Genome Sequences of Four Italian Streptococcus thermophilus Strains of Dairy Origin.

Treu, Laura; Vendramin, Veronica; Bovo, Barbara; Campanaro, Stefano; Corich, Viviana; Giacomini, Alessio

Published in:
Genome Announcements

Link to article, DOI:
10.1128/genomeA.00126-14

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
S. thermophilus is a species of great importance for the preparation of numerous dairy products (1). In fact, besides lactic acid production, several other technological characteristics are possessed by strains of this species. The first two strains of S. thermophilus, TH982 and TH985, were isolated in the Campania region from whey and curd, respectively, obtained from mozzarella of bufala campana DOP cheese (pasta filata cheese) production. S. thermophilus strains TH1477 and 1F8CT were isolated in the Veneto region from cow milk and curd, respectively, from Grana Padano DOP cheese (long-ripening cheese) production. These strains possess a variety of functional properties of technological interest for the dairy industry, such as slow (1F8CT) or rapid (TH1477, TH982, and TH985) acidifying capabilities, low fermentation and growth rates (1F8CT), and pigment biosynthesis (TH1477).

Here, we present the genome sequences of S. thermophilus strains TH982, TH985, TH1477, and 1F8CT, generated using an Illumina MiSeq platform (with 1-kb mate-pair libraries) at the Ramaciotti Centre, Sydney, Australia. Sequencing coverages of 218×, 183×, 142×, and 261× were obtained corresponding to 1,635,998, 1,410,685, 1,258,918, and 1,544,110 paired-end reads, respectively (2× 250 bp). The files generated were assembled with Velvet version 1.2.10 (2) and ABySS software version 1.3.5 (3) (optimal k-mer, 131). The consensus sequences of the two assemblies were manually compared. Between 52 and 84 scaffolds were obtained, with total sizes of 1,733,024, 1,838,250, 1,879,471, and 1,742,121 bp for strains TH982, TH985, TH1477, and 1F8CT, respectively, with a G+C content of 39%. All strains had the scaffolds assembled into a single circular chromosome by aligning them against the reference genome of S. thermophilus CNRZ1066 (assembly ASM1184v1). Several plasmid sequences were detected by BLAST analysis in two scaffolds of strain TH985, in three of TH1477, and in five of 1F8CT.

Protein-coding open reading frames (ORFs) were predicted and annotated using the RAST server (4). The numbers of predicted protein-coding genes (CDSs) in TH982, TH985, TH1477, and 1F8CT strains were 1,924, 1,952, 1,986 and 1,864, respectively. Furthermore, 47, 69, 55, and 51 RNA genes were found, respectively. These genomes contain few phage sequences and no transposase-coding genes, while 20 to 26 clusters of regularly interspaced short palindromic repeats (CRISPRs) were found.

A comparison with the reference genome of S. thermophilus CNRZ1066 highlighted numerous CDSs that were exclusive to each strain. In particular, strains from the Veneto region include numerous sequences belonging to the oxidative stress responding genes, while those from Campania contain unique genes for cholesterol and betaine uptake and betaine biosynthesis, which are important in the osmotic stress response.

These data are intended to increase the availability of genomes of S. thermophilus strains of dairy origin (5, 6) in order to better understand their biodiversity and their known and potential technological properties.

Nucleotide sequence accession numbers. The sequences of this whole-genome shotgun project have been deposited at DDBJ/EMBL/GenBank under the accession no. AZTL00000000, AZTM00000000, AZTJ00000000, and AZTK00000000 for S. thermophilus strains TH982, TH985, TH1477, and 1F8CT, respectively. The versions described in this paper are AZTL01000000, AZTM01000000, AZTJ01000000, and AZTK01000000, respectively.

ACKNOWLEDGMENTS

This study was supported in part by POR Veneto “Fondo Europeo di Sviluppo Regionale” 2007–2013—Asse 1, Azione 1.1.2. We thank Angioletta Lombardi for providing S. thermophilus strains.

REFERENCES


