



OPEN

## Antimicrobial susceptibilities and comparative whole genome analysis of two isolates of the probiotic bacterium *Lactiplantibacillus plantarum*, strain ATCC 202195

Lisa G. Pell<sup>1,11</sup>, Rachael G. Horne<sup>2,11</sup>, Stuart Huntley<sup>3</sup>, Hafizur Rahman<sup>4</sup>, Sanchita Kar<sup>4</sup>, Mohammad Shahidul Islam<sup>4</sup>, Kara C. Evans<sup>3</sup>, Samir K. Saha<sup>4</sup>, Aaron Campigotto<sup>5,6</sup>, Shaun K. Morris<sup>1,5,7,8</sup>, Daniel E. Roth<sup>1,5,7,10</sup>✉ & Philip M. Sherman<sup>2,5,9</sup>✉

A synbiotic containing *Lactiplantibacillus plantarum* [American Type Culture Collection (ATCC) strain identifier 202195] and fructooligosaccharide was reported to reduce the risk of sepsis in young infants in rural India. Here, the whole genome of two isolates of *L. plantarum* ATCC 202195, which were deposited to the ATCC approximately 20 years apart, were sequenced and analyzed to verify their taxonomic and strain-level identities, identify potential antimicrobial resistant genes and virulence factors, and identify genetic characteristics that may explain the observed clinical effects of *L. plantarum* ATCC 202195. Minimum inhibitory concentrations for selected antimicrobial agents were determined using broth dilution and gradient strip diffusion techniques. The two *L. plantarum* ATCC 202195 isolates were genetically identical with only three high-quality single nucleotide polymorphisms identified, and with an average nucleotide identity of 99.99%. In contrast to previously published reports, this study determined that each isolate contained two putative plasmids. No concerning acquired or transferable antimicrobial resistance genes or virulence factors were identified. Both isolates were sensitive to several clinically important antibiotics including penicillin, ampicillin and gentamicin, but resistant to vancomycin. Genes involved in stress response, cellular adhesion, carbohydrate metabolism and vitamin biosynthesis are consistent with features of probiotic organisms.

Accumulating evidence supports the use of probiotics to prevent morbidity and mortality during early infancy in both high- and low- to middle-income settings<sup>1–3</sup>. Despite marked heterogeneity in probiotic strains studied across randomized controlled trials, recent evidence has sparked interest in the potential of *Lactiplantibacillus plantarum* ATCC 202195<sup>4</sup>, formerly named *Lactobacillus plantarum* ATCC 202195, to improve early infant health outcomes, particularly in resource-constrained settings<sup>5</sup>.

<sup>1</sup>Centre for Global Child Health, Hospital for Sick Children, Toronto, ON, Canada. <sup>2</sup>Cell Biology Program, Research Institute, Hospital for Sick Children, Toronto, ON, Canada. <sup>3</sup>International Flavors & Fragrances Inc., Madison, WI, USA. <sup>4</sup>Child Health Research Foundation, Dhaka, Bangladesh. <sup>5</sup>Department of Paediatrics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada. <sup>6</sup>Division of Microbiology, Hospital for Sick Children, Toronto, ON, Canada. <sup>7</sup>Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada. <sup>8</sup>Division of Infectious Diseases, Hospital for Sick Children, Toronto, ON, Canada. <sup>9</sup>Gastroenterology, Hepatology and Nutrition, Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G 1X8, Canada. <sup>10</sup>Paediatric Medicine and Child Health Evaluative Sciences, Hospital for Sick Children, Peter Gilgan Centre for Research and Learning, 686 Bay Street, Toronto, ON M5G 0A4, Canada. <sup>11</sup>These authors contributed equally: Lisa G. Pell and Rachael G. Horne. ✉email: daniel.roth@sickkids.ca; philip.sherman@sickkids.ca

*Lactiplantibacillus plantarum* ATCC 202195 is a Gram-positive, non-spore forming, facultative anaerobe originally isolated from human feces<sup>6</sup>. Oral administration of *L. plantarum* ATCC 202195 plus fructooligosaccharide (FOS) to newborns has been shown to lead to sustained intestinal colonization in early infancy<sup>5,7</sup>. In a placebo-controlled randomized trial in rural India, oral administration of *L. plantarum* ATCC 202195 plus FOS to newborns (birthweight  $\geq$  2000 g and gestational age  $\geq$  35 weeks) was reported to reduce episodes of clinical sepsis within the first 60 days of life, compared to placebo, by 40% (95% CI 0.48–0.74)<sup>5</sup>. A 78% (95% CI 0.09–0.53) reduction in the incidence of blood culture-confirmed bacterial sepsis was also reported in the same study<sup>5</sup>. Lactobacilli were not grown in any blood culture, and adverse events involving the gastrointestinal tract were similar between trial arms, with the investigational product being generally well tolerated<sup>5</sup>. The direction and magnitude of the estimated effect sizes, combined with early safety data, indicate that *L. plantarum* ATCC 202195 has the potential to improve infant health outcomes, especially in settings where the burden of severe infection and rates of infant mortality are high.

Prior to the aforementioned clinical trial, an isolate of *L. plantarum* ATCC 202195, which we refer to herein as *L. plantarum* ATCC 202195-A, was deposited to the American Type Culture Collection (ATCC) in January 1999 as a patent submission deposit (U.S. Patent Number: 6,132,710)<sup>6</sup> and assigned the ATCC number 202195. Notably, the patent holder permitted select commercial sale of the isolate; however, no rights to propagate or characterize the isolate were permitted. Approximately twenty years after the initial patent deposit to the ATCC, a new isolate of the *L. plantarum* ATCC 202195 strain was submitted to the ATCC, which we refer to here as *L. plantarum* ATCC 202195-B. Commercial production and distribution by the ATCC were authorized by the patent holder. At the time of both submissions to the ATCC, whole genome sequence data were not publicly available for either *L. plantarum* ATCC 202195-A or *L. plantarum* ATCC 202195-B. Thus, the strain-level identity of both isolates was unknown and the identity of *L. plantarum* ATCC 202195-B was not genetically verified.

Strain-level identification ensures that microbes held in culture repositories are correctly identified and handled<sup>8</sup>, provides a genetic signature to monitor both probiotic engraftment and probiotic-associated bacteremia, and provides a baseline sequence to assess whether the genetic composition of the strain changes over time. ATCC, as an International Depository Authority, assumes no liability for the accuracy of the depositors' information concerning patent deposits. Since probiotics are living microorganisms, as part of their natural evolutionary process they may acquire deleterious genetic mutations that could mitigate proposed health benefits<sup>9,10</sup>, which are often strain-specific<sup>11,12</sup>. Thus, verifying the genetic identity of a probiotic strain prior to its administration and throughout the manufacturing process is of paramount importance.

Despite increasing clinical interest in *L. plantarum* ATCC 202195, there is only one peer-reviewed published report describing the draft genome of this probiotic strain<sup>13</sup>, and one additional complete genome assembly released on Genbank (accession GCA\_010586945.1), with no linked peer reviewed report. Neither genome addressed a critical concern regarding the genetic lineage of *L. plantarum* ATCC 202195.

The antimicrobial susceptibility profile of a probiotic strain and the potential for the ability to transfer antimicrobial resistance (AMR) genes or virulence factors to other microorganisms is a key consideration in determining a probiotic's suitability for use in humans and/or animals. There is general consensus that antimicrobial susceptibility profiles should be determined for all probiotics<sup>14</sup> and that a microbe must be confirmed to contain no acquired or transferrable AMR<sup>15</sup> or virulence factors<sup>16</sup> prior to its use as a probiotic agent. The need to determine the antimicrobial susceptibility profile of a probiotic strain extends beyond the potential to transfer AMR elements to other microorganisms. If a probiotic is deemed to be the cause or suspected cause of bacteremia or other invasive infection, knowledge of AMR patterns can inform clinical treatment considerations. In addition, given that there is growing evidence in support of the combined delivery of probiotics and antibiotics<sup>17,18</sup>, knowledge of the AMR phenotype of a microbe could inform the development of novel multi-component treatment regimens.

While the antimicrobial susceptibility profile of *L. plantarum* ATCC 202195-A was previously reported and it was indicated not to harbor a plasmid<sup>5</sup>, minimum inhibitory concentrations (MICs) for this probiotic strain have not been described nor have there been sequence-based reports to corroborate resistance phenotypes.

In the present study, we aimed to determine the sequence of the bacterial genome and verify the taxonomic and strain-level identity of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B. By analyzing the bacterial genome, we aimed to identify AMR genes and potential virulence factors, and to provide possible mechanistic insights underlying the clinical effects of *L. plantarum* ATCC 202195. We also aimed to determine the MICs for each isolate of *L. plantarum* ATCC 202195 by employing a panel of 20 clinically relevant antimicrobial agents. Together, these findings can be used to inform the design of future clinical studies employing *L. plantarum* ATCC 202195 either as a probiotic or in a synbiotic formulation.

## Results

**Antimicrobial susceptibility and minimum inhibitory concentrations of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B.** Based on broth dilution assays, the antimicrobial susceptibility profiles of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B were identical (Table 1). Both isolates were found to be sensitive to penicillin (MIC = 4  $\mu$ g/ml), ampicillin (MIC = 2  $\mu$ g/ml), meropenem (MIC  $\leq$  0.25  $\mu$ g/ml), clindamycin (MIC  $\leq$  0.12  $\mu$ g/ml), linezolid (MIC = 4  $\mu$ g/ml), erythromycin (MIC  $\leq$  0.25  $\mu$ g/ml), chloramphenicol (MIC = 8  $\mu$ g/ml), gentamicin (MIC  $\leq$  2  $\mu$ g/ml), piperacillin/tazobactam (MIC  $\leq$  8  $\mu$ g/ml) and daptomycin (MIC  $\leq$  0.25  $\mu$ g/ml). By contrast, both isolates were resistant to vancomycin (MIC  $\geq$  256  $\mu$ g/ml) and tetracycline (MIC  $\geq$  32  $\mu$ g/ml). MICs for ceftriaxone (MIC  $\leq$  0.5  $\mu$ g/ml), levofloxacin (MIC  $\geq$  8  $\mu$ g/ml), ciprofloxacin (MIC  $\geq$  4  $\mu$ g/ml), quinupristin/dalfopristin (MIC = 2  $\mu$ g/ml), rifampin (MIC  $\geq$  8  $\mu$ g/ml), trimethoprim/sulfamethoxazole (MIC  $\leq$  0.5/9.5  $\mu$ g/ml), gatifloxacin (MIC = 2  $\mu$ g/ml) and oxacillin (MIC = 4  $\mu$ g/ml) were determined; however, clinical breakpoints for these antimicrobials and *Lactobacillus* spp. are not available from

Antibiotic class	Antimicrobial	Broth dilution MICs and interpretations			Prior Evidence from <i>L. plantarum</i> ATCC 202195, other strains of <i>L. plantarum</i> , and other types of lactobacilli
		<i>L. plantarum</i> ATCC 201295-A	<i>L. plantarum</i> ATCC 202195-B	Interpretation	
		MIC (µg/ml)	MIC (µg/ml)		
Beta-lactam: Penicillin	Penicillin	4	4	S <sub>a</sub> _bc	Clinical breakpoint interpretation consistent with previous disc diffusion findings for the <i>L. plantarum</i> ATCC 202195 strain <sup>5</sup> Observed MIC value within one doubling dilution of the upper end of the MIC range that was previously reported for 46 other isolates of <i>L. plantarum</i> (0.5–2 mg/l) <sup>22</sup>
	Ampicillin	2	2	S <sub>a,b,c</sub>	Strain-specific susceptibility data were not previously available for ampicillin Observed MIC value within the range of values previously reported for 46 other strains of <i>L. plantarum</i> (0.125–2 mg/l) <sup>22</sup> and 10 other <i>L. plantarum</i> strains (0.5–32 µg/ml) <sup>23</sup>
	Oxacillin + 2% NaCl	4	4	_a,b,c	Strain-specific susceptibility data were not previously available for oxacillin. Oxacillin activity against other strains of <i>L. plantarum</i> has not previously been reported
Beta-lactam: Carbapenem	Meropenem	≤ 0.25	≤ 0.25	S <sub>a,c</sub> _b	Strain-specific susceptibility data were not previously available for meropenem Meropenem activity against other strains of <i>L. plantarum</i> has not previously been reported
Beta-lactam: Cephalosporin	Ceftriaxone	≤ 0.5	≤ 0.5	_a,b,c	Previous disc diffusion assays suggest that <i>L. plantarum</i> ATCC 202195 is ceftriaxone-sensitive <sup>5</sup> . Clinical breakpoint interpretations were not available for ceftriaxone and lactobacilli Ceftriaxone activity against other strains of <i>L. plantarum</i> has not previously been reported
Beta-lactam: Penicillin/Beta-lactamase inhibitor	Piperacillin/tazobactam	≤ 8	≤ 8	S <sub>c</sub> _a,b	Strain-specific susceptibility data were not previously available for piperacillin/tazobactam Piperacillin/tazobactam activity against other strains of <i>L. plantarum</i> have not been described previously
Lincosamide	Clindamycin	≤ 0.12	≤ 0.12	S <sub>a,b,c</sub>	Strain-specific susceptibility data were not previously available MIC value was within the range of values independently reported by two different groups for 46 (0.032–1 mg/L) <sup>22</sup> and 10 (0.03–32 µg/ml) <sup>23</sup> other <i>L. plantarum</i> strains
Oxazolidinone	Linezolid	4	4	S <sub>a</sub> _bc	Strain-specific susceptibility data were not previously available MIC value was within the range of values previously reported for 10 other strains of <i>L. plantarum</i> (2–8 µg/ml) <sup>23</sup> and within one doubling dilution of the upper end of the range observed for 46 other <i>L. plantarum</i> strains (1–2 mg/l) <sup>22</sup>
Continued					

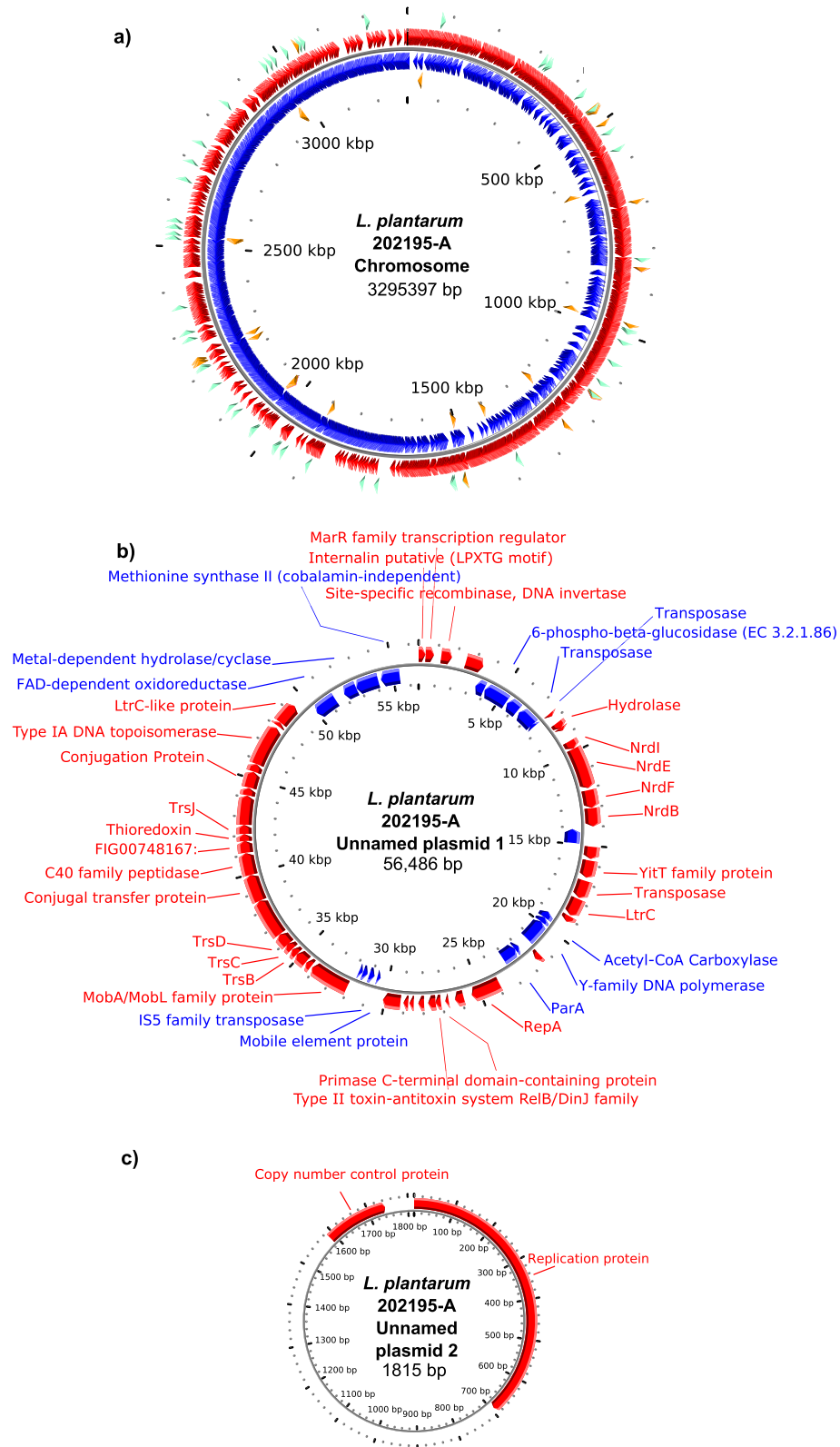
Antibiotic class	Antimicrobial	Broth dilution MICs and interpretations			Prior Evidence from <i>L. plantarum</i> ATCC 202195, other strains of <i>L. plantarum</i> , and other types of lactobacilli
		<i>L. plantarum</i> ATCC 201295-A	<i>L. plantarum</i> ATCC 202195-B	Interpretation	
		MIC (µg/ml)	MIC (µg/ml)		
Macrolide	Erythromycin	≤ 0.25	≤ 0.25	S <sup>a,b</sup> IE <sup>c</sup>	Strain-specific susceptibility data were not previously available MIC value was within the range of values independently reported by two different groups for 10 (0.25–16 µg/ml) <sup>23</sup> and 46 (0.016–0.5 mg/l) <sup>22</sup> other strains of <i>L. plantarum</i>
Glycopeptide	Vancomycin	≥ 256	≥ 256	R <sup>a,c</sup> _b	Interpretation consistent with previous disc diffusion findings for the <i>L. plantarum</i> ATCC 202195 strain <sup>5</sup>
Chloramphenicol	Chloramphenicol	8	8	S <sup>b,c</sup> _a	Strain-specific susceptibility data were not previously available MIC was within the range of values previously reported for 46 other strains of <i>L. plantarum</i> (2–8 mg/l) <sup>22</sup> and 10 other strains of <i>L. plantarum</i> (4–16 µg/ml) <sup>23</sup>
Fluoroquinolone	Levofloxacin	≥ 8	≥ 8	_a,b,c	Strain-specific susceptibility data were not previously available Levofloxacin activity against other strains of <i>L. plantarum</i> has not previously been reported
	Ciprofloxacin	≥ 4	≥ 4	_a,b,c	Using disc diffusion assays, it was previously reported that <i>L. plantarum</i> ATCC 202195 is ciprofloxacin-resistant <sup>5</sup> Clinical breakpoints were not available to interpret MICs for ciprofloxacin and <i>L. plantarum</i> Among ten <i>L. plantarum</i> type strains previously tested for susceptibility to ciprofloxacin, the MIC values ranged between 16 and 256 µg/ml <sup>23</sup> . The range of ciprofloxacin concentrations tested by microbroth dilution only went to as high as 2 µg/ml. By contrast, the range of ciprofloxacin concentrations evaluated by E-testing was 0.002–32 µg/ml. The true MIC value for <i>L. plantarum</i> ATCC 202195 and ciprofloxacin may be closer to the value described by E-testing (i.e., ≥ 32 µg/ml)
	Gatifloxacin	2	2	_a,b,c	Strain-specific susceptibility data were not previously available Gatifloxacin activity against other strains of <i>L. plantarum</i> has not previously been reported
Aminoglycoside	Gentamicin	≤ 2	≤ 2	S <sup>b</sup> _a,c	Interpretation is consistent with previous disc diffusion findings for <i>L. plantarum</i> ATCC 202195 <sup>5</sup> MIC value fell within the range published independently by two different research groups for 46 (≤ 1–8 mg/l) <sup>22</sup> and 10 other <i>L. plantarum</i> (0.5–512 µg/ml) <sup>23</sup> isolates
Tetracycline	Tetracycline	≥ 32	≥ 32	R <sup>b</sup> _a,c	Strain-specific susceptibility data were not previously available for tetracycline MIC value within the range of values previously reported for 10 other strains of <i>L. plantarum</i> (4 and 64 µg/ml) <sup>23</sup>
Continued					

Antibiotic class	Antimicrobial	Broth dilution MICs and interpretations			Prior Evidence from <i>L. plantarum</i> ATCC 202195, other strains of <i>L. plantarum</i> , and other types of lactobacilli
		<i>L. plantarum</i> ATCC 201295-A	<i>L. plantarum</i> ATCC 202195-B	Interpretation	
		MIC (µg/ml)	MIC (µg/ml)		
Macrolide-lincosamide-streptogramin	Quinu/Dalfopristin	2	2	– <sub>a,b,c</sub>	Strain-specific susceptibility data were not previously available MIC of quinupristin/dalfopristin for <i>L. plantarum</i> ATCC 202195 was within the range reported for 10 other <i>L. plantarum</i> strains (1–8 µg/ml) <sup>23</sup> and within one doubling dilution of the upper end of the MIC range reported for 46 other <i>L. plantarum</i> isolates (0.063–1 mg/L) <sup>22</sup>
Lipopeptide	Daptomycin	≤0.25	≤0.25	S <sup>a</sup> – <sub>b,c</sub>	Strain-specific susceptibility data were not previously available Daptomycin activity against other strains of <i>L. plantarum</i> have not been described previously
Rifamycin	Rifampin	≥8	≥8	– <sub>a,b,c</sub>	Strain-specific susceptibility data were not previously available MIC of rifampin was within one doubling dilution of the upper end of the MIC range reported for 10 other <i>L. plantarum</i> strains (0.12–4 µg/ml) <sup>23</sup>
Sulfonamide	Trimethoprim/Sulfamethoxazole	≤0.5/9.5	≤0.5/9.5	– <sub>a,b,c</sub>	Strain-specific susceptibility data were not previously available MIC of trimethoprim/sulfamethoxazole was within the range of MIC values previously reported for 46 other strains of <i>L. plantarum</i> (0.5–>512 mg/l) <sup>22</sup>

**Table 1.** Antimicrobial susceptibility and minimum inhibitory concentrations of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B. Clinical breakpoint interpretations were categorized as susceptible ("S"), intermediate ("I"), resistant ("R"), no interpretive criteria available ("–"), and insufficient evidence ("IE"). <sup>a</sup>Interpretation based on CLSI, <sup>b</sup>EFSA, and on <sup>c</sup>EUCAST.

the CLSI<sup>19</sup>, EUCAST<sup>20</sup>, or EFSA<sup>21</sup>. Where clinical breakpoints were available, susceptibility interpretations were identical for all antimicrobials assessed by both broth dilution and E-testing, except for tetracycline (Supplementary Table 1). The MIC for tetracycline as determined by broth dilution (≥ 32 µg/ml) was four times higher than the value determined by E-testing (8 µg/ml) for *L. plantarum* ATCC 202195-A. In microbroth dilution assays, the concentration of tetracycline only went up to 16 µg/ml (MIC reported as equal to or greater than two-times the antibiotic concentration in the well where growth occurred). The range of tetracycline concentrations evaluated by E-testing was 0.016–256.0 µg/ml and growth inhibition intersected the side of the E-strip at 8 µg/ml (Supplementary Table 1). Thus, the true MIC value for *L. plantarum* ATCC 202195 and tetracycline may lie in between the values described by microbroth dilution assays and E-testing. MICs determined by E-tests were identical (penicillin, erythromycin and vancomycin) or within one doubling dilution (ceftriaxone (MIC = 1 µg/ml), chloramphenicol (4 µg/ml) and gentamicin (1 µg/ml)) to those determined by broth dilution assays for six of the eight antimicrobials assayed (Table 1). The MIC for ciprofloxacin (≥ 32 µg/ml) as determined by E-testing was 8 times higher than the value determined using the broth dilution assay, which likely reflects the antibiotic concentration range tested in each assay. *L. plantarum* ATCC 202195-B was negative for beta-lactamase activity.

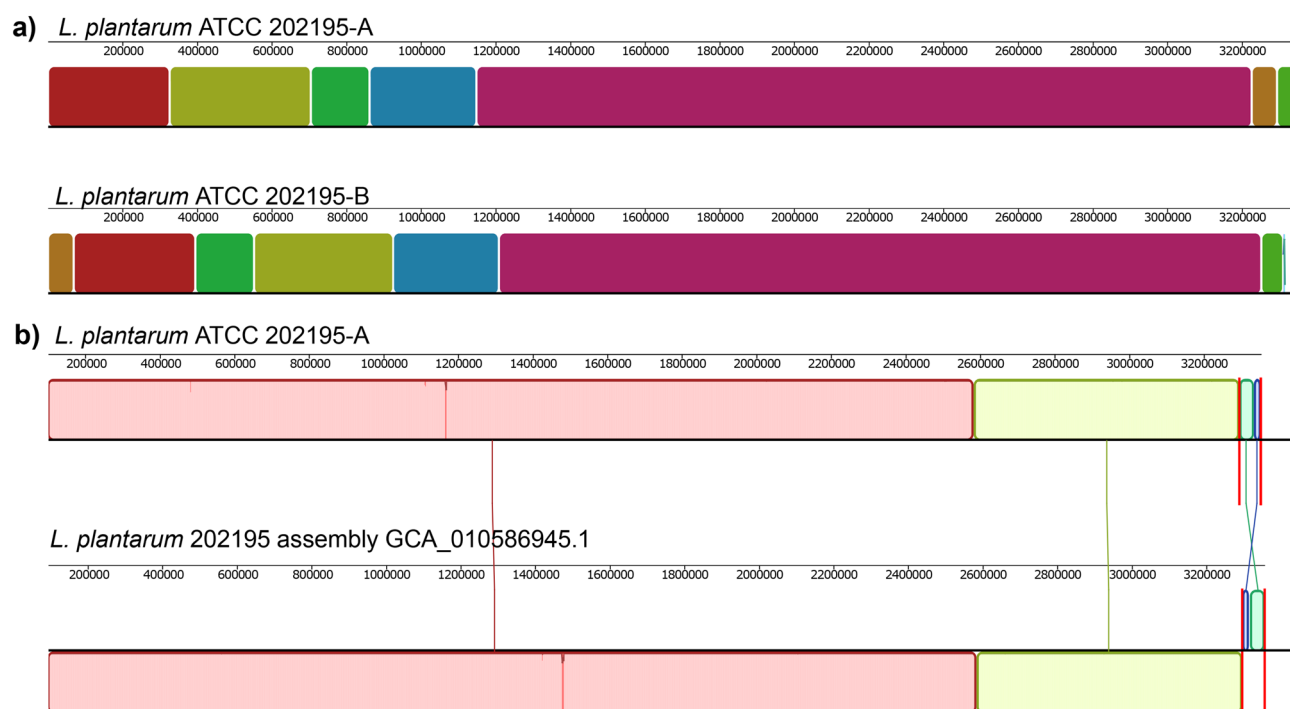
**Comparisons of the *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B genomes.** To corroborate our phenotypic data, the genomes of both isolates were sequenced and compared. The genomic sequence for isolate *L. plantarum* ATCC 202195-A was generated by International Flavors & Fragrances Inc (IFF) (formerly Dupont Nutrition & Biosciences) from DNA extracted from pure monoculture purchased from ATCC in 2017. The resulting sequencing reads assembled into a complete genome containing one chromosome (3,295,397 bp) (Fig. 1a) and two plasmids; unnamed plasmid 1 (56,486 bp) (Fig. 1b) and unnamed plasmid 2 (1815 bp) (Fig. 1b), respectively, with a total genome length of 3,353,698 base pairs (G + C content = 44.3%). The genomic sequences for *L. plantarum* ATCC 202195-B were generated using an ATCC isolate purchased in 2019 by The Hospital for Sick Children, (Toronto, ON) and reads were assembled into a draft genome with a total of 146 contigs, with 66 contigs > 150 bp, and a total genome length of 3,332,603 base pairs. There was a difference in genome length of 21,098 base pairs relative to *L. plantarum* ATCC 202195-A (G + C content = 44.3%). Two plasmids were also present in the draft genome of *L. plantarum* ATCC 202195-B. All of the sequencing reads from *L. plantarum* ATCC 202195-B mapped to the genome of *L. plantarum* ATCC 202095-



**Figure 1.** Circular genomic map of the *L. plantarum* ATCC 2021295-A. **(a)** The chromosome is 3,295,397 base pairs in size and has a G + C content of 44%. **(b)** Unnamed plasmid 1 (56,486 base pairs, G + C content = 40%); and, **(c)** Unnamed plasmid 2 (1815 base pairs, G + C content = 37.4%). Images were generated using CGViewer<sup>24</sup> and gene annotation was prepared using RASTtk<sup>25</sup>. Forward strand genes are denoted as red arrows, reverse strand genes are denoted as blue arrows, RNA genes are denoted as orange arrows, and repeat regions are denoted as aqua coloured arrows. The original image was generated using GC Viewer (<https://paulstothard.github.io/cgviewer/>) and the text was modified in Inkscape v1.0.1 (<https://inkscape.org/release/inkscape-1.0.1/>).

Nucleotide position	Variation (LP202195-A/LP202195-B)	Quality score <sup>a</sup>	Sequence depth at variant site	Strand allele probability (SAP) <sup>b</sup>	Strand reference probability (SRP) <sup>c</sup>	Putative function	SNP type
2978657	CAATG/CATG	217.678	12	29.068	0	Transposase, IS4 family	Deletion
3220411	G/A	158.242	102	31.2394	174.557	Surface protein (LPXTG motif)	Nonsynonymous
3219413	G/T	118.453	667	825.998	376.504	Surface protein (LPXTG motif)	Nonsynonymous

**Table 2.** Summary of single nucleotide variants between *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B. <sup>a</sup>Based on an estimated probability that the alternate allele is present at the loci, as calculated using Freebayes<sup>26</sup>. <sup>b</sup>Strand balance probability for the alternate allele, as calculated using Freebayes<sup>26</sup>. <sup>c</sup>Strand balance probability for the reference genome, as calculated using Freebayes<sup>26</sup>.

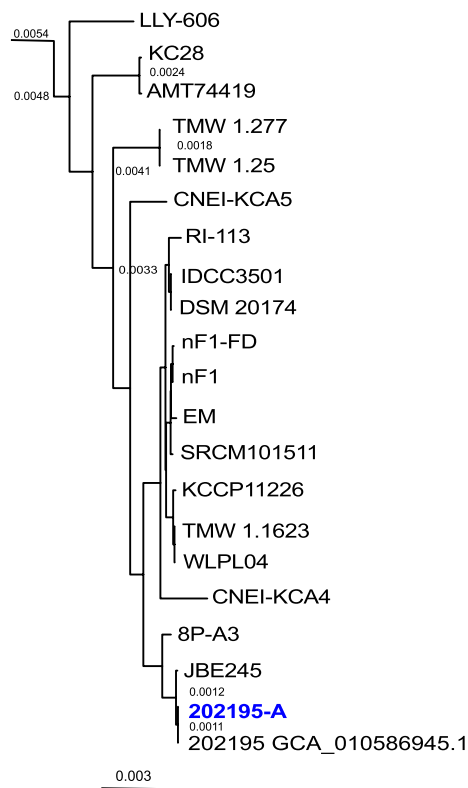


**Figure 2.** Comparison of sequenced isolates of *L. plantarum* ATCC 202195 (a) Graphical representation of a global sequence alignment of the co-linear blocks (identified using progressiveMAUVE<sup>27</sup>) conserved between the *L. plantarum* ATCC 2021295-A and *L. plantarum* ATCC 2021295-B genomes. (b) Mauve alignment of *L. plantarum* ATCC 202195-A and *Lactiplantibacillus plantarum* ATCC 202195 assembly GCA\_010586945.1. Matching coloured blocks correspond to co-linear blocks shared between the sequences. The original image was generated by Mauve software (<http://darlinglab.org/mauve/user-guide/mauvealigner.html>) and edited in Inkscape v1.0.1 (<https://inkscape.org/release/inkscape-1.0.1/>).

A, with coverage of greater than 1000 $\times$ , indicating that this discrepancy in genome length is likely a result of variation in sequence methodology and genome assemblers.

Comparison between the two bacterial genomes for single nucleotide polymorphisms (SNPs) identified just three high-quality variants (Table 2). Two of the three variants were localized within one gene encoding a putative surface-layer protein (LPXTG motif), one of which was a missense mutation and the other was a stop-gain mutation. The remaining SNP was located within a transposase, IS4 family gene and was classified as a deletion variant (Table 2).

The average nucleotide identity (ANI) between the two genomes was 99.99% and global alignment identified 10 locally colinear blocks with a minimum weight of 109 (Fig. 2a). Comparing the genome of *L. plantarum* ATCC 202195-A with a previously released complete genome assembly for *L. plantarum* strain ATCC 202195 (accession number: GCA\_010586945.1) resulted in an ANI of 99.99% and global alignment identified four locally colinear blocks with a minimum weight of 19,941 (Fig. 2b). In a comparison of our complete genome of *L. plantarum* ATCC 202195-A with the partial draft *L. plantarum* ATCC 202195 genome released by Wright et al.<sup>13</sup> (accession number: GCA\_004354995.1), we found an ANI of 99.98%. Although Wright et al. did not report any associated plasmid sequences, we found 100% sequence homology between both plasmids described here with contigs 10 and contigs 16 of the GCA\_004354995.1 genome assembly. *L. plantarum* ATCC 202195



**Figure 3.** Phylogenetic tree of the clade containing *L. plantarum* ATCC 202195-A. A phylogenetic tree was generated using only the core genome alignment that was derived from the pan-genome analysis of 133 *L. plantarum* strains. *L. plantarum* ATCC 202195 is shown in blue and *L. plantarum* ATCC 202195 assembly GCA\_010586945.1 is labelled as '202195 GCA\_010586945.1'. Branch lengths equal to or greater than 0.001 are shown. The complete phylogenetic tree of all 133 *L. plantarum* strains can be found in Supplementary Fig. 1. The phylogenetic tree was generated by Figtree (<http://tree.bio.ed.ac.uk/stats.html>) and edited in Inkscape v1.0.1 (<https://inkscape.org/release/inkscape-1.0.1/>).

complete genome assembly GCA\_010586945.1 was identified as the closest genetic relative to *L. plantarum* ATCC 202195-A (Fig. 3), followed by *L. plantarum* strain JBE245. *L. plantarum* ATCC 202195 genome assembly GCA\_010586945.1 lacked sequence homology to unnamed plasmid 2 that was identified in both *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B.

**Antimicrobial resistance and virulence genes.** As genomic comparison revealed near identical genome sequences, screening for AMR and virulence genes was limited to the most complete genome (*L. plantarum* ATCC 202195-A). Initial high stringency screening of AMR databases did not identify potential antimicrobial resistance genes. By reducing the screening threshold stringency, three partial matches to the AMR genes *LmrD*, *LmrC* and *rpoB* from CARD<sup>28</sup> and 12 partial matches to known virulence factors in VFDB<sup>29</sup> were identified (Table 3). Based on observed phenotypic resistance patterns, additional targeted screening was performed against known resistance genes for vancomycin (*ddl*), tetracycline [*tet(M)*, *tet(S)*, *tet(W)*, *tet(O)*, and *tet(Q)*], and ciprofloxacin (*gyrA*) (Table 3). *L. plantarum* ATCC 202195-A genome was found to contain an intrinsic active site mutation (F261Y) in the *ddl* gene that confers resistance to vancomycin. A single partial match to a tetracycline resistance gene, *tetM*, was found, and while *gyrA* gene was found to be present in *L. plantarum* ATCC 202195-A genome, no mutations responsible for conferring fluoroquinolone resistance were identified.

**Genome features of *L. plantarum* ATCC 202195-A.** Annotation of the genome, including two putative plasmids, using the RASTtk pipeline revealed a total of 3286 coding sequences, including 2287 and 999 sequences assigned as functional and hypothetical proteins, respectively. A total of 72 transfer RNA genes and 16 ribosomal RNA genes were identified. Unnamed plasmid 1 likely represents a conjugative plasmid with a total of 56,486 base pairs (Fig. 1b) that encode 64 genes, 50 of which have assigned putative functions. Unnamed plasmid 2, a non-conjugative plasmid (Fig. 1c), contains just 2 genes: one gene encodes a putative replication protein and the other codes for a copy number control protein. Evaluating the assembled plasmid sequences for sequence homology revealed an almost exact match of unnamed plasmid 2 with that of *Pediococcus clausenii* ATCC BAA-344 plasmid pPECL-1<sup>32</sup> (sequence identity of 99%, and 100% coverage by BLASTn). Unnamed plasmid 1 has high homology to the unnamed plasmid (60,765 base pair in size) identified in the previously released *L. plantarum* ATCC 202195 genome assembly GCA\_010586945.1 (92% query coverage and 100% identity).



Class	Gene	Product	<i>L. plantarum</i> ATCC 202195-A	
			Coverage (%)	Identity (%)
Virulence genes <sup>a</sup>	<i>cpsI</i>	UDP-galactopyranose mutase	92.81	67.28
	<i>clpC</i>	endopeptidase Clp ATP-binding chain C	27.28	69.19
	<i>clpP</i>	ATP-dependent Clp protease proteolytic subunit	93.47	70.54
	<i>hasC</i>	UDP-glucose pyrophosphorylase	87.65	71.81
	<i>htpB</i>	Hsp60 60 K heat shock protein HtpB	90.74	64.8
	<i>cpsJ</i>	ABC transporter ATP-binding protein	41.15	66.67
	<i>lap</i>	Listeria adhesion protein Lap	48.83	64.35
	<i>bsh</i>	bile salt hydrolase	89.98	71.9
	<i>cpsA</i>	undecaprenyl diphosphate synthase	33.33	74.45
	<i>clfA</i>	Clumping factor A fibrinogen-binding protein	31.26	77.7
	<i>clpE</i>	ATP-dependent protease	65.7	66.67
	<i>cap8P</i>	capsular polysaccharide synthesis enzyme	25	75.68
	Antimicrobial resistance genes	<i>lmrD<sup>b</sup></i>	Multidrug transporter subunit	31.07
<i>rpoB<sup>b</sup></i>		beta subunit of RNA polymerase	15.42	67.92
<i>lmrC<sup>b</sup></i>		Multidrug transporter subunit	41	67.57
<i>tetM<sup>c</sup></i>		Tetracycline resistance protein	95	30.76

**Table 3.** Antimicrobial resistance genes and putative virulence genes present in the genomes of *L. plantarum* ATCC 2021295-A and *L. plantarum* ATCC 202195-B. <sup>a</sup>Virulence genes were identified by comparing DNA sequences from both isolates to the Virulence Factor Database (VFDB)<sup>29</sup>. <sup>b</sup>Antimicrobial resistance gene were identified using ABRicate<sup>30</sup> and The Comprehensive Antibiotic Resistance Database (CARD)<sup>28</sup>, and initial results were verified by BLASTn<sup>31</sup>. <sup>c</sup>*L. plantarum* ATCC 202195 gene loci G9282\_00305 compared to KRM38573.1 *L. aviarius* subsp. *aviarius* DSM 20655.

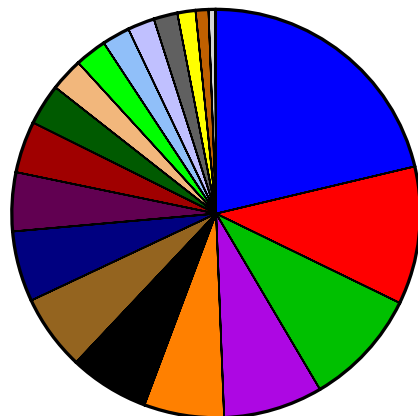
Neither plasmid contains genes that code for identified AMR or a virulence factor. Notably, unnamed plasmid 1 encodes several genes that upon transfer would be advantageous to the recipient, including a manganese transport protein, MntH, magnesium and cobalt transport protein, CorA, 6-phospho-beta-glucosidase (EC 3.2.1.86), BglA-2, adenosylhomocysteinase (EC 3.3.1.1) and methionine synthase II (cobalamin-independent).

**Functional annotation of *L. plantarum* ATCC 202195-A genome.** Using the eggNOG database<sup>33</sup>, 2860 genes were assigned to at least one of the cluster orthologous group families comprising 20 functional groups. Of those 2860 genes, 10.9% were assigned to the transcription functional group, followed by 9.3% of genes assigned to carbohydrate transport and metabolism, and 7.8% of genes assigned to amino acid transport and metabolism (Fig. 4). Further functional characterization by KEGG Pathways analysis, revealed a similar functional distribution of genes, with the majority of identified pathways attributed to nucleotide, carbohydrate and amino acid metabolism. A PATRIC subsystem analysis revealed that 32% of identified genes from *L. plantarum* ATCC 202195-A mapped to 10 of the 11 possible superclasses, including metabolism (38.4%), protein processing (14.9%), and stress response, defence and virulence (9.3%).

**Stress response and adhesion.** Based on subsystem analyses, a total of 33 unique genes were sub-categorized as stress response (Table 4). The *L. plantarum* ATCC 202195-A genome encodes genes involved in acid tolerance including four bile salt acid hydrolases (*bsh1*, *bsh2*, *bsh3* and *bsh4*), and eight sodium-proton antiport genes. A total of 16 genes were subclassified as stress response genes, including nine genes that encode putative universal stress response proteins and seven genes that play a role in the oxidative stress response including glutathione peroxidase, NADH peroxidase and catalase. Several genes responsible for responses to heat and cold stress were also identified, including a number of putative protein-folding chaperones within the well characterized Clp protein family (*clpC*, *clpP*, *clpL*, *clpX*, *clpB* and *clpE*), and four homologous members of the cold shock protein family (*cspP*, *cspL*, *cspR* and *cspC*). Two of the three small heat shock proteins (sHSP) genes encoded by the *L. plantarum* reference strain WCFS1<sup>34</sup> were also identified in *L. plantarum* ATCC 202195-A.

The *L. plantarum* ATCC 202195-A genome also contains several genes that facilitate cellular adhesion, including two genes that encode fibronectin binding proteins, one gene that encodes a mucin binding protein, two mucus adhesion promoting protein genes (*mapA*), two enolases, two surface layer LPXTG anchored proteins and four putative LPXTG internallins.

**Carbohydrate processing and utilization of fructo-oligosaccharides (FOS).** Of the 266 genes encoded by *L. plantarum* ATCC 202195-A that were identified as being involved in carbohydrate metabolism according to COG annotation, only 94 of these genes were annotated as carbohydrate-active enzymes by the CAZy database<sup>36</sup>. Specifically, *L. plantarum* ATCC 202195-A was found to encode 52 glycoside hydrolases, 35 glycosyltransferases, three carbohydrate binding modules, two carbohydrate esterases and two auxiliary activity enzymes, indicating a strong metabolic capability to degrade and process complex carbohydrates. Embedded within the pool of carbohydrate processing genes was a conserved pts1BCA operon<sup>37</sup>, which is responsible for



### Clusters of Orthologous Groups

*L. plantarum* ATCC 202195-A

- Transcription (10.9%)
- Carbohydrate metabolism and transport (9.3%)
- Amino acid metabolism and transport (7.9%)
- Replication and repair (6.3%)
- Cell wall/membrane/envelope biogenesis (6.3%)
- Translation (6.0%)
- Inorganic ion transport and metabolism (5.6%)
- Nucleotide metabolism and transport (4.7%)
- Energy production and conversion (4.1%)
- Coenzyme metabolism (3.3%)
- Intracellular trafficking and secretion (2.6%)
- Signal transduction (2.5%)
- Lipid metabolism (2.2%)
- Defense mechanisms (2.1%)
- Post-translational modification, protein turnover, chaperone functions (1.9%)
- Cell cycle control and mitosis (1.4%)
- Secondary structure (1.0%)
- Cell motility (0.5%)
- Extracellular structures (0.1%)
- Function unknown (21.3%)

**Figure 4.** Functional categorization of 2860 predicted open reading frames (ORFs) in the genome of *L. plantarum* ATCC 202195-A based on Clusters of Orthologous Groups of proteins (COGs). Image was generated in Graphpad 9.

Sub-classification	Subsystem name	Gene count	Role count
Stress response	Glutathione: Redox cycle	2	2
	Universal stress protein family	9	1
	Cluster containing Glutathione synthetase	2	2
	Hydroxy-fatty acid production as stress response	1	1
	Glutathione: Biosynthesis and gamma-glutamyl cycle	1	1
	Protection from Reactive Oxygen Species	1	1
Stress response: heat/cold shock	Cold shock proteins of CSP family	4	1
	Heat shock dnaK gene cluster extended	14	12
Stress response: osmotic stress	Osmoregulation	6	2
	Choline uptake and conversion to betaine clusters	7	8
Undefined	Hfl operon	2	2

**Table 4.** Stress response and defense features in the genome of *L. plantarum* ATCC 202195-A. Subsystems analysis of the superclasses Stress Response, Defense, and Virulence, was performed using The Pathosystems Resource Integration Center (PATRIC)<sup>35</sup>, and Rapid Annotations Using Subsystems Technology (RASTtk)<sup>25</sup> for annotations.

B vitamins	Gene symbol	Gene product	
Folate (B9)	–	Substrate-specific component FolT of folate ECF transporter	
	–	Substrate-specific component FolT of folate ECF transporter	
	<i>folB</i>	Dihydroneopterin aldolase (EC 4.1.2.25)	
	<i>folK</i>	2-Amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (EC 2.7.6.3)	
	<i>folE</i>	GTP cyclohydrolase I (EC 3.5.4.16) type 1	
	<i>folC2</i>	Dihydrofolate synthase (EC 6.3.2.12) at Folylpolyglutamate synthase (EC 6.3.2.17)	
	<i>folP</i>	Dihydropteroate synthase (EC 2.5.1.15)	
	–	Dihydrofolate reductase (EC 1.5.1.3)	
	–	Dihydrofolate reductase (EC 1.5.1.3)	
	<i>ecfT</i>	Transmembrane component of general energizing module of ECF transporters	
	<i>ecfA2</i>	ATPase component of general energizing module of ECF transporters	
	<i>ecfA</i>	ATPase component of general energizing module of ECF transporters	
	<i>tdk</i>	Thymidine kinase (EC 2.7.1.21)	
	<i>glyA</i>	Serine hydroxymethyltransferase (EC 2.1.2.1)	
	<i>folC</i>	Dihydrofolate synthase (EC 6.3.2.12) at Folylpolyglutamate synthase (EC 6.3.2.17)	
	<i>thyA</i>	Thymidylate synthase (EC 2.1.1.45)	
	<i>folA</i>	Dihydrofolate reductase (EC 1.5.1.3)	
	<i>trmFO</i>	Methylenetetrahydrofolate-tRNA-(uracil-5-)-methyltransferase TrmFO (EC 2.1.1.74)	
	<i>fhs</i>	Formate-tetrahydrofolate ligase (EC 6.3.4.3)	
	<i>fnt</i>	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	
	<i>folD</i>	Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)/Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	
	<i>fthC</i>	5-Formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)	
	<i>metK</i>	S-Adenosylmethionine synthetase (EC 2.5.1.6)	
	–	Substrate-specific component FolT of folate ECF transporter	
	Riboflavin (B2)	<i>ribC1</i>	FMN adenylyltransferase (EC 2.7.7.2)/Riboflavin kinase (EC 2.7.1.26)
		<i>ribF</i>	MN adenylyltransferase (EC 2.7.7.2)/Riboflavin kinase (EC 2.7.1.26)
<i>ribA</i>		3,4-Dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12) / GTP cyclohydrolase II (EC 3.5.4.25)	
<i>ribG</i>		Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26)/5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193)	
<i>ribD</i>		Putative FMN hydrolase (EC 3.1.3.-); 5-Amino-6-(5'-phosphoribitylamino)uracil phosphatase	
<i>ribT</i>		COG0454 Histone acetyltransferase HPA2 and related acetyltransferases	
<i>ribH</i>		6,7-Dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)	
<i>npp</i>		Alkaline phosphodiesterase I (EC 3.1.4.1)/Nucleotide pyrophosphatase (EC 3.6.1.9)	
<i>ribE</i>		Riboflavin synthase (EC 2.5.1.9)	

**Table 5.** List of putative genes in the *L. plantarum* ATCC 202195-A genome involved in B vitamin biosynthesis.

the import of short-chain FOS (scFOS) into the cytosol. The *L. plantarum* ATCC 202195-A genome also contains three gene clusters critical to the production of the short-chain fatty acid butyrate from acetyl-CoA.

**Biosynthesis of B complex vitamins.** Evaluating the *L. plantarum* ATCC 202195-A genome for genes involved in the biosynthesis of B vitamins, revealed a complete gene cluster (*folA*, *folB*, *folC1*, *folC2*, *folD*, *folE*, *folK*, *folP* and *folQ*) that is involved in folate (vitamin B9) biosynthesis and utilization (Table 5). In contrast to *L. plantarum* strain WCFS1<sup>38</sup>, the *L. plantarum* ATCC 202195-A genome contains a complete riboflavin operon, including the *ribA*, *ribB*, *ribH*, *ribE* and *ribG* genes, required for riboflavin biosynthesis (Table 5). *L. plantarum* ATCC 202195-A also encodes genes involved in thiamine and biotin utilization and salvage; however, based on the genome sequences, the microbe appears incapable of de novo synthesis of either of these vitamins.

**Bacteriocins.** The *L. plantarum* ATCC 202195-A genome encodes three plantaricin specific operons, including the regulator operon *plnABCD*, the bacteriocin operon *plnEFI* and the transport operon *plnGHTUVW*, which are required to produce the class IIb bacteriocins, plantaricin plnEF and plnA. Bacteriocins are a heterogeneous group of bioactive bacterial peptides that act as antimicrobial agents against closely related susceptible bacterial species<sup>39</sup>. Similar to the reference strain *L. plantarum* WCFS1<sup>38</sup>, *L. plantarum* ATCC 202195-A also encodes the plantaricin immunity protein *plnL*, directly upstream from the regulatory operon.

## Discussion

Prior to the widespread use of a probiotic agent in humans, it is imperative to delineate microbial susceptibility to antimicrobial agents, strain level identification and establish if antimicrobial resistance is intrinsic or has the potential to be transferred to other microorganisms<sup>40</sup>. Here, we present the first comparative genome analysis of two isolates of *L. plantarum* ATCC 202195, which were procured from separate ATCC deposits that occurred approximately 20 years apart. Our analysis revealed previously unreported features of the *L. plantarum* ATCC 202195 genome including two putative plasmid sequences. In addition, we have provided the most complete antimicrobial susceptibility characterization of this clinically important strain to date.

To be considered safe for human consumption, it has been suggested that probiotics should be susceptible to at least two major classes of currently available antibiotics<sup>41</sup>. While transferable resistance is uncommon among lactic acid-producing bacteria, acquired antibiotic resistance has been identified in isolates considered for probiotic or nutritional uses<sup>22</sup>. Herein, we evaluated the genomic sequence of *L. plantarum* ATCC 202195 for the presence of AMR genes and tested the susceptibility of each isolate against a panel of 20 antimicrobial agents across 16 classes of antibiotics and corroborated these findings with the genome sequence data. The two isolates had identical in vitro antimicrobial susceptibility profiles, were both found to be sensitive to eight of the newly tested antimicrobials (Table 1), and with two exceptions (rifampin and penicillin), the observed MICs fell within the range of values previously reported for other strains of *L. plantarum*<sup>22,23</sup>. MICs observed for rifampin and penicillin were both within one doubling dilution of the upper range of MICs previously reported for other *L. plantarum* isolates<sup>23</sup>.

Screening the genome of both *L. plantarum* ATCC 202195 isolates for sequence homology to AMR genes revealed only partial hits, with low sequence identity and coverage, to the multidrug efflux heterodimers *LmrCD* and *rpoB* ( $\beta$ -subunit of RNA polymerase) for which a mutation is known to confer resistance to rifampin<sup>22,23</sup>. In vitro susceptibility testing for rifampin using a broth dilution assay demonstrated growth at the highest rifampin concentration tested, which may be supportive of rifampin resistance; however, we were unable to make a conclusive determination in this regard due to the lack of established clinical breakpoints. Additionally, a partial match to a tetracycline resistance gene *tetM*, was identified which may be responsible for the phenotypic resistance observed in this strain. Both isolates were resistant to vancomycin, consistent with previously published findings for this strain generated by disc diffusion assay<sup>42</sup>. The signature active site mutation in the *ddl* gene, which confers resistance to vancomycin, was conserved within the *L. plantarum* ATCC 202195 genome, corroborating our phenotypic findings. While a clinical breakpoint interpretation was not available for ciprofloxacin activity against lactobacilli, previous disc diffusion assays with *L. plantarum* ATCC 202195 suggested that the strain is resistant to ciprofloxacin<sup>5</sup>. Herein, we report ciprofloxacin MICs suggestive of resistance for *L. plantarum* ATCC 202195 using the broth dilution and E-testing assays, respectively. Notably, while the target *gyrA* gene was present in the *L. plantarum* ATCC 202195 genome, no fluoroquinolone resistant mutations were identified in this gene and thus, ciprofloxacin resistance may likely be attributed to efflux mechanisms. Both isolates tested in the present study were sensitive to penicillin (MIC = 4  $\mu$ g/ml) and gentamicin (MIC  $\leq$  2  $\mu$ g/ml), consistent with previous findings for *L. plantarum* ATCC 202195 obtained using disc diffusion<sup>5</sup>. Since the World Health Organization currently recommends the use of ampicillin plus gentamicin for the initial empiric treatment of neonates with suspected sepsis<sup>43</sup>, evidence that *L. plantarum* ATCC 202195 is susceptible to these agents further supports the safety of this probiotic strain in the unlikely event that an infant administered *L. plantarum* ATCC 202195 were to develop probiotic-related bacteremia or infection.

Verifying the genetic lineage of these two isolates is of critical importance to any future clinical work using this probiotic strain. Herein, our comparative genome analyses revealed that *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B are identical with only three high-quality SNPs identified between the two genomes, and an ANI of 99.99%. Relative to other strains of *L. plantarum*, *L. plantarum* ATCC 202195 forms a distinct branch on the phylogenetic tree and its closest phylogenetic neighbours were *L. plantarum* ATCC 202195 (GCA\_010586945.1) and *L. plantarum* JBE245, a food-related bacterium<sup>44</sup>. In contrast to previous reports, which indicated that *L. plantarum* ATCC 202195 did not contain a plasmid<sup>5,13</sup>, we identified two putative plasmids in both isolates of *L. plantarum* ATCC 202195; however, these plasmids were only assembled into complete sequences in the *L. plantarum* ATCC 202195-A genome.

Plasmid sequence assembly and identification is challenging due to the presence of repeat sequences and the use of short reads for genome sequencing<sup>45</sup>. Herein, we resolved plasmid sequences by combining whole genome sequencing data with sequences that were generated using DNA obtained from a plasmid extraction. Often, plasmid sequences identified in *L. plantarum* strains do not encode any genes that functionally impact the host<sup>46</sup>. However, we annotated several plasmid-related genes that encode potentially impactful biological properties, including metal transport, amino acid synthesis and carbohydrate processing. Specifically, BglA-2 (6-phospho-beta-glucosidase) encoded on unnamed plasmid 1 is responsible for catalyzing the conversion of cellbiose-6P to glucose and glucose-6P, which then can enter the glycolic pathway to generate further energy for the cell<sup>47</sup>. Transfer of this gene through a conjugative plasmid could potentially provide an adaptive advantage to the plasmid recipient microbe<sup>47</sup>.

Evaluating the functional attributes of a probiotic strain by annotation of the complete genome sequence sheds light onto mechanistic underpinnings of documented clinical outcomes. Previous work has found that oral administration of *L. plantarum* ATCC 202195 leads to sustained colonization<sup>5</sup>. Here we report strain specific functional gene annotation related to clinical phenotypic findings. Notably, the *L. plantarum* ATCC 202195 genome was found to contain an array of genes involved in environmental stress responses, including genes related to cellular adhesion, bile acid tolerance, and responses to heat and cold stress, as well as oxidative stress. Since bile acids have known antimicrobial properties<sup>48</sup>, it is advantageous to the bacterium that *L. plantarum* ATCC 202195 encodes several bile acid hydrolase genes, the products of which can metabolize conjugated bile

salts<sup>49</sup>. The *L. plantarum* ATCC 202195 genome also contains the heat shock protein genes: *clpE*, *clpX* and *clpP*. The production of the ATPases ClpE and ClpX and the ClpP protease increase in response to heat-shock and other microenvironmental stresses<sup>50</sup>, and play an important role in maintaining protein quality by regulating proteolysis<sup>51</sup>.

Microbial production of secondary metabolites is a potential mechanism by which probiotics can improve host health. Microbial production of the short-chain fatty acid butyrate has been linked to enhanced intestinal epithelial barrier function<sup>52</sup>, modulation of inflammatory status<sup>52,53</sup>, regulation of colonic T-cell differentiation<sup>54</sup> and overall homeostasis in the intestinal tract. Multiple studies have linked the production of short-chain fatty acids to the utilization of prebiotics, including fructo-oligosaccharides and galacto-oligosaccharides<sup>55</sup>. The *L. plantarum* ATCC 202195 genome encodes three clusters of genes involved in butyrate production, as well as an operon for fructo-oligosaccharide metabolism. In other strains of *L. plantarum*, butyrate production is upregulated in the presence of fructo-oligosaccharides<sup>56,57</sup>. A potential relationship between the metabolism of fructo-oligosaccharides and the production of butyric acid by *L. plantarum* ATCC 202195 is intriguing, since FOS could well have played a role in the positive clinical outcomes in newborns reported by Panigrahi et al.<sup>5</sup> Moreover, the identification of a diverse repertoire of carbohydrate active enzymes suggests that *L. plantarum* ATCC 202195 is able to degrade, utilize and synthesize both simple and complex saccharides. The functional annotation of the complete genome of *L. plantarum* ATCC 202195 provides the foundation for future mechanistic investigations; in vitro and in vivo studies are needed to verify the impact of each highlighted functional characteristic on the microbe and host.

Advances in next generation sequencing have increased the availability of high-quality genomic data and improved our ability to identify phylogenetic relationships; however, technical differences in genome sequencing protocols and genome assembly tools can result in artifactual variance<sup>58</sup>. Here, we sequenced, assembled and compared genomes of two isolates from the same strain, *L. plantarum* ATCC 202195. Our methodology used a combination of short- and long-read sequences for ATCC 202195-A genome assembly and only short-read sequences for ATCC 202195-B, as well as two different assembly tools. To limit technical bias in our comparative analyses, we compared unassembled sequencing reads from *L. plantarum* ATCC 202195-B against the complete assembled *L. plantarum* ATCC 202195-A genome, and utilizing genome alignment tools that allow for rearrangement<sup>27</sup>; however, technical variation may have played a role in the identification of 3 SNPs and the discrepancy in genome length between the two isolates.

This study confirms that *L. plantarum* ATCC 202195-B, which was only recently made commercially available, is genetically identical to the isolate of *L. plantarum* ATCC 202195 first deposited into the ATCC over 20 years ago and presumably identical to the isolate of *L. plantarum* ATCC 202195 used in the hospital and community-based trials conducted in India<sup>5,7</sup>. *L. plantarum* ATCC 202195 does not contain any unexpected AMR patterns and it is susceptible to multiple clinically important groups of antimicrobial agents. We show that *L. plantarum* ATCC 202195 contains two plasmids, but since there are no concerning plasmid-encoded antimicrobial or virulence genes, *L. plantarum* ATCC 202195 does not pose a material threat for the transfer of AMR or virulence factors to other microorganisms. While the probiotic potential of *L. plantarum* ATCC 202195 has been described in an initial clinical trial<sup>5</sup>, the genomic characteristics that were identified in the current work, including the identification of genes involved in stress responses, cellular adhesion, carbohydrate metabolism, and vitamin biosynthesis, provide further evidence in support of the probiotic properties of *L. plantarum* ATCC 202195 and shed light on potential mechanisms by which the strain exerts its biological effects on the human host. Taken together, the findings arising from this study will inform the design of future clinical trials and programs to employ *L. plantarum* ATCC 202195 for use as either a probiotic or, together with FOS, as a synbiotic<sup>59</sup>.

## Materials and methods

**Bacterial strains and isolates.** *Lactiplantibacillus plantarum* ATCC 202195-A was procured from the ATCC (Manassas, Virginia) by International Flavors & Fragrances Inc. (formerly DuPont Nutrition & Biosciences) in October 2017. *L. plantarum* ATCC 202195-B was purchased from the ATCC by researchers at the Hospital for Sick Children (SickKids) in March 2019.

**Antimicrobial susceptibility testing.** Resistance profiles of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B were tested against the following 20 antibacterial agents: ciprofloxacin (0.5–2.0 µg/ml), erythromycin (0.25–4.0 µg/ml), meropenem (0.25–2.0 µg/ml), ceftriaxone (0.12–2.0 µg/ml), clindamycin (0.12–2.0 µg/ml), gentamicin (2.0–16.0 µg/ml), penicillin (0.06–8.0 µg/ml), ampicillin (0.12–16.0 µg/ml), chloramphenicol (1.0–32.0 µg/ml), tetracycline (2.0–16.0 µg/ml), levofloxacin (0.25–8.0 µg/ml), linezolid (0.50–8.0 µg/ml), piperacillin-tazobactam (8.0–128.0 µg/ml/4.0 µg/ml), vancomycin (1.0–128.0 µg/ml), quinupristin/dalfopristin (0.12–4.0 µg/ml), daptomycin (0.25–8.0 µg/ml), rifampin (0.50–4.0 µg/ml), trimethoprim/sulfamethoxazole (0.5–4.0 µg/ml/9.5–76.0 µg/ml), gatifloxacin (1.0–8.0 µg/ml) and oxacillin + 2% NaCl (0.25–8.0 µg/ml) using broth dilution assays in the Clinical Microbiology Laboratory at the Hospital for Sick Children (Toronto, ON, Canada).

Susceptibility testing was performed according to the Clinical and Laboratory Standards Institute (CLSI), M07, 11th edition and M45, 3rd edition, guideline for *Lactobacillus* spp.<sup>19,60</sup>. Direct colony suspensions, prepared to an equivalent of 0.5 McFarland standard in cation-adjusted Mueller–Hinton broth supplemented with lysed horse blood (CAMHB-LHB), were inoculated into three different commercially available Sensititre plates: Gram Positive MIC, Streptococcus STP6F AST and Gram Negative GN4F AST plates (ThermoFisher). Sensititre plates were incubated for 48 h at 35 °C, 5% CO<sub>2</sub> and MICs were manually recorded and interpreted using clinical breakpoints established by CLSI<sup>19</sup>, the European Committee on Antimicrobial Susceptibility Testing (EUCAST)<sup>20</sup> and the European Food Safety Authority (EFSA)<sup>21</sup>. Clinical breakpoint interpretations were categorized as susceptible

("S"), intermediate ("I"), resistant ("R"), no interpretive criteria available ("–") and insufficient evidence ("IE"). To assess the potential for variability in observed MICs, all assays performed using the Sensititre Gram Positive MIC plate were repeated in triplicate for both *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B. To minimize bias during interpretation, laboratory technicians and analysts were blinded to the 'A' or 'B' identity of each isolate.

Resistance of *L. plantarum* ATCC 202195-B to eight of the 20 antimicrobial agents assayed using broth dilution assays (ciprofloxacin, erythromycin, ceftriaxone, gentamicin, penicillin, chloramphenicol, tetracycline and vancomycin) was also tested using gradient strip diffusion (Epsilon meter test) at the Child Health Research Foundation (CHRF) in Dhaka, Bangladesh. E-strips were procured from AB Biodisk, Sweden (penicillin (0.002–32.0 µg/ml), ceftriaxone (0.002–32.0 µg/ml), gentamicin (0.016–256.0 µg/ml), tetracycline (0.016–256.0 µg/ml) and chloramphenicol (0.016–256.0 µg/ml), Hi Media, India [ciprofloxacin (0.002–32.0 µg/ml), erythromycin (0.016–256.0 µg/ml)] and Liofilchem, Italy (vancomycin (0.016–256.0 µg/ml)). E-testing was performed as recommended by the CLSI<sup>19</sup>. In brief, a single colony of *L. plantarum* ATCC 202195-B was emulsified in normal saline to achieve turbidity to an equivalent of 0.5 McFarland standard. The inoculum was swabbed over the entire surface of a Mueller–Hinton agar plate supplemented with 5% sheep blood. After the plate had dried (approximately 10 min), E-strips were placed onto the surface of the agar. MICs, which were determined based on the edge where growth inhibition intersected the side of the E-strip, were recorded after incubation for ~24 h at 37 °C and then interpreted using breakpoints from CLSI<sup>19</sup>, EUCAST<sup>20</sup> and EFSA<sup>21</sup>. MIC interpretations were categorized using the same criteria applied to the broth dilution assays. *L. plantarum* ATCC 202195-B was also tested for beta-lactamase activity using Nitrocefin (BD BBL, New Jersey, USA).

**Bacterial propagation and nucleic acid extraction.** *Lactiplantibacillus plantarum* ATCC 202195-A was cultured anaerobically using the BD GasPak EZ container system (BD 260001) in 4 ml of Man-Rogosa-Sharpe (MRS) broth at 37 °C for 16–18 h. Genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Model: 69506) following the pre-treatment protocol optimized for Gram-positive bacteria. In brief, 2 ml aliquots of *L. plantarum* ATCC 202195-A were centrifuged at 10,000×g for 5 min to harvest bacterial cells. Cells were resuspended in 180 µl of 1× Tris/EDTA buffer at pH 8.0 (BP2473-1, ThermoFisher) containing lysozyme (20 mg/ml) and then incubated for 30 min at 37 °C. Genomic DNA was eluted into 125 µl of elution buffer. RNA was removed by adding RNase (5 µg/µl) followed by incubation for 30 min at 37 °C. DNA yield and quality were assessed by electrophoresis in a 2% agarose gel and using a Qubit 2.0 fluorometer, respectively.

*L. plantarum* ATCC 202195-B was cultured in 5 ml MRS broth, which was incubated without shaking under aerobic conditions for 16–18 h at 37 °C and aerated with 5% CO<sub>2</sub>. Genomic DNA was purified using the One-4-All Genomic DNA Miniprep Kit (Bio Basic, Markham, Ontario, Canada) according to the manufacturer's protocol, with the addition of an enzymatic lysis step where bacterial cells were resuspended in lysis buffer (20 mM Tris–HCl, pH 8.0, 2 mM Na EDTA, 1.2% Triton X-100, 20 mg/ml lysozyme) and then incubated at 37 °C for 60 min. DNA yield and quality were assessed using a ratio of absorbances measured at 260 nm and 280 nm and 230 nm and 260 nm, respectively, using a NanoDrop 2000c spectrophotometer (ThermoFisher Scientific, Wilmington, MA).

**Plasmid DNA extraction.** *Lactiplantibacillus plantarum* ATCC 202195-A was cultured in 5 ml MRS broth, which was incubated without shaking under aerobic conditions for 16–18 h at 37 °C and aerated with 5% CO<sub>2</sub>. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen, Germany), following the manufacturer's guidelines. Extracted plasmid DNA yield and quality was assessed by Nanodrop, as described above, and agarose gel electrophoresis.

**Whole genome sequencing and de novo genome assembly.** Whole-genome shotgun sequencing of *L. plantarum* ATCC 202195-A was performed at the Roy J. Carver Biotechnology Center, University of Illinois Urbana-Champaign, using methods previously described<sup>61</sup>. In brief, Illumina reads were prepared using the Hyper library preparation kit (Kapa Biosystems, Roche, Basel, Switzerland) and sequenced on Illumina MiSeq with paired-end reads 250 nucleotides in length. For Oxford Nanopore (ONT) sequencing, 1 µg of DNA was sheared in a g-Tube and barcoded with 1D Native barcoding genomic DNA kit (EXP-NBD103 and SQK-LSK108). Ten libraries were pooled and sequenced on a GridION X5 sequencer. ONT reads were base-called using Albacore v. 2.1.10 (Oxford Nanopore Technologies, Oxford, United Kingdom). Raw paired-end reads were trimmed using Sickle (v1.3)<sup>62</sup>. Hybrid assembly of the Illumina and ONT reads was performed using Unicycler (v0.4.7)<sup>63</sup>, and the initial genome assembly was refined using Pilon<sup>64</sup>. To inspect the fidelity of the assembly, Illumina reads were mapped to assembled reference contigs using the Burrows-Wheeler Aligner (BWA) (v0.7.17)<sup>65</sup> and Samtools (v1.9)<sup>66</sup>, and visualized using the Integrative Genomics Viewer (IGV) (v2.5.3)<sup>67</sup>. No manual corrections were made.

Whole-genome shotgun sequencing of *L. plantarum* ATCC 202195-B was performed at the Clinical Genomics Centre at Mount Sinai Hospital (Toronto, Ontario, Canada) using the NexteraXT platform (Illumina) with sequencing libraries prepared with the Nextera XT DNA library preparation kit. Paired-read sequence libraries were generated at 150 nucleotides in length, according to the manufacturer's instructions. Raw sequenced reads were trimmed using Trimmomatic<sup>68</sup>; nucleotides at the beginning and end of each sequencing read were discarded if the quality score was below 20 Phred. Reads shorter than 30 bases in length were also discarded. FastQ Screen<sup>69</sup> was used to identify human genomic DNA contamination using PhiX as a calibration control. Reads associated with the human genome were removed from all subsequent analyses. BBNorm (v37.02)<sup>70</sup> was used to normalize coverage in high-depth regions of the genome. SPAdes genome assembler (v3.13.3)<sup>71</sup> was used to reconstruct the genome and the quality of the genome assembly was assessed using QUAST (v5.0.2)<sup>72</sup>.

Next generation sequencing of the extracted plasmid DNA from *L. plantarum* ATCC 202195-A was performed at the Clinical Genomics Centre at Mount Sinai Hospital (Toronto, Canada) using the same methodology employed for *L. plantarum* ATCC 202195-B whole genome sequencing. A stricter filtering and trimming protocol was utilized to process plasmid sequencing reads as per Gallegos<sup>73</sup>. In brief, raw reads were trimmed using Trimmomatic<sup>68</sup>, and discarded with a quality score below 30 and a minimum length of 50 base pairs. Trimmed reads were assembled using Unicycler (v0.4.7)<sup>63</sup>, with bold contig bridging. Assembled contigs were flagged as putative plasmid sequences if the assembled multiplicity was  $\geq 10\times$  or if the contigs were denoted as complete and circular. The resulting identified putative plasmid sequences were screened against the NCBI nucleotide database using BLASTn<sup>31</sup> and PLSDb<sup>74</sup>.

**Comparative genome analysis.** Overall similarity between the assembled genomes of *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B and other publicly available assembled genomes for *L. plantarum* ATCC 202195 (accession number: GCA\_010586945.1 and GCA\_004354995.1) were assessed by comparing the average nucleotide identities (ANI) using the Orthologous Average Nucleotide Identity Tool (OAT)<sup>75</sup>. A global alignment was performed using progressiveMauve<sup>27</sup> and the contig mover feature was employed to account for variation introduced due to differing genome assembly tools used for *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B. The two bacterial genomes were then compared for SNPs. Paired-end normalized read libraries from the genome of *L. plantarum* ATCC 202195-B were aligned with the assembled genome of *L. plantarum* ATCC 202195-A using the BWA-MEM algorithm<sup>65</sup> with default parameters. SNPs were identified using the variant caller FreeBayes (v1.2.0)<sup>26</sup>. For aligned reads to be evaluated as a potential variant, a minimum mapping quality of 20 was used and ploidy was set to 1. Filtering thresholds were employed to detect variants with a quality score  $> 10$  and a read depth  $> 5$ . Variants that remained after filtering were assessed for their potential functional impact using SnpEFF<sup>76</sup>.

**Functional genome annotation of *L. plantarum* ATCC 202195-A and identification of antimicrobial resistance and virulence genes in *L. plantarum* ATCC 202195-A and *L. plantarum* ATCC 202195-B.** The de novo assembled *L. plantarum* ATCC 202195-A genome was annotated using RASTtk (v1.3.0)<sup>25</sup> and the evolutionary genealogy of genes: Non-supervised Orthologous Groups ( ) Database<sup>33</sup>. Annotated genes were compared against the Clusters of Orthologous Groups (COG)<sup>77</sup> and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases<sup>78</sup> to make inferences about higher order biological functions. Carbohydrate Active Enzymes (CAZymes) were identified using the meta server for automated carbohydrate-active enzyme annotation (dbCAN)<sup>79</sup>. CAZyme annotation of a putative *L. plantarum* ATCC 202195 gene was based on agreement between at least two of the three annotation tools employed by dbCAN, including HMMER which searched against the CAZY hidden Markov models database, DIAMOND which screened against the pre-annotated CAZY database, and Hotpep, which screened against a conserved CAZyme short peptide sequence database.

*Lactiplantibacillus plantarum* ATCC 202195-A was screened against five antimicrobial and virulence factor databases: Comprehensive Antibiotic Resistance Database (CARD)<sup>28</sup>, ResFinder<sup>80</sup>, Antibiotic Resistance Gene-ANNOTation (ARG-annot)<sup>81</sup>, Virulence Factor Database (VFDB)<sup>29</sup> and the NCBI Bacterial antimicrobial resistance reference gene database using ABRicate (v0.5)<sup>30,82</sup>. Two different stringency thresholds for AMR gene detection were used: a high stringency threshold, which used a sequence identity and query coverage cut-offs of  $> 80\%$  and a low stringency threshold employing a sequence identity cut-off of  $> 50\%$  and a query coverage cut-off of  $> 10\%$ . Targeted screening for resistance genes related to phenotypic resistance patterns was performed using BLASTp.

**Phylogenetic analyses.** A total of 133 complete *L. plantarum* genomes were retrieved from NCBI in October 2020 (Accession numbers listed in Supplemental Table 2) and reannotated using Prokka (v1.14.5)<sup>83</sup>. The rapid large-scale prokaryote pan genome analysis (Roary) tool (v3.13.0)<sup>84</sup> was used to identify the pan-genome of *L. plantarum* ATCC 202195-A and the other 133 *L. plantarum* genomes, including a previously submitted complete genome for *Lactiplantibacillus plantarum* ATCC 202195 (accession number: GCA\_010586945.1), which was released in February 2020. The core pan-genome alignment was used to construct a phylogenetic tree with the maximum likelihood method using Randomized Axelerated Maximum Likelihood (RaxML) (v2.0.6)<sup>85</sup> with fast bootstrapping, which was visualized using FigTree<sup>86</sup>.

**Genome accession numbers.** The sequence of *L. plantarum* ATCC 202195-A has been deposited in the GenBank database under accession number accession numbers CP063750.1, CP063751.1, and CP063752.1 (<https://www.ncbi.nlm.nih.gov/nucleotide/CP063750.1>). The unassembled reads of *L. plantarum* ATCC 202195-B have also been deposited in the NCBI Sequence Read Archive (SRA) under the accession number SRR13686146.

Received: 12 February 2021; Accepted: 13 July 2021

Published online: 05 August 2021

## References

1. Sun, J. *et al.* Effects of probiotics on necrotizing enterocolitis, sepsis, intraventricular hemorrhage, mortality, length of hospital stay, and weight gain in very preterm infants: A meta-analysis. *Adv. Nutr.* **8**, 749–763. <https://doi.org/10.3945/an.116.014605> (2017).

2. Deshpande, G., Jape, G., Rao, S. & Patole, S. Benefits of probiotics in preterm neonates in low-income and medium-income countries: A systematic review of randomised controlled trials. *BMJ Open* **7**, e017638. <https://doi.org/10.1136/bmjopen-2017-017638> (2017).
3. Aceti, A. *et al.* Probiotics prevent late-onset sepsis in human milk-fed, very low birth weight preterm infants: Systematic review and meta-analysis. *Nutrients* <https://doi.org/10.3390/nu9080904> (2017).
4. Zheng, J. *et al.* A taxonomic note on the genus *Lactobacillus*: Description of 23 Novel Genera, Emended Description of Genus *Lactobacillus* Beijerinck 1901, and Union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* **70**, 2782–2858. <https://doi.org/10.1099/ijsem.0.004107> (2020).
5. Panigrahi, P. *et al.* A randomized synbiotic trial to prevent sepsis among infants in rural India. *Nature* **548**, 407–412. <https://doi.org/10.1038/nature23480> (2017).
6. Panigrahi, P., Gewolb, I. H. & Morris, J. G. Jr. Preventing/treating neonatal NEC by administering *Lactobacillus salivarius* and *Lactobacillus plantarum* or a combination thereof. United States patent (2000).
7. Panigrahi, P. *et al.* Long-term colonization of a *Lactobacillus plantarum* synbiotic preparation in the neonatal gut. *J. Pediatr. Gastroenterol. Nutr.* **47**, 45–53. <https://doi.org/10.1097/MPG.0b013e31815a5f2c> (2008).
8. International Probiotics Association. *IPA Guidelines to Qualify a Microorganism to be Termed as “Probiotic”* (International Probiotic Association, 2017).
9. Sakurai, K. & Kawasaki, H. Genetic variation during long-term preservation of bacteria in public culture collections. *Int. J. Syst. Evol. Microbiol.* **68**, 1815–1821. <https://doi.org/10.1099/ijsem.0.002717> (2018).
10. Yelin, I. *et al.* Genomic and epidemiological evidence of bacterial transmission from probiotic capsule to blood in ICU patients. *Nat. Med.* **25**, 1728–1732. <https://doi.org/10.1038/s41591-019-0626-9> (2019).
11. McFarland, L. V., Evans, C. T. & Goldstein, E. J. C. Strain-specificity and disease-specificity of probiotic efficacy: A systematic review and meta-analysis. *Front. Med. (Lausanne)* **5**, 124. <https://doi.org/10.3389/fmed.2018.00124> (2018).
12. McFarland, L. V. Efficacy of single-strain probiotics versus multi-strain mixtures: Systematic review of strain and disease specificity. *Dig. Dis. Sci.* <https://doi.org/10.1007/s10620-020-06244-z> (2020).
13. Wright, M. E., Yu, A. O., Marco, M. L. & Panigrahi, P. Genome sequence of *Lactiplantibacillus plantarum* ATCC 202195, a probiotic strain that reduces sepsis and other infections during early infancy. *Microbiol. Resour. Announc.* <https://doi.org/10.1128/MRA.00741-20> (2020).
14. Joint FAO WHO Working Group. *Guidelines for the Evaluation of Probiotics in Food* (World Health Organization, 2002).
15. European Food Safety Authority. Qualified Presumption of Safety. April 13, 2020. Website: <https://www.efsa.europa.eu/en/topics/topic/qualified-presumption-safety-qps>. Accessed 13 April 2020.
16. Vankerckhoven, V. *et al.* Biosafety assessment of probiotics used for human consumption recommendations from the EU-PROSAFE project. *Trends Food Sci. Technol.* **19**, 102–114 (2008).
17. Neut, C., Mahieux, S. & Dubreuil, L. J. Antibiotic susceptibility of probiotic strains: Is it reasonable to combine probiotics with antibiotics?. *Med. Mal. Infect.* **47**, 477–483. <https://doi.org/10.1016/j.medmal.2017.07.001> (2017).
18. Guo, Q., Goldenberg, J. Z., Humphrey, C., El Dib, R. & Johnston, B. C. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst. Rev.* **4**, CD004827. <https://doi.org/10.1002/14651858.CD004827.pub5> (2019).
19. Clinical and Laboratory Standards Institute. *Methods for Antimicrobial Dilution and Disk Susceptibility Testing of Infrequently Isolated or Fastidious Bacteria* 3rd edn. (Clinical and Laboratory Standards Institute, 2015).
20. The European Committee on Antimicrobial Susceptibility Testing. *Breakpoint tables for interpretation of MICs and zone diameters*. Version 10.0 (2020). <https://www.eucast.org>. Accessed 23 Sept 2020.
21. EFSA FEEDAP Panel (EFSA Panel on Additives and Products or Substances used in Animal Feed) *et al.* Guidance on the characterization of microorganisms used as feed additives or as production organisms. *EFSA J.* <https://doi.org/10.2903/j.efsa.2018.5206> (2018).
22. Klare, I. *et al.* Antimicrobial susceptibilities of *Lactobacillus*, *Pediococcus* and *Lactococcus* human isolates and cultures intended for probiotic or nutritional use. *J. Antimicrob. Chemother.* **59**, 900–912. <https://doi.org/10.1093/jac/dkm035> (2007).
23. Campedelli, I. *et al.* Genus-wide assessment of antibiotic resistance in. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/AEM.01738-18> (2019).
24. Stothard, P. & Wishart, D. S. Circular genome visualization and exploration using CGView. *Bioinformatics* **21**, 537–539. <https://doi.org/10.1093/bioinformatics/bti054> (2005).
25. Brettin, T. *et al.* RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365. <https://doi.org/10.1038/srep08365> (2015).
26. Garrison E., Marth, G. Haplotype-based variant detection from short-read sequencing. arXiv preprint [arXiv.1207.3907](https://arxiv.org/abs/1207.3907) [q-bio.GN] (2012).
27. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**, e11147. <https://doi.org/10.1371/journal.pone.0011147> (2010).
28. Jia, B. *et al.* CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573. <https://doi.org/10.1093/nar/gkw1004> (2017).
29. Chen, L. *et al.* VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* **33**, D325–328. <https://doi.org/10.1093/nar/gki008> (2005).
30. Seemann, T. ABRicate. Github. <https://github.com/tseemann/abricate>. Accessed July 2020.
31. Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. *Bioinformatics* **24**, 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322> (2008).
32. Pittet, V. *et al.* Complete genome sequence of the beer spoilage organism *Pediococcus clausenii* ATCC BAA-344T. *J. Bacteriol.* **194**, 1271–1272. <https://doi.org/10.1128/JB.06759-11> (2012).
33. Huerta-Cepas, J. *et al.* eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314. <https://doi.org/10.1093/nar/gky1085> (2019).
34. Arena, M. P. *et al.* The phenotypic analysis of *Lactobacillus plantarum* shp mutants reveals a potential role for hsp1 in cryotolerance. *Front. Microbiol.* **10**, 838. <https://doi.org/10.3389/fmicb.2019.00838> (2019).
35. Snyder, E. E. *et al.* PATRIC: The VBI PathoSystems resource integration center. *Nucleic Acids Res.* **35**, D401–406. <https://doi.org/10.1093/nar/gkl858> (2007).
36. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res.* **42**, D490–495. <https://doi.org/10.1093/nar/gkt1178> (2014).
37. Saulnier, D. M., Molenaar, D., de Vos, W. M., Gibson, G. R. & Kolida, S. Identification of prebiotic fructooligosaccharide metabolism in *Lactobacillus plantarum* WCFS1 through microarrays. *Appl. Environ. Microbiol.* **73**, 1753–1765. <https://doi.org/10.1128/AEM.01151-06> (2007).
38. Kleerebezem, M. *et al.* Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1990–1995. <https://doi.org/10.1073/pnas.0337704100> (2003).
39. Todorov, S. D. Bacteriocins from *Lactobacillus plantarum* - production, genetic organization and mode of action: producao, organizacao genetica e modo de acao. *Braz J Microbiol* **40**, 209–221. <https://doi.org/10.1590/S1517-83822009000200001> (2009).
40. Rouanet, A. *et al.* Live biotherapeutic products, a road map for safety assessment. *Front. Med.* <https://doi.org/10.3389/fmed.2020.00237> (2020).



41. Borriello, S. P. *et al.* Safety of probiotics that contain lactobacilli or bifidobacteria. *Clin. Infect. Dis.* **36**, 775–780. <https://doi.org/10.1086/368080> (2003).
42. Mariam, D. H., Mengistu, Y., Hoffner, S. E. & Andersson, D. I. Effect of rpoB mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **48**, 1289–1294. <https://doi.org/10.1128/aac.48.4.1289-1294.2004> (2004).
43. World Health Organization. *Integrated Management of Childhood Illness: Chart Booklet* (World Health Organization, 2014).
44. Heo, J. & Uhm, T. B. Complete genome sequence of *Lactobacillus plantarum* JBE245 isolated from Meju. *Korean J. Microbiol.* **53**, 344–346. <https://doi.org/10.7845/kjm.2017.7070> (2017).
45. Arredondo-Alonso, S., Willems, R. J., van Schaik, W. & Schurch, A. C. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb. Genom.* **3**, e000128. <https://doi.org/10.1099/mgen.0.000128> (2017).
46. Cui, Y. *et al.* Plasmids from food lactic acid bacteria: Diversity, similarity, and new developments. *Int. J. Mol. Sci.* **16**, 13172–13202. <https://doi.org/10.3390/ijms160613172> (2015).
47. Yu, W. L. *et al.* Structural insights into the substrate specificity of a 6-phospho-beta-glucosidase BglA-2 from *Streptococcus pneumoniae* TIGR4. *J. Biol. Chem.* **288**, 14949–14958. <https://doi.org/10.1074/jbc.M113.454751> (2013).
48. Allain, T. *et al.* Bile-salt-hydrolases from the probiotic strain *Lactobacillus johnsonii* La1 mediate anti-giardial activity in vitro and in vivo. *Front. Microbiol.* **8**, 2707. <https://doi.org/10.3389/fmicb.2017.02707> (2017).
49. Jones, B. V., Begley, M., Hill, C., Gahan, C. G. & Marchesi, J. R. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13580–13585. <https://doi.org/10.1073/pnas.0804437105> (2008).
50. Ricciardi, A. *et al.* Genotypic diversity of stress response in *Lactobacillus plantarum*, *Lactobacillus paraplantarum* and *Lactobacillus pentosus*. *Int. J. Food Microbiol.* **157**, 278–285. <https://doi.org/10.1016/j.ijfoodmicro.2012.05.018> (2012).
51. Frees, D., Savijoki, K., Varmanen, P. & Ingmer, H. Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol. Microbiol.* **63**, 1285–1295. <https://doi.org/10.1111/j.1365-2958.2007.05598.x> (2007).
52. Yan, H. & Ajuwon, K. M. Butyrate modifies intestinal barrier function in IPEC-J2 cells through a selective upregulation of tight junction proteins and activation of the Akt signaling pathway. *PLoS One* **12**, e0179586. <https://doi.org/10.1371/journal.pone.0179586> (2017).
53. Bach Knudsen, K. E. *et al.* Impact of diet-modulated butyrate production on intestinal barrier function and inflammation. *Nutrients* <https://doi.org/10.3390/nu10101499> (2018).
54. Luu, M. *et al.* Regulation of the effector function of CD8(+) T cells by gut microbiota-derived metabolite butyrate. *Sci. Rep.* **8**, 14430. <https://doi.org/10.1038/s41598-018-32860-x> (2018).
55. Pan, X. D., Chen, F. Q., Wu, T. X., Tang, H. G. & Zhao, Z. Y. Prebiotic oligosaccharides change the concentrations of short-chain fatty acids and the microbial population of mouse bowel. *J. Zhejiang Univ. Sci. B* **10**, 258–263. <https://doi.org/10.1631/jzus.B0820261> (2009).
56. Botta, C. *et al.* Genomic assessment in *Lactobacillus plantarum* links the butyrogenic pathway with glutamine metabolism. *Sci. Rep.* **7**, 15975. <https://doi.org/10.1038/s41598-017-16186-8> (2017).
57. Bergillos-Meca, T. *et al.* In vitro evaluation of the fermentation properties and potential probiotic activity of *Lactobacillus plantarum* C4 in batch culture systems. *LWT Food Sci. Technol.* **60**, 420–426. <https://doi.org/10.1016/j.lwt.2014.08.006> (2015).
58. Magoc, T. *et al.* GAGE-B: An evaluation of genome assemblers for bacterial organisms. *Bioinformatics* **29**, 1718–1725. <https://doi.org/10.1093/bioinformatics/btt273> (2013).
59. Swanson, K. S. *et al.* The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of synbiotics. *Nat. Rev. Gastroenterol. Hepatol.* <https://doi.org/10.1038/s41575-020-0344-2> (2020).
60. Clinical and Laboratory Standards Institute. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically*, M07 11th edn. (Clinical and Laboratory Standards Institute, 2018).
61. Morovic, W. *et al.* Transcriptional and functional analysis of *Bifidobacterium animalis* subsp. *lactis* exposure to tetracycline. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/AEM.01999-18> (2018).
62. Josh, N. & Fass, J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. <https://github.com/najoshi/sickle> (2011). Accessed Jan 2020.
63. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595. <https://doi.org/10.1371/journal.pcbi.1005595> (2017).
64. Walker, B. J. *et al.* 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> (2014).
65. Li, H. Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics* **28**, 1838–1844. <https://doi.org/10.1093/bioinformatics/bts280> (2012).
66. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509> (2011).
67. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26. <https://doi.org/10.1038/nbt.1754> (2011).
68. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).
69. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338. <https://doi.org/10.12688/f1000research.15931.2> (2018).
70. Bushnell, B. BBMap: A Fast, Accurate, Splice-Aware Aligner. <https://sourceforge.net/projects/bbmap/> (2014). Accessed Jan 2020.
71. Bankevich, A. *et al.* 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> (2012).
72. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086> (2013).
73. Gallegos, J. E., Rogers, M. F., Cialek, C. A. & Peccoud, J. Rapid, robust plasmid verification by de novo assembly of short sequencing reads. *Nucleic Acids Res.* **48**, e106. <https://doi.org/10.1093/nar/gkaa727> (2020).
74. Galata, V., Fehlmann, T., Backes, C. & Keller, A. PLSDB: A resource of complete bacterial plasmids. *Nucleic Acids Res.* **47**, D195–D202. <https://doi.org/10.1093/nar/gky1050> (2019).
75. Lee, I., Ouk Kim, Y., Park, S. C. & Chun, J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* **66**, 1100–1103. <https://doi.org/10.1099/ijsem.0.000760> (2016).
76. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92. <https://doi.org/10.4161/fly.19695> (2012).
77. Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36. <https://doi.org/10.1093/nar/28.1.33> (2000).
78. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–114. <https://doi.org/10.1093/nar/gkr988> (2012).
79. Huang, L. *et al.* dbCAN-seq: A database of carbohydrate-active enzyme (CAZyme) sequence and annotation. *Nucleic Acids Res.* **46**, D516–D521. <https://doi.org/10.1093/nar/gkx894> (2018).
80. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644. <https://doi.org/10.1093/jac/dks261> (2012).

81. Gupta, S. K. *et al.* ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* **58**, 212–220. <https://doi.org/10.1128/AAC.01310-13> (2014).
82. Hoyles, L. *et al.* Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat. Med.* **24**, 1070–1080. <https://doi.org/10.1038/s41591-018-0061-3> (2018).
83. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153> (2014).
84. Page, A. J. *et al.* Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693. <https://doi.org/10.1093/bioinformatics/btv421> (2015).
85. Kozlov, A. M., Darrriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAXML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305> (2019).
86. Rambaut, A. FigTree v1.4.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. (2012). <https://tree.bio.ed.ac.uk/software/figtree>. Accessed June 2020.

## Acknowledgements

The authors acknowledge The Clinical Genomics Centre at Mount Sinai Hospital, Toronto, Ontario, Canada for sequencing *L. plantarum* ATCC 202195-B. PMS is the recipient of a Canada Research Chair in Gastrointestinal Disease. DER receives financial support from the Bill and Melinda Gates Foundation.

## Author contributions

K.C. prepared *L. plantarum* ATCC 202195-A for sequencing at International Flavors & Fragrances Inc. (formerly DuPont Nutrition & Biosciences). S.H. generated genome sequencing data for *L. plantarum* ATCC 202195-A. RGH prepared *L. plantarum* ATCC 202195-B for sequencing, compared the genomic sequences of the two isolates, and annotated the *L. plantarum* ATCC 202195-A genome. L.G.P., A.C., H.R., M.S.I., S.K., S.K.M., D.E.R. and S.S. generated or interpreted antimicrobial susceptibility data. L.G.P. and R.G.H. drafted the initial manuscript. L.G.P. edited the manuscript. A.C., S.S., D.E.R. and P.M.S. provided laboratory facilities. L.G.P., R.G.H., D.E.R. and P.M.S. conceived and designed the experiments. D.E.R. and P.M.S. oversaw the project. All authors have reviewed and approved the submitted manuscript.

## Competing interests

This work was supported by the Bill & Melinda Gates Foundation [INV-007389]. DER has received funding from the Bill and Melinda Gates Foundation to conduct phase II and III trials of *L. plantarum* ATCC 202195, for which International Flavors & Fragrances Inc. will donate probiotic/synbiotic preparations of *L. plantarum* ATCC 202195 for use as investigational products. PMS has received financial and research support from Lallemand Human Nutrition, honoraria from Abbott Nutrition and Mead Johnson Nutritionals for CME activities and financial compensation for serving on Advisory Boards for Nestle-Gerber and Cargill Inc. related to the use of prebiotics and probiotics in infants and children. SH and KCE are employees of International Flavors & Fragrances Inc. which will manufacture and supply *L. plantarum* ATCC 202195-containing probiotic/synbiotic preparations. The remaining authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94997-6>.

**Correspondence** and requests for materials should be addressed to D.E.R. or P.M.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021