



Whole-Genome Sequence of *Streptococcus iniae* Strain AH1, Isolated from Hybrid Tilapia (*Oreochromis niloticus* × *Oreochromis aureus*)

 Ahmed H. Al-Harbi^a

^aLife Science and Environment Research Institute, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia

ABSTRACT Here, we describe the whole-genome sequence of *Streptococcus iniae* strain AH1, which was isolated from moribund farmed hybrid tilapia (*Oreochromis niloticus* × *Oreochromis aureus*) in Saudi Arabia. The genome is composed of a single linear chromosome of 2,068,661 bp, with a G+C content of 36.8%.

Streptococcus iniae is a hemolytic, non-Lancefield-classified Gram-positive pathogenic bacterium that was isolated originally from a subcutaneous abscess of a captive Amazon freshwater dolphin, *Inia geoffrensis*, in 1976 (1). *S. iniae* affects numerous freshwater and marine fish species, including tilapia, causing significant economic losses in the aquaculture industry worldwide (2–4). In addition, *S. iniae* is also known as an opportunistic zoonotic pathogen, causing serious infectious disease in humans (5–8).

S. iniae AH1 was isolated from moribund farmed hybrid tilapia (*Oreochromis niloticus* × *Oreochromis aureus*) in March 1992 in Saudi Arabia as described previously (9, 10).

A single colony of *S. iniae* AH1 was grown aerobically overnight in tryptone soya broth (Oxoid), at 30°C. Genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen), following the manufacturer's instructions. DNA quantity and quality were checked with a NanoDrop spectrophotometer (Thermo Fisher Scientific) and gel electrophoresis, respectively. The whole genome was sequenced using a combination of the PacBio Sequel II platform (Pacific Biosciences, USA) and the NovaSeq 6000 platform (150-bp paired-end reads) (Illumina, USA). For PacBio sequencing, genomic DNA was sheared using the g-TUBEs (Covaris, Woburn, MA) and then size selected to greater than 10 kb using BluePippin, end repaired and ligated with universal hairpin adapters using the SMRTbell template prepare kit v1.0 (Pacific Biosciences) according to the manufacturer's instructions, and sequenced using PacBio P6-C4 chemistry and 240-min movies. The Illumina sequencing library was prepared using the Nextera XT library preparation kit (Illumina) following the manufacturer's instructions.

The reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (11). The library quality was analyzed by Qubit and quantitative PCR (qPCR), and the average fragment size was estimated using an Agilent (Santa Clara, CA, USA) 2100 bioanalyzer. PacBio sequencing yielded a total of 184,455 subreads (mean subread length, 12,636 bp) with an N_{50} value of 14,277 bp, totaling 2.33 Gbp. *De novo* assembly was performed using Flye v2.8 (12) and Mecat2 vJAN-2020 software. A total of 7,892,050 raw Illumina paired-end reads were generated, and 7,761,264 clean reads were used for error correction with Pilon v1.22 (13) for the final genome assembly. The quality and completeness of the assembled genome were assessed by benchmarking universal single-copy orthologs (BUSCO) (v4.1.4) (14) with the bacteria_odb10 data set (created 6 March 2020). This resulted in 95.2% complete genes, with 95.2% of the genome with single copy genes, 0.0% duplicated genes, 0.8% fragmented genes, and 4.0% genes missing. The average genome coverage was 938.54×. The genomes were

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

Copyright © 2022 Al-Harbi. 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 aalharbi1377@gmail.com, aalharbi@kacst.edu.sa.

The authors declare no conflict of interest.

Received 17 February 2022

Accepted 9 April 2022

Published 26 May 2022

TABLE 1 Main genome features of *Streptococcus iniae* AH1

Genetic element	Genome size (bp)	G+C content (%)	Topology	No. of coding sequences	No. of rRNAs	No. of tRNAs	No. of proteins	No. of pseudogenes
Chromosome	2,068,661	36.8	Linear	1,951	18	68	1,864	53
Contig 2	4,202	0	Linear	1	0	0	0	0
Contig 4	1,143	0	Linear	1	0	0	1	0
Contig 6	97,500	52.2	Linear	2	0	0	1	1
Contig 7	6,528	6.1	Linear	1	0	0	1	0
Total or average ^a	2,178,034	37.3		1956	18	68	1,867	54

^a An average value for G+C content (%) is indicated.

annotated by NCBI using the Prokaryotic Genome Annotation Pipeline (PGAP) v4.2 (15). Default parameters were used for all software unless otherwise specified.

The whole-genome sequence of *S. iniae* AH1 consists of 5 contigs ranging from 1,143 bp to 2,068,661 bp, which had a combined length of 2,178,034 bp; with a G+C content of 37.3%; and with 1,956 protein-coding sequences (CDSs), 68 tRNAs, and 18 rRNAs (Table 1).

The resulting chromosome was 2,068,661 bp, with a G+C content of 36.8%, and was predicted to carry 1,951 protein-coding sequences (CDSs), 68 tRNAs, and 18 rRNAs, which is consistent with the other *S. iniae* genomovar strains sequenced previously, namely, *S. iniae* 89353 (16) and *S. iniae* YM011 (17).

Data availability. The whole-genome sequence of *S. iniae* strain AH1, reported in the manuscript, has been deposited at DDBJ/ENA/GenBank under the accession number [JAHZUZ000000000](https://www.ncbi.nlm.nih.gov/nuccore/JAHZUZ000000000), BioProject accession number [PRJNA742166](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA742166), and BioSample accession number [SAMN19929365](https://www.ncbi.nlm.nih.gov/biosample/SAMN19929365). The version described in this paper is the first version, [JAHZUZ010000000](https://www.ncbi.nlm.nih.gov/nuccore/JAHZUZ010000000). The raw Illumina and PacBio sequence reads were deposited in the Sequence Read Archive (SRA) under accession number [SRR17134084](https://www.ncbi.nlm.nih.gov/sra/SRR17134084) and [SRR17157250](https://www.ncbi.nlm.nih.gov/sra/SRR17157250), respectively.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- Pier GB, Madin SH. 1976. *Streptococcus iniae* sp. nov., a beta-hemolytic streptococcus isolated from an Amazon freshwater dolphin, *Inia geoffrensis*. *Int J Syst Bacteriol* 26:545–553. <https://doi.org/10.1099/00207713-26-4-545>.
- Agnew W, Barnes AC. 2007. *Streptococcus iniae*: an aquatic pathogen of global veterinary significance and a challenging candidate for reliable vaccination. *Vet Microbiol* 122:1–15. <https://doi.org/10.1016/j.vetmic.2007.03.002>.
- Austin B, Austin DA. 2016. *Bacterial fish pathogens*, sixth ed. Springer, Chichester, UK.
- Plumb JA, Hanson LA. 2011. *Health maintenance and principal microbial diseases of cultured fishes*, third ed. Wiley-Blackwell, Hoboken, NJ.
- Weinstein MR, Litt M, Kertesz DA, Wyper P, Rose D, Coulter M, McGeer A, Facklam R, Ostach C, Willey BM, Borczyk A, Low DE. 1997. Invasive infections due to a fish pathogen, *Streptococcus iniae*. *S. iniae* study group. *N Engl J Med* 337:589–594. <https://doi.org/10.1056/NEJM199708283370902>.
- Lau SK, Woo PCY, Tse H, Leung KW, Wong SSY, Yuen KY. 2003. Invasive *Streptococcus iniae* infections outside North America. *J Clin Microbiol* 41:1004–1009. <https://doi.org/10.1128/JCM.41.3.1004-1009.2003>.
- Koh TH, Kurup A, Chen J. 2004. *Streptococcus iniae* discitis in Singapore. *Emerg Infect Dis* 10:1694–1696. <https://doi.org/10.3201/eid1009.040029>.
- Sun JR, Yan JC, Yeh CY, Lee SY, Lu JJ. 2007. Invasive infection with *Streptococcus iniae* in Taiwan. *J Med Microbiol* 56:1246–1249. <https://doi.org/10.1099/jmm.0.47180-0>.
- Al-Harbi AH. 1994. First isolation of *Streptococcus* sp. from hybrid tilapia (*Oreochromis niloticus* × *O. aureus*) in Saudi Arabia. *Aquaculture* 128:195–201. [https://doi.org/10.1016/0044-8486\(94\)90308-5](https://doi.org/10.1016/0044-8486(94)90308-5).
- Al-Harbi AH. 2011. Molecular characterization of *Streptococcus iniae* isolated from hybrid tilapia (*Oreochromis niloticus* × *Oreochromis aureus*). *Aquaculture* 312:15–18. <https://doi.org/10.1016/j.aquaculture.2010.12.014>.
- 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>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Seppely M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 1962:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.
- 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>.
- Gong HY, Wu SH, Chen CY, Huang CW, Lu JK, Chou HY. 2017. Complete genome sequence of *Streptococcus iniae* 89353, a virulent strain isolated from diseased tilapia in Taiwan. *Genome Announcements* 5:e01524-16. <https://doi.org/10.1128/genomeA.01524-16>.
- Liu Y, Li L, Yu F, Luo Y, Liang W, Yang Q, Wang R, Li M, Tang J, Gu Q, Luo Z, Chen M. 2020. Genome-wide analysis revealed the virulence attenuation mechanism of the fish-derived oral attenuated *Streptococcus iniae* vaccine strain YM011. *Fish Shellfish Immunol* 106:546–554. <https://doi.org/10.1016/j.fsi.2020.07.046>.