AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Lactobacillus fermentum* Strain AGR1485, a Human Oral Isolate

Marc A. Bailie,[a,b,c] Eric Altermann,[b,c,d] Wayne Young,[b,c,d] Nicole C. Roy,[b,c,d,e,f] Warren C. McNabb[c,d]

[a]School of Food and Advanced Technology, Massey University, Palmerston North, New Zealand
[b]Food Nutrition & Health Team, AgResearch, Grasslands Research Centre, Palmerston North, New Zealand
[c]Riddet Institute, Massey University, Palmerston North, New Zealand
[d]The High-Value Nutrition National Science Challenge, Auckland, New Zealand
[e]Department of Human Nutrition, University of Otago, Dunedin, New Zealand
[f]Liggins Institute, University of Auckland, Auckland, New Zealand

ABSTRACT *Lactobacillus fermentum* is found in food products and is generally considered safe. *L. fermentum* AGR1485 promotes barrier integrity in Caco-2 cells and has genetic similarities to other known probiotic *L. fermentum* strains. *L. fermentum* AGR1485 has potential as a probiotic and was sequenced to explore these probiotic properties. The genome is a 2.2-Mbp circular chromosome with no plasmids and a GC content of 51.15%.

Lactobacillus fermentum is commonly found in food products (1, 2) and is used as a probiotic therapeutic agent for the treatment of intestinal and vaginal pathogens (3–5). *L. fermentum* AGR1485 increases the barrier integrity of Caco-2 monolayers but does not induce colonic inflammation in germfree mice (6, 7). *L. fermentum* AGR1485 has genetic similarities to other known probiotic *L. fermentum* strains, but little is known about the probiotic mechanisms of action of these strains. *L. fermentum* AGR1485 was isolated by oral swab from a healthy human volunteer and identified as a member of the species *L. fermentum* using 16S rRNA gene sequencing (8). The genome of *L. fermentum* AGR1485 was sequenced to determine its probiotic properties and genetic characteristics.

L. fermentum AGR1485 was cultured in de Man-Rogosa-Sharpe (MRS) broth (Merck Ltd., Auckland, New Zealand) and incubated overnight at 37°C. The cells were lysed by grinding under liquid nitrogen followed by the Qiagen bacterial lysis method (9). High-molecular-weight DNA was purified on Qiagen Genomic-tip/100G columns (Bio-Strategy, Auckland, New Zealand) following the manufacturer's specifications (9).

Whole-genome sequencing was done using both Illumina and PacBio single-molecule real-time (SMRT) sequencing for hybrid assembly. Illumina library creation and sequencing were carried out by BGI (Beijing Genomic Institute, Shenzhen, China), and PacBio SMRTbell library creation and sequencing were done by Novogene (Hong Kong, China). Illumina template DNA was sheared into 500-bp fragments for library creation using a TruSeq library preparation kit from Illumina. The resulting template DNA was sequenced on a HiSeq 2000 genome analyzer that generated 2,738,965 100-bp paired-end reads. The PacBio SMRTbell library was created from sheared template DNA, and the hairpin dimers were purified by magnetic beads with size selection conditions. The adapters were removed using PacBio's MagBead kit. SMRT sequencing of the SMRTbell templates was carried out on a PacBio Sequel platform, creating 249,954 subreads with an average length of 8,790 bp and an N_{50} value of 10,057 bp.

Default parameters were applied for all software packages unless otherwise speci-

Citation Bailie MA, Altermann E, Young W, Roy NC, McNabb WC. 2020. Complete genome sequence of *Lactobacillus fermentum* strain AGR1485, a human oral isolate. Microbiol Resour Announc 9:e00841-20. https://doi.org/10.1128/MRA.00841-20.

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

Copyright © 2020 Bailie et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Eric Altermann, Eric.Altermann@agresearch.co.nz.

Received 20 July 2020
Accepted 8 August 2020
Published 3 September 2020

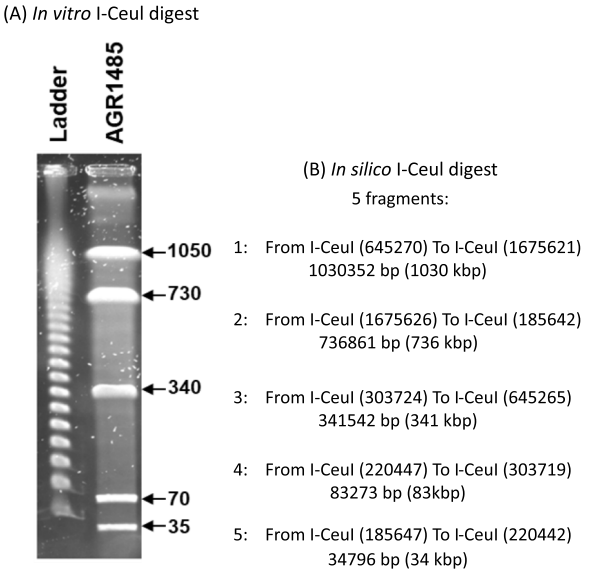

FIG 1 (A) Pulsed-field gel electrophoresis of AGR1485 genomic DNA digested with restriction enzyme I-CeuI. The marker ladder contained lambda DNA, where the fragments were multimers of 48.5 kb. The values given are the sizes (kb) of the DNA fragments from the bacterial strains. Image and caption adapted from Sengupta et al. (6) under a CC BY-NC 4.0 license. (B) *In silico* digest results of the AGR1485 genome assembly using I-CeuI restriction sites processed by UGENE (17) and presented as the range from one restriction site to the next (fragment sizes are in base pairs and kilobase pairs).

fied. Illumina reads were assessed for length, quality, and adapters before and after trimming using FastQC v0.11.9 (10). Illumina reads were quality controlled and trimmed with Trimmomatic v0.39 (11). Hybrid *de novo* assembly was performed with Unicycler v0.4.7 (12) using the trimmed Illumina short reads and uncorrected PacBio long reads. The assembly graph was checked for errors and completeness using Bandage v0.8.1 (13) before final assembly and polishing. Polishing was done with Pilon v1.22 (14) using all the sequencing reads until no further improvements to the assembly could be made.

The polished assembly was assessed with CheckM v1.0.18 to have 99.18% completeness and 0.546% contamination (15). QUAST v4.6.3 (16) confirmed that the genome sequence was a single 2,226,862-bp chromosome with a GC content of 51.15% and contained no missing or ambiguous nucleotides. Read coverage for the assembly was calculated to be 989.2×. During assembly, Unicycler v0.4.7 predicted the chromosome to be circular. UGENE v34.0 (17) was used to overlap the ends of the *L. fermentum* AGR1485 sequence and calculate an *in silico* digest at the I-CeuI restriction sites. The *in silico* restriction digest fragment pattern and fragment sizes closely resembled previously published pulsed-field gel electrophoresis results of the same genome using a commercial I-CeuI restriction enzyme (Fig. 1) (6).

The assembly was annotated with PGAP v4.11 (18) and GAMOLA2 v16.0 (19), which found 2,367 open reading frames (ORFs), 1,839 Clusters of Orthologous Groups (COGs), and 1,920 conserved domains. The genetic characteristics that give rise to this organism's unique phenotypes are likely harbored in this chromosome, because no plasmids were identified during DNA purification or assembly.

Data availability. The PacBio long reads (SRX7669246) and Illumina reads (SRX7669245) described here have been deposited at NCBI/GenBank under BioProject number PRJNA588334. The whole-genome sequence (accession number CP047584.1) is available from NCBI/GenBank under BioSample accession number SAMN13241836.

ACKNOWLEDGMENTS

We thank Kerri Reilly and Ruy Jauregui for the helpful advice, training, and expertise they provided during this work.

M.A.B. was supported by a Ph.D. fellowship from the Riddet Institute through

funding provided by the New Zealand Tertiary Education Commission. E.A. and W.Y. were partly funded by the AgResearch Strategic Science Investment Fund (contract number A21246).

REFERENCES

1. Tulumoğlu Ş, Kaya Hİ, Şimşek Ö. 2014. Probiotic characteristics of Lactobacillus fermentum strains isolated from tulum cheese. Anaerobe 30:120–125. https://doi.org/10.1016/j.anaerobe.2014.09.015.
2. Swain MR, Anandharaj M, Ray RC, Parveen Rani R. 2014. 2014. Fermented fruits and vegetables of Asia: a potential source of probiotics. Biotechnol Res Int 2014:250424. https://doi.org/10.1155/2014/250424.
3. Shieh M-J, Shang H-f, Liao F-H, Zhu J-S, Chien Y-W. 2011. Lactobacillus fermentum improved intestinal bacteria flora by reducing Clostridium perfringens. E Spen Eur E J Clin Nutr Metab 6:e59–e63. https://doi.org/10.1016/j.eclnm.2011.01.001.
4. Kaewnopparat S, Dangmanee N, Kaewnopparat N, Srichana T, Chulasiri M, Settharaksa S. 2013. In vitro probiotic properties of Lactobacillus fermentum SK5 isolated from vagina of a healthy woman. Anaerobe 22:6–13. https://doi.org/10.1016/j.anaerobe.2013.04.009.
5. Kaur B, Balgir P, Mittu B, Chauhan A, Kumar B. 2013. Purification and physicochemical characterization of anti-Gardnerella vaginalis bacteriocin HV6b produced by Lactobacillus fermentum isolate from human vaginal ecosystem. Am J Biochem Mol Biol 3:91–100. https://doi.org/10.3923/ajbmb.2013.91.100.
6. Sengupta R, Anderson RC, Altermann E, McNabb WC, Ganesh S, Armstrong KM, Moughan PJ, Roy NC. 2015. Lactobacillus fermentum AGR1487 cell surface structures and supernatant increase paracellular permeability through different pathways. Microbiologyopen 4:541–552. https://doi.org/10.1002/mbo3.260.
7. Anderson RC, Ulluwishewa D, Young W, Ryan LJ, Henderson G, Meijerink M, Maier E, Wells JM, Roy NC. 2016. Human oral isolate Lactobacillus fermentum AGR1487 induces a pro-inflammatory response in germ-free rat colons. Sci Rep 6:20318–20318. https://doi.org/10.1038/srep20318.
8. Anderson RC, Cookson AL, McNabb WC, Kelly WJ, Roy NC. 2010. Lactobacillus plantarum DSM 2648 is a potential probiotic that enhances intestinal barrier function. FEMS Microbiol Lett 309:184–192. https://doi.org/10.1111/j.1574-6968.2010.02038.x.
9. Qiagen. 2015. Qiagen genomic DNA handbook. Qiagen, Hilden, Germany. https://www.qiagen.com/us/resources/resourcedetail?id=d2b85b26-16dd-4259-a3a7-a08cbd2a08a3&lang=en.
10. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
11. 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.
12. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
13. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383.
14. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
15. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
16. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
17. Okonechnikov K, Fursov M, Golosova O, UGENE Team. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. https://doi.org/10.1093/bioinformatics/bts091.
18. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.
19. Altermann E, Lu J, McCulloch A. 2017. GAMOLA2, a comprehensive software package for the annotation and curation of draft and complete microbial genomes. Front Microbiol 8:346. https://doi.org/10.3389/fmicb.2017.00346.