AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Annotated Genome Sequence of *Limosilactobacillus fermentum* AGR1487

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 *Limosilactobacillus fermentum* is a probiotic species; however, *L. fermentum* AGR1487 increases colon inflammation in germfree mice and decreases barrier integrity in Caco-2 cells. The AGR1487 genome was sequenced to explore these phenotypes. The genome is a single, circular, 1,939,032-bp chromosome with a G+C content of 52.17% and no plasmids.

Limosilactobacillus fermentum strains are regularly used for fermented food production and preservation as acid-producing starter cultures (1, 2). In humans, strains of *L. fermentum* have been shown to improve the ratio of beneficial microorganisms of the large intestine and have been used as probiotic treatments for intestinal and vaginal diseases (1–3). However, *L. fermentum* AGR1487 has been found to increase colon inflammation in germfree mice and decreased the barrier integrity of Caco-2 monolayers (4, 5). *L. fermentum* AGR1487 was isolated from an oral swab of a healthy human and identified using 16S rRNA gene sequencing (6). AGR1487 was sequenced to explore this unique barrier disruptive phenotype and its genetic characteristics.

L. fermentum AGR1487 cells were grown in de Man-Rogosa-Sharpe (MRS) broth (Merck Ltd., Auckland, New Zealand) to stationary phase overnight at 37°C. Genome extraction, purification, and Illumina and PacBio shotgun sequencing were carried out as previously described (7). The Illumina library was created using the TruSeq library kit with genomic DNA sheared into 500-bp fragments and sequenced on a HiSeq 2000 genome analyzer. Illumina sequencing generated 2,523,872 2×100-bp paired-end (200-bp combined) Illumina reads. The sheared genomic DNA was used for the creation of a 10-kb PacBio SMRTbell library. Ten-kilobyte size selection conditions were used when purifying the hairpin dimers by magnetic bead, and the adapters were removed using PacBio's MagBead kit. Sequencing was carried out on the PacBio Sequel platform, generating 344,060 subreads with an average length of 8,498 bp and an N_{50} value of 9,837 bp.

Default parameters were applied for all software packages unless otherwise specified. Illumina short-read quality control was done using FastQC v0.11.9 (8) before and after trimming with Trimmomatic v0.39 (9). Assembly graphs were assessed for errors using Bandage v0.8.1 (10). A single circular genome assembly was produced by Unicycler v0.4.7 (11) using the trimmed Illumina short reads along with uncorrected PacBio long reads. The final genome assembly was polished for three rounds using Pilon v1.22 (12).

CheckM v.1.0.18 (13) reported genome completeness rates of 99.18% and 0.55%

Citation Bailie MA, Altermann E, Young W, Roy NC, McNabb WC. 2021. Complete annotated genome sequence of *Limosilactobacillus fermentum* AGR1487. Microbiol Resour Announc 10:e01056-20. https://doi.org/10.1128/MRA.01056-20.

Editor Catherine Putonti, Loyola University Chicago

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

Received 11 September 2020
Accepted 23 November 2020
Published 7 January 2021

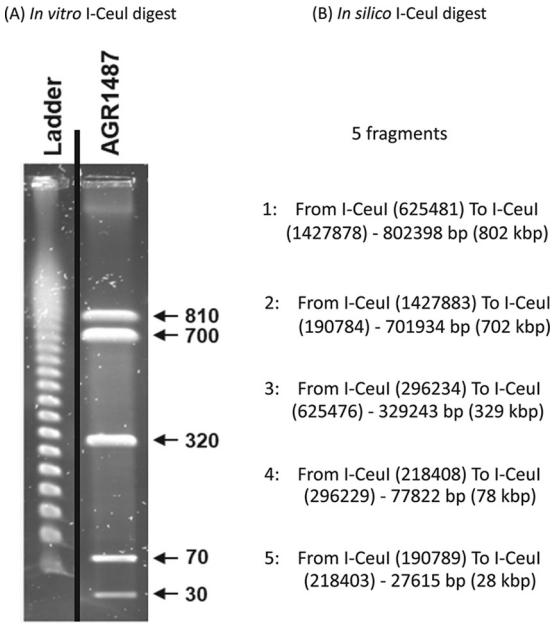

FIG 1 (A) Pulsed-field gel electrophoresis of AGR1487 genomic DNA digested with restriction enzyme I-CeuI. The marker ladder contained lambda DNA, where the fragments were multimers of 48.5 kb. The values given are the sizes (kb) of the DNA fragments from the bacterial strains. The graph depicts the relevant parts of the original published gel image for conciseness, and the vertical black line indicates the boundaries between image slices. (Adapted from *Microbiologyopen* [4].) (B) *In silico* digest results of the AGR1487 genome assembly using I-CeuI restriction sites processed by UGENE (11) and presented as the range from one restriction site to the next (fragment size in bp and Kbp).

contamination. The basic statistics were calculated using QUAST v4.6.3 (14), which found that the resulting assembly was a single 1,939,032-bp contig with a G+C content of 52.17% and no ambiguous bases or gaps filled with arbitrary place holders (Ns). The expected average read depth was calculated to be 1,510.51×. The genome assembly was uploaded to UGENE v34.0 (15), the ends of the sequence were digitally overlapped, and an *in silico* digest at the I-CeuI restriction sites was calculated. The resulting fragment pattern from the *in silico* digest matched a previously published restriction digest of AGR1487 that used a commercial I-CeuI restriction enzyme (Fig. 1) (4).

PGAP v4.10 (16) and GAMOLA2 v16.0 (17) were used to annotate the AGR1487 genome assembly, which was found to harbor 2,065 open reading frames (ORFs), 1,743 conserved domains, and 1,666 clusters of orthologous groups (COGs). The genetic origin of the barrier disruptive phenotype for this strain is likely found in this chromosome, as no plasmids were found during genomic DNA purification and sequence assembly.

Data availability. The PacBio long reads and Illumina MiSeq sequence reads described here have been deposited at NCBI/GenBank under BioProject accession number PRJNA596816. The whole-genome sequence is available from NCBI/GenBank under BioSample accession number SAMN13639333 or directly using the assembly accession number CP047585.

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 Centre of Research Excellence, 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. 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.
2. 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.
3. Kaur B, Balgir P, Mittu B, Chauhan A, Kumar B. 2012. 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.
4. 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.
5. 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. https://doi.org/10.1038/srep20318.
6. 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.
7. 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.
8. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. Okonechnikov K, Golosova O, Fursov M, the UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. https://doi.org/10.1093/bioinformatics/bts091.
16. 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.
17. 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.