

Comparative genomics of *Streptococcus macedonicus* ACA-DC 198 against related species within the *Streptococcus bovis*/*Streptococcus equinus* complex

Papadimitriou K.^{1,*}, Anastasiou R.¹, Georgalaki M.¹, Ferreira S.³, Supply P.^{3,4}, Papandreou N.C.², Pot B.⁴ and Tsakalidou E.¹

¹ Agricultural University of Athens, Department of Food Science and Technology, Laboratory of Dairy Research, Athens, Greece

² University of Athens, Faculty of Biology, Department of Cell Biology and Biophysics, Athens, Greece

³ Genoscreen, Genomic Platform and R&D, Campus de l'Institut Pasteur, Lille, France,

⁴ Institut Pasteur de Lille, Center for Infection and Immunity of Lille (CIIL), Lille, France

* Correspondence to: kpapadimitriou@aua.gr

Abstract

Apart from *Streptococcus thermophilus* other streptococci that can be found growing in milk belong to the *Streptococcus bovis*/*Streptococcus equinus* complex (SBSEC). Interestingly, *Streptococcus macedonicus*, which is a member of SBSEC, has been suggested to be adapted to milk and to be nonpathogenic. However, the species is phylogenetically related to *Streptococcus gallolyticus* and *Streptococcus pasteurianus* (formerly known as *S. bovis* biotypes I and II.2, respectively), which in turn are considered pathogenic, since they have been implicated in endocarditis and colon cancer in humans. Comparative analysis of the *S. macedonicus* genome with the complete genomes of its related streptococci (including that of *S. infantarius*, which is also a dairy isolate) indicated that a significant portion of the genomic organization has been conserved overall. Following a gene presence/absence strategy, we determined that *S. macedonicus* shows a reduced capacity to reside in the gastrointestinal tract of ruminants when compared to *S. gallolyticus* since it misses important genes for metabolizing complex carbohydrates of plant origin and for detoxifying this environment. *S. macedonicus* also lacks several pathogenicity traits found in *S. gallolyticus*. For example from the three pilus gene clusters (*pil1*, *pil2*, *pil3*), which may mediate the binding of *S. gallolyticus* to the extracellular matrix, *S. macedonicus* carries only one (i.e. the *pil3*). Gene gain events are also evident in the *S. macedonicus* genome sometimes originating from dairy bacteria, like the acquisition of the lactococcal plasmid pSMA198. Functional analysis of the *S. macedonicus* genome is necessary to further assess its pathogenic and technological potential.

Results and Discussion

Figure 1 - The circular map of the genome of *Streptococcus macedonicus* ACA-DC 198. Genomic features appearing from the periphery to the centre of the map: 1. Forward CDSs (blue); 2. Reverse CDSs (red); 3. Putative pseudogenes (cyan); 4. rRNA genes (orange); 5. tRNA genes (green); 6. %GC plot; 7. GC skew.

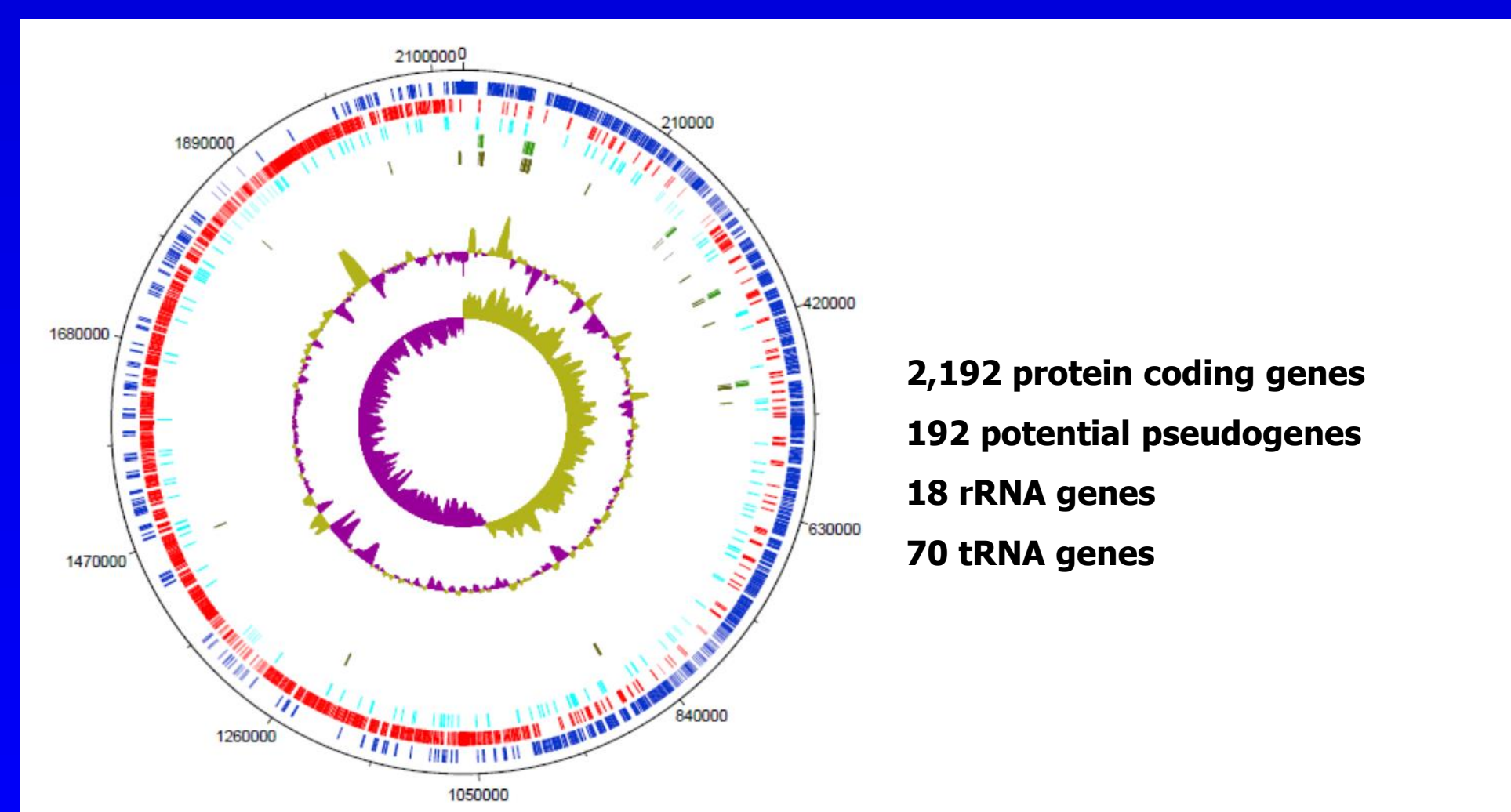


Figure 2 - Chromosome alignments of the *Streptococcus bovis*/*Streptococcus equinus* complex members as calculated by progressiveMauve. Chromosome alignments among *Streptococcus gallolyticus*, *Streptococcus macedonicus* and *Streptococcus pasteurianus* (A) and all the aforementioned streptococci and *Streptococcus infantarius* (B). Local collinear blocks (LCBs) of conserved sequences among the strains are represented by rectangles of the same colour. Connecting lines can be used to visualize synteny or rearrangement. LCBs positioned above or under the chromosome (black line) correspond to the forward and reverse orientation, respectively. The level of conservation is equivalent to the level of vertical colour filling within the LCBs (e.g. white regions are strain-specific). Sequences not placed within an LCB are unique for the particular strain.

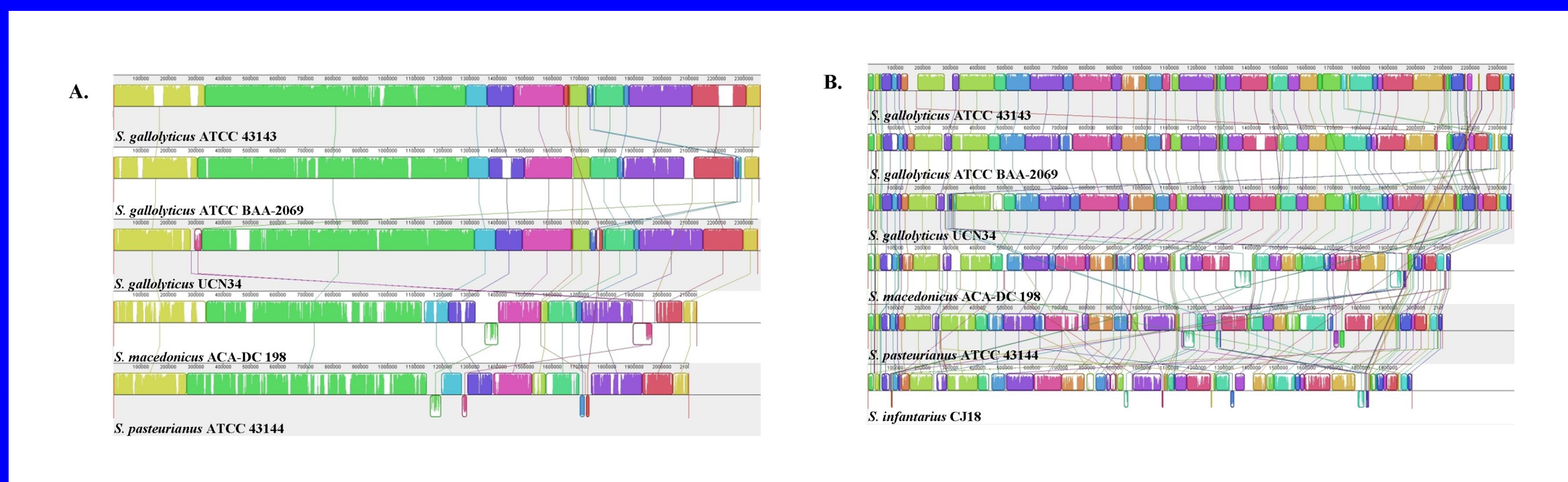


Figure 3 - Core genome analysis of members of the *Streptococcus bovis*/*Streptococcus equinus* complex. Whole CDS Venn diagrams of *Streptococcus gallolyticus*, *Streptococcus macedonicus* and *Streptococcus pasteurianus* (A) or *Streptococcus gallolyticus*, *Streptococcus infantarius*, *Streptococcus macedonicus* and *Streptococcus pasteurianus* (B). In (B) *Streptococcus gallolyticus* ATCC 43143 was selected as a representative of the *Streptococcus gallolyticus* species, since it has the longest genome size among the three sequenced strains.

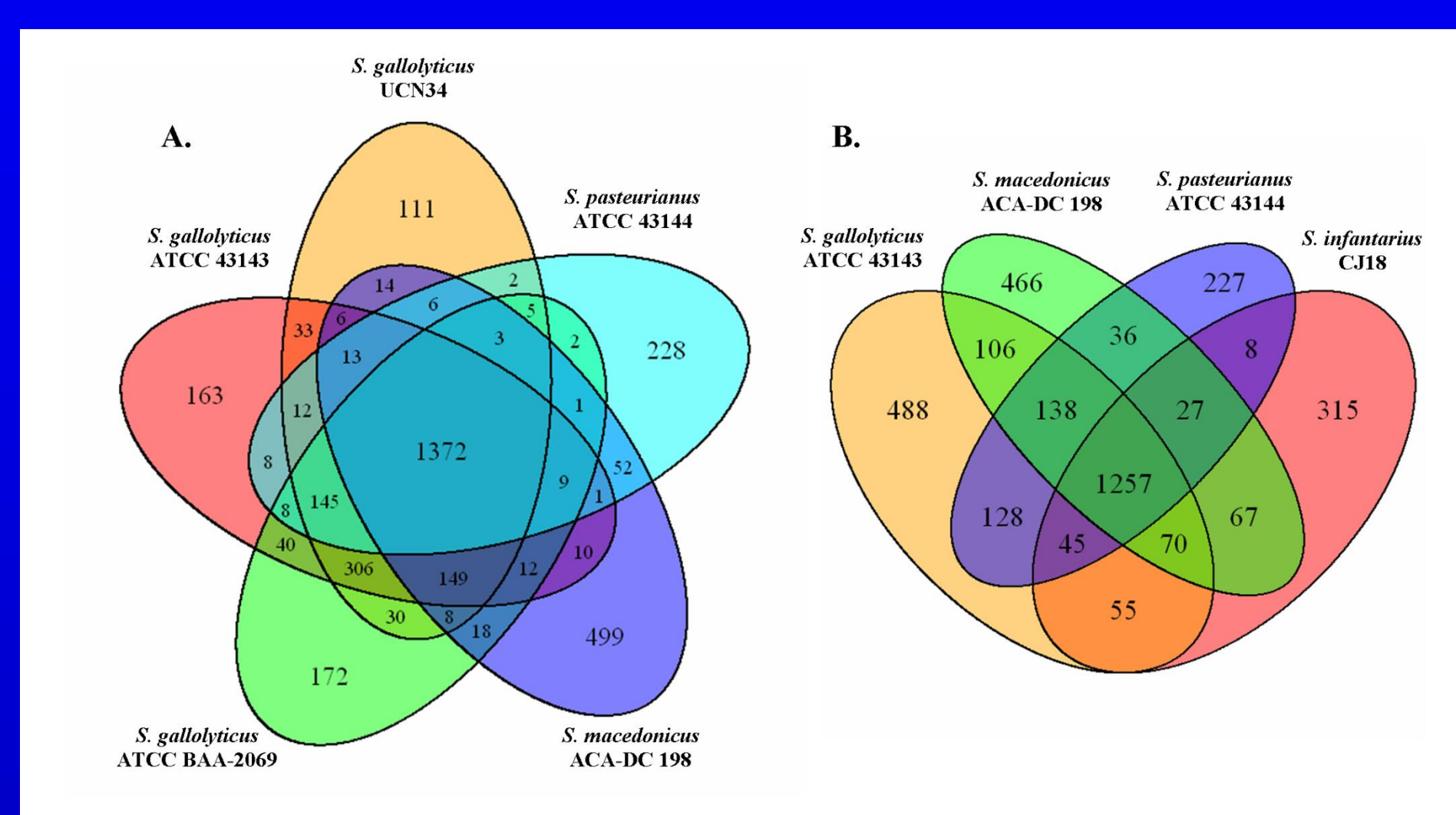


Figure 4 - The extra gene cluster for lactose and galactose metabolism found solely in *Streptococcus macedonicus* among the SBSEC members.

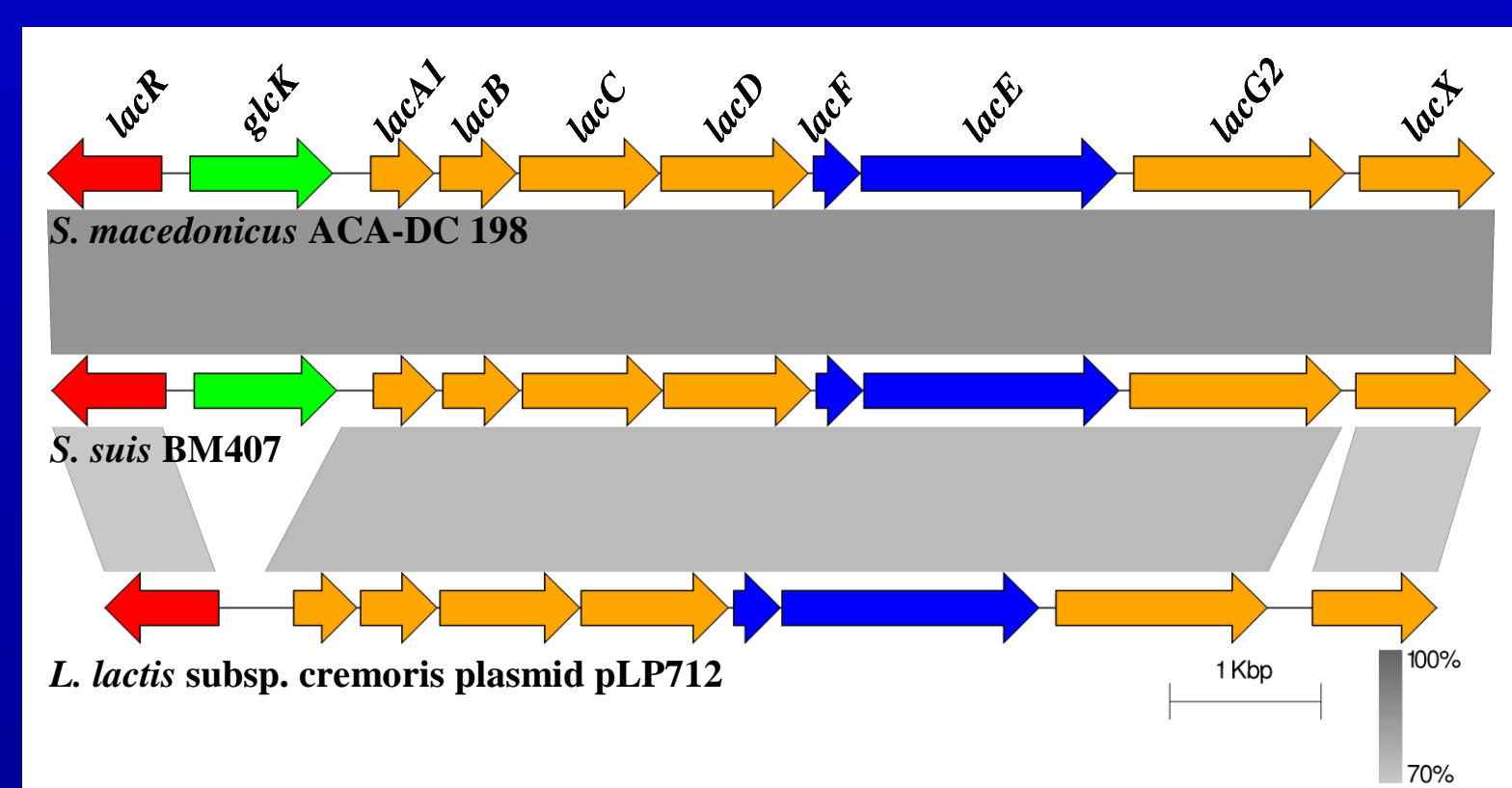


Table 1 - Genes within the *Streptococcus bovis*/*Streptococcus equinus* complex potentially involved in adaptation to the rumen.

| S. gallolyticus UCN 34 locus_tag | gene | function | S. gallolyticus ATCC BAA 2069 | S. gallolyticus ATCC 43143 | S. macedonicus ACA-DC 198 | S. pasteurianus ATCC 43144 | S. infantarius C718 |
|----------------------------------|-----------------|------------------------------------|-------------------------------|----------------------------|---------------------------|----------------------------|---------------------|
| gallo_0112 | fruA | fructan hydrolase | / | / | / | / | / |
| gallo_0330 | - | beta-1,4-endoglucanase (cellulase) | / | / | / | / | / |
| gallo_0757 | - | alpha-amylase | / | / | / | / | / |
| gallo_0162 | - | mannase | / | / | / | / | / |
| gallo_0189 | - | endo-beta-1,4-galactanase | / | / | pseudo | / | pseudo |
| gallo_1577 | - | pectate lyase | / | / | / | / | / |
| gallo_1578 | - | pectate lyase | / | / | / | / | / |
| gallo_1632 | amyE | alpha-amylase | / | / | / | / | / |
| gallo_0933 | tanA | tanins degradation | / | / | / | / | / |
| gallo_1609 | similar to tanA | tanins degradation | / | / | / | / | / |
| gallo_2106 | padC | gallic acid decarboxylation | / | / | / | / | / |
| gallo_0906 | - | gallic acid decarboxylation | / | / | / | / | / |
| gallo_0818 | bsh | bile salt hydrolase | / | / | pseudo | / | / |

Figure 5 - Map of plasmid pSMA198 isolated from *Streptococcus macedonicus* and its annotation.

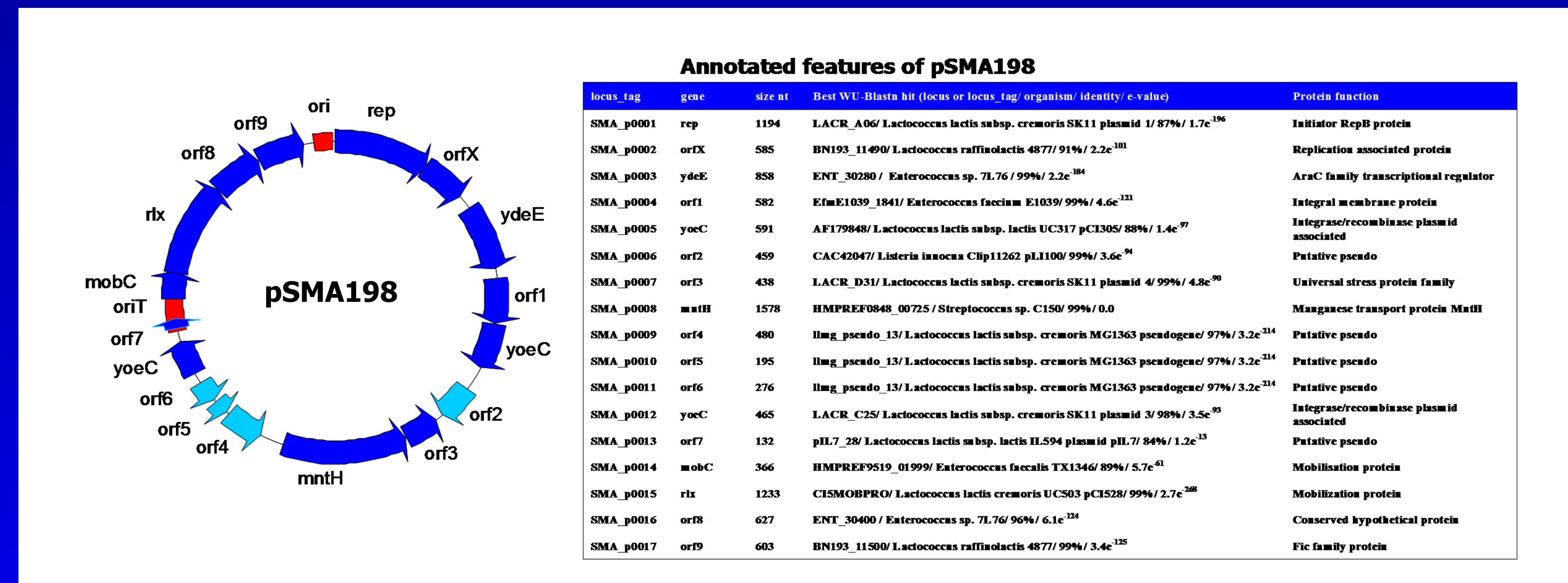


Figure 6 - Sequence alignment in a circular fashion of pSMA198 against the lactococcal pSK11b (A), pVF22 (B) and pIL5 (C) of dairy origin. Local alignments produced by BLAST are presented using ribbons whose color corresponds to four quartiles of the alignment's bitscore (red: top 25%, orange: second 25%, green: third 25% and blue: worst 25%). In order to aid orientation, the position of the ori or oriT of pSMA198 has been added in the figures. (D) Maximum likelihood tree of the pSMA198 Rep generated using the Phylogeny.fr pipeline.

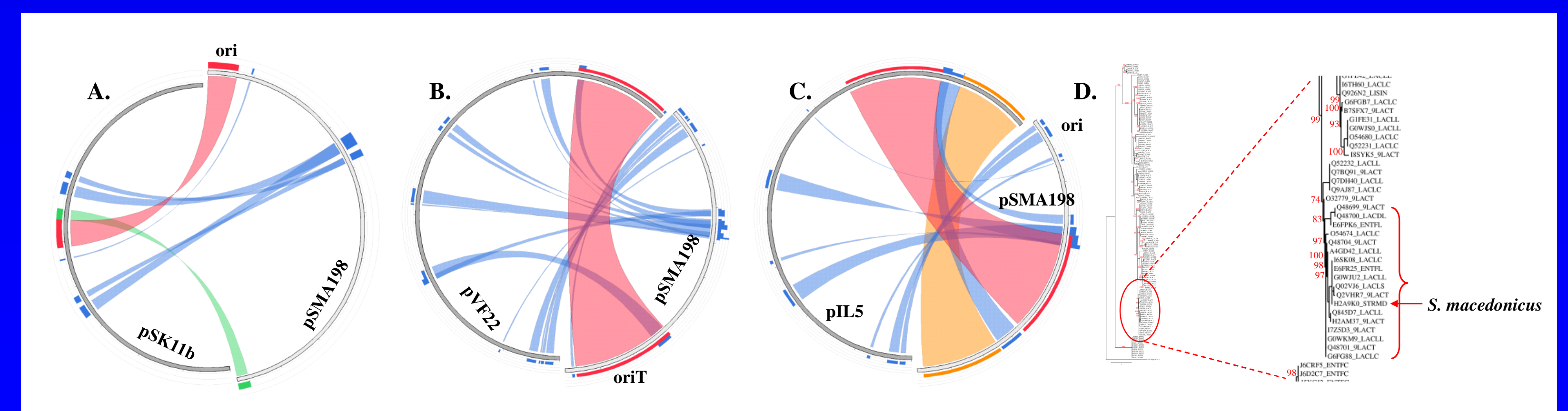


Figure 7 - Sequence alignment of chromosomal regions of *Streptococcus macedonicus* ACA-DC 198 against pIL5 (A), pVF22 (B) and pGdh442 (C) performed using Kodon. The flanking transposase gene showing high identity to orf2 (SMA_p0006) is underlined. Colored areas between the sequences correspond to different levels of identity that is depicted within the areas.

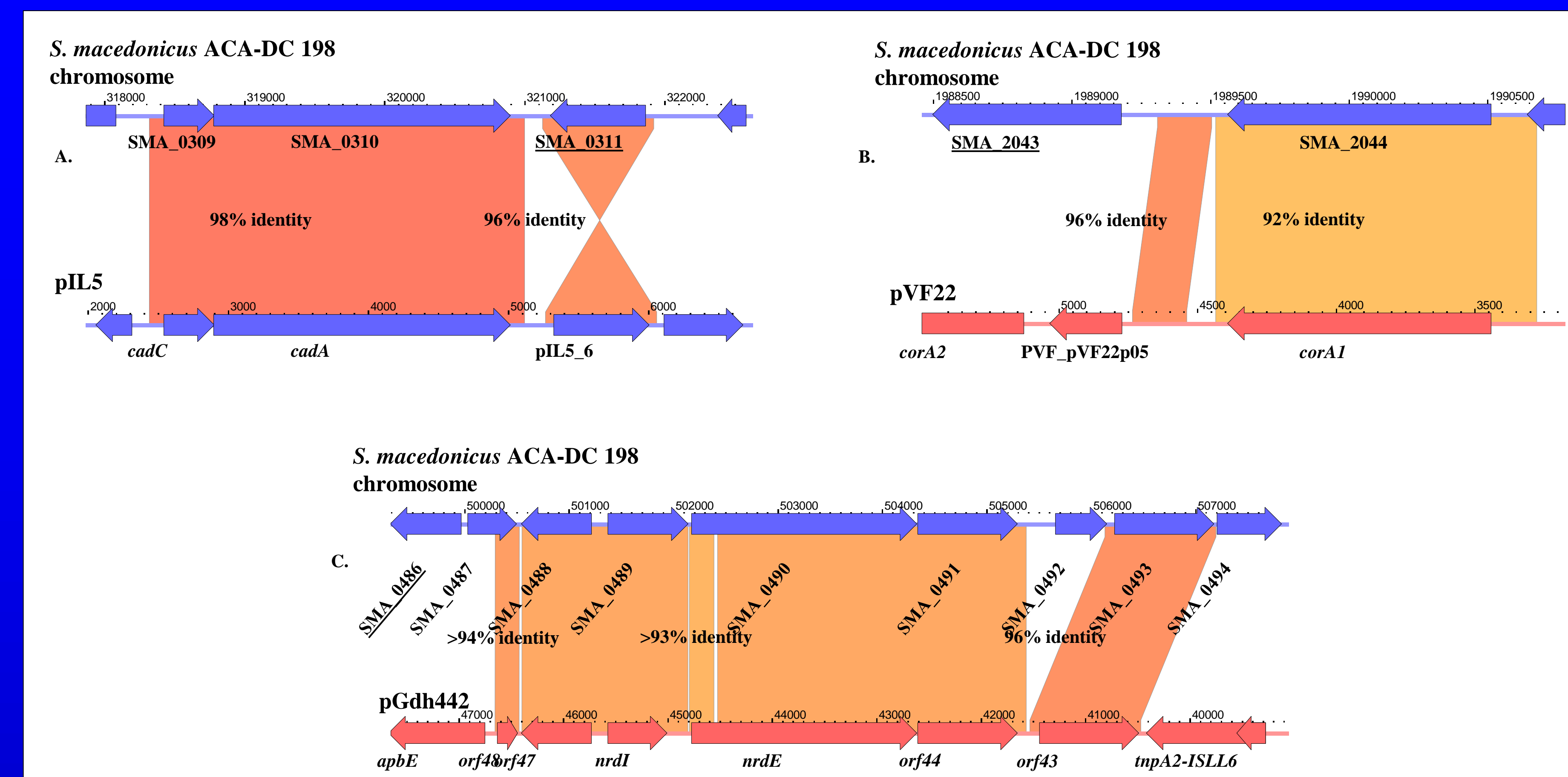


Table 2 - Genes in the *Streptococcus bovis*/*Streptococcus equinus* complex identified as putative virulence factors

| S. gallolyticus UCN 34 locus_tag | gene | function | S. gallolyticus ATCC BAA 2069 | S. gallolyticus ATCC 43143 | S. macedonicus ACA-DC 198 | S. pasteurianus ATCC 43144 | S. infantarius C718 |
|----------------------------------|------|------------------------|-------------------------------|----------------------------|---------------------------|----------------------------|---------------------|
| gallo_2179 | - | accessory pilin (pil1) | / | / | / | / | / |
| gallo_2178 | - | major pilin (pil1) | / | / | / | / | / |
| gallo_2177 | - | sortase C (pil1) | / | / | / | / | / |
| gallo_1570 | - | accessory pilin (pil2) | / | / | / | / | / |
| gallo_1569 | - | major pilin (pil2) | / | / | / | / | / |
| gallo_1568 | - | sortase C (pil2) | / | / | / | / | / |
| gallo_2040 | - | accessory pilin (pil3) | / | / | / | / | / |
| gallo_2039 | - | major pilin (pil3) | / | / | / | / | / |
| gallo_2038 | - | sortase C (pil3) | / | / | / | / | / |

Bibliography

- Papadimitriou K., Anastasiou R., Mavrogonatou E., Blom J., Papandreou N.C., Hamodrakas S.J., Ferreira S., Renault P., Supply P., Pot B., and Tsakalidou E. (2014) Comparative genomics of the dairy isolate *Streptococcus macedonicus* ACA-DC 198 against related members of the *Streptococcus bovis*/*Streptococcus equinus* complex. BMC Genomics 8:272
- Papadimitriou K., S. Ferreira, N. C. Papandreou, E. Mavrogonatou, P. Supply, B. Pot, and E. Tsakalidou (2012) Complete genome sequence of the dairy isolate *Streptococcus macedonicus* ACA-DC 198. J Bacteriol 194:1838-9.
- Papadimitriou K., T. Plakas, R. Anastasiou, S. Ferreira, P. Supply, P. Renault, N. C. Papandreou, B. Pot, and E. Tsakalidou (Under review) Analysis of the lactococcal plasmid pSMA198 found in *Streptococcus macedonicus* ACA-DC 198 points towards the habituation of the strain to the dairy environment.

Acknowledgments

The present work was cofinanced by the European Social Fund and the National resources EPEAEK and YYPEPTH through the Thales project.

