Garment Sequence of *Parascardovia denticolens* IPLA 20019, Isolated from Human Breast Milk

Miguel Gueimonde, a Francesca Bottacini, b Douwe van Sinderen, b Marco Ventura, a Abelardo Margolles, a and Borja Sánchez a

Instituto de Productos Lácteos de Asturias–Consejo Superior de Investigaciones Científicas (IPLA-CSIC), Department of Microbiology and Biochemistry of Dairy Products, Villaviciosa, Asturias, Spain; a Alimentary Pharmabiotic Centre and Department of Microbiology, Bioscience Institute, National University of Ireland, Western Road, Cork, Ireland; and Laboratory of Probiogenomics, Department of Genetics, Biology of Microorganisms, Anthropology and Evolution, University of Parma, Parma, Italy

This work describes the draft genome of *Parascardovia denticolens* IPLA 20019, isolated from human milk. This species, usually isolated from caries lesions, is taxonomically related to the genus *Bifidobacterium*. The genetic information of IPLA 20019 enhances our understanding of the adaptation of this *P. denticolens* strain from human breast milk.

*P. denticolens* (6), formerly *Bifidobacterium denticolens*, was first described in 1996 (3) and represents the only member of the genus *Parascardovia*. This member of the family *Bifidobacteriaceae* is rod shaped, possesses a high G+C content (55% ± 1%) genome, and produces acetic and lactic acids as main fermentation products. Although first described from human dental caries (3), it has now been isolated from different gastrointestinal places, including saliva and feces of mammals (2, 4). In a recent study, we isolated *P. denticolens* IPLA 20019 from breast milk (7).

We sequenced a total number of 7,672,966,364 bp, which had been generated by 75,969,964 paired-end reads, with an average length of 101 bp, using an Illumina HiSeq2000 sequencer, resulting in a 200-fold genome sequence coverage. The sequence reads were assembled using Velvet de novo sequence assembler 1.2.05 software (8), followed by N50 contig value optimization employing VelvetOptimiser 2.2.0 software, and obtained from the Victorian Bioinformatics Consortium of the Monash University. These programs were run at the “Cluster de Modelización Científica” of the Oviedo University (http://cms.uniovi.es). The final assembly result was a set of 38 unoriented contigs, with a total length of 2,003,311 bp.

Contigs were ordered using the draft genome of *P. denticolens* strain DSM 10105 T, isolated from a caries lesion (NCBI reference sequence, NZ_AEON00000000.1) using CONTIGuator 2 (5). Thirteen contigs, with a total length of 72,719 bp, could not be aligned with the genome of the type strain. Around 130 kb of DNA present in IPLA 20019 appeared to be absent in DSM 10105 T, and although both are unfinished draft genomes, this finding suggests that the extra genetic information of the isolate from breast milk might be of ecological importance for the strain from breast milk.

We analyzed the draft genome of *P. denticolens* IPLA 20019 for the presence of open reading frames (ORFs) and tRNAs using the RAST server 4.0 (1). A total of 1,682 ORFs were identified, which possessed a G+C content of 56.76%. There were 68 tRNAs for the IPLA 20019 contigs, which was in contrast with the 51 tRNAs predicted for DSM 10105 T (assembled in two scaffolds). The RAST server was able to functionally annotate 992 deduced proteins (the remaining 690 predicted protein products were annotated as hypothetical proteins), 535 of which could be functionally assigned to 212 subsystems.

The IPLA 20019 strain encodes a complete ATP synthase and a complete “bifido shunt,” the typical glucose fermentation path-way of the genus *Bifidobacterium*. Neither bile salt hydrolase nor proteins participating in extracellular polysaccharide biosynthesis were predicted by the RAST server. The genome encoded homologs of pilus-like structures, including a type IV secretion system and a pilus assembly protein, potentially involved in colonization, and *P. denticolens* IPLA 20019 encodes few, if any, (antibiotic) resistance determinants. This, together with its source of isolation, makes this strain a potential interest from a technological point of view.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession number AKII00000000. The version described in this paper is the first version, AKII01000000.

Acknowledgments

Borja Sánchez is the recipient of a Juan de la Cierva postdoctoral contract from the Spanish Ministerio de Ciencia e Innovación. This work was supported by grants AGL2010-14952 and RM2010-00012-00-00 from the Spanish Ministerio de Ciencia e Innovación and by an IRCSET Embark postgraduate fellowship to F.B. D.V.S. is a member of The Alimentary Pharmabiotic Centre, which is a Centre for Science and Technology (CSET) funded by Science Foundation Ireland (SFI), through the Irish Government’s National Development Plan.

Maria Fernández García is especially acknowledged for her excellent technical assistance.

References

5. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A. 2011. CONTIGuator:

