



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

# Draft genome sequence data of *Streptomyces* sp. SS1-1, an endophytic strain showing cytotoxicity against the human lung cancer A549 cell line

Quan Dang Nguyen<sup>a</sup>, Phung Minh Truong<sup>b</sup>,  
Thao Nguyen Thanh Vo<sup>a</sup>, Truc Dao Xuan Chu<sup>a</sup>,  
Chuong Hoang Nguyen<sup>c,\*</sup>

<sup>a</sup> Biotechnology Center of Ho Chi Minh City, Ho Chi Minh City, Vietnam

<sup>b</sup> Center for Research and Application in Bioscience, Ho Chi Minh City, Vietnam

<sup>c</sup> University of Science, Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam

## ARTICLE INFO

## Article history:

Received 25 November 2019

Revised 14 March 2020

Accepted 20 March 2020

Available online 8 April 2020

## Keywords:

Cytotoxicity

*Streptomyces*

NGS

A549 cell line

SS1-1

## ABSTRACT

We report in this article the cytotoxicity of *Streptomyces* sp. SS1-1 against the human lung cancer A549 cell line, its draft genome sequence and a total of 20 predicted secondary metabolite biosynthetic gene clusters. *Streptomyces* sp. SS1-1 was an endophytic bacterial strain isolated from the plant *Catharanthus roseus* in Ho Chi Minh City, Vietnam. When cultured in the PY medium, this strain shows a cytotoxic effect on the A549 cell line. The draft genome of *Streptomyces* sp. SS1-1 has four contigs of total 7,815,656 bp and the GC content of this genome is 72.2%. AntiSMASH analysis reveals 20 putative biosynthetic gene clusters for the largest contig. The genome sequencing of *Streptomyces* sp. SS1-1 is essential for the molecular identification of gene cluster(s) responsible for secondary metabolite(s) with cytotoxic activity.

© 2020 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY license.

(<http://creativecommons.org/licenses/by/4.0/>)

\* Corresponding author.

E-mail address: [nhchuong@hcmus.edu.vn](mailto:nhchuong@hcmus.edu.vn) (C.H. Nguyen).

## Specifications Table

<i>Subject</i>	Biology
<i>Specific subject area</i>	Microbiology, Genomics, Biotechnology
<i>Type of data</i>	Figures, tables, draft genome sequence, raw sequencing data
<i>How data were acquired</i>	MTT (3-[4,5-methylthiazol-2-yl]-2,5-diphenyl-tetrazolium bromide) assay, flow cytometry analysis, genome sequencing by PacBio Sequel and Illumina HiSeq 4000
<i>Data format</i>	Raw and Analyzed
<i>Parameters for data collection</i>	Cytotoxic effect on the A549 cell line was evaluated by MTT assay and flow cytometry analysis. Genome sequencing of the strain was performed by PacBioSequel and Illumina HiSeq4000. Gene annotation and analysis were done by the PGAP, Busco, BMAP's MinHash Sketch programs. Putative biosynthetic gene clusters were predicted by the AntiSMASH program.
<i>Description of data collection</i>	<i>Streptomyces</i> sp. SS1-1 was isolated from <i>Catharanthus roseus</i> and cultured in the PY medium for expressing cytotoxic activity. MTT assay and flow cytometry analysis were performed on the A549 cell line to evaluate the cytotoxicity. Genomic DNA of <i>Streptomyces</i> sp. SS1-1 was extracted and sequenced by PacBio Sequel and HiSeq4000. The genome sequence was used for gene prediction and annotation as well as analysis to find putative biosynthetic gene clusters for secondary metabolites in <i>Streptomyces</i> sp. SS1-1.
<i>Data source location</i>	Center for Research and Application in Bioscience, Ho Chi Minh City, Vietnam
<i>Data accessibility</i>	Data are available within this article. The draft genome sequence of <i>Streptomyces</i> sp. SS1-1 is available in GenBank under the accession number NZ_WBXN000000001. The raw sequencing data are available in the Sequence Read Archive (SRA) database under the accession number SRX6989938 at <a href="https://www.ncbi.nlm.nih.gov/sra/SRX6989938[accn]">https://www.ncbi.nlm.nih.gov/sra/SRX6989938[accn]</a> .

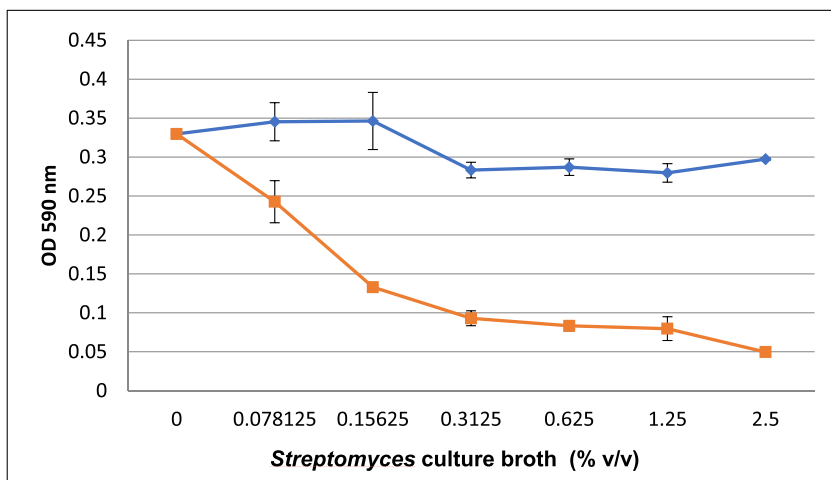
## Value of the data

- *Streptomyces* sp. SS1-1 isolated from *Catharanthus roseus* expresses the cytotoxic effect on the human lung cancer A549 cell line.
- The draft genome sequence data of *Streptomyces* sp. SS1-1 will be useful for finding biosynthetic gene cluster(s) responsible for secondary metabolite(s) with cytotoxic activity.
- In *Streptomyces* sp. SS1-1 genome, 20 putative biosynthetic gene clusters for secondary metabolites were identified.
- Data presented here could enrich the database of biosynthetic gene clusters responsible for cytotoxicity in *Streptomyces* bacteria.

## 1. Data

*Streptomyces* sp. SS1-1 was isolated from the stem of *Catharanthus roseus* using the protocol of Qin et al. with small modifications. No microbial growth was detected on the SFM agar and the nutrient agar when the surface-sterilized stem sample was imprinted on these culture media. Moreover, the spore of *Streptomyces* sp. SS1-1 could not grow on the SFM agar after sequential treatment with the solutions used for surface sterilization of the stem sample. These data supported the endophytic origin of *Streptomyces* sp. SS1-1. In literature, *Catharanthus roseus* was used to isolate the endophytic microorganisms including *Streptomyces* species which were associated with antibacterial, antifungal, anticancer activities [1–4].

*Streptomyces* sp. SS1-1 was tested for cytotoxic activity on the human lung cancer A549 cell line. The PY culture broth inoculated with *Streptomyces* sp. SS1-1 showed a remarkable reduction in the A549 cell count in comparison with the unseeded PY culture broth revealed by the MTT (3-[4,5-methylthiazol-2-yl]-2,5-diphenyl-tetrazolium bromide) assay (Fig. 1). Moreover, cell apoptosis analyzed by flow cytometry showed a 5.23% increase of apoptotic cells when the A549 cell line was treated with 1.25% (v/v) of the *Streptomyces* sp. SS1-1 culture broth compared with that of the untreated cells. The apoptotic cell rate increased by up to 14.45% when the cancer



**Fig. 1.** Cytotoxic activity of the *Streptomyces* sp. SS1-1 culture broth on the A549 cell line revealed by MTT assay. Curves show the 48 h of treatment of the A549 cells with the PY medium (the upper curve) and with the PY medium inoculated with *Streptomyces* sp. SS1-1 (the lower curve). Values were represented as mean  $\pm$  SD ( $n=3$ ).

**Table 1**

Features of the draft genome of *Streptomyces* sp. SS1-1.

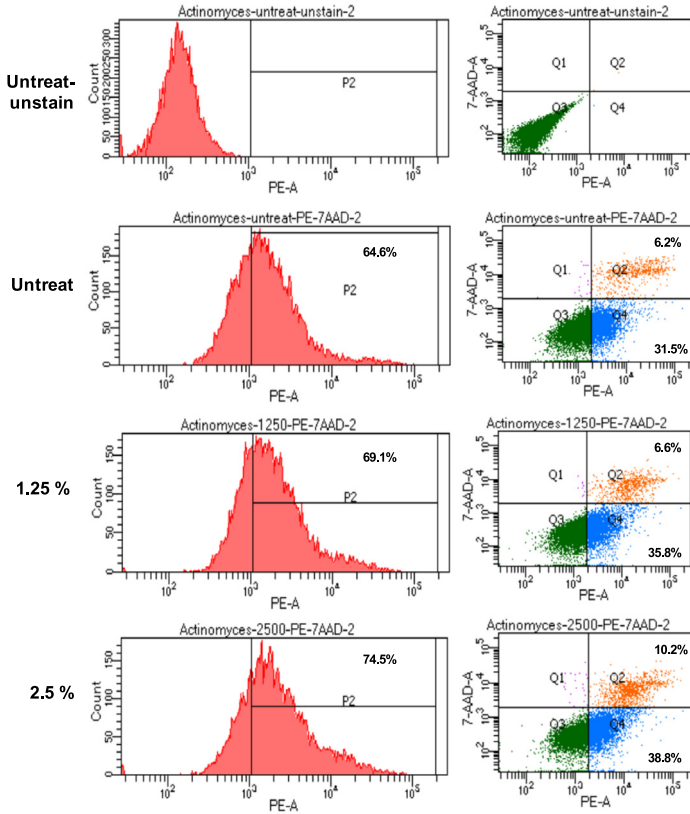
Feature	<i>Streptomyces</i> sp. SS1-1
Source of isolation	<i>Catharanthus roseus</i>
Genome size (bp)	7,815,656
Contig	4 (14,134 bp; 49,440 bp; 199,786 bp; 7,552,296 bp)
GC content (%)	72.2
Gene total	7205
Protein coding sequences	6831
tRNA	68
rRNA	6 (5S), 6 (16S), 6 (23S)
ncRNA	3
Pseudogene	285

cells were treated with 2.5% (v/v) of the *Streptomyces* sp. SS1-1 culture broth. The *Streptomyces* culture broth induced the increase of both early apoptotic cells (PE Annexin V positive, 7-AAD negative) and late apoptotic cells (PE Annexin V and 7-AAD positive) in the A549 cell population (Fig. 2). The preliminary data showed that the cytotoxic effect of *Streptomyces* sp. SS1-1 on the human lung cancer A549 cell line could be due to apoptosis but other cytotoxic mechanisms could not be excluded.

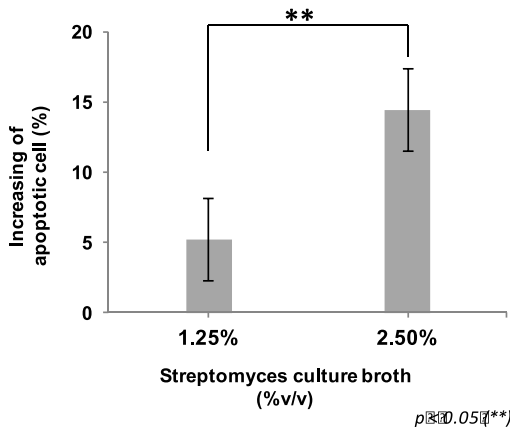
Genome sequencing was performed for *Streptomyces* sp. SS1-1 to identify putative gene clusters for secondary metabolites, one(s) of which could be responsible for the cytotoxic activity of the strain. The assembled genome of *Streptomyces* sp. SS1-1 produced four contigs of total length 7,815,656 bp and the coverage of 147-fold. The genome completeness was checked by Busco and was 95.3% complete, with 95.3% of the genome single copy and 0.0% duplicated. The GC content of the draft genome was 72.2%. An assessment of the assembly by BBMap's MinHash Sketch suggested it is most similar to *Streptomyces* sp. NRRL WC-3604. A total of 7205 genes was predicted in the genome of *Streptomyces* sp. SS1-1 by the NCBI Prokaryotic Genome Annotation Pipeline. Of 7205 genes, there are 6831 protein-coding genes, 68 tRNA genes, 3 ncRNA genes, 18 rRNA (5S (6), 16S (6), 23S (6)) genes and a total of 285 pseudogenes (Table 1).

Prediction of putative gene clusters for secondary metabolites performed with the Anti-SMASH program showed 20 putative gene clusters for secondary metabolites in the largest con-

**A**



**B**



**Fig. 2.** Flow cytometry analysis of the A549 cell apoptosis caused by the 24 h of treatment with the *Streptomyces* sp. SS1-1 culture broth. A: Representative flow cytometry result showing the apoptotic cells of the A549 cell line treated with 0%, 1.25% and 2.5% (v/v) of the *Streptomyces* sp. SS1-1 culture broth; B: Increasing of apoptotic cells treated with 1.25% and 2.5% (v/v) of the *Streptomyces* sp. SS1-1 culture broth compared with that of the untreated cells. Values were represented as mean  $\pm$  SD ( $n = 3$ ) with a *t*-test analysis using Microsoft Excel.

tig of 7,552,296 bp while the other three smaller contigs had no predicted gene cluster for secondary metabolites. Among 20 predicted gene clusters of *Streptomyces* sp. SS1-1, five gene clusters had 100% similarity with known gene clusters in *Streptomyces* species. They were the gene clusters for ectoine (with the ectoine gene cluster of *Streptomyces anulatus*), melanin (with the melanin gene cluster of *Streptomyces coelicolor* A3(2)), terpene (with the albaflavenone gene cluster of *Streptomyces coelicolor* A3(2)), terpene/butyrolactone (with the gamma-butyrolactone gene cluster of *Streptomyces acidiscabies*), terpene (with the isorenieratene gene cluster of *Streptomyces argillaceus*). Five gene clusters in *Streptomyces* sp. SS1-1 showing moderate to high similarities with known gene clusters in *Streptomyces* species were terpene (with hopene of *Streptomyces coelicolor* A3(2) with 92% similarity), siderophore (with desferrioxamine B of *Streptomyces coelicolor* A3(2) with 83% similarity), type 2 polyketide (with spore pigment of *Streptomyces avermitilis* with 75% similarity), melanin (with melanin of *Streptomyces avermitilis* with 57% similarity), and type 1 PKS (with hitachimycin of *Streptomyces scarisporus* with 54% similarity). These gene clusters in *Streptomyces* sp. SS1-1 shared most of the similar genes with known gene clusters in *Streptomyces* species (terpene, siderophore, type 2 polyketide gene clusters) or at least contained core biosynthetic genes and additional biosynthetic genes with known gene clusters in *Streptomyces* species (melanin, hitachimycin gene clusters). Six gene clusters in *Streptomyces* sp. SS1-1 for siderophores, NRPS/NRPS-like, type 3 PKS, type 1 PKS/NRPS, butyrolactone, bacteriocin (Table 2) had very low similarity with known gene clusters for secondary metabolites in *Streptomyces* species. In addition, they did not share core biosynthetic genes and additional biosynthetic genes with known gene clusters for secondary metabolites in *Streptomyces* species. There were three gene clusters for two terpenes and one bacteriocin in the genome of *Streptomyces* sp. SS1-1 which had no match with known gene clusters for secondary metabolites in bacteria and fungi in the database of the AntiSMASH program (Table 2). A search in literature for cytotoxic activity of ectoine, melanin, albaflavenone, gamma-butyrolactone, isorenieratene, hopene, desferrioxamine B, spore pigment, hitachimycin revealed that hitachimycin, ectoine, melanin possess cytotoxic effects on cancer cell lines and brine shrimp model [5–7]. Further experiments will be performed to identify the involvement of the potential biosynthetic gene clusters in the cytotoxic activity of *Streptomyces* sp. SS1-1.

The draft genome sequence of *Streptomyces* sp. SS1-1 has been deposited in GenBank under the accession number NZ\_WBXN000000001. The raw sequencing data were submitted to the SRA database under the accession number SRX6989938 at [https://www.ncbi.nlm.nih.gov/sra/SRX6989938\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX6989938[accn]).

## 2. Experimental design, materials and methods

For isolating *Streptomyces* sp. SS1-1 from *Catharanthus roseus*, the plant was collected, transported to the laboratory and processed immediately within 2–3 h. Stems were cut with a size of 3 cm and thoroughly washed in running tap water to remove epiphytes and other adhered materials on the surface of the stem. Then, surface sterilization of the stem sample was performed according to the protocol of Qin et al. [8] with modifications. Firstly, stems were rinsed with 0.1% Tween 20 for a few seconds and washed with sterile distilled water. Then, the stem sample was immersed in 5% NaOCl in 5 min and then transferred to 5% Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> for 10 min to neutralize NaOCl. In the following steps, the stem sample was soaked in 70% ethanol for 5 min and then washed with sterile distilled water to remove residual ethanol. Finally, the stem sample was rinsed in 10% NaHCO<sub>3</sub> for 10 min and allowed to air-dry in the cabinet. The stem sample was then ground in a mortar with sterile 0.9% NaCl and the plant extract was spread on the Soya Flour Mannitol (SFM) agar plate containing 30 µg/ml nalidixic acid, 50 µg/ml cycloheximide, and 50 µg/ml nystatin to suppress the growth of Gram-negative bacteria and fungi. The SFM plate was incubated at 28 °C for 7 days to promote the growth of *Streptomyces* species. For checking the effectiveness of the surface sterilization protocol, the surface-sterilized tissue was imprinted on the SFM agar and the nutrient agar to observe the microbial growth. Alternatively, the spore of *Streptomyces* sp. SS1-1 was sequentially treated with the disinfecting solutions as

**Table 2**  
Putative gene clusters for secondary metabolites in the 7,552,296 bp contig of the *Streptomyces* sp. SS1-1 draft genome.

Number	Position	Type	Most similar known cluster	Type	Similarity (%)
1	1,36,466..246,539	Type 1 PKS	Hitachimycin ( <i>Streptomyces scarisporus</i> )	Polyketide	54
2	686,910..706,180	Terpene	ND	ND	ND
3	1,301,113..1,319,856	Terpene	ND	ND	ND
4	1,605,024..1,615,422	Ectoine	Ectoine ( <i>Streptomyces anulatus</i> )	Other	100
5	2,525,072..2,534,456	Melanin	Melanin ( <i>Streptomyces coelicolor</i> A3(2))	Other	100
6	2,614,349..2,605,105	Siderophore	Desferrioxamine B ( <i>Streptomyces coelicolor</i> A3(2))	Other	83
7	2,796,432..2,806,015	Butyrolactone	Fluostatin ( <i>uncultured bacterium</i> BAC AB649/1850)	Type 2 PKS	12
8	4,772,131..4,792,445	Terpene	Albaflavenone ( <i>Streptomyces coelicolor</i> A3(2))	Terpene	100
9	4,916,120..4,988,638	Type 2 PKS	Spore pigment ( <i>Streptomyces avermitilis</i> )	Type 2 PKS	75
10	5,424,194..5,434,924	Siderophore	Ficellomycin ( <i>Streptomyces ficellus</i> )	NRPS	3
11	5,501,459..5,585,717	NRPS, NRPS-like	Auroramycin ( <i>Streptomyces filamentosus</i> )	Type 1 PKS	10
12	5,652,601..5,662,898	Bacteriocin	ND	ND	ND
13	5,695,869..5,716,543	Terpene, butyrolactone	Gamma-butyrolactone ( <i>Streptomyces acidiscabies</i> )	Other	100
14	5,842,791..5,855,911	Siderophore	Grincamycin ( <i>Streptomyces lusitanus</i> )	Type 2 PKS-saccharide	5
15	6,300,385..6,326,310	Terpene	Hopene ( <i>Streptomyces coelicolor</i> A3(2))	Terpene	92
16	6,844,582..6,854,797	Bacteriocin	Informatipeptin ( <i>Streptomyces viridochromogenes</i> DSM 40,736)	Lanthipeptide	42
17	6,886,817..6,897,191	Melanin	Melanin ( <i>Streptomyces avermitilis</i> )	Other	57
18	6,940,607..7,019,736	Type 1 PKS, NRPS	Foxicins A-D ( <i>Streptomyces diastatochromogenes</i> )	NRPS-Type 1 PKS	12
19	7,129,058..7,154,715	Terpene	Isorenieratene ( <i>Streptomyces argillaceus</i> )	Terpene	100
20	7,278,601..7,319,785	Type 3 PKS	Lipopeptide 8D1-1/ Lipopeptide 8D1-1 ( <i>Streptomyces rochei</i> )	NRPS	11

ND: not determined.

above and plated on SFM agar to check whether they can grow after the surface sterilization protocol.

For testing cytotoxicity activity, *Streptomyces* sp. SS1-1 was cultured in the PY broth [9] using baffled Erlenmeyer at 28 °C. The Erlenmeyer was shaken at 180rpm for 5 days. Then, the cultured broth was filtered through a 0.45 μm filter unit and was used directly for the MTT assay and the flow cytometry analysis. For MTT assay, a quantity of  $2 \times 10^3$  A549 cells was distributed in each well of a 96-well plate containing 100 μl of DMEM (Dulbecco's modified Eagle's medium) supplemented with 10% FBS (Fetal bovine serum), 100 units/ml penicillin and 100 μg/ml streptomycin. The cells were cultured for 24 h at 37 °C in a humidified 5% CO<sub>2</sub> incubator for attachment. After that, the cell culture medium was replaced with the fresh medium containing various concentrations of the PY broth and the PY broth inoculated with *Streptomyces* sp. SS1-1 and left for 48 h. Then, the medium was discarded and 100 μl of serum-free medium and 10 μl of 5 mg/ml MTT solution (Biobasic) were added into each well. After incubation at 37 °C for 3 h, the MTT medium was replaced by 100 μl of MTT solvent and the absorbance was read at 590 nm. For flow cytometry analysis,  $1.5 \times 10^5$  of A549 cell was seeded in each well of a 6-well plate containing 5 ml of DMEM supplemented with 10% FBS and the plate was incubated for 48 h. Then, the medium was replaced by 3 ml of fresh medium containing 0%, 1.25% and 2.5% (v/v) of the *Streptomyces* sp. SS1-1 culture broth and the plate was incubated for 24 h. Cells were detached from the plate and stained with PE Annexin V and 7-AAD (Apoptosis detection kit, BD) before cell apoptosis was detected by flow cytometry.

For genome sequencing using NGS technologies, the mycelium of *Streptomyces* sp. SS1-1 was harvested from 3 days old culture in Tryptone Soya Broth and subjected for genomic DNA extraction with the Qiagen MagAttract HMW kit (Qiagen). Genomic DNA was converted into sequencing libraries using the PacBio Multiplex kit (Pacific Biosciences, Menlo Park, CA) and subsequently, sequenced with a PacBio Sequel using Sequencing Reagent Kit v2.1 by the University of Oregon GC3F facility. In parallel, genomic DNA was converted to Illumina libraries using the Illumina Nextera DNA Flex kit (Illumina, San Diego, CA) and sequenced on a HiSeq4000 with paired-end 150 bp reads (University of Oregon GC3F facility). PacBio Sequel reads were assembled with Canu 1.7 [10] with a genome size of 8 Mbp and option corOutCoverage=60. The Canu assembly was polished with the PacBio raw reads using the arrow program from PacBio. This consensus was then polished using Pilon [11] and the Illumina reads.

Genome completeness of *Streptomyces* sp. SS1-1 was checked by BUSCO [12]. Genome sequence comparison was performed by BMap's MinHash Sketch program. Gene prediction and annotation were carried out using the NCBI Prokaryotic Genome Annotation Pipeline [13]. AntiSMASH program [14] was used to search putative gene clusters for secondary metabolites in *Streptomyces* sp. SS1-1 genome.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRedit authorship contribution statement

**Quan Dang Nguyen:** Investigation, Writing - original draft, Writing - review & editing. **Phung Minh Truong:** Investigation. **Thao Nguyen Thanh Vo:** Investigation, Formal analysis. **Truc Dao Xuan Chu:** Investigation. **Chuong Hoang Nguyen:** Conceptualization, Supervision, Writing - original draft, Writing - review & editing.

## Acknowledgments

The authors wish to thank Center for Research and Application in Bioscience for the financial support for this work.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.dib.2020.105497](https://doi.org/10.1016/j.dib.2020.105497).

## References

- [1] S. Roy, D. Banerjee, Broad spectrum antibacterial activity of granaticinic acid, isolated from *streptomyces thermoviolaceus* NT1; an endophyte in *catharanthus roseus* (L.) G. Don, J. Appl. Pharm. 5 (2015) 006–011.
- [2] P.P.C. Palem, G.C. Kuriakose, C. Jayabaskaran, An endophytic fungus, *talaromyces radicus*, isolated from *catharanthus roseus*, produces vincristine and vinblastine, which induce apoptotic cell death, PLoS ONE 10 (2015) e0144476.
- [3] D.H. Yan, X. Song, H. Li, T. Luo, G. Dou, G. Strobel, Antifungal activities of volatile secondary metabolites of four *diaporthe* strains isolated from *catharanthus roseus*, J. Fungi 4 (2018), doi:10.3390/jof4020065.
- [4] E.F. Rakotoniriana, G. Chataigné, G. Raoelison, C. Rabemanantsoa, F. Munaut, M. El Jaziri, S. Urveg-Ratsimamanga, J. Marchand-Brynaert, A.M. Corbisier, S. Declerck, J. Quetin-Leclercq, Characterization of an endophytic whorl-forming *streptomyces* from *catharanthus roseus* stems producing polyene macrolide antibiotic, Can. J. Microbiol. 58 (2012) 617–627.
- [5] I. Umezawa, H. Takeshima, K. Komiyama, Y. Koh, H. Yamamoto, M. Kawaguchi, A new antitumor antibiotic, stubomycin, J. Antibiot. 34 (1981) 259–265.
- [6] M. Sheikhpour, A. Sadeghi, F. Yazdian, A. Movafagh, A. Mansoori, Anticancer and apoptotic effects of ectoine and hydroxyectoine on non-small cell lung cancer cells: an in-vitro investigation, Multidiscip. Cancer Investig. 3 (2019) 14–19.
- [7] D.N. Madhusudhan, B.B. Mazhari, S.G. Dastager, D. Agsar, Production and cytotoxicity of extracellular insoluble and droplets of soluble melanin by *streptomyces* DMZ-3, Biomed. Res. Int. (2014), doi:10.1155/2014/306895.
- [8] S. Qin, J. Li, H.H. Chen, G.Z. Zhao, W.Y. Zhu, C.L. Jiang, L.H. Xu, W.J. Li, Isolation, diversity, and antimicrobial activity of rare actinobacteria from medicinal plants of tropical rain forests in Xishuangbanna, China, Appl. Environ. Microbiol. 75 (2009) 6176–6186.
- [9] K.H. Lee, K.W. Kim, K.H. Rhee, Identification of *streptomyces* sp. KH29, which produces an antibiotic substance processing an inhibitory activity against multidrug-resistant *acinetobacter baumannii*, J. Microbiol. Biotechnol. 20 (2010) 1672–1676.
- [10] S. Koren, B.P. Walenz, K. Berlin, J.R. Miller, A.M. Phillippy, Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation, Genome Res. 27 (2017) 722–736.
- [11] B.J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C.A. Cuomo, Q. Zeng, J. Wortman, S.K. Young, A.M. Earl, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, PLoS ONE 9 (2014) e112963.
- [12] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics 31 (2017) 321–3212.
- [13] T. Tatusova, M. DiCuccio, A. Badretdin, V. Chetvernin, E.P. Nawrocki, L. Zaslavsky, A. Lomsadze, K.D. Pruitt, M. Borodovsky, J. Ostell, NCBI prokaryotic genome annotation pipeline, Nucleic Acids Res. 44 (2016) 6614–6624.
- [14] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S.Y. Lee, M.H. Medema, T. Weber, antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline, Nucleic Acids Res. 47 (2019) W81–W87.