AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Streptococcus halitosis* sp. nov., Isolated from the Dorsal Surface of the Tongue of a Patient with Halitosis

George Tetz,[a,b] Daria Vikina,[a] Stuart Brown,[c] Paul Zappile,[f] Igor Dolgalev,[c] Aristotelis Tsirigos,[c,d,e] Adriana Heguy,[d,e,f] Victor Tetz[a]

[a]Human Microbiology Institute, New York, New York, USA
[b]Tetz Laboratories, New York, New York, USA
[c]Applied Bioinformatics Laboratories, New York University Medical Center, New York, New York, USA
[d]Department of Pathology, New York University School of Medicine, New York, New York, USA
[e]Laura and Isaac Perlmutter Cancer Center, New York University School of Medicine, New York, New York, USA
[f]Genome Technology Center, Division of Advanced Research Technologies, NYU School of Medicine, New York, New York, USA

ABSTRACT Here, we report the draft genome of *Streptococcus halitosis* sp. nov. strain VT-4, a novel bacterium isolated from the dorsal part of the tongue of a patient with halitosis. The genome comprised 1,880,608 bp with a GC content of 41.0%. There were 1,782 predicted protein-coding genes, including those associated with virulence and antibiotic resistance.

Streptococcus is a genus of nonmotile, Gram-positive, coccus-shaped bacteria. These bacteria cause a variety of human pathologies, including some associated with the oral cavity, such as dental caries, tartar, and gingivitis (1, 2). However, none of the *Streptococcus* species are associated with halitosis that is characterized with an increase in the number of volatile sulfur compounds (VSC), particularly due to H_2S-producing microorganisms on the tongue (3).

Streptococcus halitosis sp. nov. strain VT-4 was isolated from the dorsal part of the tongue of a patient with halitosis using an original workflow as previously described (4–6). For genomic DNA isolation, a single colony grown on Columbia agar at 37°C. Genomic DNA was extracted using a genomic DNA purification kit (QIAamp). The 16S rRNA gene was amplified with the universal bacterial primers 27F and 1492R and assembled using SeqMan v7 software (7, 8). The 16S rRNA gene of *Streptococcus halitosis* VT-4 possesses 99% sequence identity with *S. oralis* and 98% sequence identity with *S. infantis*, *S. mitis*, and *S. pneumoniae* (9).

Paired-end libraries (300-bp length) were prepared using a TruSeq DNA sample prep kit and then sequenced using a HiSeq 2500 instrument (Illumina, USA). Samples underwent library preparation and sequencing according to the manufacturer's instructions.

Raw sequence reads were cleaned with Trimmomatic (v0.38) to remove Illumina adapters and trimmed within a sliding 4-base window, cutting when the average quality per base dropped below 15 (10). Contaminating human sequences were removed by alignment to the human reference genome (GRCH38) with Bowtie 2 (v2.3.4.2), collecting only unaligned read pairs (11). A draft genome was assembled using SPAdes v3.7.1 with default parameters (12). Contigs with an average kmer coverage of less than 10-fold were deleted.

The assembled 18 contigs had an average kmer coverage of 130-fold, with a total length of 1,880,608 bp, a GC content of 41.0%, and an N_{50} value of 1,331,771 bp. The assembled sequences were annotated using the NCBI Prokaryotic Genome

Citation Tetz G, Vikina D, Brown S, Zappile P, Dolgalev I, Tsirigos A, Heguy A, Tetz V. 2019. Draft genome sequence of *Streptococcus halitosis* sp. nov., isolated from the dorsal surface of the tongue of a patient with halitosis. Microbiol Resour Announc 8:e01704-18. https://doi.org/10.1128/MRA.01704-18.

Editor J. Cameron Thrash, University of Southern California

Address correspondence to George Tetz, georgetets@gmail.com.

Received 18 December 2018
Accepted 4 January 2019
Published 24 January 2019

Annotation Pipeline (13). The genome of *S. halitosis* VT-4 harbored 49 tRNA genes, 4 rRNAs, and 3 noncoding RNA (ncRNA) operons, and it had 1,782 protein-coding sequences.

We revealed multidrug resistance transporters of the ATP-binding cassette, multidrug and toxic compound extrusion, and major facilitator superfamily families. Furthermore, resistance determinants, such as an aminoglycoside phosphotransferase and a BlaEC family class C beta-lactamase, were found. Virulence factors, including hemolysin III, metalloproteases, serine proteases, endopeptidases, deoxyribonucleases, ribonucleases, and adhesins, were identified in the genome.

We discovered the presence of cystathionine beta-lyase and L-lactate dehydrogenase genes, associated with the formation of hydrogen sulfide (H_2S), known as a predominant VSC associated with halitosis (3, 14, 15).

An *in silico* DNA-DNA hybridization (dDDH) was carried out using the Genome-to-Genome Distance Calculator (16–18). The results revealed that the *S. halitosis* VT-4 genome was distinct from the genomes of representative strains of related species (i.e., *S. oralis*, *S. infantis*, *S. mitis*, and *S. pneumoniae*, with similarity values of 60.00, 25.80, 25.80, and 31.40%, respectively) and were below the 70% cutoff for dDDH (16, 17).

Further studies of *S. halitosis* VT-4 and bacteriophages associated with this bacterium will enhance our understanding of its implications in human pathologies of the oral cavity (19, 20).

Data availability. The complete genome sequence of *Streptococcus halitosis* sp. nov. strain VT-4 has been deposited in the NCBI database under accession number QEMY00000000. Run data are available from the Sequence Read Archive (SRA) under the accession number SRR8172788.

ACKNOWLEDGMENTS

We thank the Applied Bioinformatics Center (BFX) at the NYU School of Medicine for providing bioinformatics support and helping with the analysis and interpretation of the data. This work used computing resources at the High-Performance Computing Facility (HPCF) of the Center for Health Informatics and Bioinformatics at the NYU Langone Medical Center.

REFERENCES

1. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. 2005. Defining the normal bacterial flora of the oral cavity. J Clin Microbiol 43:5721–5732. https://doi.org/10.1128/JCM.43.11.5721-5732.2005.
2. Kennedy HF, Morrison D, Tomlinson D, Gibson BE, Bagg J, Gemmell CG. 2003. Gingivitis and toothbrushes: potential roles in viridans streptococcal bacteraemia. J Infect 46:67–70.
3. Washio J, Sato T, Koseki T, Takahashi N. 2005. Hydrogen sulfide-producing bacteria in tongue biofilm and their relationship with oral malodour. J Med Microbiol 54:889–895. https://doi.org/10.1099/jmm.0.46118-0.
4. Tetz G, Tetz V. 2017. Introducing the sporobiota and sporobiome. Gut Pathog 9:38. https://doi.org/10.1186/s13099-017-0187-8.
5. Tetz G, Tetz V, Vecherkovskaya M. 2016. Genomic characterization and assessment of the virulence and antibiotic resistance of the novel species *Paenibacillus* sp. strain VT-400, a potentially pathogenic bacterium in the oral cavity of patients with hematological malignancies. Gut Pathog 8:6. https://doi.org/10.1186/s13099-016-0089-1.
6. Tetz G, Tetz V. 2015. Complete genome sequence of *Bacilli bacterium* strain VT-13-104 isolated from the intestine of a patient with duodenal cancer. Genome Announc 3:e00705-15. https://doi.org/10.1128/genomeA.00705-15.
7. Klepac-Ceraj V, Lemon KP, Martin TR, Allgaier M, Kembel SW, Knapp AA, Lory S, Brodie EL, Lynch SV, Bohannan BJ, Green JL, Maurer BA, Kolter R. 2010. Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and *Pseudomonas aeruginosa*. Environ Microbiol 12:1293–1303. https://doi.org/10.1111/j.1462-2920.2010.02173.x.
8. Swindell SR, Plasterer TN. 1997. SeqMan: contig assembly. Methods Mol Bio 70:75–89.
9. Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J. 2014. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. PLoS One 10:e94249. https://doi.org/10.1371/journal.pone.0094249.
10. 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.
11. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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.
13. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W. 2013. Prokaryotic genome annotation pipeline. *In* Beck J, Benson D, Coleman J, Hoeppner M, Johnson M, Maglott D, Mizrachi I, Morris R, Ostell J, Pruitt K, Rubinstein W, Sayers E, Sirotkin K, Tatusova T (ed), The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
14. Nishiya Y, Yoshida Y, Yoshimura M, Fukamachi H, Nakano Y. 2005. Homogeneous enzymatic assay for L-cysteine with βC-S lyase. Biosci Biotechnol Biochem 69:2244–2246.
15. Yoshida Y, Ito S, Kamo M, Kezuka Y, Tamura H, Kunimatsu K, Kato H. 2010. Production of hydrogen sulfide by two enzymes associated with biosynthesis of homocysteine and lanthionine in *Fusobacterium nucleatum* subsp. *nucleatum* ATCC 25586. Microbiology 156:2260–2269. https://doi.org/10.1099/mic.0.039180-0.

16. Auch AF, von Jan M, Klenk H-P, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134. https://doi.org/10.4056/sigs.531120.
17. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
18. Chun J, Rainey FA. 2014. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. Int J Syst Evol Microbiol 64:316–324. https://doi.org/10.1099/ijs.0.054171-0.
19. Tetz GV, Ruggles KV, Zhou H, Heguy A, Tsirigos A, Tetz V. 2017. Bacteriophages as potential new mammalian pathogens. Sci Rep 7:7043. https://doi.org/10.1038/s41598-017-07278-6.
20. Tetz G, Tetz V. 2018. Bacteriophages as new human viral pathogens. Microorganisms 16:54.