Journal of Biotechnology 267 (2018) 50–54

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Short communication

Complete genome sequence of *Streptomyces peucetius* ATCC 27952, the producer of anticancer anthracyclines and diverse secondary metabolites

Dipesh Dhakal[a], Si-Kyu Lim[c], Dae Hee Kim[d], Byung-Gee Kim[e], Tokutaro Yamaguchi[b], Jae Kyung Sohng[a,b,*]

[a] Department of Life Science and Biochemical Engineering, Sun Moon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam, 31460, Republic of Korea
[b] Department of BT-Convergent Pharmaceutical Engineering, Sun Moon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam, 31460, Republic of Korea
[c] GenoTech Corporation, 26-69 Gajeongbuk-ro, Daejeon 34113, Republic of Korea
[d] GeneChem Corporation, 59-5 Jang-Dong, Yuseong-Gu, Daejeon, 34025, Republic of Korea
[e] School of Chemical and Biological Engineering, Seoul National Univeristy, 1 Gwanak-ro Gwanak-gu, Seoul, 08826, Republic of Korea

ARTICLE INFO

Chemical compounds studied in this article:
Doxorubicin (Pubchem CID: 31703)
Daunorubicin (Pubchem CID: 30323)
Hopene (Pubchem CID: 9978948)
T1-pks (Pubchem CID: NA)
Geosmin (Pubchem CID: 9978948)

Keywords:
Streptomyces
Genome sequence
Gene cluster
Bioactive compounds
Enzymes

ABSTRACT

Streptomyces peucetius ATCC 27952 is a filamentous soil bacterium with potential to produce anthracyclines such as doxorubicin (DXR) and daunorubicin (DNR), which are potent chemotherapeutic agents for the treatment of cancer. Here we present the complete genome sequence of *S. peucetius* ATCC 27952, which consists of 8,023,114 bp with a linear chromosome, 7187 protein-coding genes, 18 rRNA operons and 66 tRNAs. Bioinformatic analysis of the genome sequence revealed ~68 putative gene clusters involved in the biosynthesis of secondary metabolites, including diverse classes of natural products. Diverse secondary metabolites of PKS (polyketide synthase) type II (doxorubicin and daunorubicin), NRPS (non-ribosomal peptide synthase) (T1-pks), terpene (hopene) etc. have already been reported for this strain. In addition, *in silico* analysis suggests the potential to produce diverse compound classes such as lantipeptides, lassopeptides, NRPS and polyketides. Furthermore, many catalytically-efficient enzymes involved in hydroxylation, methylation etc. have been characterized in this strain. The availability of genomic information provides valuable insight for devising rational strategies for the production and isolation of diverse bioactive compounds as well as for the industrial application of efficient enzymes.

1. Introduction

Actinobacteria are well known for their ability to produce a variety of structurally complex but functionally important bioactive natural products. Among the *Actinobacteria,* members of the genus *Streptomyces* are common sources of bioactive molecules, with clinical applications such as antibacterial, antitumor, immunosuppressant, etc. (Chaudhary et al., 2013). Structurally divergent but clinically useful antitumor drugs, such as anthracyclines (aclarubicin, daunomycin, and doxorubicin), glycopeptides (bleomycin and actinomycin D), aureolic acids (mithramycin), enediynes (neocarzinostatin), antimetabolites (pentostatin), and many other compound classes have been isolated from different *Streptomyces* spp. (Olano et al., 2009).

Streptomyces peucetius var. caesius (ATCC 27952) was obtained from *S. peucetius* ATCC 29050, after mutagenic treatment with *N*-nitroso-*N*-methyl urethane. It differs from the parent culture by the color of the vegetative and aerial mycelia. *S. peucetius* var. caesius (ATCC 27952) is a well-known producer of two anthracyclines, daunorubicin and doxorubicin, which are already approved chemotherapies for cancer patients (Arcamone et al., 1969; Arcamone and Cassinelli, 1998). It has been reported that a number of organisms (including *S. peucetius* ATCC 29050) have the ability to produce DNR, but *S. peucetius* ATCC 27952 is the only organism reported to produce DXR (Niraula et al., 2010). The anticancer activities of these anthracyclines results from DNA damage via the inhibition of DNA topoisomerase II and the generation of free radicals, DNA binding, alkylation and intercalation (Minotti et al., 2004). DNR and DXR derivatives have also been reported to possess antibacterial activities (Gumpert et al., 1982). The study of biosynthetic gene clusters and analysis of biosynthetic mechanisms have established that ε-rhodomycinone is formed through a pathway involving a type II PKS, which is accompanied with decoration by thymidine diphospho (TDP)-L-daunosamine via glycosyltransferase DnrS (Grimm et al., 1994; Singh et al., 2010) as shown in **Fig. S1**.

S. peucetius ATCC 27952 has been shown to produce a diverse range

* Corresponding author at: Department of Life Science and Biochemical Engineering, Sun Moon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam, 31460, Republic of Korea.
E-mail address: sohng@sunmoon.ac.kr (J.K. Sohng).

https://doi.org/10.1016/j.jbiotec.2017.12.024
Received 11 September 2017; Received in revised form 20 December 2017; Accepted 30 December 2017
Available online 04 January 2018
0168-1656/

Table 1
Versatile catalytic enzymes sourced from *Streptomyces peucetius* var. caesius and their biochemical properties.

Enzyme Name	Gene Accession No	Biochemical characterization	References
CYP107AJ1	AJ605547	Dealkylation of 7-ethoxycoumarin	Niraula et al. (2011)
CYP105P2.	AJ605540	Hydroxylation of flavone	Niraula et al. (2012)
CYP105F2	AJ605543	Hydroxylation of oleandomycin	Shrestha et al. (2008a)
CYP166B1	AJ605542	Dealkylation of 7-ethoxycoumarin	Shrestha et al. (2010)
CYP107P3	AJ605551	Dealkylation of 7-ethoxycoumarin	Shrestha et al.(2008b)
CYP147F1	AJ605536	Hydroxylation of fatty acids	Bhattarai et al. (2013)
CYP107N3	AJ605544	Epoxidation of oleic acid	Bhattarai et al. (2012)
CYP157C4	AJ605537	Dealkylation of 7-ethoxycoumarin	Rimal et al. (2015a)
CYP129A2	AAD04715	Hydroxylation of daunorubicin	Rimal et al. (2015b)
DnrK	AAA99001	Methylation of carminomycin	Madduri et al. (1993)
SpOMT2884	KF420279	Methylation of 7,8-dihydroflavone	Koirala et al. (2014)
SpOMT2884	KF420279	Methylation of 8-hydroxydaidzein	Chiang et al. (2015)
SpOMT2884	KF420279	Methylation of 3′-hydroxygenistein	Chiang et al. (2017)
SpOMT7740	KU745626	Methylation of diverse natural products	Parajuli et al. (2017)
DnrS	L47164	Glycosylation of anthracyclines	Han et al. (2011)

of microbial chemicals (Ghimire et al., 2008a,b; Singh et al., 2009; Park et al., 2013; Ghimire et al., 2014), as shown in **Fig. S2**. Moreover, it has been speculated that this strain not only harbors the capability of producing diversified secondary metabolites, but is also a prolific resource for catalytically-efficient enzymes. The successful heterologous expression of enzymes sourced from *S. peucetius* ATCC 27952 has provided insight into the catalytic efficiencies of the versatile enzymes found in this strain (Shrestha et al., 2008a,b; Niraula et al., 2011; Koirala et al., 2014; Chiang et al., 2015) (see Table 1). In particular, cytochrome P450 enzymes have versatile biocatalytic actions such as aliphatic and aromatic bond hydroxylation and double-bond epoxidation, all with precise chemo-, regio-, and stereoselectivity (Podust and Sherman, 2012), hence they are of particular interest. These versatile enzymes, in turn, require specific electron transfer partners such as Fdx (Ferrodoxin) and Fdr (Ferrodoxin reductase) for catalyzing particular enzymatic reactions (Chun et al., 2007).

The phylogentic tree based on 16S rRNA sequence, constructed using MEGA (version 7) and neighbor joining method exhibited distinct branch of *S. peucetius* var. caesius (ATCC 27952) with other *Streptomyces* strains (Fig. 1). Furthermore, it was anticipated that the complete genome sequence of *S. peucetius* var. caesius (ATCC 27952), can provide information about different biosynthetic gene clusters and diverse enzyme classes involved in biosynthesis and the maintenance of metabolism. This information provides deep insight into the diversity of secondary metabolites and versatile biocatalysts in this strain. Sequencing was performed using an Applied Biosystems 3730xl DNA Analyzer and an Ion Torrent PGM™ sequencer by GenoTech Corporation (Daejeon, Korea) to obtain information about the genome sequence of *S. peucetius* var. caesius (ATCC 27952). Genomic libraries with diverse insert sizes (~30–45 kb) were constructed and nucleotide sequences were generated with 100-fold coverage of the *S. peucetius* ATCC 27952 genome. Hybrid genome assembly was achieved using the Phred/Phrap/Consed software package (http://www.phrap.org) and CLC Genomics Workbench (CLC bio, Inc.). Sequence gaps between contigs were filled using PCR amplification followed by Sanger sequencing. Genome annotation was performed using an in-house pipeline (Song et al., 2010; Song et al., 2016). RNAmmer and tRNAscan-SE were used to detect rRNA and tRNA genes, respectively. Protein-coding genes were predicted by Glimmer, HMMER, GeneMark.hmm and Prodigal. AutoFACT was used for automatic functional assignment of the coding sequences with public databases, GenBank, COG, UniRef90 and KEGG. Manual curation of the predicted results was performed using Artemis.

The complete genome of *S. peucetius* ATCC 27952 is composed of a single linear chromosome of 8,023,114 bp with 70.6% GC content (Fig. 2). The genomic DNA contains 7187 protein-coding genes, 66 tRNAs and 18 rRNAs. The phylogenetic analysis established relatedness of *S. peucetius* ATCC 27952 with different *S. peucetius* strains, but the complete genome information of most of them are not available. However it was observed that the genomic architecture of *S. peucetius* ATCC 27952 is significantly different from the genome sequence of *S. peucetius* strain NRRL WC-3868, which was reported previously (Gene Bank accession no: NZ_JOCK00000000) (Ju et al., 2015). Based on analysis of 16srRNA sequence by NCBI-BLAST we could observe distinct disparity in sequence composition (Fig. S**3**).The specific genomic features of *S. peucetius* ATCC 27952 and *S. peucetius* strain NRRL WC-3868 are presented in Table 2. Genome mining has been an influential approach for the discovery of novel natural products from sequenced

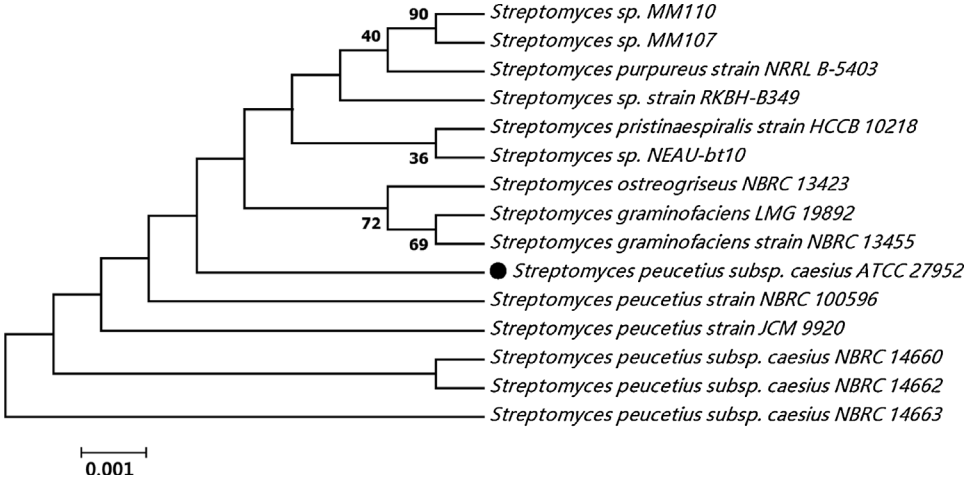

Fig 1. Phylogenetic tree of *S. peucetius* var. *caesius* ATCC 27952. MEGA 7.0 was used to construct the tree by neighbor-joining method based on 16 s RNA sequence with 1000 replications in bootstrap test. Bootstrap confidence levels > 30% are indicated at the internodes. The scale bar indicates nucleotide substitutions per nucleotide position.