AMERICAN SOCIETY FOR MICROBIOLOGY

Microbiology®
Resource Announcements

Check for updates

Whole-Genome Sequencing of *Lactobacillus johnsonii* MT4, a Novel Strain Isolated from the Oral Cavity of C57BL/6 Mice

Roberto Vazquez-Munoz,[a] Jordan T. Russell,[b] Angela Thompson,[a] Yanjiao Zhou,[b] Anna Dongari-Bagtzoglou[a]

aDepartment of General Dentistry, University of Connecticut Health Center, Farmington, Connecticut, USA
bDepartment of Medicine, University of Connecticut Health Center, Farmington, Connecticut, USA

ABSTRACT   *Lactobacillus johnsonii* strain MT4, isolated from the oral cavity of C57BL/6 mice, elicits antimicrobial activity against disease-associated microorganisms. Short-read sequencing of the whole genome revealed a single genome of 1,883,026 bp, with a GC content of 34.4%, and no plasmids.

*L*actobacillus johnsonii is associated with probiotic properties, host immunomodulation, and mitigation of certain metabolic syndromes (1–3). MT4 displays antibacterial and antifungal properties (4, 5).

MT4 was initially isolated by our group from C57BL/6 mice tongues and identified as *L. johnsonii*, as described by Bertolini et al. (4). Pure cultures were incubated overnight in de Man, Rogosa, and Sharpe broth under static, anaerobic conditions at 37°C. Cells were lysed using a custom lysis buffer (6) and zirconia beads. Total DNA was extracted using the DNeasy blood and tissue purification kit (Qiagen).

Purified genomic DNA (gDNA) was quantified using the double-stranded DNA (dsDNA) high-sensitivity assay for Qubit 3.0 (Life Technologies, USA). gDNA fragmentation was analyzed on the Agilent 4200 TapeStation using the genomic DNA assay (Agilent Technologies). gDNA (1 ng) was normalized to 0.2 ng/μL for whole-genome shotgun library preparation using the Illumina Nextera XT library preparation kit (Illumina), according to the manufacturer's instructions. The libraries were validated for length (average length, 450 bp; average insert size, 315 bp) and adapter dimer removal using the high-sensitivity D5000 ScreenTape assay (Agilent Technologies). The libraries were then quantified and normalized using the dsDNA high-sensitivity assay for Qubit 3.0 and prepared for Illumina sequencing by denaturing and dilution. The sample was run on the Illumina NovaSeq 6000 SP 300-cycle sequencing kit with v1.5 chemistry. The target read depth was achieved with paired-end (PE), 150-bp reads. The sequencing reads were filtered based on Illumina base-calling software algorithms. Quality control was assessed by analyzing forward and reverse FASTQ sequences in the Computational Biology Core High Performance Computing (HPC) facility and phiX reads were filtered from both paired-end FASTQ files using the "filter_phix" command with default parameters in USEARCH v10.0.240 (7). FASTQC was used to assess the sequence data quality and adapter content. The sequence data were high quality (>Q30 throughout) and did not require additional quality filtering. Approximately 10% of paired-end reads contained Nextera adapter sequences, which were trimmed using Trimmomatic v0.39 in PE mode using the provided Nextera adapter file and keeping a minimum trimmed read length of >36 bases (8).

In total, 3,026,429 read pairs were obtained, of which 278,192 forward and 278,217 reverse reads were trimmed to remove Nextera sequences. The trimmed reads were assembled *de novo* using SPAdes v3.15.2 (9) with default parameters through Unicycler v0.4.8 to assemble the genome from both paired-end FASTQ files. The final assembly consisted of 68 contigs (N_{50}, 90.96 kb) and had a total length of 1,883,026 bp, with a GC content of 34.4%, in a single genome with no plasmids. The genome assembly averaged a

sequencing depth of 240×. The MT4 assembled genome was annotated using Prokka v1.14.6 (10), followed by PGAP, after submission to the NCBI GenBank database. Using PGAP, 1,826 genes were predicted, including 1,740 protein-coding genes, 60 RNA genes (3 noncoding RNAs [ncRNAs], 3 rRNAs, and 54 tRNAs), and 26 pseudogenes. The MT4 genes include antimicrobial-associated bacillomycin D, surfactin, and MspI/p75 (5).

Genome-based taxonomy, using Kraken2 v2.0.8-beta (11), confirmed MT4 as an *L. johnsonii* strain (average nucleotide identity, 98.53%). A phylogenetic analysis on the core genes of 17 *L. johnsonii* strains found at NCBI (5) using Roary v3.13.0 (12) revealed that MT4 is closely related to the probiotic strain NCC533 (La1). Roary was run with default parameters, specifying core gene alignment performed using PRANK, core gene presence in 99% of included strains, and 95% blastp protein identity.

Data availability. The assembled MT4 genome sequence and related data can be found under GenBank accession number JAJQJG000000000, BioProject accession number PRJNA787656, and BioSample accession number SAMN23838460. The raw FASTQ sequencing data are available under SRA accession number SRR17309641.

REFERENCES

1. Vazquez-Munoz R, Dongari-Bagtzoglou A. 2021. Anticandidal activities by Lactobacillus species: an update on mechanisms of action. Front Oral Health 2:689382. https://doi.org/10.3389/froh.2021.689382.
2. Marcial GE, Ford AL, Haller MJ, Gezan SA, Harrison NA, Cai D, Meyer JL, Perry DJ, Atkinson MA, Wasserfall CH, Garrett T, Gonzalez CF, Brusko TM, Dahl WJ, Lorca GL. 2017. Lactobacillus johnsonii N6.2 modulates the host immune responses: a double-blind, randomized trial in healthy adults. Front Immunol 8:655. https://doi.org/10.3389/FIMMU.2017.00655.
3. Yamano T, Iino H, Takada M, Blum S, Rochat F, Fukushima Y. 2006. Improvement of the human intestinal flora by ingestion of the probiotic strain Lactobacillus johnsonii La1. Br J Nutr 95:303–312. https://doi.org/10.1079/bjn20051507.
4. Bertolini M, Vazquez Munoz R, Archambault L, Shah S, Souza JGS, Costa RC, Thompson A, Zhou Y, Sobue T, Dongari-Bagtzoglou A. 2021. Mucosal bacteria modulate Candida albicans virulence in oropharyngeal candidiasis. mBio 12:e0193721. https://doi.org/10.1128/mBio.01937-21.
5. Vazquez-Munoz R, Thompson A, Russell JT, Sobue T, Zhou Y, Dongari-Bagtzoglou A. 2022. Insights from the Lactobacillus johnsonii genome suggest the production of metabolites with antibiofilm activity against the pathobiont Candida albicans. Front Microbiol 13:853762. https://doi.org/10.3389/fmicb.2022.853762.
6. Abusleme L, Hong B-Y, Hoare A, Konkel JE, Diaz PI, Moutsopoulos NM. 2017. Oral microbiome characterization in murine models. Bio Protoc 7:e2655. https://doi.org/10.21769/BioProtoc.2655.
7. Edgar RC, Bateman A. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461.
8. 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.
9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. 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.
10. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
11. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0.
12. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.