AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of "*Candidatus* Nanosynbacter" Strain HMT-348_TM7c-JB, a Member of *Saccharibacteria* Clade G1

Jonathon L. Baker[a,b]

[a]Genomic Medicine Group, J. Craig Venter Institute, La Jolla, California, USA
[b]Department of Pediatrics, U.C. San Diego School of Medicine, La Jolla, California, USA

ABSTRACT *Saccharibacteria* are abundant and diverse members of the human oral microbiome; however, they are poorly understood and appear to exhibit an epibiont/parasitic lifestyle dependent on host bacteria. Here, a complete metagenome-assembled genome (MAG) sequence of an organism from *Saccharibacteria* clade G1 human microbial taxon (HMT) 348 is reported, strain HMT-348_TM7c-JB.

Saccharibacteria have a small cell size, reduced genomes, and appear to have an epiparasitic lifestyle, dependent on host bacteria (1–3). *Saccharibacteria* are common constituents of the oral microbiota, and although they have been correlated with inflammation and disease, their relationship to human health and their overall physiology and lifestyle are poorly understood (4–7). There are at least 6 major clades of oral *Saccharibacteria*, with the clade G1 group, human microbial taxon (HMT) 348, being one of the most abundant *Saccharibacteria* groups detected in supragingival and subgingival plaque, and on the buccal mucosa (6, 8, 9). Draft genomes of HMT-348 have been binned out of saliva and plaque sequencing libraries (4, 9) or assembled as single-cell amplified genomes (SAGs) (10, 11). Although one study was able to isolate an HMT-348 organism along with a putative host species (10), at the time of this work, no complete genomes of HMT-348 were available. Obtaining complete genomes is of special importance for *Saccharibacteria*, as they frequently lack the "essential" core genes and pathways that are used for determining draft genome completeness (3, 8). Furthermore, complete *Saccharibacteria* genomes are also helpful in guiding the isolation, culture, and subsequent study of *Saccharibacteria*, which has proven to be an immense challenge (12). In this study, Nanopore sequencing was used to obtain the complete genome sequence of a *Saccharibacteria* HMT-348 organism, HMT-348_TM7c-JB.

The draft assembly of HMT-348_TM7c-JB, "*Candidatus* Nanosynbacter sp." isolate JCVI_32_bin.19, was reported in 2021, obtained from human saliva in Los Angeles, CA, USA, and fragmented into 7 contigs (4). From the same saliva sample used to obtain the original draft genome sequence, high-molecular-weight genomic DNA was extracted using a phenol:chloroform-based protocol (13) and examined for purity, size, and concentration using a TapeStation instrument (Agilent Technologies). The DNA was not sheared or size selected. A long-read library was prepared using a ligation sequencing kit (Oxford Nanopore Technologies) and sequenced on a GridION using an R9.4.1 flow cell (Oxford Nanopore Technologies). Base calling, quality control, error correction, and adapter trimming were performed using Guppy v4.0.11/MinKNOW v20.06.9 (Oxford Nanopore Technologies), resulting in 9.7 million reads (N_{50}, 6,360 bp). Human reads were removed using minimap2 v2.17-r941 (14), and the remaining long reads were assembled using meta-Flye v2.9-b1768 (15). Among the contigs generated in the Flye metagenomic assembly, a circular 841,116-bp fragment was obtained with >99% average nucleotide identity (ANI; determined using Anvi'o [16]) to "*Candidatus* Nanosynbacter sp." isolate JCVI_32_bin.19. Illumina reads from the original short-read library (used to generate the

Editor Irene L. G. Newton, Indiana University, Bloomington

Address correspondence to jobaker@jcvi.org.

The author declares no conflict of interest.

Received 19 January 2022
Accepted 24 March 2022

sequence for "*Candidatus* Nanosynbacter sp." isolate JCVI_32_bin.19) were mapped to the Flye contig using BWA-MEM v0.7.17-r1188 (17) and used with Polypolish v0.4.3 to remove errors remaining in the circular Flye draft genome. HMT-348_TM7c-JB was annotated using the NCBI Prokaryotic Genome Annotation Pipeline v5.1. Circulator v1.5.5 (18) was used to rotate the genome start to *dnaA*. For all software, default parameters were used unless otherwise noted. The resulting chromosome was 841,302 bp long with a GC content of 38.21%, and is predicted to encode 852 genes. This complete metagenome-assembled genome (MAG) will provide valuable information regarding the lifestyle and evolution of "*Candidatus* Nanosynbacter sp." HMT-348.

Data availability. The complete genome sequence of HMT-348_TM7c-JB is available via GenBank under the accession number CP090820.1. The BioProject accession number for the genome is PRJNA624185. The short reads used to polish the draft genome are available in the Sequence Read Archive (SRA) database under the accession number SRX4318835, and the long reads used to generate the draft assembly are available under SRA accession number SRX13639916.

ACKNOWLEDGMENTS

I thank Karrie Goglin-Alemeida, Jelena Jablanovic, and Kara Riggsbee for performing the library preparation and sequencing.

This research was supported by NIH/NIDCR K99-DE029228.

REFERENCES

1. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523:208–211. https://doi.org/10.1038/nature14486.
2. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF. 2016. A new view of the tree of life. Nat Microbiol 1:16048. https://doi.org/10.1038/nmicrobiol.2016.48.
3. He X, McLean JS, Edlund A, Yooseph S, Hall AP, Liu SY, Dorrestein PC, Esquenazi E, Hunter RC, Cheng G, Nelson KE, Lux R, Shi W. 2015. Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. Proc Natl Acad Sci U S A 112:244–249. https://doi.org/10.1073/pnas.1419038112.
4. Baker JL, Morton JT, Dinis M, Alvarez R, Tran NC, Knight R, Edlund A. 2021. Deep metagenomics examines the oral microbiome during dental caries, revealing novel taxa and co-occurrences with host molecules. Genome Res 31:64–74. https://doi.org/10.1101/gr.265645.120.
5. Abu Fanas S, Brigi C, Varma SR, Desai V, Senok A, D'Souza J. 2021. The prevalence of novel periodontal pathogens and bacterial complexes in Stage II generalized periodontitis based on 16S rRNA next generation sequencing. J Appl Oral Sci 29:e20200787. https://doi.org/10.1590/1678-7757-2020-0787.
6. Bor B, Bedree JK, Shi W, McLean JS, He X. 2019. Saccharibacteria (TM7) in the human oral microbiome. J Dent Res 98:500–509. https://doi.org/10.1177/0022034519831671.
7. Chipashvili O, Utter DR, Bedree JK, Ma Y, Schulte F, Mascarin G, Alayyoubi Y, Chouhan D, Hardt M, Bidlack F, Hasturk H, He X, McLean JS, Bor B. 2021. Episymbiotic Saccharibacteria suppresses gingival inflammation and bone loss in mice through host bacterial modulation. Cell Host Microbe 29:1649–1662.e7. https://doi.org/10.1016/j.chom.2021.09.009.
8. McLean JS, Bor B, Kerns KA, Liu Q, To TT, Solden L, Hendrickson EL, Wrighton K, Shi W, He X. 2020. Acquisition and adaptation of ultra-small parasitic reduced genome bacteria to mammalian hosts. Cell Rep 32:107939. https://doi.org/10.1016/j.celrep.2020.107939.
9. Shaiber A, Willis AD, Delmont TO, Roux S, Chen LX, Schmid AC, Yousef M, Watson AR, Lolans K, Esen OC, Lee STM, Downey N, Morrison HG, Dewhirst FE, Mark Welch JL, Eren AM. 2020. Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. Genome Biol 21:292. https://doi.org/10.1186/s13059-020-02195-w.
10. Cross KL, Campbell JH, Balachandran M, Campbell AG, Cooper SJ, Griffen A, Heaton M, Joshi S, Klingeman D, Leys E, Yang Z, Parks JM, Podar M. 2019. Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. Nat Biotechnol 37:1314–1321. https://doi.org/10.1038/s41587-019-0260-6.
11. Marcy Y, Ouverney C, Bik EM, Losekann T, Ivanova N, Martin HG, Szeto E, Platt D, Hugenholtz P, Relman DA, Quake SR. 2007. Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proc Natl Acad Sci U S A 104:11889–11894. https://doi.org/10.1073/pnas.0704662104.
12. Bor B, Collins AJ, Murugkar PP, Balasubramanian S, To TT, Hendrickson EL, Bedree JK, Bidlack FB, Johnston CD, Shi W, McLean JS, He X, Dewhirst FE. 2020. Insights obtained by culturing saccharibacteria with their bacterial hosts. J Dent Res 99:685–694. https://doi.org/10.1177/0022034520905792.
13. Baker JL, Edlund A. 2020. Composite long- and short-read sequencing delivers a complete genome sequence of B04Sm5, a reutericyclin- and mutanocyclin-producing strain of *Streptococcus mutans*. Microbiol Resour Announc 9:e01067-20. https://doi.org/10.1128/MRA.01067-20.
14. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
15. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x.
16. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen OC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjodin A, Scott JJ, Vazquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 6:3–6. https://doi.org/10.1038/s41564-020-00834-3.
17. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
18. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.