Genome sequence of *Lactococcus garvieae* IPLA 31405 isolated from "Casín", a traditional Spanish cheese made from raw cow's milk

Baltasar Mayo[1]*, Ana B. Flórez[1], Pilar Reimundo[2], Ángel Alegría[1], José A. Guijarro[2] and Susana Delgado[1]

[1] Departamento de Microbiología y Bioquímica, Instituto de Productos Lácteos de Asturias (IPLA), Consejo Superior de Investigaciones Científicas (CSIC), Carretera de Infiesto, s/n, 33300-Villaviciosa, Asturias
[2] Departamento de Biología Funcional, Área de Microbiología, Facultad de Medicina, IUBA, Universidad de Oviedo, 33006 Oviedo, Asturias

*baltasar.mayo@ipla.csic.es

Introduction and Objectives

Lactococcus garvieae is the etiological agent of lactococcosis in marine and freshwater fish species. This bacterium has also been retrieved from mastitis in cows and other ruminants, and implicated in an increasing number of human infections. Additionally, *L. garvieae* has been repeatedly reported as a majority component of the native microbiota of farmhouse dairy products manufactured from raw milk. As, association between consumption of raw milk cheese and human infections have never been reported, *L. garvieae* might contribute to the overall quality of dairy products.

Here we report the draft genome sequence of *L. garvieae* IPLA 31405, isolated among the dominant microbiota of Casín, a Spanish traditional raw milk cheese. The genome sequence is expected to give clues on the safety of the strain and on its technological potential. In addition, the genome sequence of strain IPLA 31405 will provide further insights into the intra-specific variation of *L. garvieae* species. Comparison of its genome with those of other sequenced strains may supply information on the safety of the strains and on niche(dairy)-specific genes.

Material and Methods

A genomic library of 0.5 kbp was constructed and paired-end sequenced to approximately 155-fold coverage with a HiSeq 1000 System sequencer (Illumina, Inc., San Diego, CA, USA). Quality-filtered reads were assembled in contigs using Velvet software. Gaps within the contigs were closed by direct sequencing of amplicons obtained by PCR with oligonucleotide primers designed to anneal in the flanking regions. Annotation was done by merging the results obtained from the RAST server, BG7 (Era7 Information Technologies, Granada, Spain), and BLAST analysis (http://blast.ncbi.nlm.nih.gov). KEGG Pathway (http://www.genome.jp/kegg/pathway.html), Uniprot (http://www.uniprot.org) and COG (http://www.ncbi.nlm.nih.gov/COG) databases were consulted for description of specific genes.

Results and Discussion

The draft genome sequence of IPLA 31405 includes 23 contigs from 598 to 1,017,382 base pairs (bp) and is composed of 2,052,308 bp with a GC content of 38.53%. It encodes 1,874 predicted coding sequences, which were classified into 23 classes and 308 subsystems by the RAST server. Single predicted copies of the 16S, 23S and 5S rRNA genes were found, as well as 46 predicted genes for tRNAs. The RAST server allocated 40 genes within the virulence subsystem, of which 29 seem to be involved in resistance to heavy metals and toxic compounds. However, sequence homology and a similar gene context to equivalent gene clusters encoding orthologous groups of proteins in *Lactococcus lactis* and other lactic acid bacteria species argue against a role of these genes in pathogenesis. Finally, several glycosyltransferases scattered through the genome were observed, and a gene cluster likely to be involved in the synthesis of cell wall exopolysaccharides with a rhamnosyl backbone similar to that in *L. lactis* IL 1403 was also identified. However, gene clusters with the typical genetic organization of capsule operons were not encountered.

Two theta-type plasmid replicons were identified on the IPLA 31405 sequence, although the plasmid profile of this strain consist of only a single plasmid band of around 45 kbp (Figure 1). The other replicon must therefore be integrated into the plasmid molecule or into the chromosome. Two integrated phages belonging to the P335 group of *L. lactis* phages, were also recorded. In addition, a transposon highly similar to *Tn*6086 from *Enterococcus faecium* TC 6 was located in the vicinity of one of the phages. This transposon harbors an active *tet*(M) gene encoding tetracycline resistance (Figure 2). IPLA 31045 utilizes lactose through a phosphotransferase operon (*lacXGEFDCBA*) identified in the genome (Figure 3). At the DNA level, this region is 99% identical to that on the *L. lactis* plasmid pVF50, and it is likely associated to the *L. garvieae* IPLA 31405 plasmid. However, in congruence with its limited growth in milk, extracellular proteinases equivalent to the *L. lactis* PrtP were not found. A four-gene operon-like structure was identified, encoding production, resistance and secretion of a class IIb bacteriocin identical to the recently reported garvieacin Q (Figure 4).

Figure 1.- Panel A, plasmid profile of *L. garvieae* IPLA 31405. Panel B, digestion profile of the IPLA 31405 plasmid with several restriction enzymes. Order: 1, *Xho*I/*Bam*HI; 2, *Xho*I/*Stu*I; 3, *Xho*I/*Pst*I; 4, *Xho*I/N*he*I. M1, lambda DNA digested with *Eco*RI; M2, Gene ruler DNA ladder (Fermentas).

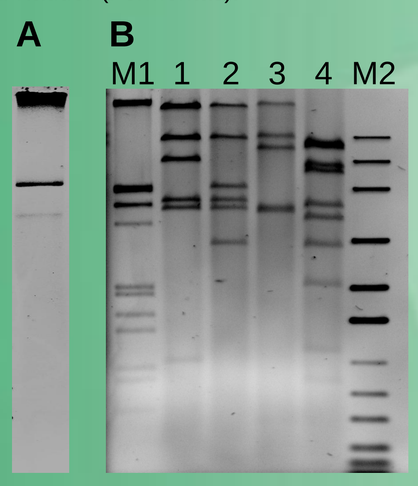


Figure 2.- Genetic organization around the *tet*(M) locus in *L. garvieae* IPLA 31405 and comparison with the same loci in other lactic acid bacteria species. Equivalent genes are denoted by the same colour.

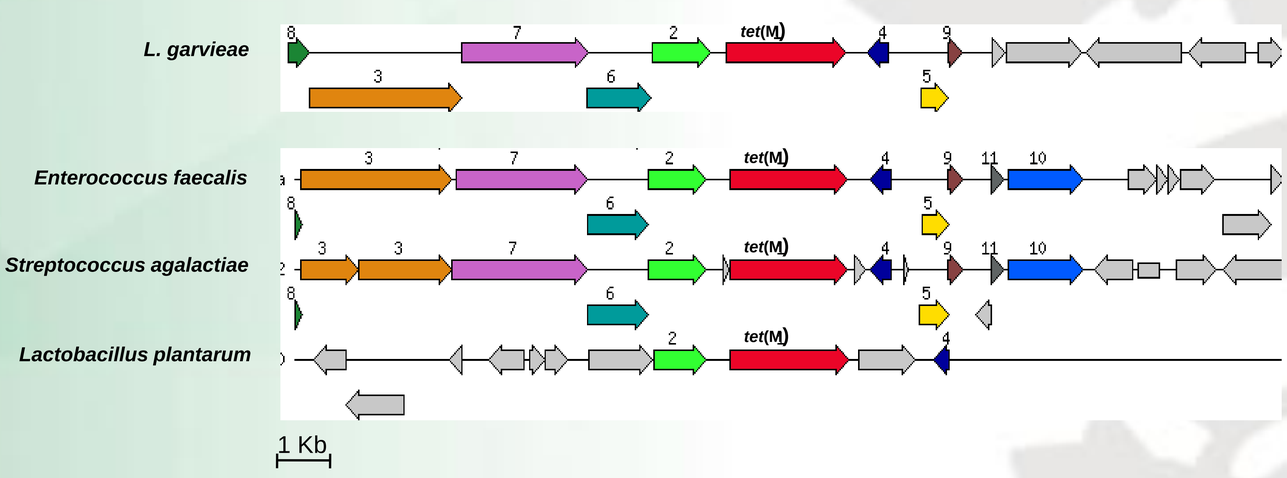


Figure 3.- Genetic organization of the lactose operon *lacXGEFDCBA* operon and its divergently transcribed represor operon in *L. garvieae* IPLA 31405. Gene order: 1, β-phospho-galactosidase (LacG); 2, tagatose-1,6-diphophate aldolase; 3, lactose-specific IIB PTS component; 4, lactose IIA PTS component; 5, tagatose-6-phosphate kinase; 6, galactose-6-phosphate isomerase (LacB); 7, galactose-6-phosphate isomerase (LacA); 8, *lacX*; 9, lactose represor (LacR).

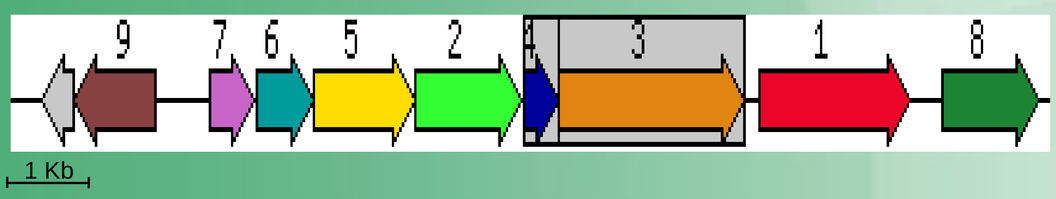
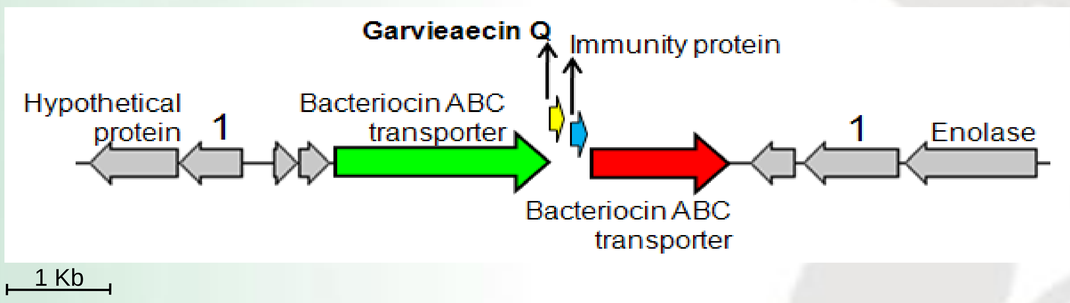


Figure 4.- Genetic organization of the garvieacin Q operon. 1, integrase-recombinase genes.



Conclusions

1.- The draft genome sequence of *L. garvieae* IPLA 31405 has been obtained and analyzed. Large rearrangements breaking the chromosome colinearity were shown when the genome sequences of IPLA 31405 and ATCC 49156 were compared.

2.- Comparison of the genome sequence of IPLA 3104 with another sequenced dairy strain (TB25), identified 220 strain-specific genes, mostly of plasmid or phage origin. In addition, 86 further genes present in both strains sheared percentages of identity lower than 60%, thus suggesting a different phylogenetic origin.

3.- In accordance with IPLA 31405 negative phenotypes (haemolysin and gelatinase activity, production the biogenic amines tyramine and histamine), evidence for virulence-related or harmful genes was not recorded, but the presence of a *tet*(M) gene in a presumably mobile element makes IPLA 31405 unacceptable as a starter.