -galactosidase and an ABC transporter permease (14), which further accentuates the differences from *S. agalactiae* serotype Ia strains isolated from cases of human infection.

Data availability. All data have been deposited in the NCBI Sequence Read Archive (SRA). Raw sequence reads are available under accession number [PRJNA627590](https://www.ncbi.nlm.nih.gov/sra/PRJNA627590), and the assembled genome is available under GenBank accession number [CP053027.1](https://www.ncbi.nlm.nih.gov/genbank/CP053027.1).

ACKNOWLEDGMENT

Morena Santi and this research were supported by a Biotechnology and Biological Sciences Research Council (BBSRC) doctoral studentship.

REFERENCES

- Delannoy CMJ, Crumlish M, Fontaine MC, Pollock J, Foster G, Dagleish MP, Turnbull JF, Zadoks RN. 2013. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. *BMC Microbiol* 13:41. <https://doi.org/10.1186/1471-2180-13-41>.
- Farley MM. 2001. Group B streptococcal disease in nonpregnant adults. *Clin Infect Dis* 33:556–561. <https://doi.org/10.1086/322696>.
- Pradeep PJ, Suebsing R, Sirthammajak S, Kampeera J, Jitrakorn S, Saksmerprome V, Turner W, Palang I, Vanichviriyakit R, Senapin S, Jeffs A, Kiatpathomchai W, Withyachumanarnkul B. 2016. Evidence of vertical transmission and tissue tropism of streptococcosis from naturally infected red tilapia (*Oreochromis* spp.). *Aquacult Rep* 3:58–66. <https://doi.org/10.1016/j.aqrep.2015.12.002>.
- Al-Harbi AH. 2016. Phenotypic and genotypic characterization of *Streptococcus agalactiae* isolated from hybrid tilapia (*Oreochromis niloticus* × *O. aureus*). *Aquaculture* 464:515–520. <https://doi.org/10.1016/j.aquaculture.2016.07.036>.
- Verner-Jeffreys DW, Roberts E, Driscoll J, Bayley AE, Algoët M. 2011. Evaluation of the Biolog Microlog 1 System for the identification of Gram positive cocci pathogenic for fish. *Bull Eur Assoc Fish Pathol* 31:171–181.
- Leigh JA, Egan SA, Ward PN, Field TR, Coffey TJ. 2010. Sortase anchored proteins of *Streptococcus uberis* play major roles in the pathogenesis of bovine mastitis in dairy cattle. *Vet Res* 41:63. <https://doi.org/10.1051/vetres/20100036>.
- 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>.
- 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>.
- 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>.
- Nobbs AH, Rosini R, Rinaudo CD, Maione D, Grandi G, Telford JL. 2008. Sortase A utilizes an ancillary protein anchor for efficient cell wall anchoring of pili in *Streptococcus agalactiae*. *Infect Immun* 76:3550–3560. <https://doi.org/10.1128/IAI.01613-07>.
- Rajagopal L. 2009. Understanding the regulation of group B streptococcal virulence factors. *Future Microbiol* 4:201–221. <https://doi.org/10.2217/17460913.4.2.201>.
- Toro N. 2003. Bacteria and Archaea group II introns: additional mobile genetic elements in the environment. *Environ Microbiol* 5:143–151. <https://doi.org/10.1046/j.1462-2920.2003.00398.x>.
- Titus Brown C, Irber L. 2016. sourmash: a library for MinHash sketching of DNA. *J Open Source Softw* 1:27. <https://doi.org/10.21105/joss.00027>.
- Delannoy CMJ, Zadoks RN, Crumlish M, Rodgers D, Lainson FA, Ferguson HW, Turnbull J, Fontaine MC. 2016. Genomic comparison of virulent and non-virulent *Streptococcus agalactiae* in fish. *J Fish Dis* 39:13–29. <https://doi.org/10.1111/jfd.12319>.