

Genome Sequence of the Probiotic Strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494

Christian Chervaux,¹ Christine Grimaldi,¹ Alexander Bolotin,² Benoit Quinquis,²
Sophie Legrain-Raspaud,¹ Johan E. T. van Hylckama Vlieg,¹
Gerard Denariáz,¹ and Tamara Smokvina^{1*}

Danone Research, R.D. 128, 91767 Palaiseau Cedex, France,¹ and INRA, Micalis,
Domaine de Vilvert, 78352 Jouy en Josas, France²

Received 30 June 2011/Accepted 18 July 2011

***Bifidobacterium animalis* subsp. *lactis* CNCM I-2494 is part of a commercialized fermented dairy product with documented health benefits revealed by multiple randomized placebo-controlled clinical trials. Here we report the complete genome sequence of this strain, which has a circular genome of 1,943,113 bp with 1,660 open reading frames and 4 ribosomal operons.**

Bifidobacteria are known to be the first bacteria to colonize the human intestine after birth. Various strains are frequently associated with positive effects on human health (12) and are used as probiotics. The *Bifidobacterium* genus is comprised of 34 species with considerable phylogenetic heterogeneity. Multilocus sequence typing (4) revealed that *Bifidobacterium animalis* is phylogenetically distant from other species of the genus, such as *B. breve*, *B. longum*, and *B. bifidum*. Within *B. animalis*, *B. animalis* subsp. *lactis* has drawn particular attention, mainly due to its capacity to grow and survive in milk. Several strains are applied as probiotics in dairy products, and the strain CNCM I-2494 is part of a fermented dairy product commercialized since 1987. Studies have demonstrated the ability of this product to reduce severe colitis in mice (13) and improve gastrointestinal well-being in humans (1, 7).

A total of 34,000 Sanger shotgun sequencing reads (ABI 3700) were realized on cosmid and plasmid libraries and assembled by using Phred-Phrap-Consed. Gap closure by primer walking and Sanger sequencing indicated one circular contig with an average 13.6× coverage. Subsequently, all instances of potential frameshifts and low-quality regions of the sequence were manually inspected and corrected, leading to the complete bona fide genome sequence of 1,943,113 bp.

Open reading frames (ORFs) were predicted using a combination of public (CRITICA) and proprietary similarity-based gene identification algorithms (9). The 1,660 identified ORFs were loaded into the ERGO Genome Analysis Suite (Integrated Genomics, Inc.) for annotation. The annotation was inspected manually and improved where relevant by using sequence similarity tools, such as Pfam, ProSite, ProDom, and COG. For genes without adequate similarity, tools and algorithms associated with the ERGO platform that exploit gene context information, gene fusion events, and regulatory gene clusters were used for function assignment. After manual annotation, functions were assigned to 1,214 ORFs. A genome-wide prediction for protein localization elements was also per-

formed; SignalP v 3.0 (3) predicted 16 proteins carrying signal peptides, and the TMHMM 2.0 algorithm identified the presence of one or more putative transmembrane helices in 396 ORFs (25%). Psort analysis (5) and detection of anchor motifs (LPxTG or WxL) suggested that several proteins can be anchored to the cell wall and potentially play a role in bacterium-host interactions. These include large ORFs encoding predicted proteins homologous to proteins involved in binding to eukaryotic cells. Notably, two of them, a predicted collagen adhesion protein and a predicted fibronectin binding protein, carry a von Willebrand factor type A (vWA) domain, which has been shown to promote adhesion to extracellular matrices (10).

Complete genome sequences of several *B. animalis* subsp. *lactis* strains have been reported and compared (2, 6, 8, 11) and have revealed a high genome similarity and almost complete synteny in the subspecies. Comparison of these publically available genome sequences for strain AD011, DSMZ10140, BI04, and BB12 to the genome sequence of strain CNCM I-2494 indicated several single-nucleotide polymorphisms (69 to 319) and from 622 to 1,227 single-base-pair indels, confirming the high level of homogeneity of *B. animalis* subsp. *lactis*.

Nucleotide sequence accession number. The complete genome sequence of *B. animalis* subsp. *lactis* CNCM I-2494 has been deposited in GenBank under accession number CP002915.

We thank Integrated Genomics (Chicago, IL) for sequencing efforts, Claudine Médigue (Genoscope, France) for genome comparison expertise, and Eric Johansen and colleagues (Christian Hansen, Denmark) for fruitful discussion. We acknowledge Leyla Boumghar-Bourthai, Nathalie Goupil-Feuillerat, and Nathan McNulty for their help in genome analysis.

This work was funded by Danone Research.

REFERENCES

1. Agrawal, A., et al. 2009. Clinical trial: the effects of a fermented milk product containing *Bifidobacterium lactis* DN-173010 on abdominal distension and gastrointestinal transit in irritable bowel syndrome with constipation. *Aliment. Pharmacol. Ther.* **29**:104–114.
2. Barrangou, R., et al. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and BI-04. *J. Bacteriol.* **191**:4144–4151.

* Corresponding author. Mailing address: Danone Research, R.D. 128, 91767 Palaiseau Cedex, France. Phone: 33-1-69357320. Fax: 33-1-69357683. E-mail: tamara.smokvina@danone.com.

3. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**:783–795.
4. **Deletoile, A. F., et al.** 2010. Species delineation and clonal diversity in four *Bifidobacterium* species as revealed by multilocus sequencing. *Res. Microbiol.* **161**:82–90.
5. **Gardy, J. L., et al.** 2005. PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* **21**:617–623.
6. **Garrigues, C., E. Johansen, and M. B. Pedersen.** 2010. Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. *J. Bacteriol.* **192**:2467–2468.
7. **Guyonnet, D., A. Schlumberger, L. Mhamdi, S. Jakob, and O. Chassany.** 2009. Fermented milk containing *Bifidobacterium lactis* DN-173010 improves gastrointestinal well-being and digestive symptoms in women reporting minor digestive symptoms: a randomised, double-blind, parallel, controlled study. *Br. J. Nutr.* **102**:1654–1662.
8. **Kim, J. F., et al.** 2009. Genome sequence of the probiotic bacterium *Bifidobacterium animalis* subsp. *lactis* AD011. *J. Bacteriol.* **191**:678–679.
9. **Overbeek, R., et al.** 2003. The ERGOTM genome analysis and discovery system. *Nucleic Acids Res.* **31**:164–171.
10. **Pareti, F. I., et al.** 1986. Isolation and characterization of a collagen binding domain in human von Willebrand factor. *J. Biol. Chem.* **261**:15310–15315.
11. **Sun, Z., et al.** 2010. Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. *J. Bacteriol.* **192**:4080–4081.
12. **Turroni, F., A. Ribbera, E. Foroni, D. van Sinderen, and M. Ventura.** 2008. Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie Van Leeuwenhoek* **94**:35–50.
13. **Veiga, P., et al.** 2010. *Bifidobacterium animalis* subsp. *lactis* fermented milk product reduces inflammation by altering a niche for colitogenic microbes. *Proc. Natl. Acad. Sci. U. S. A.* **107**:18132–18137.