



# Complete Genome Sequence of *Streptococcus ratti* JH145

Sandra C. Garrett,<sup>a</sup> Sara Olson,<sup>a</sup> Brenton R. Graveley<sup>a</sup>

<sup>a</sup>Department of Genetics and Genome Sciences, Institute for Systems Genomics, UConn Health, Farmington, Connecticut, USA

**ABSTRACT** We report here the complete genome sequence of *Streptococcus ratti* strain JH145. *Streptococcus ratti* is a cariogenic species of mutans streptococcus that has been isolated from rat and human teeth. The strain JH145, derived from strain BHT-2, is interesting for oral health because it does not produce cariogenic lactic acid but shows robust biofilm production.

*Streptococcus ratti* is a species of mutans streptococcus that has been isolated from both rat and human teeth. *Streptococcus ratti* is closely related to *Streptococcus mutans* (1) and, like *S. mutans*, it has been isolated from carious lesions in humans and can cause experimental caries in rodent models (2–4). Strain JH145 was isolated from the parent strain BHT-2 following mutagenesis (5, 6); BHT-2 is a spontaneous streptomycin-resistant derivative of BHT (7). Although *S. ratti* BHT was isolated from human dental caries and was shown to be cariogenic (2), the mutant strain JH145 did not produce lactic acid or cause caries in a rodent model (6). To provide a genetic baseline for understanding tooth colonization and the formation of caries, we sequenced the complete genome of *S. ratti* JH145. To date, a complete genome sequence for *S. ratti* has not been cataloged in GenBank.

*Streptococcus ratti* JH145 was obtained from the ATCC (deposit number 31377) and grown in Todd Hewitt broth under microaerobic conditions (candle jar with ~5% CO<sub>2</sub> [8]). Genomic DNA was isolated using the Qiagen DNeasy PowerBiofilm kit (Qiagen, Valencia, CA). Illumina platform libraries were prepared using the NEBNext Ultra II library kit (New England Biolabs, Ipswich, MA). Libraries were sequenced on a MiSeq instrument (Illumina, San Diego, CA) with a 240 × 60-bp paired-end protocol, followed by fastq generation and adapter trimming with the generateFASTQ module within the MiSeq Reporter analysis package provided with the instrument. Following adapter trimming, the observed mean read lengths were 228 bases and 60 bases, and the mean Q scores were 37 and 36 for reads 1 and 2, respectively. In addition, long-read Oxford Nanopore libraries were prepared using the Ligation sequencing kit (Oxford Nanopore Technologies, Oxford, UK) and sequenced on a MinION R9.4.1 flow cell for 24 h. After sequencing, nanopore reads were base called using Guppy version 3.1.5 (9) with Q score filtering set to a quality threshold of 7. Porechop version 0.2.4 with default settings was used to trim adapter sequences (<https://github.com/rwick/Porechop>). In total, 2,169,276 MiSeq reads and 602,986 Oxford Nanopore reads were generated. The mean read length for Oxford Nanopore reads was 4.9 kb. *De novo* assembly was carried out using SPAdes version 3.9.0 with BayesHammer error correction and default assembly settings (10). A single contig with 170× coverage was returned, and this was annotated using the NCBI Prokaryotic Genome Annotation Pipeline version 4.9 (11). To check that this contig represented the full, closed genome, we created a relinearized reference genome from it. In this reference genome, the first 10 kb of sequence was moved from the beginning of the contig to the end so that the original ends of the assembled contig now presented a continuous, closed sequence. The MiSeq short reads were then aligned to the relinearized reference genome with Bowtie 2 version 2.3.5.1 using default settings (12). Overall, 99.29% of the short reads aligned. In addition, Oxford Nanopore long reads with a Q score greater than

**Citation** Garrett SC, Olson S, Graveley BR. 2020. Complete genome sequence of *Streptococcus ratti* JH145. Microbiol Resour Announc 9:e00144-20. <https://doi.org/10.1128/MRA.00144-20>.

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

**Copyright** © 2020 Garrett 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 Sandra C. Garrett, [garrett@uchc.edu](mailto:garrett@uchc.edu), or Brenton R. Graveley, [graveley@uchc.edu](mailto:graveley@uchc.edu).

**Received** 13 February 2020

**Accepted** 18 February 2020

**Published** 5 March 2020

10 were aligned with minimap2 using default settings (13). Both alignment files were used to generate custom genome coverage tracks (14, 15) which were examined manually on the UCSC genome browser (16). Tracks showed even, uninterrupted coverage over the reference, including over the closed linearization point, which supports the conclusion that the genome is closed.

The complete *S. rattii* JH145 genome is 2,096,943 bp long with a GC content of 40.9%. A total of 1,941 protein coding genes was predicted (51 pseudogenes), along with 86 RNA genes and two CRISPR arrays. The CRISPR arrays were annotated as a type II-A and a type I-E system and contain 11 and 8 spacers, respectively. Unexpectedly, spacer content of the type II-A CRISPR array was identical to that found in *S. mutans* NCTC10920 (GenBank assembly number [GCA\\_900638045](https://doi.org/10.1093/genbank/GCA_900638045)), suggesting a close phylogenetic relationship (17). We used Mashree (18) with default settings to analyze genome distances among *S. rattii* JH145, *S. rattii* FA-1 (type strain, genome in six contigs; [GCA\\_000286075](https://doi.org/10.1093/genbank/GCA_000286075)), and the following three strains of *S. mutans*: *S. mutans* UA159 (type strain; [GCF\\_000007465.2](https://doi.org/10.1093/genbank/GCF_000007465.2)), *S. mutans* GS-5 ([GCA\\_000271865](https://doi.org/10.1093/genbank/GCA_000271865)), and *S. mutans* NCTC10920 ([GCA\\_900638045](https://doi.org/10.1093/genbank/GCA_900638045)). JH145 showed greatest similarity to *S. mutans* NCTC10920, followed by *S. rattii* FA-1. *Streptococcus mutans* NCTC10920 is reported as *S. mutans* on GenBank but elsewhere is identified as *S. rattii*, a designation supported by our comparison. Our complete assembly for *S. rattii* JH145 will provide a resource for comparative genomics and dental health studies.

**Data availability.** The complete genome sequence of *Streptococcus rattii* JH145 has been deposited in GenBank under accession number [CP043405](https://doi.org/10.1093/genbank/CP043405) (assembly number [GCA\\_008803015](https://doi.org/10.1093/genbank/GCA_008803015)) and BioProject identifier [PRJNA573065](https://doi.org/10.1093/bioinformatics/PRJNA573065); raw reads were deposited in the Sequence Read Archive under accession numbers [SRX6958171](https://doi.org/10.1093/bioinformatics/SRX6958171) (MiSeq) and [SRX6958172](https://doi.org/10.1093/bioinformatics/SRX6958172) (Oxford Nanopore). Raw output files for the Oxford Nanopore sequencing (fast5 format) were also deposited under accession number [SRR10977074](https://doi.org/10.1093/bioinformatics/SRR10977074).

## ACKNOWLEDGMENT

This work was supported by National Institutes of Health grant R35GM118140 (B.R.G.).

## REFERENCES

- Coykendall AL. 1974. Four types of *Streptococcus mutans* based on their genetic, antigenic and biochemical characteristics. *J Gen Microbiol* 83: 327–338. <https://doi.org/10.1099/00221287-83-2-327>.
- Jablon JM, Zinner DD. 1966. Differentiation of cariogenic streptococci by fluorescent antibody. *J Bacteriol* 92:1590–1596. <https://doi.org/10.1128/JB.92.6.1590-1596.1966>.
- Zinner DD, Aran AP, Saslaw MS, Jablon JM, Fitzgerald RJ. 1967. Induction of dental caries in gnotobiotic rats by streptococci of human origin. *Nature* 213:200–201. <https://doi.org/10.1038/213200a0>.
- Zinner DD, Jablon JM, Aran AP, Saslaw MS, Fitzgerald RJ. 1966. Comparative pathogenicity of streptococci of human origin in hamster caries. *Arch Oral Biol* 11:1419–1420. [https://doi.org/10.1016/0003-9969\(66\)90031-8](https://doi.org/10.1016/0003-9969(66)90031-8).
- Hillman JD. 1978. Lactate dehydrogenase mutants of *Streptococcus mutans*: isolation and preliminary characterization. *Infect Immun* 21: 206–212. <https://doi.org/10.1128/IAI.21.1.206-212.1978>.
- Hillman JD, McDonell E, Cramm T, Hillman CH, Zahradnik RT. 2009. A spontaneous lactate dehydrogenase deficient mutant of *Streptococcus rattus* for use as a probiotic in the prevention of dental caries. *J Appl Microbiol* 107:1551–1558. <https://doi.org/10.1111/j.1365-2672.2009.04333.x>.
- Johnson CP, Gross SM, Hillman JD. 1980. Cariogenic potential *in vitro* in man and *in vivo* in the rat of lactate dehydrogenase mutants of *Streptococcus mutans*. *Arch Oral Biol* 25:707–713. [https://doi.org/10.1016/0003-9969\(80\)90124-7](https://doi.org/10.1016/0003-9969(80)90124-7).
- Hillman JD. 9 January 1979. Method of controlling dental caries with *Streptococcus mutans* mutant strains. US patent no. 4,133,875.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network base-calling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- 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>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- 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. <https://doi.org/10.1093/bioinformatics/btp352>.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* 12:996–1006. <https://doi.org/10.1101/gr.229102>.
- Shariat N, Dudley EG. 2014. CRISPRs: molecular signatures used for pathogen subtyping. *Appl Environ Microbiol* 80:430–439. <https://doi.org/10.1128/AEM.02790-13>.
- Katz L, Griswold T, Morrison S, Caravas J, Zhang S, Bakker H, Deng X, Carleton H. 2019. Mashree: a rapid comparison of whole genome sequence files. *J Open Source Softw* 4:1762. <https://doi.org/10.21105/joss.01762>.