AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Draft Genome Sequence of *Weissella paramesenteroides* STCH-BD1, Isolated from Ensiled *Sorghum bicolor*

**Ashley V. Baugh,[a] Thomas M. Howarth,[b] Katrina L. West,[b] Lydia E. J. Kerr,[b] John Love,[b] David A. Parker,[a,b] Jeffrey R. Fedenko,[a] Richard K. Tennant[a,b]**

[a]Shell International Exploration and Production, Inc., Biodomain, Shell Technology Centre Houston, Houston, Texas, USA
[b]Biosciences, University of Exeter, Exeter, Devon, United Kingdom

**ABSTRACT** *Weissella paramesenteroides* has potential as an industrial biocatalyst due to its ability to produce lactic acid. A novel strain of *W. paramesenteroides* was isolated from ensiled sorghum. The genome was sequenced using a hybrid assembly of Oxford Nanopore and Illumina data to produce a 2-Mbp genome and 22-kbp plasmid sequence.

*Weissella* spp. are Gram-positive coccobacillus-shaped bacteria (1, 2) which were reclassified from the genus *Leuconostoc* (2). *Weissella* spp. have received industrial interest due to their probiotic nature and ability to ferment a range of carbohydrates to lactic and acetic acids (3). Specifically, *Weissella paramesenteroides* is able to produce D-lactic acid (2, 4) and is commonly identified in silage material (5).

*Weissella paramesenteroides* STCH-BD1 was isolated from ensiled *Sorghum bicolor*. Fresh sorghum, cultivated in Florida, was ensiled for 180 days at approximately 21°C in 5-gallon buckets fitted with a 3-piece airlock to maintain anaerobic conditions. Ensiled sorghum was squeezed using a garlic press, and the pressate was spread onto an MRS agar plate (6), which was incubated aerobically at 30°C. A single colony of the isolated bacteria was cultured in MRS broth and incubated aerobically at 30°C and 180 rpm for 16 h. Cells were lysed in lysis tubes containing lysing matrix E (MP Bio, USA) and were placed in the MP Bio FastPrep instrument and operated at 6.0 $ms^{-1}$ for 40 s. Lysates were centrifuged, and DNA was purified using the GeneJET genomic DNA purification kit (Thermo Scientific, USA). Oxford Nanopore Technologies (UK) libraries were prepared using the SQK-RBK004 rapid sequencing kit and sequenced on a MinION instrument attached to a MinIT device (Oxford Nanopore Technologies) using an R9.4a flow cell (Oxford Nanopore Technologies). Oxford Nanopore sequence reads were base called using Guppy v4.2.2 operating in high-accuracy mode and yielded 398,617 DNA sequence reads with an $N_{50}$ value of 3,091 bp. Illumina DNA sequencing libraries were prepared using the Nextera XT library preparation kit (Illumina, USA) and sequenced on the Illumina MiSeq platform, using a 250-bp paired-end sequencing flowcell which yielded 273,888 DNA sequence reads. Default parameters were used for subsequent analysis except where otherwise noted. A *de novo* hybrid assembly using the raw Illumina and Oxford Nanopore reads was performed using MaSuRCA v3.4.2 (7) which was configured as part of the pipeline to use Flye (8) as the final assembler of the corrected reads.

The genome sequence was assembled to a single, linear 2,052,436-bp contig with a GC content of 38% and 210-fold coverage. A circular 22,825-bp contig with a GC content of 33% and 870-fold coverage was identified by Flye and designated a plasmid. The assembled genome was verified using BWA-MEM v0.7.17 (9) and validated in Tablet v1.19.09.03 (10) to ensure complete coverage. Taxonomic classification of the isolate was performed using Kraken 2 v2.0.7 (11) against the standard bacterial database,

**Citation** Baugh AV, Howarth TM, West KL, Kerr LEJ, Love J, Parker DA, Fedenko JR, Tennant RK. 2021. Draft genome sequence of *Weissella paramesenteroides* STCH-BD1, isolated from ensiled *Sorghum bicolor*. Microbiol Resour Announc 10:e01328-20. https://doi.org/10.1128/MRA.01328-20.

**Editor** J. Cameron Thrash, University of Southern California



Address correspondence to Richard K. Tennant, R.K.Tennant@exeter.ac.uk.

**Received** 25 November 2020
**Accepted** 15 January 2021
**Published** 18 February 2021

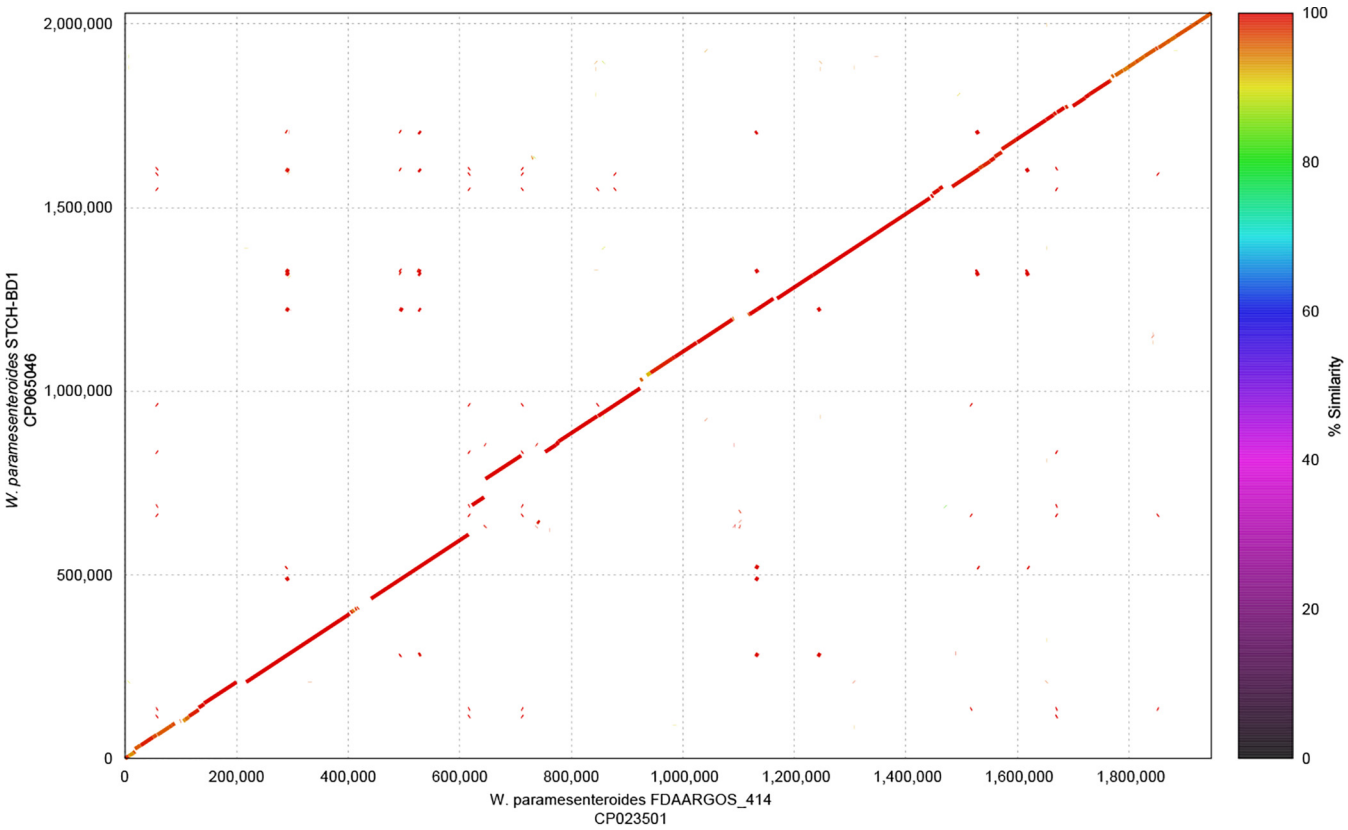


**FIG 1** Genome alignment between *Weissella paramesenteroides* STCH-BD1 (GenBank accession number CP065046) and *Weissella paramesenteroides* FDAARGOS_414 (CP023501) chromosomes. Percentage similarities are displayed using a rainbow color scale.

and fragments of the completed genome were taxonomically verified using NCBI blastn (12). The assembled *W. paramesenteroides* STCH-BD 1 genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (13), which identified 1,934 coding sequences. The completed genome of *W. paramesenteroides* FDAARGOS_414 which was available from NCBI, under accession number CP023501, was compared with the genome of *W. paramesenteroides* STCH-BD 1 using the dnadiff package v1.3 within the MUMmer package v4.0.0rc1 (14) and visualized using mummerplot with the --color option (Fig. 1). The two genomes of *W. paramesenteroides* shared 98.7% homology between the two chromosomes.

**Data availability.** The draft genome sequence of *W. paramesenteroides* STCH-BD1 is deposited in GenBank under the accession numbers CP065045 and CP065046. Oxford Nanopore and Illumina DNA sequence reads have been deposited in the NCBI Sequence Read Archive under accession numbers SRR13083241 and SRR13083242, respectively.

## ACKNOWLEDGMENTS

This work was supported by internal research funding at Shell International Exploration and Production, Inc.

We declare no competing interests.

## REFERENCES

1. Kamboj K, Vasquez A, Balada-Llasat JM. 2015. Identification and significance of Weissella species infections. Front Microbiol 6:1204. https://doi.org/10.3389/fmicb.2015.01204.
2. Collins MD, Samelis J, Metaxopoulos J, Wallbanks S. 1993. Taxonomic studies on some leuconostoc-like organisms from fermented sausages: description of a new genus Weissella for the Leuconostoc paramesenteroides group of species. J Appl Bacteriol 75:595–603. https://doi.org/10.1111/j.1365-2672.1993.tb01600.x.
3. Abriouel H, Lerma LL, Casado Muñoz MC, Montoro BP, Kabisch J, Pichner R, Cho G-S, Neve H, Fusco V, Franz CMAP, Gálvez A, Benomar N. 2015. The controversial nature of the Weissella genus: technological and functional aspects versus whole genome analysis-based pathogenic potential for their application in food and health. Front Microbiol 6:1197. https://doi.org/10.3389/fmicb.2015.01197.
4. Fusco V, Quero GM, Cho G-S, Kabisch J, Meske D, Neve H, Bockelmann W, Franz CMAP. 2015. The genus Weissella: taxonomy, ecology and

biotechnological potential. Front Microbiol 6:155. https://doi.org/10.3389/fmicb.2015.00155.
5. Tohno M, Kobayashi H, Nomura M, Uegaki R, Cai Y. 2012. Identification and characterization of lactic acid bacteria isolated from mixed pasture of timothy and orchardgrass, and its badly preserved silage. Anim Sci J 83:318–330. https://doi.org/10.1111/j.1740-0929.2011.00955.x.
6. De Man J, Rogosa D, Sharpe ME. 1960. A medium for the cultivation of lactobacilli. J Appl Bacteriol 23:130–135. https://doi.org/10.1111/j.1365-2672.1960.tb00188.x.
7. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. Bioinformatics 29:2669–2677. https://doi.org/10.1093/bioinformatics/btt476.
8. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
9. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.
10. Milne I, Bayer M, Stephen G, Cardle L, Marshall D. 2016. Tablet: visualizing next-generation sequence assemblies and mappings, p 253–268. *In* Edwards D (ed), Plant bioinformatics: methods and protocols. Springer, New York, NY.
11. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0.
12. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
13. 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.
14. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. PLoS Comput Biol 14:e1005944. https://doi.org/10.1371/journal.pcbi.1005944.