

Complete Genome Sequence of *Streptococcus* sp. Strain NPS 308

Yoshio Kondo,^a Yoshitoshi Ogura,^b Keiko Sato,^c Keigo Imamura,^a Tomonori Hoshino,^a Miyuki Nishiguchi,^a Tomoyuki Hasuwa,^d Hiroyuki Moriuchi,^d Tetsuya Hayashi,^b Taku Fujiwara^a

Department of Pediatric Dentistry, Nagasaki University Graduate School of Biochemical Sciences, Nagasaki, Japan^a; Department of Bacteriology, Faculty of Medical Sciences, Kyushu University, Fukuoka, Japan^b; Division of Microbiology and Oral Infection, Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Biochemical Sciences, Nagasaki, Japan^c; Department of Pediatrics, Nagasaki University Graduate School of Biochemical Sciences, Nagasaki, Japan^d

***Streptococcus* sp. strain NPS 308, isolated from an 8-year-old diagnosed with infective endocarditis, likely presents a novel species of *Streptococcus*. Here, we present a complete genome sequence of this species, which will contribute to better understanding of the pathogenesis of infective endocarditis.**

Received 10 October 2016 Accepted 24 October 2016 Published 22 December 2016

Citation Kondo Y, Ogura Y, Sato K, Imamura K, Hoshino T, Nishiguchi M, Hasuwa T, Moriuchi H, Hayashi T, Fujiwara T. 2016. Complete genome sequence of *Streptococcus* sp. strain NPS 308. *Genome Announc* 4(6):e01349-16. doi:10.1128/genomeA.01349-16.

Copyright © 2016 Kondo 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 Yoshio Kondo, yosioji@nagasaki-u.ac.jp.

Streptococcus species are Gram-positive facultatively anaerobic cocci that are widely present in the oral cavity, occupying an established part of commensal flora (1). Infective endocarditis (IE) is characterized by formation of the vegetation that is a mass of platelets, fibrin, microcolonies of microorganisms, and inflammatory cells on the surfaces of the endocardium, particularly heart valves, and oral viridans streptococci are one of the major pathogens of IE (2).

The sequenced strain, *Streptococcus* sp. NPS 308, was isolated from blood drawn from an 8-year-old girl who had a history of ventricular septal defect since the neonatal period and was hospitalized for fever of unknown origin in Nagasaki University Hospital, Japan. The echocardiography showed the vegetation near the apex of the tricuspid valve, and she was diagnosed with IE in accordance with Duke diagnostic criteria. She had undergone no invasive dental treatment before the onset of IE. However, since oral streptococci were detected in blood culture, an intraoral route of infection was suspected in this case.

The physiological and biochemical features indicated the strain as *Streptococcus* sp. Moreover, phylogenetic analysis based on 16S rRNA gene sequence placed the strain among streptococci; however, classification of the strain to the species level was still inconclusive.

Bacterial cells of *Streptococcus* sp. NPS 308 were inoculated into brain heart infusion broth and incubated under anaerobic conditions. Genomic DNA was extracted using MasterPure Gram-positive DNA purification kit (Epicentre), according to the manufacturer's introduction. Single-molecule real-time sequencing reads were generated using the PacBio RSII sequencer. Whole-genome sequencing produced a total of 162,753 reads with a mean length of 6,974 bp. Subsequent *de novo* assembly utilizing the HGAP3 protocol yielded a single scaffold with 314-fold average reference coverage. To correct sequence errors, Illumina short-read sequencing was performed using the illumina GAIIx (50-bp single-end run). By mapping the obtained 5,829,284 reads to the initial genome assembly using BWA (3) and SAMtools (4), a total of eight errors (one mismatch and seven insertions/deletions)

were corrected. The final assembly contains 1,924,728 bp, with a G+C content of 41.09%. The genome sequence was annotated by Prokka version 1.11 (5) and identified 1,828 coding sequences, 62 tRNAs, four rRNA clusters, and one repeat region.

The average nucleotide identity (ANI) (6, 7) demonstrated that *Streptococcus* sp. NPS 308 differed slightly from mitis phylogenetic group of *Streptococcus*, showing ANI values of <95% compared to the genomes of the closest species, such as *Streptococcus oralis* ATCC 35037 (90.9%) and Uo5 (90.8%) (8). The alignment between *S. oralis* Uo5 and NPS 308 showed that there was no significant genome rearrangement. PHAST (9) program identified two prophage regions within *Streptococcus* sp. NPS 308.

The complete genome sequence of *Streptococcus* sp. NPS 308 will help in establishing a new species of *Streptococcus* and will provide insight into the pathogenicity of this species and host interaction.

Accession number(s). The complete genome sequence has been deposited in the DDBJ/EMBL/GenBank database under GenBank accession no. [AP017652](https://www.ncbi.nlm.nih.gov/nuccore/6017652).

FUNDING INFORMATION

This work, including the efforts of Yoshio Kondo, was funded by Japan Society for the Promotion of Science (JSPS) (15K2059). This work, including the efforts of Miyuki Nishiguchi, was funded by Japan Society for the Promotion of Science (JSPS) (21592602).

REFERENCES

1. Norder Grusell E, Dahlén G, Ruth M, Ny L, Quiding-Järbrink M, Bergquist H, Bove M. 2013. Bacterial flora of the human oral cavity, and the upper and lower esophagus. *Dis Esophagus* 26:84–90. <http://dx.doi.org/10.1111/j.1442-2050.2012.01328.x>.
2. Murdoch DR, Corey GR, Hoen B, Miró JM, Fowler VG, Jr, Bayer AS, Karchmer AW, Olaison L, Pappas PA, Moreillon P, Chambers ST, Chu VH, Falcó V, Holland DJ, Jones P, Klein JL, Raymond NJ, Read KM, Tripodi MF, Uthiri R, Wang A, Woods CW, Cabell CH, International Collaboration on Endocarditis-Prospective Cohort Study (ICE-PCS) Investigators. 2009. Clinical presentation, etiology, and outcome of infective endocarditis in the 21st century: the International Collaboration on Endocarditis-Prospective Cohort Study. *Arch Intern Med* 169:463–473. <http://dx.doi.org/10.1001/archinternmed.2008.603>.

3. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <http://dx.doi.org/10.1093/bioinformatics/btp324>.
4. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.
5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.
6. Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* 4:e1900v1. <http://dx.doi.org/10.7287/peerj.preprints.1900v1>.
7. Kim M, Oh HS, Park SC, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 64: 346–351. <http://dx.doi.org/10.1099/ijs.0.059774-0>.
8. Reichmann P, Nuhn M, Denapate D, Brückner R, Henrich B, Maurer P, Rieger M, Klages S, Reinhard R, Hakenbeck R. 2011. Genome of *Streptococcus oralis* strain Uo5. *J Bacteriol* 193:2888–2889. <http://dx.doi.org/10.1128/JB.00321-11>.
9. 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>.