

# Draft Genome Sequence of Beta-Hemolytic *Streptococcus iniae* KCTC 11634

Hye Sung Choi,<sup>a</sup> Mun Gyeong Kwon,<sup>a</sup> Myoung Sug Kim,<sup>a</sup> Myoung Ae Park,<sup>a</sup> Dong-Wook Kim,<sup>c</sup> Jin-Young Park,<sup>c</sup> Ji-Sun Kim,<sup>c</sup> Yun-Jeong Na,<sup>c</sup> Min-Young Kim,<sup>c</sup> Dae-Soo Kim,<sup>b</sup> Sung-Hwa Chae,<sup>c</sup> Jung Soo Seo<sup>a</sup>

Pathology Division, National Fisheries Research and Development Institute (NFRDI), Busan, Republic of Korea<sup>a</sup>; Human Derived Material Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea<sup>b</sup>; Research Institute of GnCIBIO Co., Ltd., Daejeon, Republic of Korea<sup>c</sup>

***Streptococcus iniae* is a beta-hemolytic, Gram-positive coccus, which affects a broad range of freshwater and marine fish species, causing substantial economic losses in the aquaculture industry worldwide. Thus, it is very important to derive a complete genome sequence of the bacterium to aid in the development of vaccines and methods for preventing fish streptococcosis and zoonotic infections in humans. Here, we present the draft genome sequence of *S. iniae* KCTC 11634 (1,955,615 bp, with a G+C content of 36.6%), which contains 1,868 putative coding sequences.**

Received 26 September 2013 Accepted 4 October 2013 Published 7 November 2013

**Citation** Choi HS, Kwon MG, Kim MS, Park MA, Kim D-W, Park J-Y, Kim J-S, Na Y-J, Kim M-Y, Kim D-S, Chae S-H, Seo JS. 2013. Draft genome sequence of beta-hemolytic *Streptococcus iniae* KCTC 11634. *Genome Announc.* 1(6):e00897-13. doi:10.1128/genomeA.00897-13.

**Copyright** © 2013 Choi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Sung-Hwa Chae, [chaekitt@chol.com](mailto:chaekitt@chol.com), or Jung Soo Seo, [jsseosoo@korea.kr](mailto:jsseosoo@korea.kr).

*Streptococcus iniae* is a beta-hemolytic, Gram-positive coccus, which affects a broad range of freshwater and marine fish species, causing substantial economic losses in the aquaculture industry worldwide (1). Through originally isolated from the subcutaneous lesions of a captive Amazon freshwater dolphin, *Inia geoffrensis*, in 1976 (2), *S. iniae* is predominantly a fish pathogen with a broad host range of fresh and saltwater species, such as trout, tilapia, salmon, barramundi, yellowtail, flounder, and hybrid striped bass (1). Streptococcosis caused by *S. iniae* has increased mortality over a period of several weeks in the dominant aquaculture industry of *Paralichthys olivaceus* in South Korea. More importantly, humans can also become infected by *S. iniae*, and a number of human cases have been reported in North America and Asia (3, 4).

Genomic DNA was extracted from the cultured bacteria using the alkaline lysis method (5). We sequenced the genome of this species because it had not been sequenced at the time our sequencing project began, according to the Genomes OnLine Database (GOLD) (6).

We report the genome sequence of *S. iniae* KCTC 11634, obtained using a whole-genome shotgun strategy (7, 8) with a Roche 454 GS (FLX Titanium) pyrosequencing system (760,988 reads totaling ~199.9 Mb, for ~28.9-fold coverage of the genome). Pyrosequencing was processed using Roche's software, according to the manufacturer's instructions. All of the paired reads were assembled using the Newbler assembler 2.6 (454 Life Science), which generated 174 contigs (accession no. BANM01000001 to BANM01000174) (largest contig, ~170 kb) of >100 bp in size. The initial draft assembly contains 69 contigs in 8 scaffolds. The predicted proteins were annotated using Basic Local Alignment Search Tool (BLAST) (9) and the Rapid Annotations using Subsystems Technology (RAST) server (10). In addition, open reading frame (ORF) prediction was merged to use the CD-HIT software to search contigs against the Glimmer 3.02 modeling

software package and Genemark version 2.5 (11), tRNAscan-SE 1.21 (12), RNAmmer 1.2 (13), and Clusters of Orthologous Groups (COG) (14) databases to annotate the gene descriptions.

The *S. iniae* draft genome includes 1,955,615 bp and is composed of 1,868 predicted coding sequences (CDSs), with a G+C content of 36.6%. The genome contains representatives of 307 subsystems, and we used this information to reconstruct the metabolic network (determined using the RAST server). A distinguishing subsystem feature was the absence of a gene corresponding to ABC transporter ATP-binding protein, DNA polymerase, DNA repair protein, membrane protein, repeat-containing cell surface protein precursor, transcriptional regulator. The CDSs annotated by the COG database were classified into 8 major categories (R, E, S, K, G, P, J, and L) from among the 46 COG groups. The enzymes identified included the phosphotransferase system (PTS) (EC 2.7.1.69), DNA polymerase III (EC 2.7.7.7), and RNA pseudouridylyl synthase (EC 5.4.99.12). A more detailed analysis of this genome and a comparative analysis with other *S. iniae* genomes may identify other genes upon finalization of the genome.

**Nucleotide sequence accession numbers.** The draft genome sequence of *S. iniae* KCTC 11634 is available in GenBank under the accession no. [BANM000000000](https://www.ncbi.nlm.nih.gov/nuccore/BANM000000000) (DF266798 to DF266805).

## ACKNOWLEDGMENT

This work was supported by a grant from National Fisheries Research & Development Institute.

## REFERENCES

1. 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.
2. 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.

3. Lau SK, Woo PC, Tse H, Leung KW, Wong SS, Yuen KY. 2003. Invasive *Streptococcus iniae* infections outside North America. *J. Clin. Microbiol.* 41:1004–1009.
4. 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*. *N. Engl. J. Med.* 337:589–594.
5. Birnboim HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513–1523.
6. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. 2010. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 38:D346–D354. doi:10.1093/nar/gkp848.
7. Fullwood MJ, Wei CL, Liu ET, Ruan Y. 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* 19:521–531.
8. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, McKenney K, Sutton G, FitzHugh W, Fields C, Gocyne JD, Scott J, Shirley R, Liu Li-Ing, Glodek A, Kelley JM, Weidman JF, Phillips CA, Spriggs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Fuhrmann JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512.
9. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
10. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi:10.1186/1471-2164-9-75.
11. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679.
12. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
13. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
14. Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families. *Science* 278:631–637.