Genomics Data 2 (2014) 44–45

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Genome sequencing and annotation of *Amycolatopsis azurea* DSM 43854^T

Indu Khatri [a], Srikrishna Subramanian [a,]*, Shanmugam Mayilraj [b,]*

[a] *Protein Science and Engineering, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*
[b] *Microbial Type Culture Collection and Gene Bank (MTCC), CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

ARTICLE INFO

Article history:
Received 30 October 2013
Accepted 9 December 2013
Available online 12 March 2014

Keywords:
Amycolatopsis azurea strain DSM 43854^T
Illumina-HiSeq
NGS QC tool kit
Rapid Annotations using Subsystems Technology (RAST)

ABSTRACT

We report the 9.2 Mb genome of the azureomycin A and B antibiotic producing strain *Amycolatopsis azurea* isolated from a Japanese soil sample. The draft genome of strain DSM 43854^T consists of 9,223,451 bp with a G + C content of 69.0% and the genome contains 3 rRNA genes (5S–23S–16S) and 58 aminoacyl-tRNA synthetase genes. The homology searches revealed that the PKS gene clusters are supposed to be responsible for the biosynthesis of naptomycin, macbecin, rifamycin, mitomycin, maduropeptin enediyne, neocarzinostatin enediyne, C-1027 enediyne, calicheamicin enediyne, landomycin, simocyclinone, medermycin, granaticin, polyketomycin, teicoplanin, balhimycin, vancomycin, staurosporine, rubradirin and complestatin.

Specifications	
Organism/cell line/tissue	*Amycolatopsis azurea*
Strain(s)	DSM 43854^T
Sequencer or array type	Sequencer; the Illumina-HiSeq 1000
Data format	Processed
Experimental factors	Microbial strain
Experimental features	Draft genome sequence of *A. azurea* strain DSM 43854^T, assembly and annotation
Consent	n/a

Direct link to deposited data

Deposited data can be found here: http://www.ncbi.nlm.nih.gov/nuccore/ANMG00000000.

The genus *Amycolatopsis* was proposed by Lechevalier et al. in 1986 [1]. *A. azurea* was proposed by Omura et al. in 1983 [2] and later on emended by Henssen et al. in 1987 [3] and was isolated from soil that produced two water soluble antibiotics, namely, azureomycin A and B. *A. azurea* strain DSM 43854^T was obtained from the open collection of the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany. The strain *A. azurea* DSM 43854^T is a Gram-positive bacteria and contains aerial mycelium with white to blue color. 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. azurea* was sequenced using the Illumina-HiSeq 1000 technology. Sequencing resulted in 22,170,928 paired-end reads (insert size of 350 bp) of 101 bp length. A total of 21,919,041 high-quality reads with approximately 480× coverage were assembled using CLCbio wb5.5 (http://www.clcbio.com) (word size 55 and bubble size 60) to obtain 154 contigs (N_{50}, 127,182 bp). The functional annotation was carried out by RAST (Rapid Annotation using Subsystem Technology) [4], tRNA was predicted by tRNAscan-SE-1.23 [6] and rRNA genes by RNAmmer 1.2 [5]. The genome contains 3 rRNA genes (5S–23S–16S) and 58 aminoacyl-tRNA synthetase genes. The draft genome of *A. azurea* consists 154 contigs of 9,223,451 bp with an average G + C content of 69.0%.

A total of 8603 coding regions (4406 from a positive strand and 4197 from a negative strand) were found in the genome of which 5396 (63%) could be functionally annotated. The genome coding density is 89% with an average gene length of 948 bp. The annotated genome has 117 genes involved in virulence, disease and defense including 61 genes for resistance to antibiotics and toxic compounds such as tetracycline, vancomycin, arsenic and beta-lactamase (Fig. 1). Five genes are involved in dormancy and sporulation of bacteria. The genome comprises of 55 genes for sulfur metabolism and 47 genes for phosphorus metabolism. The homology searches revealed that the PKS gene clusters are supposed to be responsible for the biosynthesis of naptomycin, macbecin, rifamycin, mitomycin, maduropeptin enediyne, neocarzinostatin enediyne, C-1027 enediyne, calicheamicin enediyne, landomycin, simocyclinone, medermycin, granaticin, polyketomycin, teicoplanin, balhimycin, vancomycin, staurosporine, rubradirin and complestatin.

---

* Corresponding authors at: CSIR-Institute of Microbial Technology (IMTECH), Sector 39-A, Chandigarh 160036, India. Tel.: +91 1726665483, +91 172 6665166; fax: +91 172 2695215.

E-mail addresses: krishna@imtech.res.in (S. Subramanian), mayil@imtech.res.in (S. Mayilraj).

http://dx.doi.org/10.1016/j.gdata.2013.12.003

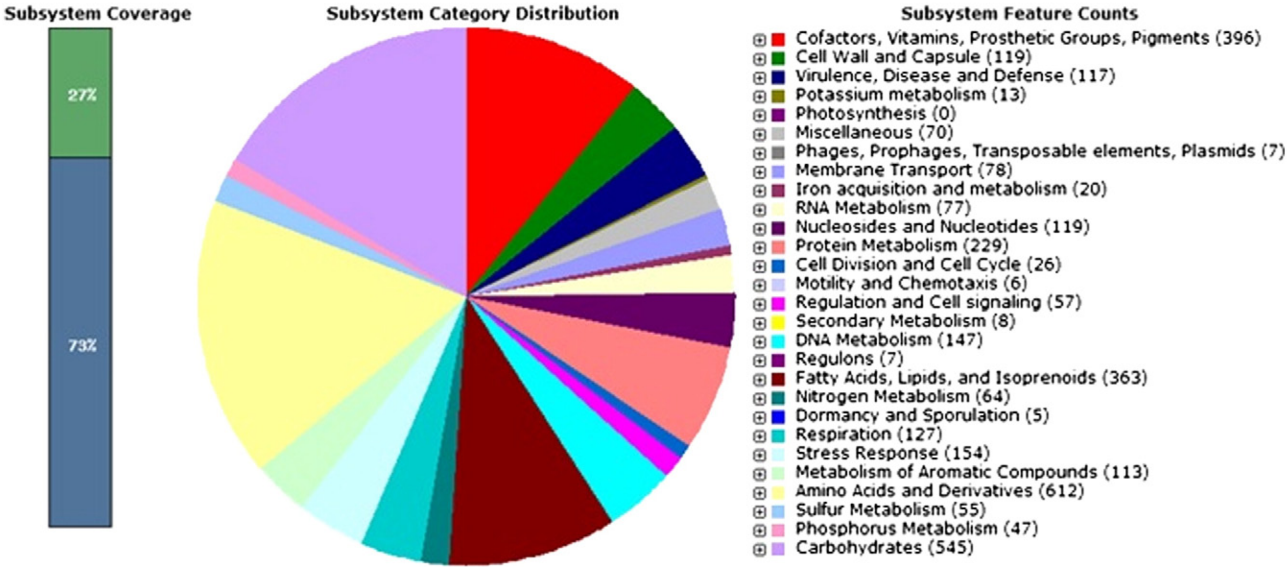

Fig. 1. Sub-system distribution of strain *Amycolatopsis azurea* DSM 43854^T (based on RAST annotation server).

The functional comparison of genome sequences available on the RAST server revealed the closest neighbors of *A. azurea* as *Actinosynnema mirum* DSM 43827 (score 536) followed by *Streptomyces* sp. AA4 (score 432), *Saccharomonospora viridis* DSM 43017 (score 376) and *Rhodococcus jostii* RHA1 (score 357).

Nucleotide sequence accession number

The *A. azurea* DSM 43854^T whole genome shot gun (WGS) project has been deposited at DDBJ/EMBL/GenBank under the project accession number ANMG00000000 of the project (01) and has the accession numbers ANMG01000000 and consists of sequences ANMG01000001–ANMG01000154.

Conflict of interest

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

Acknowledgment

This work was funded by CSIR-IMTECH (BSC0402). I.K. is supported by a research fellowship from the University Grant Commission. We thank the C-CAMP next-generation genomics facility for their help in obtaining the NGS data. This is IMTECH communication number 059/2013.

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] S. Omura, H. Tanaka, Y. Tanaka, P. Spiri-Nakagawa, R. Oiwa, Y. Takahashi, K. Matsuyama, Y. Iwai, Studies on bacterial cell wall inhibitors. VII. Azureomycins A and B, new antibiotics produced by *Pseudonocardia azurea* nov. sp. Taxonomy of the producing organisms, isolation, characterization and biological properties. J. Antibiot. 32 (1979) 985–994.

[3] A. Henssen, H.W. Kothe, R.M. Kroppenstedt, Transfer of *Pseudonocardia azurea* and "*Pseudonocardia fastidiosa*" to the genus *Amycolatopsis*, with emended species description. Int. J. Syst. Bacteriol. 37 (1987) 292–295.

[4] 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.

[5] K. Lagesen, P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35 (2007) 3100–3108.

[6] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25 (1997) 955–964.