Complete Genome Sequence of *Lactococcus lactis* subsp. *cremoris* A76

Alexander Bolotin, Benoît Quinquis, Stanislas Dusko Ehrlich, and Alexei Sorokin

INRA, AgroParisTech, UMR1319 Micalis, Jouy-en-Josas, France; and INRA, AgroParisTech, Department MICA, Jouy-en-Josas, France

We report the complete genome sequence of *Lactococcus lactis* subsp. *cremoris* A76, a dairy strain isolated from a cheese production outfit. Genome analysis detected two contiguous islands fitting to the *L. lactis* subsp. *lactis* rather than to the *L. lactis* subsp. *cremoris* lineage. This indicates the existence of genetic exchange between the diverse subspecies, presumably related to the technological process.

*Lactococcus lactis* is commonly used in the production of fermented dairy products. The *L. lactis* strains are divided in two lineages, *L. lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris*, based on pheno- and genotyping (12, 13, 15). Strains of both lineages are used in mixed industrial starters. Such complex mixtures are of practical importance (6, 8), but this could lead to genetic transfer and emergence of strains with new features, raising questions of safety of the recombinants in food production (4). The strain *L. lactis* subsp. *cremoris* A76 was isolated in 1965. It was shown to have a characteristic pattern of phage sensitivity (5) and a complex multilocus sequence typing (MLST) profile with one loci of seven (ycdB) associated with *L. lactis* subsp. *lactis* rather than with *L. lactis* subsp. *cremoris* (2).

The complete genome sequence of *L. lactis* A76 was determined by whole-genome shotgun sequencing, using Sanger technology, of a clone library with insert sizes of 2 to 3 kb. The genome was assembled using STADEN software (http://staden.sourceforge.net) into 445 contigs longer than 1,000 bp, the gaps were closed by multiplex PCR (3, 14), and low-covered regions were resolved by primer walking sequencing of PCR products. The completed sequence had a 5-fold mean coverage. Annotation was done using the RAST suite (1).

The genome of *L. lactis* A76 consists of one circular chromosome (2,452,618 bp) containing 2,643 predicted coding sequences (CDS), 7% of which are interrupted or truncated by frame shifts, and of four plasmids (pQA554, 53,632 bp and 69 CDS; pQA549, 49,221 bp and 42 CDS; pQA518, 17,600 bp and 12 CDS; pQA505, 3,978 bp and 3 CDS). pQA549 has a 98% nucleotide identity over 80% and 20% of the length to the plasmids 4 and 3, respectively, of *L. lactis* SK11 (11).

Comparative analysis of the *L. lactis* A76 strain and six other *L. lactis* genomes (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi) was performed using Mummer (10), SplitTree4 (9), and Mauve (7). The genomes share 1.84 Mb of common core with about 10% of overall nucleotide divergence. About 95% of the A76 chromosome is of the *L. lactis* subsp. *cremoris* lineage, having 99.2% sequence identity to the SK11 strain (excluding the repeated regions). However, two contiguous regions of the A76 chromosome are of the *L. lactis* subsp. *lactis* lineage. The first (60 kb) encodes mainly genes involved in cell wall biosynthesis and also carries the hypothetical ycbB gene, shown to have inter- and intraspecies mobility (2). This region is present, albeit rearranged, in other *L. lactis* strains. The second (40 kb) corresponds to a prophage highly similar to other prophages of the *L. lactis* subsp. *lactis* lineage. The presence of regions of two lineages in the genome of a single strain shows the recombination origin of the isolate and raises the question of risks related to the emergence of new recombinant strains associated with technological processes.

**Nucleotide sequence accession numbers.** The complete genome of *L. lactis* subsp. *cremoris* A76 has been deposited in GenBank under accession numbers CP003132 (chromosome), CP003133 (pQA554 plasmid), CP003134 (pQA549 plasmid), CP003135 (pQA518 plasmid), and CP003136 (pQA505 plasmid).

**REFERENCES**