

Complete Genome Sequence of *Bifidobacterium longum* GT15: Unique Genes for Russian Strains

Natalia V. Zakharevich,^a Olga V. Averina,^a Ksenia M. Klimina,^a Anna V. Kudryavtseva,^b Artem S. Kasianov,^a Vsevolod J. Makeev,^{a,b} Valery N. Danilenko^a

Vavilov Institute of General Genetics, Moscow, Russia^a; Engelhardt Institute of Molecular Biology, Moscow, Russia^b

In this study, we report the first completely annotated genome sequence of the Russian-origin *Bifidobacterium longum* subsp. *longum* strain GT15. We discovered 35 unique genes (UGs) which were detected from only the *B. longum* GT15 genome and were absent from other *B. longum* strain genomes (not of Russian origin).

Received 13 November 2014 Accepted 14 November 2014 Published 18 December 2014

Citation Zakharevich NV, Averina OV, Klimina KM, Kudryavtseva AV, Kasianov AS, Makeev VJ, Danilenko VN. 2014. Complete genome sequence of *Bifidobacterium longum* GT15: unique genes for Russian strains. *Genome Announc.* 2(6):e01348-14. doi:10.1128/genomeA.01348-14.

Copyright © 2014 Zakharevich et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Valery N. Danilenko, valeridanilenko@yandex.ru.

Bifidobacteria represent an important group of the human intestinal microbiota (1). There is an increasing interest in the positive effects of bifidobacteria on the central nervous system (CNS) via neural, neuroendocrine (2), neuroimmune, and humoral mechanisms (3–5). Here, we present the first complete annotated genome sequence for Russian-origin strain *Bifidobacterium longum* subsp. *longum* GT15. This strain was isolated from the feces of a healthy adult inhabiting Central Russia and demonstrated good probiotic properties.

The complete genome of *B. longum* GT15 is composed of a 60% G+C circular chromosome of 2,337,521 bp with no plasmids. A total of 1,893 coding sequences (CDS) were predicted in the genome. Based on a Clusters of Orthologous Groups (COG) functional classification (6, 7), the highest number of genes associated with metabolism in strain GT15 belonged to the carbohydrate metabolism [G] category, which amounted to more than 10%. The GT15 genome contains 5 rRNA operons, 56 tRNA genes, 26 pseudogenes, and 114 tandem repeats. Nine phage-related fragments were predicted in the genome of *B. longum* GT15, but no complete prophages were found. The GT15 chromosome also possesses 13 complete or disrupted insertion sequence (IS) elements (belonging to IS21, IS256, IS3, IS30, and ISL3 families). One confirmed cluster of regularly interspaced short palindromic repeat (CRISPR) (24-bp CRISPR repeat) and three potential CRISPR-related systems were discovered in the GT15 genome. In our analysis, we also examined genes for global regulatory systems; we annotated six genes of toxin-antitoxin (TA) systems type II (8), six serine/threonine protein kinases (STPKs) of eukaryotic type (9), and three genes of the WhiB-like family (10, 11). In general, comparative analysis indicated that the *B. longum* GT15 genome was similar to the available genomes. However, in the *B. longum* GT15 genome, we detected unique nucleotide sequences (more than 150 nucleotides in length) present only in the GT15 genome, and not found in any of the sequenced members of the *B. longum* species (not of Russian origin). We identified open reading frames (ORFs) within such sequences and called them unique genes (UGs). The GT15 ge-

nome contained 35 ORFs of such genes, of a total length of 39,066 bp, with G+C composition in the range of 37% to 65%. It is worth noting that all unique genes, except one, have a G+C composition lower than that of the overall genome of strain GT15. The genes encoded proteins of lengths varying from 57 to 1,218 amino acid residues. The 21 genes were similar to genes of *B. longum* subsp. *longum* strains of Russian origin: 1-6B, 2-2B, 44B, and 35B (12). Some of the unique genes are adjacent to each other, forming islands (clusters). In the genome of GT15, we found seven such gene clusters consisting of unique genes. These clusters include from two to seven genes. In addition, some of the clusters are flanked by different mobile elements and display significant divergence from the average G+C genome content or atypical codon usage, suggesting acquisition through horizontal gene transfer (HGT). A large proportion of these identified UGs encode hypothetical proteins with unknown function; thus, they may represent novel biosynthetic or human gut commensal interaction features.

Nucleotide sequence accession number. This complete genome project has been deposited in GenBank under the accession no. [CP006741](https://ncbi.nlm.nih.gov/nucl/CP006741) and submitted for automated annotation at the National Center for Biotechnology Information (NCBI). The genome annotation was conducted with the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/).

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education and Science of the Russian Federation under state contracts (grants) 14.N08.12.0021 and 14.621.21.0001.

Part of this work was performed at the EIMB RAS “Genome” Center, http://www.eimb.ru/RUSSIAN_NEW/INSTITUTE/ccu_genome_c.php.

REFERENCES

1. Turroni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, Kerr C, Hourihane J, Murray D, Fuligni F, Gueimonde M, Margolles A, De Bellis G, O’Toole PW, van Sinderen D, Marchesi JR, Ventura M.

2012. Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* 7:e36957. <http://dx.doi.org/10.1371/journal.pone.0036957>.
2. Lyte M. 2014. Microbial endocrinology: host-microbiota neuroendocrine interactions influencing brain and behavior. *Gut Microbes* 5:381–389. <http://dx.doi.org/10.4161/gmic.28682>.
 3. Clarke G, Stilling RM, Kennedy PJ, Stanton C, Cryan JF, Dinan TG. 2014. Minireview: gut microbiota: the neglected endocrine organ. *Mol. Endocrinol.* 28:1221–1238. <http://dx.doi.org/10.1210/me.2014-1108>.
 4. Aziz Q, Doré J, Emmanuel A, Guarner F, Quigley EM. 2013. Gut microbiota and gastrointestinal health: current concepts and future directions. *Neurogastroenterol. Motil.* 25:4–15. <http://dx.doi.org/10.1111/nmo.12046>.
 5. Montiel-Castro AJ, González-Cervantes RM, Bravo-Ruiseco G, Pacheco-López G. 2013. The microbiota–gut–brain axis: neurobehavioral correlates, health and sociality. *Front Integr. Neurosci.* 7:70. <http://dx.doi.org/10.3389/fnint.2013.00070>.
 6. Koonin EV. 2002. The Clusters of Orthologous Groups (COGs) database: phylogenetic classification of proteins from complete genomes, p 1–7. *In* McEntyre J, Ostell J (ed), *The NCBI handbook*, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
 7. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29:22–28. <http://dx.doi.org/10.1093/nar/29.1.22>.
 8. Averina OV, Alekseeva MG, Abilev SK, Il'in VK, Danilenko VN. 2013. Distribution of genes of toxin-antitoxin systems of mazEF and relBE families in bifidobacteria from human intestinal microbiota. *Genetika* 49: 315–327. (In Russian.)
 9. Nezametdinova VZ, Zakharevich NV, Alekseeva MG, Averina OV, Mavletova DA, Danilenko VN. 2014. Identification and characterization of the serine/threonine protein kinases in *Bifidobacterium*. *Arch. Microbiol.* 196:125–136. <http://dx.doi.org/10.1007/s00203-013-0949-8>.
 10. Ramón-García S, Ng C, Jensen PR, Dosanjh M, Burian J, Morris RP, Folcher M, Eltis LD, Grzesiek S, Nguyen L, Thompson CJ. 2013. WhiB7, an Fe-S-dependent transcription factor that activates species-specific repertoires of drug resistance determinants in Actinobacteria. *J. Biol. Chem.* 288:34514–34528. <http://dx.doi.org/10.1074/jbc.M113.516385>.
 11. Averina OV, Zakharevich NV, Danilenko VN. 2012. Identification and characterization of WhiB-like family proteins of the *Bifidobacterium* genus. *Anaerobe* 18:421–429. <http://dx.doi.org/10.1016/j.anaerobe.2012.04.011>.
 12. Shkorporov AN, Efimov BA, Khokhlova EV, Chaplin AV, Kafarskaya LI, Durkin AS, McCorrison J, Torralba M, Gillis M, Sutton G, Weibel DB, Nelson KE, Smeianov VV. 2013. Draft genome sequences of two pairs of human intestinal *Bifidobacterium longum* subsp. *longum* strains, 44B and 1–6B and 35B and 2–2B, consecutively isolated from two children after a 5-year time period. *Genome Announc.* 1(3):e00234-13. <http://dx.doi.org/10.1128/genomeA.00234-13>.