Whole-Genome Sequence of the Cheese Isolate *Streptococcus macedonicus* 679

Konstantinos Papadimitriou,[a] Eleni Mavrogonatou,[b] Alexander Bolotin,[c] Effie Tsakalidou,[a] Pierre Renault[c]

Department of Food Science and Human Nutrition, Laboratory of Dairy Research, Agricultural University of Athens, Athens, Greece[a]; Laboratory of Cell Proliferation and Ageing, Institute of Biosciences and Applications, National Centre for Scientific Research "Demokritos", Athens, Greece[b]; Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France[c]

It is well recognized that *Streptococcus macedonicus* can populate artisanal fermented foods, especially those of dairy origin. However, the safety of *S. macedonicus* remains to be established. Here, we present the whole-genome sequence of strain 679, which was isolated from a French uncooked semihard cheese made with cow milk.

*S*treptococcus macedonicus has been identified as part of the fermenting flora of dairy foods around the world, and it presents some common traits to the well-established dairy starter *Streptococcus thermophilus* (1, 2). However, *S. macedonicus* belongs to the *Streptococcus bovis/Streptococcus equinus* complex, which includes species with a zoonotic potential that have been associated, among other conditions, with endocarditis, meningitis, and colon cancer (2, 3). Here, we present the genome sequence of *S. macedonicus* 679 isolated from French uncooked semihard cheese made with cow milk (4) that may facilitate the assessment of its safety when ingested with foods.

The genome of *S. macedonicus* 679 was sequenced by mate-pair Illumina sequencing (Mr DNA, Shallowater, TX). The library was prepared using the Nextera mate-pair sample prep kit (Illumina, San Diego, CA). Genome DNA (gDNA) was isolated and quantified. Subsequently, the sample underwent fragmentation, strand displacement, circularization, shearing, streptavidin purification, end repair, adenylation, and adapter ligation. The ligated adapters were utilized during a limited-cycle (10 cycles) PCR. Following the library preparation, the average library size was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The library was sequenced by using the 600-cycle version 3 reagent kit (Illumina) in MiSeq (Illumina). Reads were assembled using SPAdes (5) against the published sequences of the chromosome and plasmid pSMA198 of *S. macedonicus* ACA-DC 198 as templates (6, 7). This reference-driven assembly resulted in two chromosomal contigs (contig_1, 79,307 bp; contig_2, 2,014,050 bp), as well as one plasmid contig (contig_3, 14,059 bp).

We employed an annotation transfer strategy using PANNOTATOR (8) and as reference the annotation of the *S. macedonicus* ACA-DC 198 genome. PANNOTATOR annotations were further enriched with RAST (9). The annotation transfer resulted in a total of 2,246 coding sequences (CDSs) in *S. macedonicus* 679 (84 in contig_1, 2,143 in contig_2, and 19 in contig_3). Among these CDSs, 75.1%, 7.7%, and 17.2% were 100%, 100 to 70%, and <70% identical to genes of *S. macedonicus*

ACA-DC 198, respectively. We then inspected manually all annotation transfers of pseudogenes from strain ACA-DC 198 to strain 679. We identified 155 potential pseudogenes. This increased percentage of pseudogenes in the genome of *S. macedonicus* 679 is in agreement with previous observations for *S. macedonicus* ACA-DC 198, suggesting a genome decay evolution pattern that may be related to adaptation of the species to the rich in nutrients environment of milk (2, 6). Even though the majority of pseudogenes were common between the two strains, we could identify some that were present in one but not the other strain. This clearly indicates that genome decay took place early in the evolution of the species, but strain-specific selective pressures also generated distinct pseudogenes. The *S. macedonicus* 679 genome needs further investigation to better establish *in silico* its safety, its technological potential, and its adaptation traits to the dairy environment.

Accession number(s). The annotated genome sequence is available at the European Nucleotide Archive under the accession numbers FLZS01000001 to FLZS01000003.

REFERENCES

1. **De Vuyst L, Tsakalidou E.** 2008. *Streptococcus macedonicus*, a multifunctional and promising species for dairy fermentations. Int Dairy J **18:** 476–485. http://dx.doi.org/10.1016/j.idairyj.2007.10.006.
2. **Papadimitriou K, Anastasiou R, Mavrogonatou E, Blom J, Papandreou NC, Hamodrakas SJ, Ferreira S, Renault P, Supply P, Pot B, 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 **15:**272. http://dx.doi.org/10.1186/1471-2164-15-272.
3. **Jans C, de Wouters T, Bonfoh B, Lacroix C, Kaindi DW, Anderegg J, Böck D, Vitali S, Schmid T, Isenring J, Kurt F, Kogi-Makau W, Meile L.** 2016. Phylogenetic, epidemiological and functional analyses of the *Streptococcus bovis/Streptococcus equinus* complex through an overarching MLST scheme. BMC Microbiol **16:**117. http://dx.doi.org/10.1186/s12866-016-0735-2.
4. **Almeida M, Hébert A, Abraham AL, Rasmussen S, Monnet C, Pons N, Delbès C, Loux V, Batto JM, Leonard P, Kennedy S, Ehrlich SD, Pop M, Montel MC, Irlinger F, Renault P.** 2014. Construction of a dairy microbial genome catalog opens new perspectives for the metagenomic analysis of

dairy fermented products. BMC Genomics **15**:1101. http://dx.doi.org/10.1186/1471-2164-15-1101.

5. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

6. **Papadimitriou K, Anastasiou R, Maistrou E, Plakas T, Papandreou NC, Hamodrakas SJ, Ferreira S, Supply P, Renault P, Pot B, Tsakalidou E.** 2015. Acquisition through horizontal gene transfer of plasmid pSMA198 by *Streptococcus macedonicus* ACA-DC 198 points towards the dairy origin of the species PLoS One **10**:e0116337. http://dx.doi.org/10.1371/journal.pone.0116337.

7. **Papadimitriou K, Ferreira S, Papandreou NC, Mavrogonatou E, Supply P, Pot B, Tsakalidou E.** 2012. Complete genome sequence of the dairy isolate *Streptococcus macedonicus* ACA-DC 198. J Bacteriol **194**:1838–1839.

8. **Santos AR, Barbosa E, Fiaux K, Zurita-Turk M, Chaitankar V, Kamapantula B, Abdelzaher A, Ghosh P, Tiwari S, Barve N, Jain N, Barh D, Silva A, Miyoshi A, Azevedo V.** 2013. PANNOTATOR: an automated tool for annotation of pan-genomes. Genet Mol Res **12**:2982–2989. http://dx.doi.org/10.4238/2013.August.16.2.

9. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R.** 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res **42**:D206–D214. http://dx.doi.org/10.1093/nar/gkt1226.