



Data in Brief

Genome sequencing and annotation of *Amycolatopsis vancoresmycina* strain DSM 44592^T☆Navjot Kaur^{a,1}, Shailesh Kumar^{b,1}, Shanmugam Mayilraj^{a,*}^a Microbial Type Culture Collection and Gene Bank (MTCC), CSIR-Institute of Microbial Technology, Chandigarh 160036, India^b Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh 160036, India

ARTICLE INFO

Article history:

Received 18 September 2013

Received in revised form 23 October 2013

Accepted 23 October 2013

Available online 28 November 2013

Keywords:

Amycolatopsis vancoresmycina strain DSM 44592^T

Illumina-HiSeq

NGS QC toolkit

Rapid annotations using subsystems technology (RAST)

ABSTRACT

We report the 9.0-Mb draft genome of *Amycolatopsis vancoresmycina* strain DSM 44592^T, isolated from Indian soil sample; produces antibiotic vancoresmycin. Draft genome of strain DSM44592T consists of 9,037,069 bp with a G+C content of 71.79% and 8340 predicted protein coding genes and 57 RNAs. RAST annotation indicates that strains *Streptomyces* sp. AA4 (score 521), *Saccharomonospora viridis* DSM 43017 (score 400) and *Actinosynnema mirum* DSM 43827 (score 372) are the closest neighbors of the strain DSM 44592^T.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

Specifications	
Organism/cell line/tissue	<i>Amycolatopsis vancoresmycina</i>
Strain(s)	DSM 44592 ^T
Sequencer or array type	Sequencer; the Illumina-HiSeq 1000
Data format	Processed
Experimental factors	Microbial strain
Experimental features	Draft genome sequence of <i>A. vancoresmycina</i> strain DSM 44592 ^T , assembly and annotation
Consent	n/a

Direct link to the data

Direct link: <http://www.ncbi.nlm.nih.gov/nuccore/AOU000000000>.

The genus *Amycolatopsis* was proposed by Lechevalier et al. [1]. *Amycolatopsis vancoresmycina* was proposed by Wink et al. [2]. *A. vancoresmycina* was isolated from Indian soil that produces novel glycopeptide antibiotics and a polyenic compound. *A. vancoresmycina*

strain DSM 44592^T which was isolated from soil [2], was obtained from an open collection of the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany, deposited by Prof. Joachim M. Wink. Strain *A. vancoresmycina* DSM 44592^T is Gram-positive bacteria, aerial mycelium with color ranged from brown beige to saffron yellow. Its peptidoglycan contains meso-diaminopimelic acid; diagnostic sugars are arabinose and galactose, mycolic acids are absent; MK-9 (H4) is the principal menaquinone; phospholipids are phosphatidylethanolamine, hydroxyl phosphatidylethanolamine, diphosphatidylglycerol and phosphatidylinositol. Genomic DNA was extracted from a 48 hour old culture using ZR Fungal/Bacterial DNA MiniPrep™ as per manufacturer's instructions. The genome of *A. vancoresmycina* strain DSM 44592^T was sequenced using the Illumina-HiSeq 1000 paired-end technology produced a total of 62,283,914 paired-end reads (insert size of 350 bp) of 101 bp. We have used NGS QC toolkit v2.3 [3] to filter the data for high quality (cut off read length for HQ = 70%, cut off quality score = 20), vector/adaptor free reads for genome assembly. A total of 57,704,188 high quality, vector filtered reads (~647.6 times coverage) were used for assembly with SOAPdenovo v1.05 (at hash length of 69) followed by GapCloser (at hash length 15) software [4]. The final assembly contains 786 contigs of total size 9,037,069 bp, with N₅₀ contig length of 18.4 kb; the largest contigs assembled measures 87.7 kb. This draft genome comprising 9,037,069 nt was annotated with the help of RAST (Rapid Annotation using Subsystem Technology) system [5] server. A total of 8340 predicted coding regions (CDSs), 3 rRNAs and 54 tRNAs were predicted.

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author at: Institute of Microbial Technology (IMTECH), Sector 39-A, Chandigarh 160036, India. Tel.: +91 172 6665166; fax: +91 172 2695215.

E-mail address: mayil@imtech.res.in (S. Mayilraj).

¹ Both are first authors, equally contributed.

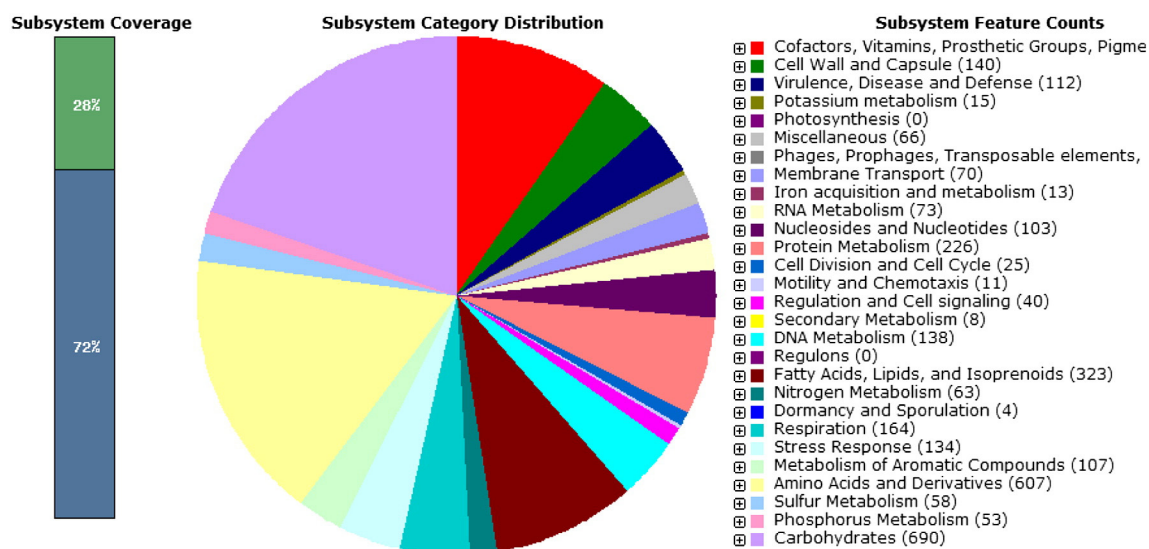


Fig. 1. Sub-system distribution of strain *Amycolatopsis vancoresmycina* strain DSM 44592T (based on RAST annotation server).

RAST annotation indicates that strain *Streptomyces* sp. AA4 (score 521), *Saccharomonospora viridis* DSM 43017 (score 400) and *Actinosynnema mirum* DSM 43827 (score 372) are the closest neighbors of the strain DSM 44592^T. Genome annotation available at RAST indicates that the strain DSM 44592^T contains the genes for glycolysis and gluconeogenesis, TCA cycle and pentose phosphate pathway. Genes of alkaline phosphatase, galactosidase (alpha and beta subunit), mannosidase (alpha and beta subunits), catalase (EC 1.11.1.6) and ferroxidase (EC 1.16.3.1) are also present in the genome annotation of strain 44592^T. We have mapped all predicted 8340 CDSs to KEGG pathways [6] with the help of KASS server [7] and found that genes of *rifA*, *rifB*, *rifC*, *rifD*, *rifE*, *rifK*, *rifL*, *rifM* and *rifN* are present in the genome, mapped on the ansamycins biosynthesis pathway. Genes of glucose-1-phosphate thymidyltransferase [EC:2.7.7.24], dTDP-glucose 4,6-dehydratase [EC:4.2.1.46], dTDP-4-dehydrorhamnose reductase [EC:1.1.1.133] and dTDP-4-dehydrorhamnose 3,5-epimerase [EC:5.1.3.13] are present and mapped at the pathway of polyketide sugar unit biosynthesis (Fig. 1).

Nucleotide sequence accession number

The *A. vancoresmycina* DSM 44592^T whole genome shot gun (WGS) project has been deposited at DDBJ/EMBL/GenBank under the project accession AOU000000000 of the project (01) has the accession numbers AOU001000000 and consists of sequences AOU001000001 AOU001000784.

Conflict of interest

The authors declare that there is no conflict of interest on any work published in this paper.

Acknowledgments

This work was funded by IMTECH-CSIR. NK is supported by a research fellowship from the University Grant Commission (UGC) Govt. of India. SK is supported by a research fellowship from the Council of Scientific and Industrial Research. This is IMTECH communication number 108/2013.

References

- [1] M.P. Lechevalier, H. Prauser, D.P. Labeda, J.S. Ruan, Two new genera of nocardioform actinomycetes: *Amycolata* gen. nov. and *Amycolatopsis* gen. nov. *Int. J. Syst. Bacteriol.* 36 (1986) 29–37.
- [2] J.M. Wink, R.M. Kroppenstedt, B.N. Ganguli, R. Suesh, S.R. Nadkarni, P. Schuman, G. Seibert, E. Stackebrandt, Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimycina* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. *Syst. Appl. Microbiol.* 26 (2003) 38–46.
- [3] R.K. Patel, M. Jain, NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7 (2012) e30619.
- [4] R. Li, H. Zhu, J. Ruan, W. Qian, X. Fang, Z. Shi, Y. Li, S. Li, G. Shan, K. Kristiansen, H. Yang, J. Wang, De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20 (2010) 265–272.
- [5] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9 (2008) 75.
- [6] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40 (2012) D109–D114.
- [7] Y. Moriya, M. Itoh, S. Okuda, A.C. Yoshizawa, M. Kanehisa, KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35 (2007) W182–W185.