AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Genomics and Proteomics | Announcement

Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BI040, a probiotic strain with multiple health benefits from the NORDBIOTIC collection

Agnieszka K. Szczepankowska,[1] Bożena Cukrowska,[2] Tamara Aleksandrzak-Piekarczyk[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT The complete genome of *Bifidobacterium animalis* subsp. *lactis* BI040—a human stool isolate, was sequenced using Illumina and Oxford Nanopore technologies. The BI040 genome is composed of a circular 1,944,141-bp chromosome which carries genes potentially involved in vitamin synthesis and gut health.

KEYWORDS *Bifidobacterium*, stool sample, vitamin B6 synthesis, gut health, probiotics, irritable bowel syndrome, respiratory viral infections, succinate, anti-obesity effects, bile salt hydrolase

Bifidobacterium animalis strains are commonly recognized for their probiotic potential and ability to colonize the human intestine (1). *B. animalis* subsp. *lactis* BI040 from the NORDBIOTIC collection was originally isolated from a human stool sample and has been deposited in the DSMZ Collection (DSM33812). In clinical trials, BI040 demonstrated beneficial effects for patients with irritable bowel syndrome (IBS) (2) and respiratory viral infections (3).

The strain was provided from the NORDBIOTIC collection and cultured overnight from a single colony at 37°C in MRS medium (Oxoid) under anaerobic conditions. DNA was isolated using a cetyltrimethylammonium bromide/lysozyme method (4) with a lysozyme/mutanolysin pre-digestion step (5) and sequenced independently with Illumina and Oxford Nanopore technologies.

Illumina library was prepared with the NEB Ultra II FS kit (New England Biolabs) and sequenced in paired-end mode (2 × 300 bp) on the MiSeq platform (Illumina). Raw data were trimmed and filtered with FASTQC (v.0.12.0) (6) and fastp (v.0.23.2) (7), resulting in 666,502 final reads of average 311 × coverage and 296,897,927 nt.

Long-read sequencing library was created using sheared and size-selected (>10 kb) DNA fragments (Short Read Eliminator kit; Circulomics) and the SQK-LSK109 native barcoding expansion kit (EXP-NBD103), loaded on an R9.4.1 flowcell and sequenced on the GridION sequencer (Oxford Nanopore Technologies). Raw ONT reads were generated using Guppy (v.6.1.3) (ONT) in super accuracy mode and filtered with NanoFilt (v.2.8.0) (8) to eliminate short (<1 kb) and low-quality (QS <12) reads, and with Porechop (v.0.2.4) (https://github.com/rrwick/Porechop) to remove residual adapters. Sequence quality was checked using NanoPlot (v.1.41.6) (8). Final Nanopore sequencing resulted in 90,371 reads with an average length of 8,657 nt, N_{50} value of 12,410, and 198 × coverage. Default parameters were used for all software.

ONT long-reads were assembled using Trycycler (v.0.5.3) (9) and Flye (v.2.9) (10), Unicycler (v.0.4.8) (11), Raven (v.1.8.1) (12), and Miniasm (v.0.3-r179) (13), reconciled and circularized with Trycycler, and polished using Racon (v.1.5.0) (14) and Medaka (v.1.7.2) (15). To obtain a high-quality consensus sequence, the ONT assembly was corrected with Illumina data using Polypolish (v.0.5.0) (16) and POLCA (v.4.0.5) (17), resulting in

Editor David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to Tamara Aleksandrzak-Piekarczyk, tamara@ibb.waw.pl.

The authors declare no conflict of interest.

Received 11 October 2024
Accepted 19 December 2024
Published 17 January 2025

a circular, 1,944,141 bp chromosome with 509.0 coverage and a GC content of 60.5%. Coding sequence (CDS) prediction and gene annotation were performed using the NCBI Prokaryotic Genome Annotation Pipeline (v.6.6) (18), identifying 1,543 predicted CDSs, 52 tRNAs, 12 rRNAs, and 3 ncRNAs. A complete set of genes for *de novo* vitamin B6 synthesis via deoxyxylulose 5-phosphate-independent pathway was identified, indicating potential nutritional benefits for BI040-fermented food products. Additionally, the presence of *sucCD* genes encoding succinyl-CoA synthase for the conversion of succinyl-CoA to succinic acid was observed. Succinate, a microbiome-produced by-product, plays a role in maintaining gut equilibrium and in inducing pro-inflammatory responses to manage local stress (19), potentially aiding in constipation relief, a condition inversely correlated with obesity-related metabolic disturbances. We identified gene encoding bile salt hydrolase [EC 3.5.1.24], which catalyzes the deconjugation of glycine- and taurine-linked bile salts, reducing their reabsorption efficiency compared to conjugated bile salts, and potentially contributing to lower serum cholesterol levels (20).

ACKNOWLEDGMENTS

Sequencing was carried out at the DNA Sequencing and Synthesis Facility of the Institute of Biochemistry and Biophysics of the Polish Academy of Science.

AUTHOR AFFILIATIONS

[1]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland
[2]The Children Memorial Health Institute, Department of Pathomorphology, Warsaw, Poland

AUTHOR ORCIDs

Agnieszka K. Szczepankowska http://orcid.org/0000-0002-8733-3283
Tamara Aleksandrzak-Piekarczyk http://orcid.org/0000-0002-4725-760X

AUTHOR CONTRIBUTIONS

Agnieszka K. Szczepankowska, Conceptualization, Investigation, Validation, Writing – original draft, Writing – review and editing | Bożena Cukrowska, Supervision, Validation, Writing – review and editing | Tamara Aleksandrzak-Piekarczyk, Investigation, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

Complete sequencing data has been deposited in GenBank under BioProject PRJNA1071652 and BioSample SAMN39639087. Whole-genome data are available under accession number CP147973. Illumina SRA reads are available under accession number SRX23536629. Oxford Nanopore SRA reads are available under accession number SRX23536628.

REFERENCES

1. Skrzydło-Radomańska B, Prozorow-Król B, Cichoż-Lach H, Majsiak E, Bierła JB, Kanarek E, Sowińska A, Cukrowska B. 2021. The effectiveness and safety of multi-strain probiotic preparation in patients with diarrhea-predominant irritable bowel syndrome: a randomized controlled study. Nutrients 13:756. https://doi.org/10.3390/nu13030756
2. Kolesnyk PO, Paliy IH, Sydorchuk LP, Hoda ZP, Ivanchenko NO, Lych OS, Huley NR, Matsyura OI, Slyuzar ZL, Gerasymov SV. 2024. The role of nutritional support with probiotics in outpatients with symptomatic acute respiratory tract infections: a multicenter, randomized, double-blind, placebo-controlled dietary study. BMC Nutr 10:4. https://doi.org/10.1186/s40795-023-00816-8
3. Sharma M, Wasan A, Sharma RK. 2021. Recent developments in probiotics: an emphasis on *Bifidobacterium*. Food Biosci 41:100993. https://doi.org/10.1016/j.fbio.2021.100993
4. Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol Chapter 2:Unit 2.4. https://doi.org/10.1002/0471142727.mb0204s56
5. Szczepankowska AK, Cukrowska B, Aleksandrzak-Piekarczyk T. 2024. Complete genome sequence of the probiotic *Lacticaseibacillus paracasei* LPC100 strain from the NORDBIOTIC collection isolated from a human fecal sample. Microbiol Resour Announc 13:e0034424. https://doi.org/10.1128/mra.00344-24

6. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
7. Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560
8. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149
9. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biol 22:266. https://doi.org/10.1186/s13059-021-02483-z
10. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8
11. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595
12. Vaser R, Šikić M. 2021. Time- and memory-efficient genome assembly with Raven. Nat Comput Sci 1:332–336. https://doi.org/10.1038/s43588-021-00073-4
13. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152
14. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116
15. Wright C, Wykes M. 2023. Medaka: sequence correction provided by ONT research. https://github.com/nanoporetech/medaka.
16. Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. PLoS Comput Biol 18:e1009802. https://doi.org/10.1371/journal.pcbi.1009802
17. Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLoS Comput Biol 16:e1007981. https://doi.org/10.1371/journal.pcbi.1007981
18. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569
19. Fernández-Veledo S, Vendrell J. 2019. Gut microbiota-derived succinate: friend or foe in human metabolic diseases? Rev Endocr Metab Disord 20:439–447. https://doi.org/10.1007/s11154-019-09513-z
20. Agolino G, Pino A, Vaccalluzzo A, Cristofolini M, Solieri L, Caggia C, Randazzo CL. 2024. Bile salt hydrolase: the complexity behind its mechanism in relation to lowering-cholesterol lactobacilli probiotics. J Funct Foods 120:106357. https://doi.org/10.1016/j.jff.2024.106357