

## Genome Sequences of Three Tunicamycin-Producing *Streptomyces* Strains, *S. chartreusis* NRRL 12338, *S. chartreusis* NRRL 3882, and *S. lysosuperificus* ATCC 31396

James R. Doroghazi,<sup>2</sup> Kou-San Ju,<sup>2</sup> Daren W. Brown,<sup>1</sup> David P. Labeda,<sup>1</sup> Zixin Deng,<sup>4</sup> William W. Metcalf,<sup>2</sup> Wenqing Chen,<sup>3,4\*</sup> and Neil P. J. Price<sup>1\*</sup>

USDA, ARS, National Center for Agricultural Utilization Research, 1815 N. University St., Peoria, Illinois 61604<sup>1</sup>; Department of Microbiology and Institute for Genomic Biology at University of Illinois, Urbana-Champaign, Illinois<sup>2</sup>; Key Laboratory of Combinatorial Biosynthesis & New Drug Discovery, Ministry of Education, and School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, People's Republic of China<sup>3</sup>; and Laboratory of Microbial Metabolism and School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai 200030, People's Republic of China<sup>4</sup>

Received 27 September 2011/Accepted 28 September 2011

**We announce the sequencing of *Streptomyces chartreusis* NRRL 12338 and NRRL 3882 and *Streptomyces lysosuperificus* ATCC 31396. These are producers of tunicamycins, chartreusins, cephalosporins, holomycins, and calcimycin. The announced genomes, together with the published *Streptomyces clavuligerus* genome, will facilitate data mining of these secondary metabolites.**

The tunicamycins are structurally unique compounds that inhibit *N*-acetylhexosamine-1-phosphate–prenol phosphate transferases, blocking formation of *N*-acetylmuramylpentapeptide-undecanol pyrophosphate in bacteria or *N*-acetylglucosamine-dolichol pyrophosphate in eukaryotes. These are essential intermediates in the biosynthesis of eukaryotic *N*-glycoproteins and of bacterial peptidoglycan. Tunicamycins were discovered by Tamura from *Streptomyces lysosuperificus* ATCC 31396 (10). They are nucleotide-type antibiotics that contain two unusual features: an 11-carbon dialdose sugar called tunicamine and an anomeric-to-anomeric  $\alpha,\beta$ -1,11-glycosidic linkage between the tunicamine and *N*-acetylglucosamine. Tunicamycins are also produced by two strains of *Streptomyces chartreusis* (NRRL 3882 and NRRL 12338), *Streptomyces clavuligerus* NRRL 3585, *Clavibacter toxicus* NCPPB 3552, and *Clavibacter michiganensis* subsp. *michiganensis* (11). Similar compounds are from *Streptomyces bobilliae* (mycospocidins), *Streptomyces griseoflavus* subsp. *thuringiensis* (streptoviridins), and antibiotic 24010 (4). Twelve tunicamycin biosynthetic genes (*tunA* to *tunL*) have been identified in *S. chartreusis* NRRL 3882 and *S. clavuligerus* NRRL 3585 and are also present on the sequenced genome of *Actinosynnema mirum* DSM 44827, although with a truncated *tunL* (3, 13). *S. clavuligerus* NRRL 3585 also produces clavulinic acid (a lactamase inhibitor), cephalosporin-type antibiotics, and

holomycin, and *S. chartreusis* strains also produce chartreusin and calcimycin (calcium ionophore A23187). Gene clusters that code for biosynthesis of these natural products have also been identified (9, 6, 12, 14).

Genomic sequences were generated using paired-end sequencing with a 75-bp run on an Illumina Genome Analyzer II× at the Wisconsin Biotechnology Center. Libraries prepared using the Nextera DNA sample prep kit were tagged and pooled. The NRRL 12338 library was covered by 2,417,521 paired reads, NRRL 3882 by 618,150 paired reads, and ATCC 31396 by 571,352 paired reads. Assembly was performed using a custom pipeline that combined Velvet version 1.0.15 (15), SOAPdenovo version 1.04 (7), and EULER-SR version 1.1.2 (2) assemblies in gsAssembler version 2.5.3 (8) using 400,000 paired-end reads reformatted for gsAssembler as additional input, followed by final scaffolding with SSPACE version 1.1 (1). This resulted in a total of 8,859,518 bp, 9,028,970 bp, and 8,505,115 bp assembled into 112, 3,703, and 2,823 scaffolds, 351, 5,817, and 5,885 contigs, and N50 scaffold values of 228,498 bp, 5,008 bp, and 7,704 bp for NRRL 12338, NRRL 3882, and ATCC 31396, respectively. Prodigal (5) was used for open reading frame (ORF) prediction, resulting in 8,016 (NRRL 12338), 11,795 (NRRL 3882), and 10,616 (ATCC 31396) ORFs. G+C contents for each strain were 71.3%, NRRL 12338; 70.8%, NRRL 3882; and 72.0%, ATCC 31396.

Here, we announce the genomic sequencing of three tunicamycin-producing streptomycetes, *S. chartreusis* NRRL 12338 and NRRL 3882 and Tamura's original *S. lysosuperificus* isolate. Taken together with the previously completed genomes of *S. clavuligerus* NRRL 3585 and *Actinosynnema mirum* DSM 44827, these will aid our understanding of tunicamycin genetics and biosynthesis and of several other commercially important natural products.

**Nucleotide sequence accession numbers.** Nucleotide sequences are available in GenBank: project number 72669,

\* Corresponding author. Mailing address for N. P. J. Price: Renewable Product Technology Research Unit, National Center for Agricultural Utilization Research (NCAUR), Agricultural Research Service, U.S. Department of Agriculture, 1815 N. University St., Peoria, IL 61604. Phone: (309) 681-6246. Fax: (309) 681-6040. E-mail: neil.price@ars.usda.gov. Mailing address for W. Chen: Key Laboratory of Combinatorial Biosynthesis & New Drug Discovery, Ministry of Education, and School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China. Phone: 86-27-68756713. Fax: 86-27-68759850. E-mail: wenqingchen2010@gmail.com.

*Streptomyces lysosuperificus* ATCC 31396, accession number AGDC000000000; project number 72671, *Streptomyces chartreusis* NRRL 3882, accession number AGDD000000000; and project number 72673, *Streptomyces chartreusis* NRRL 12338, accession number AGDE000000000.

This work was supported by the National Institutes of Health (GM PO1 GM077596). J.R.D. was supported by a postdoctoral fellowship from the Institute for Genomic Biology at the University of Illinois at Urbana-Champaign.

The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the NIGMS or NIH. The mention of firm names or trade products does not imply that they are endorsed or recommended by the U.S. Department of Agriculture over other firms or similar products not mentioned.

#### REFERENCES

1. Boetzer, M., et al. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**:578–579.
2. Chaisson, M. J. P., D. Brinza, and P. A. Pevzner. 2009. De novo fragment assembly with short mate-paired reads: does the read length matter? *Genome Res.* **19**:336–346.
3. Chen, W., et al. 2010. Characterization of the tunicamycin gene cluster unveiling unique steps involved in its biosynthesis. *Protein Cell* **1**:1093–1105.
4. Eckardt, K. 1983. Tunicamycins, streptoviridins, and corynetoxins, a special subclass of nucleoside antibiotics. *J. Nat. Prod.* **46**:544–550.
5. Hyatt, D., et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119.
6. Li, B., and C. T. Walsh. 2010. Identification of the gene cluster for the dithiolopyrrolone antibiotic holomycin in *Streptomyces clavuligerus*. *Proc. Natl. Acad. Sci. U. S. A.* **46**:9731–19735.
7. Li, R., et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**:265–272.
8. Margulies, M., et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**:376–380.
9. Mellado, E., et al. 2002. The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: genetic organization of the region upstream of the *car* gene. *Microbiology* **148**:1427–1438.
10. Tamura, G. 1982. Tunicamycins, p. 10–31. Japan Scientific Press, Tokyo, Japan.
11. Tsvetanova, B. C., D. J. Kiemle, and N. P. J. Price. 2002. Biosynthesis of tunicamycin and metabolic origin of the 11-carbon dialdose sugar, tunicamine. *J. Biol. Chem.* **38**:35289–35296.
12. Wu, Q., et al. 2011. Characterization of the biosynthesis gene cluster for the pyrrole polyether antibiotic calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrob. Agents Chemother.* **55**:974–982.
13. Wyszynski, F. J., et al. 2010. Dissecting tunicamycin biosynthesis by genome mining: cloning and heterologous expression of a minimal gene cluster. *Chem. Sci.* **1**:581–589.
14. Xu, Z., et al. 2005. Biosynthesis of the antitumor agent chartreusin involves the oxidative rearrangement of an anthracyclic polyketide. *Chem. Biol.* **12**:579–588.
15. Zerbino, D. R., and E. Birney. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**:821–829.