JOURNAL OF BACTERIOLOGY, Dec. 2009, p. 7630–7631 Vol. 191, No. 24
0021-9193/09/$12.00 doi:10.1128/JB.01287-09

Complete Genome Sequence of the Probiotic *Lactobacillus rhamnosus* ATCC 53103▽

Hidetoshi Morita,[1]† Hidehiro Toh,[2]† Kenshiro Oshima,[3] Masaru Murakami,[1] Todd D. Taylor,[2] Shizunobu Igimi,[4] and Masahira Hattori[3]*

School of Veterinary Medicine, Azabu University, 1-17-71 Fuchinobe, Sagamihara, Kanagawa 229-8501, Japan[1]; Advanced Science Institute, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan[2]; Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan[3]; and Division of Biomedical Food Research, National Institute of Health Sciences, Kamiyoga, Setagaya, Tokyo 158-8501, Japan[4]

Received 26 September 2009/Accepted 28 September 2009

Lactobacillus rhamnosus **is a facultatively heterofermentative lactic acid bacterium and is frequently isolated from human gastrointestinal mucosa of healthy individuals. *L. rhamnosus* ATCC 53103, isolated from a healthy human intestinal flora, is one of the most widely used and well-documented probiotics. Here, we report the finished and annotated genome sequence of this organism.**

The complete genome sequence of *Lactobacillus rhamnosus* ATCC 53103 was determined by a whole-genome shotgun strategy with the Sanger method. Genomic libraries containing 2-kb inserts were constructed and sequenced, and 39,936 sequences were generated, giving 8.6-fold coverage from both ends of the genomic clones. Sequence reads were assembled with the Phred-Phrap-Consed program (2). Remaining gaps between contigs were closed by direct sequencing of clones. Prediction and annotation of protein-coding genes were performed as described previously (5).

The genome of *L. rhamnosus* ATCC 53103 consists of a circular 3,005,051-bp chromosome containing 2,834 predicted protein-coding genes and has no plasmid. Of all predicted protein-coding genes, we could assign 1,939 (68%) to known functions, 610 (22%) as conserved hypothetical genes, and 285 (10%) as novel hypothetical genes. This strain has a relatively high number of proteins involved in carbohydrate and amino acid metabolism and transport and defense mechanisms, compared with other sequenced intestinal lactobacilli. The genome encodes 28 complete phosphoenolpyruvate-carbohydrate phosphotransferase-type transporter systems (PTSs) and 25 putative glycosyl hydrolases, which are classified into 12 different carbohydrate-active enzyme families (http://www.cazy.org/). Of the 12 families, alpha-L-fucosidase (GH29) and alpha-mannosidase (GH38) are not found in other sequenced intestinal lactobacilli. Of the 28 PTSs, 12 are encoded by genes adjacent to glycosyl hydrolase genes and transcriptional regulator genes, allowing localized transcriptional control. This organism carries 22 multidrug ABC transporters, eight antimicrobial peptide ABC transporters, and seven beta-lactamases, suggesting its broad range of antibiotic resistance. The genome contains 17 complete two-component regulatory systems, which are most abundant among sequenced lactobacilli. Of the 17 sensor-responder pairs, 1 appears to be potentially associated with bacteriocin production, and 7 are located adjacent to genes for multidrug ABC transporters. Furthermore, the genome contains >90 putative transcriptional regulators.

As expected, extensive similarity at the sequence level is observed between *L. rhamnosus* ATCC 53103 and its closely related strain *Lactobacillus casei* ATCC 334 (4), with overall genome synteny. However, a reciprocal BLASTP search reveals 755 (27%) protein-coding genes that are present in *L. rhamnosus* ATCC 53103 but absent in *L. casei* ATCC 334. These include six carbohydrate utilization gene clusters, which contain the genes for PTSs, glycoside hydrolases, transcriptional regulators, and other carbohydrate-related proteins. Four of the six gene clusters are completely or partially present in *L. casei* BL23 (GenBank accession no. FM177140), suggesting that these gene clusters may have been lost in the lineage to *L. casei* ATCC 334. Thus, the genes specific to *L. rhamnosus* ATCC 53103 may reflect niche differences between *L. rhamnosus* ATCC 53103 (a human isolate) and *L. casei* ATCC 334 (a cheese isolate), suggesting that *L. rhamnosus* ATCC 53103 may have newly acquired these carbohydrate utilization proteins to adapt to the human gastrointestinal tract.

The genome has three gene clusters (LRHM_0182 to LRHM_0184, LRHM_0555 to LRHM_0564, and LRHM_1699 to LRHM_1702) encoding proteins with a C-terminal WxL domain, which attaches to the peptidoglycan on the cell surface (1). The proteins with the WxL domain are present together with the proteins having the DUF916 domain (Pfam PF06030) of unknown function and the small proteins with the LPXTG-like sorting motif, and their gene organizations are similar to that in *L. plantarum* WCFS1 (6). The WxL protein cluster is not found in other sequenced intestinal lactobacilli. The proteins LRHM_1529 (3,275 amino acids; the largest protein encoded by this genome) and LRHM_2193 (1,653 amino acids) contain imperfect repeats consisting of serine and alanine. The genes for both proteins could encode mucin-like cell surface adhesives, because both genes are located adjacent to glycosyltransferase genes (7). The presence of genes encoding proteins for a diverse number of fermentable sugars, a variety of cell surface adherence proteins, bacteriocin biosynthetic proteins (LRHM_2289 to LRHM_2312),

* Corresponding author. Mailing address: Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan. Phone and fax: (81)-47-136-4080. E-mail: hattori@k.u-tokyo.ac.jp.

† The first two authors contributed equally to this work.

▽ Published ahead of print on 9 October 2009.

and bile salt hydrolase (LRHM_0484) is likely to contribute to the organisms' gastric survival and promote interactions with the intestinal mucosa and microbiota.

During the preparation of this article, the genome sequence (3,010,111 bp) of *L. rhamnosus* GG, the original strain of *L. rhamnosus* ATCC 53103, was deposited in a public database (GenBank accession no. FM179322) (3). The genome of *L. rhamnosus* ATCC 53103 is 5 kb shorter than that of *L. rhamnosus* GG. Furthermore, an alignment analysis of both genome sequences shows that the 8.9-kb region (genome coordinates 618415 to 627294) of *L. rhamnosus* ATCC 53103 is inverted.

Nucleotide sequence accession number. The sequence data for the *L. rhamnosus* ATCC 53103 genome are available in DDBJ/GenBank/EMBL under accession no. AP011548.

This research was supported by the "Academic Frontier" Project for Private Universities: Matching Fund Subsidy, 2007–2009 (Azabu University), and Grant-in-Aid for Scientific Research on Priority Areas "Comprehensive Genomics" (M.H.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

We are grateful to the consortium members of the Azabu University Academic Frontier Project and thank K. Furuya, C. Yoshino, H. Inaba, K. Motomura, and Y. Hattori (University of Tokyo) and A. Yamashita, A. Tamura, and N. Itoh (Kitasato University) for technical assistance.

REFERENCES

1. **Brinster, S., S. Furlan, and P. Serror.** 2007. C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other gram-positive bacteria. J. Bacteriol. **189:**1244–1253.
2. **Gordon, D., C. Desmarais, and P. Green.** 2001. Automated finishing with Autofinish. Genome Res. **11:**614–625.
3. **Kankainen, M., L. Paulin, S. Tynkkynen, I. von Ossowski, J. Reunanen, P. Partanen, R. Satokari, S. Vesterlund, A. P. A. Hendrickx, S. Lebeer, S. C. J. De Keersmaecker, J. Vanderleyden, T. Ämäläinen, S. Laukkanen, N. Salovuori, J. Ritari, E. Alatalo, R. Korpela, T. Mattila-Sandholm, A. Lassig, K. Hatakka, K. T. Kinnunen, H. Karjalainen, M. Saxelin, K. Laakso, A. Surakka, A. Palva, T. Salusjärvi, P. Auvinen, and W. M. de Vos.** 2009. Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. Proc. Natl. Acad. Sci. USA. **106:**17193–17198.
4. **Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J. H. Lee, I. Díaz-Muñiz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer, and D. Mills.** 2006. Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. USA **103:**15611–15616.
5. **Morita, H., H. Toh, S. Fukuda, H. Horikawa, K. Oshima, T. Suzuki, M. Murakami, S. Hisamatsu, Y. Kato, T. Takizawa, H. Fukuoka, T. Yoshimura, K. Itoh, D. J. O'Sullivan, L. L. McKay, H. Ohno, J. Kikuchi, T. Masaoka, and M. Hattori.** 2008. Comparative genome analysis of *Lactobacillus reuteri* and *Lactobacillus fermentum* reveal a genomic island for reuterin and cobalamin production. DNA Res. **15:**151–161.
6. **Siezen, R., J. Boekhorst, L. Muscariello, D. Molenaar, B. Renckens, and M. Kleerebezem.** 2006. *Lactobacillus plantarum* gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria. BMC Genomics **7:**126.
7. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498–506.