

Genome Sequence of the Vancomycin-Producing *Amycolatopsis orientalis* subsp. *orientalis* Strain KCTC 9412^T

Haeyoung Jeong,^a Young Mi Sim,^a Hyun Ju Kim,^b Dong-Woo Lee,^c Si-Kyu Lim,^d Sang Jun Lee^b

Korean Bioinformation Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea^a; Infection and Immunity Research Center, KRIBB, Daejeon, Republic of Korea^b; School of Applied Biosciences, Kyungpook National University, Daegu, Republic of Korea^c; GenoTech Corporation, Daejeon, Republic of Korea^d

***Amycolatopsis orientalis* is the producer of vancomycin, a glycopeptide antibiotic that is used for the treatment of serious infections with Gram-positive bacteria. Here we present the next-generation sequencing (NGS)-based 9.06-Mb draft genome sequence of the type strain *Amycolatopsis orientalis* subsp. *orientalis* KCTC 9412 (DSM 40040; ATCC 19795).**

Received 13 May 2013 Accepted 14 May 2013 Published 27 June 2013

Citation Jeong H, Sim YM, Kim HJ, Lee D-W, Lim S-K, Lee SJ. 2013. Genome sequence of the vancomycin-producing *Amycolatopsis orientalis* subsp. *orientalis* strain KCTC 9412^T. *Genome Announc.* 1(3):e00408-13. doi:10.1128/genomeA.00408-13.

Copyright © 2013 Jeong et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Haeyoung Jeong, hyjeong@kribb.re.kr, or Sang Jun Lee, leesj@kribb.re.kr.

Amycolatopsis orientalis is the producer of vancomycin (1), which is a glycopeptide antibiotic of last resort against methicillin-resistant *Staphylococcus aureus*, which causes serious clinical infections (2). For several decades, industrial strains have been developed through random mutagenesis or optimization of culture conditions (3) from *A. orientalis* subsp. *orientalis* KCTC 9412, the type strain of this species. However, genome sequence information for *A. orientalis* species is not yet publicly available. For the application of genome-scale metabolic engineering and synthetic biology to obtain higher yields of antibiotic production, we determined the genome sequence of the strain.

A. orientalis subsp. *orientalis* KCTC 9412^T was grown in tryptic soy broth containing wheat starch (0.1%) at 30°C. Cells were disrupted by lysozyme plus achromopeptidase (4), and chromosomal DNA was purified by a genomic DNA isolation kit (Promega). Genome sequencing was carried out using an Illumina HiSeq 2000 system. From the genomic library, 101-nucleotide paired-end reads (4.58 Gb, at 506× coverage) were produced. Pretreatment of reads and *de novo* assembly were done using the CLC Genomics Workbench ver. 6.0.1 (CLC bio). Quality-trimmed reads, consisting of 71.7% of initial bases, were assembled into 99 contigs (>200 bp, 69.0% G+C) using a word size of 64. Total contig length and N_{50} were 9,062,467 bp and 91,540 bp, respectively. The largest contig was 233,061 bp. We tested Velvet ver. 1.2.07 (5) and SOAPdenovo ver. 1.0.5 (<http://soap.genomics.org.cn/>) using a range of *k*-mers, but the CLC Genomics Workbench showed the best results in terms of N_{50} .

The genome sequence was automatically annotated using the RAST server (6). It contains 8,249 protein-coding genes and 52 tRNA genes. Subsystem and clusters of orthologous groups (COG) coverages were 26.8% and 59.3%, respectively. Sequences related to the integrative plasmid pMEA100, which is known from *Amycolatopsis mediterranei* S699 (7), were not found. We identified a vancomycin biosynthetic gene cluster from the genome sequence, which was identical to the previously known gene cluster from *A. orientalis* ATCC 19795 (HE589771; 67,422 bp) at the nu-

cleotide level, with an ~50-bp discrepancy in the intergenic region between *vpsA* and *vpsB*. We also sequenced the genomes of *A. orientalis* DSM 43388 and DSM 46075 and compared them with KCTC 9412^T sequences. Results obtained from average nucleotide identity (8) and ortholog clustering (9) suggested that the two DSM strains were very different from the type strain. In particular, among the eight nonribosomal peptide-synthetase (NRPS) gene clusters and two polyketide synthase gene clusters in KCTC 9412^T, one NRPS gene homolog (>90% amino acid identity) was found in both DSM 43388 and DSM 46075. Vancomycin synthetic gene clusters were not found in the genomes of DSM 43388 and DSM 46075. However, we identified vancomycin resistance genes (10) *vanH*, *vanA*, and *vanX* in the genome of DSM 46075, which were also found in KCTC 9412^T. Comparative genome analysis using all available *Amycolatopsis* species indicated that KCTC 9412^T is closest to *A. decaplanina* DSM 44594^T.

Nucleotide sequence accession numbers. This Whole-Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [ASJB000000000](https://www.ncbi.nlm.nih.gov/nuccore/ASJB000000000). The version described in this paper is the first version, [ASJB010000000](https://www.ncbi.nlm.nih.gov/nuccore/ASJB010000000).

ACKNOWLEDGMENTS

This study was supported by the Small and Medium Business Administration (S2060609); the Ministry of Trade, Industry and Energy; and the KRIBB Research Initiative Program, Ministry of Science, ICT and Future Planning, Republic of Korea.

REFERENCES

1. Levine DP. 2006. Vancomycin: a history. *Clin. Infect. Dis.* 42(Suppl 1): S5–S12.
2. Moellering RC, Jr. 2006. Vancomycin: a 50-year reassessment. *Clin. Infect. Dis.* 42(Suppl 1):S3–S4.
3. Jung HM, Kim SY, Moon HJ, Oh DK, Lee JK. 2007. Optimization of culture conditions and scale-up to pilot and plant scales for vancomycin production by *Amycolatopsis orientalis*. *Appl. Microbiol. Biotechnol.* 77: 789–795.

4. Nikodinovic J, Barrow KD, Chuck JA. 2003. High yield preparation of genomic DNA from *Streptomyces*. *Biotechniques* 35:932–934, 936.
5. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.
7. Zhao W, Zhong Y, Yuan H, Wang J, Zheng H, Wang Y, Cen X, Xu F, Bai J, Han X, Lu G, Zhu Y, Shao Z, Yan H, Li C, Peng N, Zhang Z, Zhang Y, Lin W, Fan Y, Qin Z, Hu Y, Zhu B, Wang S, Ding X, Zhao GP. 2010. Complete genome sequence of the Rifamycin SV-producing *Amycolatopsis mediterranei* U32 revealed its genetic characteristics in phylogeny and metabolism. *Cell Res.* 20:1096–1108.
8. Richter M, Rossello-Mora R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U. S. A.* 106: 19126–19131.
9. Li L, Stoeckert CJ, Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13:2178–2189.
10. Marshall CG, Lessard IA, Park I, Wright GD. 1998. Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. *Antimicrob. Agents Chemother.* 42:2215–2220.