CrossMark
click for updates

Complete Genome Sequence of *Streptococcus thermophilus* SMQ-301, a Model Strain for Phage-Host Interactions

Simon J. Labrie,[a] Denise M. Tremblay,[a] Pier-Luc Plante,[b] Jessica Wasserscheid,[c] Ken Dewar,[c] Jacques Corbeil,[b] Sylvain Moineau[a]

Département de Biochimie, de Microbiologie, et de Bio-Informatique, Faculté des Sciences et de Génie, Groupe de Recherche en Écologie Buccale, Faculté de Médecine Dentaire, Félix d'Hérelle Reference Center for Bacterial Viruses, Université Laval, Québec, Canada[a]; Département de Médecine Moléculaire, Faculté de Médecine, Université Laval, Québec, Canada[b]; Department of Human Genetics, McGill University and Génome Québec Innovation Centre, Montréal, Canada[c]

Streptococcus thermophilus is used by the dairy industry to manufacture yogurt and several cheeses. Using PacBio and Illumina platforms, we sequenced the genome of *S. thermophilus* SMQ-301, the host of several virulent phages. The genome is composed of 1,861,792 bp and contains 2,037 genes, 67 tRNAs, and 18 rRNAs.

Address correspondence to Sylvain Moineau, sylvain.moineau@bcm.ulaval.ca.

*S*treptococcus thermophilus is a low-GC Gram-positive bacterium widely used by the dairy industry to obtain high-quality fermented products, such as yogurt and cheeses (1). As such, it is a rare nonpathogenic streptococcal species (2). Another distinctive feature of *S. thermophilus* is its production of lactic acid from only a few sugars, including glucose, lactose, and sucrose (3). Selecting a suitable industrial *S. thermophilus* strain is a long process, encouraging extensive use and thorough characterization of the available strains (4).

S. thermophilus SMQ-301 is an industrial strain used to make cheese that is also sensitive to several virulent phages, including the reference phage DT1 (5). Phage infection of bacterial cultures is still the main risk factor for slowed milk fermentation, and the dairy industry relies on various strategies to control this phenomenon (4). *S. thermophilus* SMQ-301 has been used in several studies on phage biology (5–10), clustered regularly interspaced short palindromic repeat (CRISPR)-Cas systems (11–13), and sugar metabolism (3, 14, 15). It has a similar pulsed-field gel electrophoresis restriction profile (data not shown) to that of *S. thermophilus* LMD-9 (1,856,368 bp, 39.1% G+C content, 1,834 genes, two plasmids; GenBank accession no. NC_008532), which is also sensitive to DT1 (10).

The genome of *S. thermophilus* SMQ-301 was sequenced using MiSeq (Illumina) and PacBio (Pacific Biosciences) platforms. DNA extraction, library preparation, and assembly were performed as described previously (16). Briefly, DNA was purified using Genomic-tip 20/G columns, according to the manufacturer's instructions. The genome was assembled into a single contig with the PacBio reads, according to the manufacturer's instructions, and base calling accuracy was verified and corrected using Illumina reads aligned with BLAT (17) to the PacBio assembly. The Illumina and PacBio data were highly concordant, with the exception of 27 loci that were corrected with the Illumina reads. The genome of 1,861,792 bp has a G+C content of 39.1%. SMQ-301 contains no plasmids. The RAST annotation server (18) was used to annotate the genome, which encodes 2,037 proteins, 67 tRNAs, and 18 rRNAs.

We identified three CRISPR-Cas loci in the genome of *S. thermophilus* SMQ-301. According to the classification of CRISPR-Cas systems (19), two of the loci are type II-A systems (CRISPR 1 and CRISPR 3), while the CRISPR-Cas 2 locus is a type III-A system. Moreover, SMQ-301 has 16 spacers in CRISPR 1, 3 spacers in CRISPR 2, and 15 spacers in CRISPR 3. The last 10 spacers (3′) of CRISPR 1 are identical to those of LMD-9, which also has 16 spacers. Both strains have the same 3 spacers in CRISPR 2, and they share 5 spacers in CRISPR 3, while LMD-9 has only 8 spacers in this locus. The genome of SMQ-301 also encodes two type-I restriction-modification systems. Although prophage genes were detected by Phast (20), no complete prophage is present.

No known toxins were identified in the coding sequence of *S. thermophilus* SMQ-301 by the Web server VirulenceFinder (21) or by comparing the protein sequences with the VFDB (22) and DBETH (23) toxin databases using BLASTp (24). Finally, no known antibiotic resistance genes were found in the genome of SMQ-301 in comparison with the antibiotic resistance database (ARG-ANNOT) (25).

Nucleotide sequence accession number. The complete annotated genome sequence of *S. thermophilus* SMQ-301 was deposited in GenBank under accession no. CP011217.

REFERENCES

1. **Quiberoni A, Moineau S, Rousseau GM, Reinheimer J, Ackermann H.** 2010. *Streptococcus thermophilus* bacteriophages. Int Dairy J **20:**657–664. http://dx.doi.org/10.1016/j.idairyj.2010.03.012.
2. **Delorme C.** 2008. Safety assessment of dairy microorganisms: *Streptococ-*

cus thermophilus. Int J Food Microbiol **126:**274–277. http://dx.doi.org/10.1016/j.ijfoodmicro.2007.08.014.

3. **Vaillancourt K, LeMay JD, Lamoureux M, Frenette M, Moineau S, Vadeboncoeur C.** 2004. Characterization of a galactokinase-positive recombinant strain of *Streptococcus thermophilus*. Appl Environ Microbiol **70:**4596–4603. http://dx.doi.org/10.1128/AEM.70.8.4596-4603.2004.

4. **Samson JE, Moineau S.** 2013. Bacteriophages in food fermentations: new frontiers in a continuous arms race. Annu Rev Food Sci Technol **4:**347–368. http://dx.doi.org/10.1146/annurev-food-030212-182541.

5. **Tremblay DM, Moineau S.** 1999. Complete genomic sequence of the lytic bacteriophage DT1 of *Streptococcus thermophilus*. Virology **255:**63–76. http://dx.doi.org/10.1006/viro.1998.9525.

6. **Duplessis M, Moineau S.** 2001. Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. Mol Microbiol **41:**325–336. http://dx.doi.org/10.1046/j.1365-2958.2001.02521.x.

7. **Duplessis M, Russell WM, Romero DA, Moineau S.** 2005. Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. Virology **340:**192–208. http://dx.doi.org/10.1016/j.virol.2005.05.033.

8. **Lamothe G, Lévesque C, Bissonnette F, Cochu A, Vadeboncoeur C, Frenette M, Duplessis M, Tremblay D, Moineau S.** 2005. Characterization of the *cro-ori* region of the *Streptococcus thermophilus* virulent bacteriophage DT1. Appl Environ Microbiol **71:**1237–1246. http://dx.doi.org/10.1128/AEM.71.3.1237-1246.2005.

9. **Duplessis M, Lévesque CM, Moineau S.** 2006. Characterization of *Streptococcus thermophilus* host range phage mutants. Appl Environ Microbiol **72:**3036–3041. http://dx.doi.org/10.1128/AEM.72.4.3036-3041.2006.

10. **Goh YJ, Goin C, O'Flaherty S, Altermann E, Hutkins R.** 2011. Specialized adaptation of a lactic acid bacterium to the milk environment: the comparative genomics of *Streptococcus thermophilus* LMD-9. Microb Cell Fact **10**(Suppl 1):S22. http://dx.doi.org/10.1186/1475-2859-10-S1-S22.

11. **Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S.** 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J Bacteriol **190:**1390–1400. http://dx.doi.org/10.1128/JB.01412-07.

12. **Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R.** 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. J Bacteriol **190:**1401–1412. http://dx.doi.org/10.1128/JB.01415-07.

13. **Levin BR, Moineau S, Bushman M, Barrangou R.** 2013. The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. PLoS Genet **9:**e1003312. http://dx.doi.org/10.1371/journal.pgen.1003312.

14. **Cochu A, Roy D, Vaillancourt K, LeMay J.-D, Casabon I, Frenette M, Moineau S, Vadeboncoeur C.** 2005. The doubly phosphorylated form of HPr, HPr(Ser-P)(His~P), is abundant in exponentially growing cells of *Streptococcus thermophilus* and phosphorylates the lactose transporter LacS as efficiently as HPr(His~P). Appl Environ Microbiol **71:**1364–1372. http://dx.doi.org/10.1128/AEM.71.3.1364-1372.2005.

15. **Vaillancourt K, Bédard N, Bart C, Tessier M, Robitaille G, Turgeon N, Frenette M, Moineau S, Vadeboncoeur C.** 2008. Role of *galK* and *galM* in galactose metabolism by *Streptococcus thermophilus*. Appl Environ Microbiol **74:**1264–1267. http://dx.doi.org/10.1128/AEM.01585-07.

16. **Labrie SJ, El Haddad L, Tremblay DM, Plante PL, Wasserscheid J, Dumaresq J, Dewar K, Corbeil J, Moineau S.** 2014. First complete genome sequence of *Staphylococcus xylosus*, a meat starter culture and a host to propagate *Staphylococcus aureus* phages. Genome Announc **2**(4):e00671-14. http://dx.doi.org/10.1128/genomeA.00671-14.

17. **Kent WJ.** 2002. BLAT–the BLAST-like alignment tool. Genome Res **12:**656–664. http://dx.doi.org/10.1101/gr.229202.

18. **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.

19. **Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV.** 2011. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol **9:**467–477. http://dx.doi.org/10.1038/nrmicro2577.

20. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39:**W347–W352. http://dx.doi.org/10.1093/nar/gkr485.

21. **Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM.** 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol **52:**1501–1510. http://dx.doi.org/10.1128/JCM.03617-13.

22. **Chen L, Xiong Z, Sun L, Yang J, Jin Q.** 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res **40:**D641–D645. http://dx.doi.org/10.1093/nar/gkr989.

23. **Chakraborty A, Ghosh S, Chowdhary G, Maulik U, Chakrabarti S.** 2012. DBETH: a database of bacterial exotoxins for human. Nucleic Acids Res **40:**D615–D620. http://dx.doi.org/10.1093/nar/gkr942.

24. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.

25. **Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM.** 2014. Arg-Annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother **58:**212–220. http://dx.doi.org/10.1128/AAC.01310-13.