AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequences of *emm111* Type *Streptococcus pyogenes* Strain GUR, with Antitumor Activity, and Its Derivative Strain GURSA1 with an Inactivated *emm* Gene

Maria A. Suvorova,[a] Anna N. Tsapieva,[a] Emilie Glad Bak,[b] Valery A. Chereshnev,[c] Ekaterina P. Kiseleva,[a] Alexander N. Suvorov,[a,d] Manimozhiyan Arumugam[b]

Department of Molecular Microbiology, Institute of Experimental Medicine, Saint Petersburg, Russian Federation[a]; The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark[b]; Institute of Immunology and Physiology, Russian Academy of Sciences, Ural Branch, Ekaterinburg, Russian Federation[c]; Saint Petersburg State University, Saint Petersburg, Russian Federation[d]

ABSTRACT We present here the complete genome sequence of *Streptococcus pyogenes* type *emm111* strain GUR, a throat isolate from a scarlet fever patient, which has been used to treat cancer patients in the former Soviet Union. We also present the complete genome sequence of its derivative strain GURSA1 with an inactivated *emm* gene.

Streptococcus pyogenes is a Gram-positive pathogenic bacterium causing a wide variety of human diseases (1). It has also been used for bacteriolytic cancer therapy starting with "Coley's toxins" (2) and followed by OK-432 (picibanil) (3), but the mechanisms behind its antitumor activity are unknown. *S. pyogenes* strain GUR, isolated from a scarlet fever patient in the former Soviet Union in 1938, has been used clinically for anticancer treatment in the former Soviet Union for more than 20 years (4) and showed therapeutic effects in several murine cancer models (M. Suvorova, E. P. Kiseleva, and A. N. Suvorov, unpublished data). To identify the role of the antiphagocytic M protein in the cytotoxic activity of this strain, we generated a mutant (strain GURSA1) by inactivating the *emm* gene by insertional mutagenesis (5). Unexpectedly, strain GURSA1 showed better cytotoxic activity in both *in vitro* and *in vivo* experiments on murine malignant tumor cells (M. Suvorova, E. P. Kiseleva, and A. N. Suvorov, unpublished data). To understand the mechanisms of how strains GUR and GURSA1 affect tumor cells, we sequenced their genomes.

Genomic DNA was extracted from overnight cultures of *S. pyogenes* strains GUR and GURSA1 using the MagAttract pathogen kit (Qiagen, USA). Fragment libraries were prepared using the Nextera XT DNA library preparation kit (Illumina, San Diego, CA), followed by 251-bp paired-end sequencing on a MiSeq instrument (Illumina), which This generated 496 Mb and 234 Mb of high-quality paired-end sequence data for strains GUR and GURSA1, respectively.

We first assembled the genome of strain GURSA1 using a custom-built SPAdes assembler version 3.10.0 (6), with the "-k 55,77,99,121,143,151,155" option to specify iterative k-mers. For scaffolding, we used full chromosome sequences from three closely related *S. pyogenes* strains (ATCC 19615, AP1, and AP53, with GenBank accession numbers NZ_CP008926, NZ_CP007537, and NZ_CP013672, respectively) using the "--untrusted-contigs" option. The resulting single scaffold was improved using Pilon version 1.22 software (7). We circularized this gapless chromosome sequence and

Received 27 July 2017 **Accepted** 31 July 2017 **Published** 21 September 2017

Citation Suvorova MA, Tsapieva AN, Bak EG, Chereshnev VA, Kiseleva EP, Suvorov AN, Arumugam M. 2017. Complete genome sequences of *emm111* type *Streptococcus pyogenes* strain GUR, with antitumor activity, and its derivative strain GURSA1 with an inactivated *emm* gene. Genome Announc 5: e00939-17. https://doi.org/10.1128/genomeA.00939-17.

Address correspondence to Manimozhiyan Arumugam, arumugam@sund.ku.dk.

rearranged it to start at the *dnaA* gene using the Circlator software (8). We then assembled the genome of strain GUR using an identical procedure, with the following modifications: different k-mers with "-k 33,55,77,99,121,143" and providing the chromosome sequence of GURSA1 with the "--untrusted-contigs" option. The resulting single scaffold was also improved, circularized, and rearranged. The genomes of strains GUR and GURSA1 consisted of 1,890,204 bp and 1,893,171 bp, respectively. They encoded 1,820 and 1,825 proteins, respectively, predicted by NCBI's Prokaryotic Genome Annotation Pipeline (9). Both genomes had 38.5% G+C content, harbored 67 tRNA genes, and encoded 6 copies of the rRNA operon. Using the PHASTER online tool (10), we identified five intact prophage regions with high homology to prophage regions from M3 serotype *S. pyogenes* strain MGAS315. Other than the genomic changes introduced during mutagenesis, GUR and GURSA1 had four single-nucleotide changes between them. Three of these changes were synonymous mutations within the third prophage region, and the fourth change was a nonsynonymous (A284V) change in the Hpr kinase/phosphorylase gene. We assigned type *emm111* to strain GUR using the instructions provided by the Centers for Disease Control and Prevention's *Streptococcus* Laboratory (https://www.cdc.gov/streplab/).

Accession number(s). The genome sequences of strains GUR and GURSA1 were deposited in GenBank under the accession numbers CP022354 and CP022206, respectively. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work has been funded by the Novo Nordisk Foundation grant NNF10CC1016515.

We thank Statens Serum Institut, Denmark, for assistance in genome sequencing.

REFERENCES

1. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. 2014. Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. Clin Microbiol Rev 27:264–301. https://doi.org/10.1128/CMR.00101-13.
2. Starnes CO. 1992. Coley's toxins in perspective. Nature 357:11–12. https://doi.org/10.1038/357011a0.
3. Ryoma Y, Moriya Y, Okamoto M, Kanaya I, Saito M, Sato M. 2004. Biological effect of OK-432 (picibanil) and possible application to dendritic cell therapy. Anticancer Res 24:3295–3301.
4. Chereshnev VA, Morova AA, Ryamzina I. 2006. Биологические законы и жизнеспособность человека: метод многофункциональной восстановительной биотерапии, 2nd ed. Perm State Agricultural Academy Press, Perm, Russia.
5. Suvorova MA, Tsapieva AN, Duplik NV, Kramskoy TA, Grabovskaya KB, Kiseleva EP, Chereshnev VA, Suvorov AN. 2016. Конструирование штамма стрептококка, мутантного по гену М-белка. Meditsinskiy Akademicheskiy Zhurnal 16:235–236.
6. 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. https://doi.org/10.1089/cmb.2012.0021.
7. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
8. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
10. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.