Draft Genome Sequence of *Streptococcus* sp. X13SY08, Isolated from Murray Cod (*Maccullochella peelii peelii*)

Yang Liu,[a] Runying Zeng,[b] Boqi Weng,[c] Tuyan Luo,[a] Qin Luo,[a] Liwen Xu[d]

Central Laboratory of Fujian Academy of Agricultural Sciences, Fujian, China[a]; State Key Laboratory Breeding Base of Marine Genetic Resource, Third Institute of Oceanography, SOA, Xiamen, China[b]; Institute of Agricultural Ecology of Fujian Academy of Agricultural Sciences, Fujian, China[c]; Key Laboratory of Fishery Ecology and Environment, Guangdong Province, Key Laboratory of South China Sea Fishery Resources Exploitation & Utilization, Ministry of Agriculture, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, China[d]

Streptococcus sp. X13SY08, isolated from freshwater Murray cod fish, likely presents a novel species of *Streptococcus*. Here, we present an annotated draft genome sequence of this species, which will improve our understanding of its physiology and pathogenesis.

Address correspondence to Yang Liu, liuyang19860613@126.com, or Liwen Xu, lwxu@163.com.

The genus *Streptococcus* encompasses a broad range of Gram-positive catalase-negative chain-forming coccus-shaped organisms. Over the last decade, most novel streptococcal species have been isolated from animal sources (1, 2). Animal streptococci have been isolated from a wide range of environments (3) and some of them have been associated with a variety of diseases such as endometritis, respiratory infections, endocarditis, meningitis, arthritis, and mastitis (4). On 18 March 2015, we isolated *Streptococcus* sp. X13SY08 from the orbital tissue of a Murray cod with lesions of panophthalmitis using Columbia blood agar plates (bioMérieux). This strain has been deposited in the Marine Culture Collection of China (accession number MCCC 1A10884). Analysis of the 16S rRNA gene sequence (GenBank accession number KT223645), showed the highest similarity with *Streptococcus ovis* DSM 16829[T] (95.8%), and physiological and biochemical features indicated that *Streptococcus* sp. X13SY08 likely represents a new species in the *Streptococcus* group, making it the first known naturally occurring strain in this clade that can infect Murray cod. This genome sequence may provide fundamental molecular information on the growth characteristics, genetic characteristics, and pathogenesis of *Streptococcus* sp. X13SY08.

The genome of *Streptococcus* sp. X13SY08 was sequenced using Illumina/Solexa MiSeq technology at the Guangzhou Gene Denovo Bio-Technology Co., Ltd. (Guangzhou, China). A library with a fragment length of 500 bp was constructed, and a total of 269 Mbp paired-end reads of 300-bp length were generated. Approximately 200 Mbp high-quality reads, which provided a 128-fold depth of coverage were assembled with SOAPdenovo version 1.05 (5). Protein-coding sequences were predicted by Glimmer software version 3.0 (6) using default parameters and annotated using BLAST searches of nonredundant protein sequences from the NCBI, Swiss-Prot and TrEMBL, COG (7), and KEGG (8) databases. For gene annotation, only significant BLAST matches with E values $\leq 10^{-5}$ were adopted. Ribosomal RNA genes were detected using RNAmmer software version 1.2 (9), and tRNA genes were detected using tRNAscan-SE (10).

The *Streptococcus* sp. X13SY08 genome consists of 1,548,621 bases (N_{50}, 1,548,621 bp) in 5 contigs, with a G+C content of 40.7%. There are 1,756 putative coding sequences, 28 tRNA genes, and 3 rRNA clusters. Gene cluster that participates in the synthesis of Linaridin was detected by antiSMASH 2.0. Genus *Streptococcus* are often recognized as the causative agent of a highly contagious and fatal disease characterized by meningitis and panophthalmitis in animals (4), so we predict pathogenicity of strain X13SY08 through bioinformatics methods. Prophage identification using the PHAST search tool (11) showed 0 prophage regions in the *Streptococcus* sp. X13SY08 genome. No known toxins were identified in the coding sequence of *Streptococcus* sp. X13SY08 by the Web server VirulenceFinder (12) or by comparing the protein sequences with the VFDB (13) and DBETH (14) toxin databases using BLASTp (15). The genome data will represent a solid platform for further characterization and exploitation of the metabolic features linked to the physiology and pathogenesis of *Streptococcus* sp. X13SY08.

Nucleotide sequence accession numbers. The *Streptococcus* sp. X13SY08 whole-genome shotgun (WGS) project has the project accession number LFYO00000000. The version number is LFYO02000000 and consists of sequences LFYO02000001 to LFYO02000005.

ACKNOWLEDGMENTS

REFERENCES

1. **Vela AI, Casas-Diaz E, Lavin S, Dominguez L, Fernandez-Garayzabal JF.** 9 January 2015. *Streptococcus pharyngis* sp. nov., a new streptococcal species isolated from the respiratory tract of wild rabbits. Int J Syst Evol Microbiol. http://dx.doi.org/10.1099/ijs.0.000351.

2. **Saito M, Shinozaki-Kuwahara N, Hirasawa M, Takada K**. 2014. *Streptococcus loxodontisalivarius* sp. nov. and *Streptococcus saliviloxodontae* sp. nov., isolated from oral cavities of elephants. Int J Syst Evol Microbiol **64**:3288–3292. http://dx.doi.org/10.1099/ijs.0.063263-0.

3. **Kilian M**. 1998. *Streptococcus* and *Lactococcus*, p 633–667. In Balows A, Duerden BI (ed), Topley & Wilson's Microbiology and Microbial Infections, vol **2**. Arnold, London, United Kingdom.

4. **Köhler W**. 2007. The present state of species within the genera *Streptococcus* and *Enterococcus*. Int J Med Microbiol **297**:133–150. http://dx.doi.org/10.1016/j.ijmm.2006.11.008.

5. **Li R, Li Y, Kristiansen K, Wang J**. 2008. SOAP: short oligonucleotide alignment program. BioInformatics **24**:713–714. http://dx.doi.org/10.1093/bioinformatics/btn025.

6. **Delcher AL, Bratke KA, Powers EC, Salzberg SL**. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. BioInformatics **23**:673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.

7. **Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV**. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res **29**:22–28. http://dx.doi.org/10.1093/nar/29.1.22.

8. **Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M**. 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res **32**:D277–D280. http://dx.doi.org/10.1093/nar/gkh063.

9. **Lagesen K, Hallin P, Rodland EA, Staerfeldt H-H, Rognes T, Ussery DW**. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35**:3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

10. **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. http://dx.doi.org/10.1093/nar/25.5.0955.

11. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS**. 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39**:W347–W352. http://dx.doi.org/10.1093/nar/gkr485.

12. **Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM**. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol **52**:1501–1510. http://dx.doi.org/10.1128/JCM.03617-13.

13. **Chen L, Xiong Z, Sun L, Yang J, Jin Q**. 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res **40**:D641–D645. http://dx.doi.org/10.1093/nar/gkr989.

14. **Chakraborty A, Ghosh S, Chowdhary G, Maulik U, Chakrabarti S**. 2012. DBETH: a database of bacterial exotoxins for human. Nucleic Acids Res **40**:D615–D620. http://dx.doi.org/10.1093/nar/gkr942.

15. **Altschul S, 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. http://dx.doi.org/10.1093/nar/25.17.3389.