

# Genome Sequence of *Streptomyces toyocaensis* NRRL 15009, Producer of the Glycopeptide Antibiotic A47934

Min Jung Kwun, Hee-Jeon Hong

Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Here we report the draft genome sequence of *Streptomyces toyocaensis* strain NRRL 15009 which is the producer of the glycopeptide antibiotic A47934. The genome sequence is predicted to harbor a total of 26 secondary metabolite biosynthetic gene clusters including the A47934 cluster.

Received 4 July 2014 Accepted 16 July 2014 Published 31 July 2014

**Citation** Kwun MJ, Hong H-J. 2014. Genome sequence of *Streptomyces toyocaensis* NRRL 15009, producer of the glycopeptide antibiotic A47934. *Genome Announc.* 2(4): e00749-14. doi:10.1128/genomeA.00749-14.

**Copyright** © 2014 Kwun and Hong. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Hee-Jeon Hong, [hh309@cam.ac.uk](mailto:hh309@cam.ac.uk).

*Streptomyces toyocaensis* strain NRRL 15009 was isolated from a sandy soil sample in the state of Washington, USA and described as the producer of A47934 (1). A47934 belongs to the glycopeptide family of antibiotics, an important class of antibiotics which also includes vancomycin and teicoplanin which are currently reserved in the clinic as the last-resort treatment of infectious disease caused by Gram-positive pathogens that are resistant to the antibiotics in mainstream use. A47934 is a sugarless glycopeptide antibiotic with a heptapeptide backbone structure closely related to that of teicoplanin, but is unique among reported glycopeptides as it contains a sulfate ester group. The complete sequence of the A47934 synthetic gene cluster in *S. toyocaensis* NRRL 15009 has previously been determined, and consists of 34 contiguous open reading frames (ORFs) encompassing 69 kb of chromosomal DNA (GenBank accession no. STU82965) (2). Here we present the first draft of the entire genome sequence of *S. toyocaensis* NRRL 15009.

The genome was sequenced to a coverage depth of 53-fold using a 454 GS-FLX instrument (Roche) with 454 titanium chemistry, and assembled using Newbler v2.5.3. This produced 87 contigs from 1,110,442 reads, and a total estimated genomic size of 7,341,583 bp. The genome has a G+C content of 72.04%. The average contig size is 97,848 bp and the largest contig size is 891,334 bp. Genome annotation was performed by the Prokaryotic Genome Annotation Pipeline (PGAP) with GeneMarkS+ (3, 4), and identified 6,315 coding sequence (CDS) regions, 82 pseudogenes, and 65 tRNAs. A single copy of each 5S rRNA, 16S rRNA, and 23S rRNA were also predicted in the draft genome sequence. The biosynthetic cluster of A47934 is located at 254,157 bp to 323,457 bp in the draft genome (on contig 1). This region also includes genes (*vanRst*, *vanSst*, *vanHst*, *vanAst*, *vanXst*, *murX*, *staO*, and *staP*) required for conferring self-resistance to the A47934 antibiotic. Using antiSMASH (5), a further 25 biosynthetic clusters potentially capable of producing secondary metabolites were also identified. These include three predicted to encode production of lantipeptides, two bacteriocin clusters, two siderophore clusters, three butyrolactone clusters, four nonribosomal peptide synthase (NRPS) clusters, four clusters encoding ter-

penes, a phenazine cluster, the ectoine biosynthesis cluster, and one type II-polyketide synthase (PKS) cluster.

The availability of the *S. toyocaensis* NRRL 15009 genome sequence will consequently assist with the search for new bioactive compounds of microbial origin, and will facilitate studies into resistance mechanisms and the regulation of secondary metabolite production.

**Nucleotide sequence accession numbers.** The draft genome sequence of *S. toyocaensis* NRRL 15009 has been deposited in the DDBJ/EMBL/GenBank database under the accession no. JFCB00000000. The version described in this paper is version JFCB01000000.

## ACKNOWLEDGMENTS

This work was supported by grants from the Royal Society (516002.K5877/ROG) and the Medical Research Council (G0700141). Sequencing was performed by the DNA sequencing facility at the Department of Biochemistry, University of Cambridge (<http://www.bioc.cam.ac.uk/dnasequencing>).

We thank Gerry Wright for kindly supplying the *S. toyocaensis* NRRL15009 strain for sequencing.

## REFERENCES

- Boeck LD, Mertz FP. 1986. A47934, a novel glycopeptide-aglycone antibiotic produced by a strain of *Streptomyces toyocaensis* taxonomy and fermentation studies. *J. Antibiot.* 39:1533–1540. <http://dx.doi.org/10.7164/antibiotics.39.1533>.
- Pootoolal J, Thomas MG, Marshall CG, Neu JM, Hubbard BK, Walsh CT, Wright GD. 2002. Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci. U. S. A.* 99:8962–8967. <http://dx.doi.org/10.1073/pnas.102285099>.
- Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity GM, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *Omics* 12:137–141. <http://dx.doi.org/10.1089/omi.2008.0017>.
- Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115. <http://dx.doi.org/10.1093/nar/26.4.1107>.
- Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* 41:W204–W212. <http://dx.doi.org/10.1093/nar/gkt449>.