

OPEN

# *Streptomyces* sp. VN1, a producer of diverse metabolites including non-natural furan-type anticancer compound

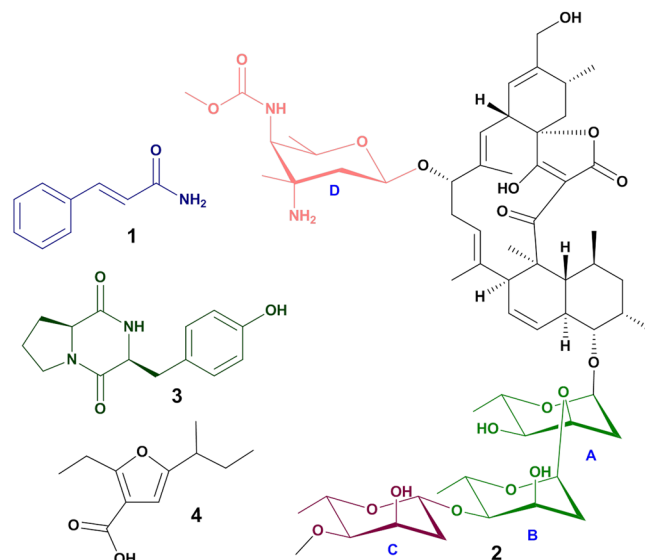
Hue Thi Nguyen<sup>1</sup>, Anaya Raj Pokhrel<sup>1</sup>, Chung Thanh Nguyen<sup>1</sup>, Van Thuy Thi Pham<sup>1</sup>, Dipesh Dhakal<sup>1</sup>, Haet Nim Lim<sup>1</sup>, Hye Jin Jung<sup>1,2</sup>, Tae-Su Kim<sup>1</sup>, Tokutaro Yamaguchi<sup>1,2</sup> & Jae Kyung Sohng<sup>1,2\*</sup>

*Streptomyces* sp. VN1 was isolated from the coastal region of Phu Yen Province (central Viet Nam). Morphological, physiological, and whole genome phylogenetic analyses suggested that strain *Streptomyces* sp. VN1 belonged to genus *Streptomyces*. Whole genome sequencing analysis showed its genome was 8,341,703 base pairs in length with GC content of 72.5%. Diverse metabolites, including cinnamamide, spiro-tetronate antibiotic lobophorin A, diketopiperazines cyclo-L-proline-L-tyrosine, and a unique furan-type compound were isolated from *Streptomyces* sp. VN1. Structures of these compounds were studied by HR-Q-TOF ESI/MS/MS and 2D NMR analyses. Bioassay-guided purification yielded a furan-type compound which exhibited *in vitro* anticancer activity against AGS, HCT116, A375M, U87MG, and A549 cell lines with IC<sub>50</sub> values of 40.5, 123.7, 84.67, 50, and 58.64 μM, respectively. *In silico* genome analysis of the isolated *Streptomyces* sp. VN1 contained 34 gene clusters responsible for the biosynthesis of known and/or novel secondary metabolites, including different types of terpene, T1PKS, T2PKS, T3PKS, NRPS, and hybrid PKS-NRPS. Genome mining with HR-Q-TOF ESI/MS/MS analysis of the crude extract confirmed the biosynthesis of lobophorin analogs. This study indicates that *Streptomyces* sp. VN1 is a promising strain for biosynthesis of novel natural products.

Natural products (NPs) have been starting points of drug discovery for several decades. Major antimicrobials and chemotherapeutics entering clinical trials are often based on NPs<sup>1,2</sup>. Drugs with a natural origin can be produced as primary or secondary metabolites from versatile living organisms. Different microorganisms such as *Streptomyces*, myxobacteria and uncultured bacteria are major sources of such beneficial NPs<sup>3,4</sup>. Moreover, different metabolic engineering approaches and sophisticated techniques employing systems biology or synthetic biology can assist in harnessing the full potential of these bacteria in terms of productivity or creating diverse products<sup>5,6</sup>. Hence, there is renewed interest in mining microorganisms for new leads owing to the remarkable success of microbial metabolites as starting points for developing effective antibiotics, anticancer agents, and agrochemicals<sup>7</sup>.

*Streptomyces* are Gram-positive, aerobic bacteria in the order of *Actinomycetales* within the class of *Actinobacteria*. Genus *Streptomyces* was first proposed by Waksman and Henrici in 1943. It was classified into the family of *Streptomycetaceae* based on its morphology and cell wall chemotype<sup>8</sup>. Previous studies have shown that more than 74% of current antibiotics are derived from the genus *Streptomyces*<sup>9</sup>. Using integrated approaches of compound screening and drug development, *Streptomyces arsenal* has been found to be able to combat antibiotic resistance<sup>5</sup>. Multiple approaches such as ribosome engineering<sup>10</sup> and genome mining<sup>11</sup> have been used to find new secondary metabolites in old *Streptomyces* strains. Besides the effort to work on old strains to explore novel biosynthetic gene clusters (BGCs), isolating new bioactive strains from the environment is also a great way to find novel BGCs<sup>12</sup>. Recently, marine bacteria have been known as sources for many novel compounds. For example,

<sup>1</sup>Department of Life Science and Biochemical Engineering, SunMoon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam, 31460, Republic of Korea. <sup>2</sup>Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam, 31460, Republic of Korea. \*email: [sohng@sunmoon.ac.kr](mailto:sohng@sunmoon.ac.kr)



**Figure 1.** Structures of compounds 1, 2, 3, and 4. Letters “A–D” indicate four sugar units in lobophorin A.

lobophorins from marine bacteria are produced by *Streptomyces* strains<sup>13,14</sup>. Streptomycin, a novel indole alkaloid compound with antibacterial activity, has been isolated from a marine *Streptomyces* sp. DA2 strain<sup>15</sup>.

In this study, we characterized a *Streptomyces* sp. VN1 isolated from the coastal region of Phu Yen Province, Da Nang, central Viet Nam. Based on phylogenetic, chemotaxonomic, and morphological characteristics, this strain belongs to genus *Streptomyces*. After a large-scale fermentation and bioassay-guided isolation, cinnamamide (1) lobophorin A (2) diketopiperazines cyclo-L-proline-L-tyrosine (3) and a furan-type compound (4) (Fig. 1) were characterized from fermentation broth of *Streptomyces* sp. VN1. Interestingly, the furan-type compound (4) exhibited anticancer activity against five types of tumor cell lines. Compound 4 effectively suppressed both migration and invasion of AGS cells. The ability of the strain to produce cinnamamide, antibacterial compounds 2 and 3<sup>16,17</sup>, anticancer compound 4, and diverse secondary metabolites make *Streptomyces* sp. VN1 an attractive target for bioactive compound screening.

## Materials and Methods

**Collection and isolation of strain.** A sea sediment sample was collected offshore at a depth of 4.7 m at Da Nang Beach in Phu Yen Province, Viet Nam (latitude: 13°3'9.972" N; longitude: 109° 11'57.444" E). *Streptomyces* sp. VN1 was isolated by the direct spreading of sediment sample onto Gause's synthetic agar plates<sup>18</sup>. After incubation at 28 °C for seven days, colonies were picked up and subcultured on modified Gause agar. The isolate strain was grown in modified Gause liquid medium containing yeast extract (3 g/L) supplemented as an additional nitrogen source. For long-term storage, pure colonies were transferred to starch-casein broth for seven days and cells were stored in 40% glycerol stock (v/v) at –80 °C. For compound production and characterization, the strain was grown on Marine broth-malt extract (MB cell).

**Morphological and phenotypic analyses.** Morphological observation of spores and mycelia of *Streptomyces* sp. VN1 was done using an FE-SEM JEOL JSM-6700 (Jeol, Japan) after culturing on starch-casein agar media for ten days at 28 °C. Sole carbon sources at a concentration of 1% (w/v) and various salts at a concentration of 0.05% were tested for growth. Growth at various temperatures (4, 10, 15, 20, 24, 28, 30, 37, 42, 50, and 55 °C) and various pH conditions (pH 3–12; 1 pH unit intervals) was also evaluated in the same culture media.

**Chemotaxonomic analyses.** Strain *Streptomyces* sp. VN1 was cultured aerobically on modified Gause liquid medium at 28 °C with shaking (160 rpm) for eight days. Cellular fatty acid compositions of the strain along with fatty acid methyl esters (FAMES) were analyzed with a 6890N Network GC system (Agilent Technology, USA) using the Sherlock Microbial Identification software package (version 6.1, database)<sup>19</sup>. Cell-wall amino acids and whole-cell carbohydrates were also extracted and analyzed as described previously<sup>20,21</sup>.

**Phylogenetic analysis.** Strain *Streptomyces* sp. VN1 was cultivated in starch-casein broth. Genomic DNA extraction was carried out according to the standard procedure<sup>22</sup>. Cloned 16S rRNA was sequenced using the Sanger sequencing method. A 16S rRNA sequence of 1,467 nucleotides in length was obtained and deposited in GenBank (accession number: KU878019). This sequence was compared with sequences obtained from the online platform EzBioCloud (<http://eztaxon-e.ezbiocloud.net/>) through BLAST<sup>23</sup>. Phylogenetic analysis of the strain was performed to determine its taxonomic position within the genus of *Streptomyces*. For the analysis of genomic sequence, automated multi-locus species tree (autoMLST)<sup>24</sup> pipeline was used to generate a phylogenetic based on alignment of default parameters (>100 core genes) present in closely related genomes using default parameters. Selected genomes were subjected to *in silico* DNA–DNA hybridization (DDH) and values were calculated using Genome-to-Genome Distance Calculator (GGDC)<sup>25</sup> version 2.1 online.

**Genome sequencing, assembly, and annotation.** Whole genome sequencing of *Streptomyces* sp. VN1 was performed using third-generation Pacific Biosciences (PacBio) RSII (Macrogen Corporation, Republic of Korea) sequencing system with single-molecule real-time (SMRT) analysis. HGAP3 was used to assemble PacBio long-reads. Filtered reads (8,558 reads, 966,726,581 bp) were then preassembled to generate long and highly accurate sequences (179,193 reads, 8,341,703 bp). For *de novo* assembly, Celera(R) Assembler was used to assemble sequences into a draft assembly. After that, the assembly was polished with a Quiver. The genome was annotated using Prokka (v1.12b). Transfer RNAs and ribosomal RNAs were analyzed using (tRNAs) tRNAscan-SE<sup>26</sup> and RNAmmer<sup>27</sup>. Their functions were predicted according to databases of COG<sup>28</sup> and EggNOG<sup>29</sup>. Secondary metabolite biosynthetic gene clusters were analyzed using antiSMASH 4.0<sup>30</sup>. The complete genome sequence of *Streptomyces* sp. VN1 was deposited in NCBI (accession number: CP036534).

**Mass spectrometric analysis of crude extraction sample.** Marine broth-malt extract was used as fermentation medium. Briefly, 5 L of fermentation broth was extracted with 10 L of methylene chloride and evaporated under reduced pressure. The concentrate was collected using 50 mL of methanol to give residue A. The supernatant was extracted for the second time with 10 L ethyl acetate and its concentrate was collected using 50 mL MeOH to give residue B. These crude extracts were analyzed by ultra-high-performance liquid chromatography electrospray ionization quadrupole time of flight high-resolution mass spectrometry (UPLC-ESI-Q-TOF-HRMS) analysis using ACQUITY UPLC<sup>®</sup> (Waters Corporations, Milford, MA, USA) coupled with SYNAPT G2-S (Waters Corporations). UPLC<sup>®</sup>/HRMS analysis was performed on an ACQUITY UPLC<sup>®</sup> BEH C18 column (1.7  $\mu$ m, 2.1  $\times$  100 mm) at a flow rate of 0.3 mL min<sup>-1</sup> using a gradient solvent mobile phase A (H<sub>2</sub>O, 0.1% trifluoroacetic acid) and B 0% to 100% ACN (0 to 12 min) at 35 °C. The injection volume was 10  $\mu$ L. Conditions used for the high-resolution mass spectrometer equipped with an electrospray ionization source were: 3 kV of capillary voltage, 300 °C of desolvation gas temperature, and 600 L/h of desolvation gas flow rate. The mass range was set from *m/z* 50 to 1,700 in positive mode. MS/MS detection mode was set to a ramp trap collision high energy from 20 V to 40 V.

**Isolation of compounds from *Streptomyces* sp. VN1 and elucidation of their structures.** Residues A and B were used as crude samples for further purification using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific) equipped with a C<sub>18</sub> column (YMC-Pack ODS-AQ, 150  $\times$  20 mm<sup>2</sup>) connected to a UV detector (220, 254, 263, and 290 nm). Residue A was subjected to HPLC using a binary gradient of solvent A, 100% water; solvent B, 100% acetonitrile; 0% B to 100% B (linear gradient, 0–20 min), 100% B (20–25 min), 100% B to 0% B (25–30 min). The flow rate was 10 mL min<sup>-1</sup> to give compound 1 (20.2 mg) and compound 2 (17.3 mg).

Residue B was subjected to HPLC using the following time program: solvent A, 100% water; solvent B, 100% acetonitrile; 0% B to 70% B (linear gradient, 0–20 min), 70% to 100% B (20–28 min), 100% B (28–30 min), 100% B to 0% B (30–40 min). The flow rate was 10 mL min<sup>-1</sup> to give compound 3 (5.2 mg) and compound 4 (6.3 mg).

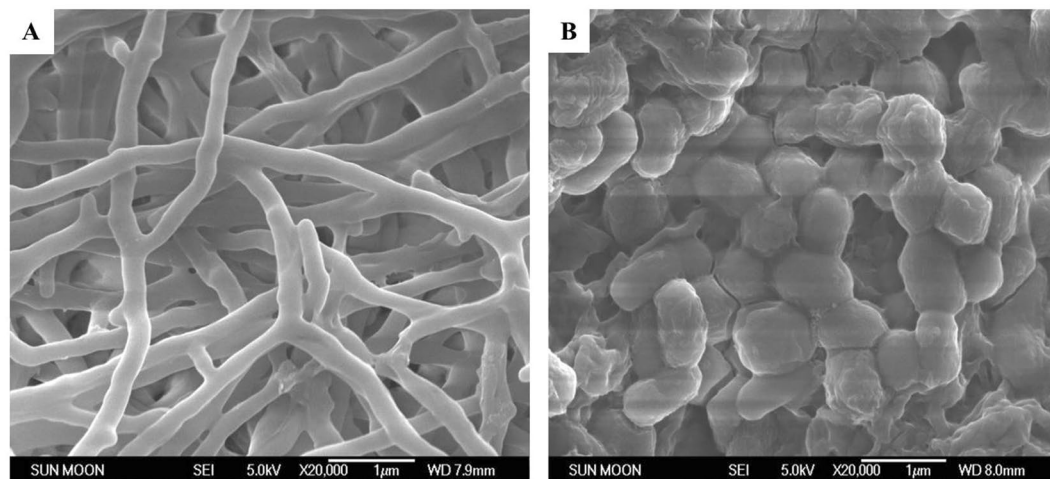
These purified compounds were dried, lyophilized, dissolved in methanol-*d*<sub>4</sub>, and subjected to 700 MHz using Bruker BioSpin nuclear magnetic resonance (NMR) (Billerica, USA) for analyses, including one-dimensional (1D) <sup>1</sup>H-NMR, <sup>13</sup>C-NMR, and two-dimensional (2D) analyses.

IR (Infrared Spectroscopy) spectra were obtained using an EQUINOX 55 FT-IR spectrometer (BRUKER Optik GmbH, Wikerstr. 13 76189 Karlsruhe, Germany). CD (Circular Dichroism) spectra were recorded using a JASCO J-715 spectrophotometer (Ochang, Republic of Korea). The CD spectrum was analyzed with a CDtoolX<sup>31</sup> software.

**Cytotoxicity assay.** Human cancer cell lines AGS, HCT116, A375SM, A549 and U87MG were obtained from the Korean Cell Line Bank (KCLB, Seoul, Korea). Human normal cell lines 267B1 and MRC-5 were purchased from the American Type Culture Collection (ATCC; Manassas, VA, USA). AGS gastric cancer and HCT116 colon cancer cells were maintained in RPMI 1640 medium containing 10% fetal bovine serum (FBS). A375SM melanoma and A549 lung cancer cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS. U87MG glioblastoma cells, normal cell lines of prostate epithelial (267B1) and lung fibroblast (MRC-5) were cultured in Minimum Essential Medium (MEM) containing 10% FBS. All cells were maintained at 37 °C in a humidified 5% CO<sub>2</sub> incubator. For cell-growth assay, various cancer cells were plated into 96-well culture plates at density of 2  $\times$  10<sup>3</sup> cells/well. After compound 4 was added to each well at various concentrations (0, 0.78, 1.56, 3.12, 6.25, 12.5, 25, 50, 100, and 200  $\mu$ M), cells were incubated at 37 °C for 72 h. Cell growth was measured using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) colorimetric assay. Briefly, 50  $\mu$ L of MTT (2 mg/mL stock solution) was added and plates were incubated at 37 °C for an additional 4 h. After removal of medium, 100  $\mu$ L of dimethyl sulfoxide (DMSO) was added to each well. The absorbance was measured at 540 nm using a microplate spectrophotometer (Thermo Scientific Multiskan<sup>®</sup> Spectrum). The cytotoxicity assays were performed on triplets and results are presented as mean values  $\pm$  standard error (SE).

**Wound-healing assay.** The migratory potential of AGS gastric cancer cells was analyzed using a wound-healing assay following published methodology<sup>32</sup>. Briefly, a confluent monolayer of AGS cells was scratched using a tip. Each well was then washed with PBS to remove nonadherent cells. Cells were then treated with compound 4 at 25  $\mu$ M or 50  $\mu$ M and then incubated at 37 °C for up to 24 h. The perimeter of the area with a central cell-free gap was confirmed under an optical microscope (Olympus). Dotted white line indicates the edge of the gap at 0 h. The wound-healing assays were performed on triplets and results are presented as mean values  $\pm$  standard error (SE).

**Cell invasion assay.** The experiment was performed following a previously described methodology<sup>32</sup>. Cell invasion was examined using Transwell chamber inserts with a pore size of 8.0  $\mu$ m. Lower and upper sides of the polycarbonate filter were coated with gelatin (1 mg/mL) and Matrigel (3 mg/mL), respectively. AGS cells were



**Figure 2.** Scanning electron microscopic image showing the morphology of *Streptomyces* sp. VN1. (A) Rectiflexible mycelium; bar, 1  $\mu\text{m}$ . (B) Smooth and oval-shaped spores; bar, 1  $\mu\text{m}$ .

seeded into the upper chamber of the filter while compound **4** at 25  $\mu\text{M}$  or 50  $\mu\text{M}$  was added to the lower chamber filled with medium. The chamber was incubated at 37  $^{\circ}\text{C}$  for 24 h. Cells were then fixed with methanol and stained with hematoxylin/eosin. The total number of cells that invaded the lower chamber of the filter was counted using an optical microscope (Olympus). The cell invasion assays were performed on triplets and results are presented as mean values  $\pm$  standard error (SE).

**Statistical analysis.** Results are presented as mean values  $\pm$  standard error (SE). Student's t-test was used to determine statistical significance between control and test groups. Statistical significance was considered at  $p < 0.05$ .

## Results

**Morphological and phenotypic characteristics.** Scanning electron microscope observations revealed that the substrate mycelium of *Streptomyces* sp. VN1 was rectiflexible. When the culture of *Streptomyces* sp. VN1 reached maturity and adequate aerial mycelium was produced, aerial hyphae differentiated into short, straight to flexuous chains (Fig. 2A) with smooth surfaces. Its spores were found to be regular and smooth (Fig. 2B). *Streptomyces* sp. VN1 displayed survival on sole carbon sources of L-arabinose, fructose, mannitol, sucrose, xylose, lactose, and starch. Starch and yeast extract were preferred carbon source and nitrogen source, respectively (Table 1). *Streptomyces* sp. VN1 produced brownish to grayish mycelium with good sporulation on the media (i.e., Marine broth-malt extract medium) used for metabolite isolation. It showed growth at temperatures of 24  $^{\circ}\text{C}$  to 40  $^{\circ}\text{C}$  with pH between 4 and 9. Various salts had different effects on the growth of *Streptomyces* sp. VN1 (Supplementary Table S1).

**Chemotaxonomic characteristics.** Predominant fatty acids in *Streptomyces* sp. VN1 were C16:0 iso (23%), C15:0 antesisio (19%), C15:0 iso (10%), and C16:0 (11%). Minor fatty acid methyl esters (FAME) such as C16:0 9-methyl (2.7%) and unsaturated fatty acids such as C16:1 iso (3.3%), C16:1 cis (7.5%), and C17:1 anteiso (2.1%) were detected (Supplementary Table S2). Fatty acids in genus *Streptomyces* are known to contain straight chains as well as iso- and anteiso-branched chains<sup>20</sup>. These FAMES have been previously found in other species of *Streptomyces*<sup>33</sup>. Cell-wall peptidoglycans of *Streptomyces* sp. VN1 contained LL-diaminopimelic acid, typical of cell-wall type I. Galactose, arabinose, and xylose were detected as major carbohydrates in its whole-cell hydrolysate.

**Phylogenetic analysis and genome annotation.** Whole genome sequencing of *Streptomyces* sp. VN1 produced a total of 179,193 sequence reads, yielding a total consensus of 8,341,703 bp with GC content of 72.5% distributed within one main contig. There are 7,176 protein-coding genes, with an average ORF length of 1,018 bp. Coding density was about 87.57%. Within *Streptomyces* sp. VN1, 86 tRNAs and 18 rRNA operons were predicted (Table 2). Predicted proteins were annotated by blasting the eggNOG database. In eggNOG functional classification, 6,987 (97.36%) of 7,176 proteins were assigned. The following four top categories were classified: transcription, carbohydrate metabolism, amino-acid metabolism, and energy production (Supplementary Fig. S1). Genome analysis with autoMLST showed that this new isolate had the highest sequence similarities with *Streptomyces* sp. FXJ7.023 (GCF\_000404005), *Streptomyces pactum* (GCF\_001767375) and *Streptomyces olivaceus* (GCF\_000721235) (Fig. 3). Estimated average nucleotide identities of *Streptomyces* sp. VN1 with genomes of *Streptomyces* sp. FXJ7.023, *Streptomyces pactum* and *Streptomyces olivaceus* were 99.4%, 99.2%, and 98.8%, respectively. Similarly, *in-silico* DDH values yielded high sequence similarities with *Streptomyces* sp. FXJ7.023 (99.87%),

Characteristics	<i>Streptomyces</i> sp. VN1
Spore morphology	Short, flexuous chains
Spore surface	Smooth
Production of diffusible pigments	–
Growth at 24°C	+
Growth at pH 5	+
Growth at pH 11	–
<b>Growth on carbon sources</b>	
L-Arabinose	+
Fructose	+
Mannitol	+
Sucrose	+
Xylose	+
Lactose	+
Starch	+
<b>Growth on nitrogen sources</b>	
Cystine	+
L-Proline	+
Glycine	+
L-Asparagine	+

**Table 1.** Morphological and physiological characteristics and carbon usage of *Streptomyces* sp. VN1.

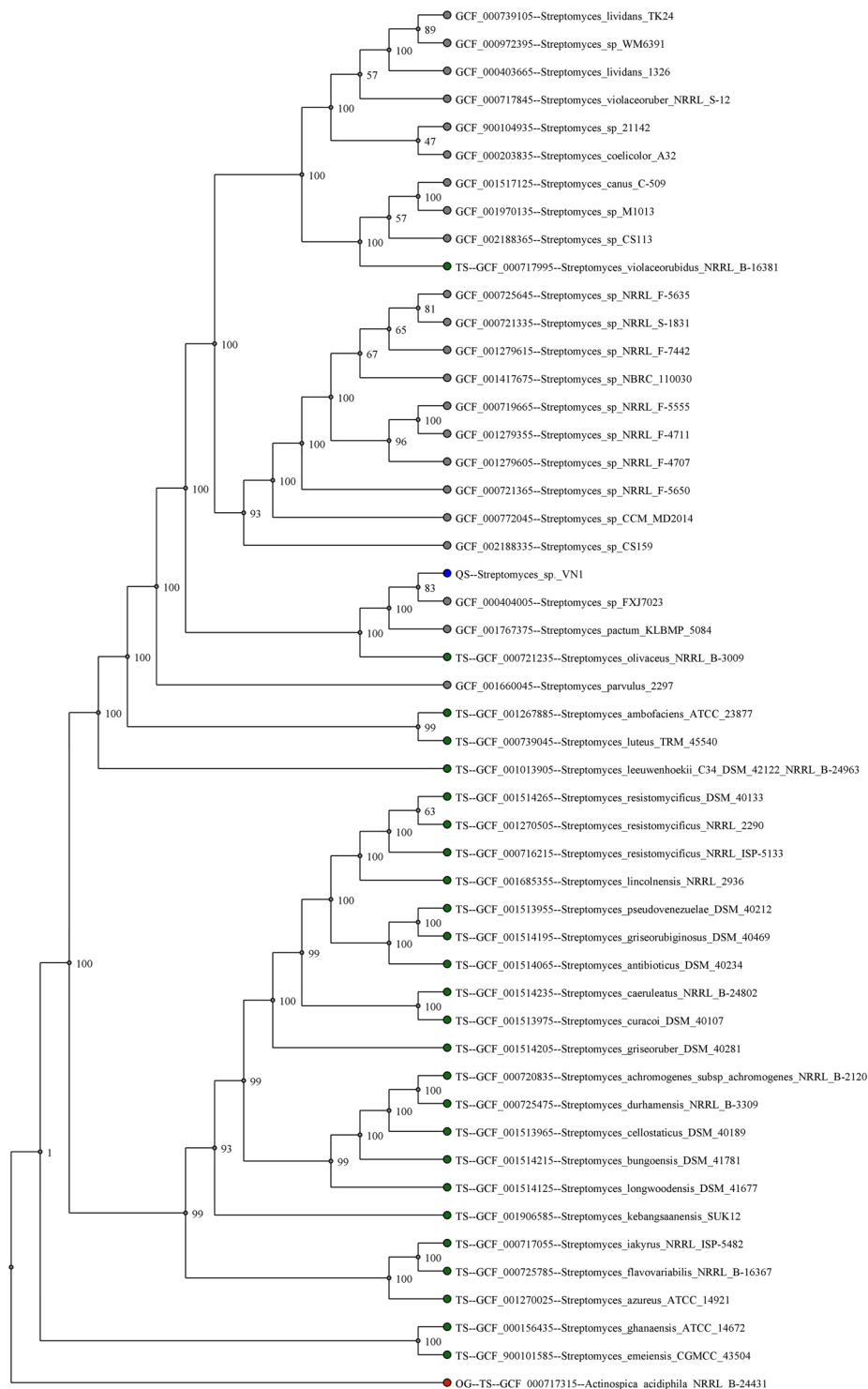
Sample	<i>Streptomyces</i> sp. VN1
Length (bp)	8,341,703
No. of reads	179,193
Coding density (%)	87.57
Average CDS length (bp)	1,018
No. of protein-coding genes	7,716
No. of tRNA genes	86
No. of rRNA	18
GC content	72.5

**Table 2.** General characteristics of the genome of isolated *Streptomyces* sp. VN1.

*Streptomyces pactum* (99.71%) and *Streptomyces olivaceus* (99.6%) (Supplementary Table S3). These data indicate that *Streptomyces* sp. VN1 is most closely related to *Streptomyces pactum* and *Streptomyces olivaceus*.

**In silico analysis of secondary metabolite biosynthesis pathways.** Thirty-four secondary metabolite biosynthetic gene clusters were identified in the *Streptomyces* sp. VN1, including type I polyketide synthases (T1PKS), type II polyketide synthases (T2PKS), type III polyketide synthases (T3PKS), non-ribosomal peptide synthetases (NRPS), terpenes, lassopeptides, thiopeptides, lanthipeptides, indoles, siderophores, bacteriocins, T3PKS-NRPS, T1PKS-NRPS, and other products (Supplementary Table S4). BGCs for common secondary metabolites found in *Streptomyces* including 2-methylisoborneol, ectoine, melanin, hopene, and coelichelin were found to be present in *Streptomyces* sp. VN1. Analysis results showed that *Streptomyces* sp. VN1 contained seven of PKS, terpene, and NRPS gene clusters related to antibiotics, displaying a similarity of more than 40%, including lobophorin A, carotenoid, friulimicin, xiamycin, enterocin, divergolide, and informatipeptin biosynthetic gene clusters (Table 3).

Cluster 1 contained an oligosaccharide-T1PKS-T3PKS-NRPS that was highly similar to lobophorin BGCs from *Streptomyces* sp. FXJ7.023 and *Streptomyces* sp. SCSIO 01127 (Table 3). A detailed analysis of the lobophorin A BGC of *Streptomyces* sp. VN1 indicated the involvement of 43 ORFs (open reading frame) and five catalytic domains related to polyketide synthases (Fig. 4). Lobophorin A biosynthesis gene clusters of *Streptomyces* sp. VN1 is highly similar to previously characterized BGCs which shared around 87–100% identities with corresponding genes in *Streptomyces* sp. FXJ7.023 and *Streptomyces* sp. SCSIO 01127 (Supplementary Table S5A). Next to the lobophorin A biosynthetic gene cluster was predicted for the T3PKS-NRPS biosynthetic gene cluster. It contained 48 ORFs, one catalytic domain related to chalcone/stilbene synthase, two catalytic domains related to polyketide synthases and three catalytic domains related to non-ribosomal peptide synthetases. Recently, this T3PKS-NRPS biosynthetic gene cluster has been characterized as being responsible for synthesizing totopotensamide<sup>34</sup>. In-depth bioinformatic analysis of the putative biosynthesis gene cluster for totopotensamide in *Streptomyces* sp. VN1 is highly similar to previously characterized BGC which showed about 96–100% similarity with corresponding genes in *Streptomyces* sp. SCSIO 02999 (Supplementary Table S5B). It has been proposed that *orf51* gene in *Streptomyces* sp. VN1, *lobR3* in *Streptomyces* sp. FXJ7.023, and *lobR1* in *Streptomyces* sp. SCSIO 01127 regulate



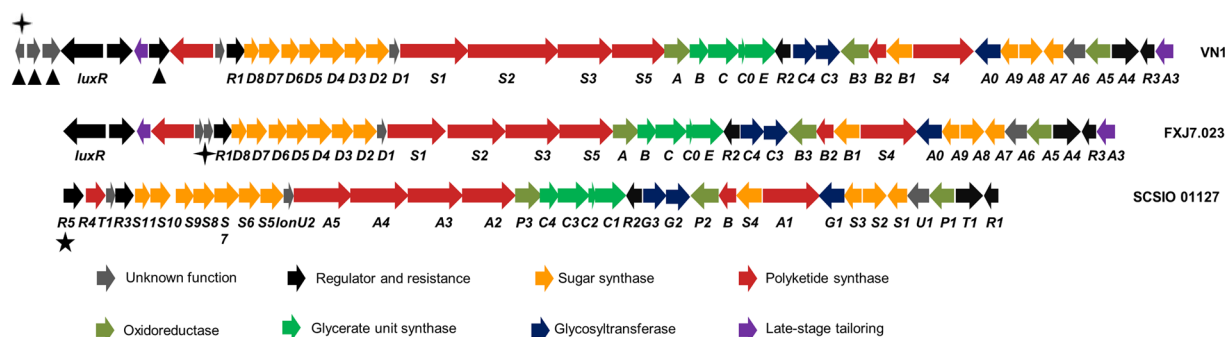
**Figure 3.** Molecular phylogenetic analysis of using default parameters (>100 core genes) by autoMLST. Bootstrap confidence levels are indicated at internodes whereas scale bar indicates nucleotide substitutions per nucleotide position.

lobophorin biosynthesis<sup>35,36</sup>. However, *lobR1* has been recently annotated as *totR5* belonging to the *tetR* family transcriptional regulator that is responsible for regulating totopotensamides synthesis<sup>37</sup> in *Streptomyces* sp. SCSIO 02999. Besides, the end of cluster 1 in *Streptomyces* sp. VN1 contained four ORFs annotated for four regulator genes and a unique ORF (*orf98*) annotated for N-acetyltransferase (Supplementary Table S5B).

Analysis results suggested that cluster five is highly similar with previously characterized carotenoid biosynthesis gene cluster. It shared 54% similarity with an existing cluster in *Streptomyces avermitilis* (Supplementary Fig. S2). This cluster contained 20 ORFs for biosynthesis of carotenoid in *Streptomyces* sp. VN1. Core ORFs

No	Cluster	Type	From	To	Most similar known biosynthetic gene cluster (percent of similarity)	Reference strain	Accession number
1	1	Oligosaccharide-T1PKS-T3PKS-NRPS	10491	225803	Lobophorin A (96%)	<i>Streptomyces</i> sp. FXJ7.023	JX306680
2	2	Terpene	237497	258555	2-Methylisoborneol (100%)	<i>Streptomyces griseus</i>	AP009493
3	5	Terpene	539486	563714	Carotenoid (54%)	<i>Streptomyces avermitilis</i>	AB070934
4	9	Ectoine	1684875	1695273	Ectoine (100%)	<i>Streptomyces anulatus</i>	AY524544
5	10	Melanin	2650272	2660898	Melanin (100%)	<i>Streptomyces coelicolor</i> A3(2)	AL645882
6	11	Lasso peptide	2720839	2743360	SSV-2083 (50%)	<i>Streptomyces svicens</i>	NZ_CM000951
7	12	Siderophore	2756614	2768407	Desferrioxamine_B (83%)	<i>Streptomyces coelicolor</i> A3(2)	AL645882
8	15	Lantipeptide	4240796	4272172	SBI-06990 alpha/SBI-06989 beta (50%)	<i>Streptomyces bingchengensis</i>	CP002047
9	17	Terpene	5321623	5342708	Albaflavenone B (100%)	<i>Streptomyces coelicolor</i> A3(2)	AL645882
10	18	T2PKS	5408372	5450926	Spore_pigment (66%)	<i>Streptomyces avermitilis</i>	AB070937
11	20	T1PKS-NRPS	6048796	6212376	Frulimicin (75%)	<i>Actinoplanes friuliensis</i>	AJ488769
12	21	T1PKS-NRPS	6260144	6309491	Xiamycin (77%)	<i>Streptomyces</i> sp. SCSIO 02999	JQ812811
13	26	T2PKS	6650019	6692435	Enterocin (95%)	<i>Streptomyces maritimus</i>	AF254925
14	28	Terpene-NRPS	7137719	7221298	Hopene (92%)	<i>Streptomyces coelicolor</i> A3(2)	AL645882
15	29	T1PKS	7275849	7358369	Divergolide (100%)	<i>Streptomyces</i> sp. HKI0576	HF563079
16	30	Bacteriocin	7687326	7697541	Informatipeptin (42%)	<i>Streptomyces viridochromogenes</i> DSM 40736	GG657757
17	32	NRPS	7898917	7949836	Coelichelin (100%)	<i>Streptomyces coelicolor</i> A3(2)	AL645882

**Table 3.** Overview of *Streptomyces* sp. VN1 genome analysis by antiSMASH of 17 secondary metabolites of biosynthetic gene clusters and seven PKS and NRPS gene clusters sharing similarity of more than 40%. The number of BGCs is determined with antiSMASH and ClusterFinder OFF.



**Figure 4.** Genetic organization of the lobophorin biosynthetic gene cluster of *Streptomyces* sp. VN1 (GenBank: SUB5241063), lobophorin gene cluster of *Streptomyces olivaceus* FXJ7.023 (GenBank: JX306680), and lobophorin gene cluster of *Streptomyces* sp. SCSIO 01127 (GenBank: KC013978). “▲”; “+”; “★” denote unique genes in *Streptomyces* sp. VN1, *Streptomyces olivaceus* FXJ7.023, and *Streptomyces* sp. SCSIO 01127, respectively.

biosynthesis such as phytoene synthase, lycopene cyclase, dehydrogenase, and methyltransferase were annotated. This cluster of *Streptomyces* sp. VN1 was predicted to have unique ORFs annotated as triacylglycerol lipases and short chain dehydrogenase/reductases (SDR), indicating that *Streptomyces* sp. VN1 might have the potential to synthesize carotenoid analogs.

Furthermore, two NRPS-hybrid synthases including a T1PKS-NRPS gene cluster showed high similarities to frulimicin and xiamycin biosynthetic gene cluster present in *Actinoplanes friuliensis* and *Streptomyces* sp. SCSIO 02999. From the frulimicin biosynthetic gene cluster, 12 core ORFs were annotated. The putative frulimicin biosynthetic gene cluster from *Streptomyces* sp. VN1 contained an additional condensation domain containing protein, a phosphopantetheine-binding domain containing protein synthetase, and an AMP-dependent protein and ligase (Supplementary Fig. S3). Therefore, *Streptomyces* sp. VN1 has high potential to produce frulimicin analogs. The putative xiamycin biosynthesis gene cluster from *Streptomyces* sp. VN1 was found to contain 33 ORFs (Supplementary Fig. S4). In this gene cluster, core and additional biosynthetic genes were annotated to polyprenyl synthetase, cytochrome P450, acyl-CoA dehydrogenase, and a crotonyl-CoA reductase/alcohol dehydrogenase including additional dehydrogenase. Compared to the most-similar known cluster of *Streptomyces* sp. SCSIO 02999, this gene cluster of *Streptomyces* sp. VN1 was found to contain a unique ORF annotated as hybrid polyketide synthetase and non-ribosomal peptide synthase. This indicates that the strain *Streptomyces* sp. VN1 might have potential to synthesize xiamycin analogs.

In bioinformatics analysis of the *Streptomyces* sp. VN1 genome, one T2PKS was annotated for an enterocin biosynthesis gene cluster that was 95% similar to the enterocin biosynthesis gene cluster from *Streptomyces*

No	<sup>13</sup> C (ppm)	<sup>1</sup> H (ppm)	Intensities
1			
2	167.15		
3	103.67		
4	99.51	5.99 (s)	1H
5	166.88		
6	15.89	2.42 (q, <i>J</i> = 7.4)	2H
7	11.50	1.05 (t, <i>J</i> = 7.4)	3H
8	167.48		
9	16.89	1.22 (d, <i>J</i> = 6.9)	3H
10	39.55	2.48 (tq, <i>J</i> = 6.9, 7.0)	1H
11	27.10	1.57 (m) 1.70 (m)	1H 1H
12	10.51	0.91 (t, <i>J</i> = 7.5)	3H

**Table 4.** <sup>1</sup>H- and <sup>13</sup>C-NMR data of compound **4** in methanol-*d*<sub>4</sub>.

*maritimus* (Supplementary Fig. S5). In this putative enterocin biosynthesis gene cluster, eight core ORFs that contained an additional methyltransferase were annotated. Thus, *Streptomyces* sp. VN1 might produce enterocin derivatives.

Additionally, *Streptomyces* sp. VN1 was predicted to harbor T1PKS pathways. The T1PKS gene cluster is related to biosynthesis of divergolide. It showed 100% homology with the existing cluster of *Streptomyces* sp. HKI0576 (Supplementary Fig. S6).

Interestingly, biosynthetic gene cluster thirty-two was predicted to be related to bacterocin informatipeptin biosynthesis. The putative informatipeptin biosynthesis cluster of *Streptomyces* sp. VN1 displayed 42% homology with an existing cluster from *Streptomyces viridochromogenes* DSM 40736 (Supplementary Fig. S7). In the genome of *Streptomyces* sp. VN1, only three ORFs (type A lantipeptide, protease, and PAS/PAC sensor protein) were found to be related to informatipeptin biosynthesis.

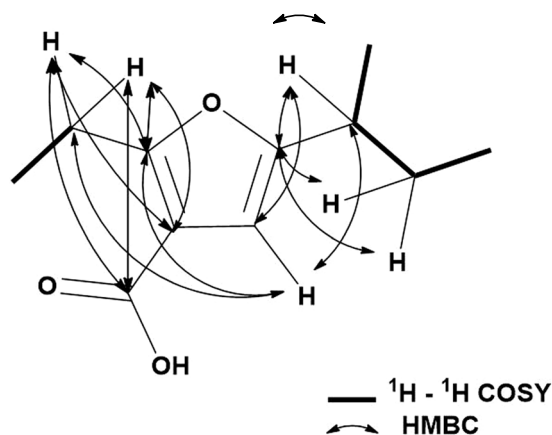
Furthermore, the T3PKS gene cluster was predicted to be related to herboxidiene biosynthesis (Supplementary Fig. S8). A comparison of the putative herboxidiene biosynthetic gene cluster with *Streptomyces chromofuscus* A7847 indicated that numerous genes in this cluster could not be annotated. This gene cluster only shared 2% homologies with the herboxidiene biosynthetic gene cluster from *Streptomyces chromofuscus* A7847.

The putative NRPS-nucleoside biosynthetic gene cluster from *Streptomyces* sp. VN1 was found to contain 34 ORFs related to nogalamycin biosynthesis (Supplementary Fig. S9). A comparison of the nogalamycin biosynthetic gene cluster from *Streptomyces nogalater* indicated that numerous genes in this cluster could not be annotated. This gene cluster only shared 40% homology with the nogalamycin biosynthetic gene cluster from *Streptomyces nogalater*.

**Structural elucidation of major compounds.** UV spectra and the HR-MS data of compounds **1** to **3** were as follows. Compound **1** (in CH<sub>3</sub>CN: H<sub>2</sub>O: trifluoroacetic acid) revealed absorption maxima at 200.73 and 273.73 nm with HR-MS *m/z* 148.0755 (calculated for C<sub>9</sub>H<sub>9</sub>NO, [M + H]<sup>+</sup> 148.0762) (Supplementary Fig. S10A). The UV spectrum of compound **2** (in CH<sub>3</sub>CN: H<sub>2</sub>O: trifluoroacetic acid) revealed absorption maxima at 210.73 and 262.73 nm with HR-MS *m/z* 1,157.6379 (calculated for C<sub>61</sub>H<sub>92</sub>N<sub>2</sub>O<sub>19</sub>, [M + H]<sup>+</sup> 1,157.6367) (Supplementary Fig. S10B). The UV spectrum of compound **3** (in CH<sub>3</sub>CN: H<sub>2</sub>O: trifluoroacetic acid) revealed absorption maxima at 201.73 and 266.73 nm with HR-MS *m/z* 261.1249 (calculated for C<sub>14</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>, [M + H]<sup>+</sup> 261.1243) (Supplementary Fig. S10C). Compounds **1**, **2**, and **3** were found to be cinnamamide, lobophorin A, and cyclo-L-proline-L-tyrosine, respectively, by comparing <sup>1</sup>H (700 MHz, methanol-*d*<sub>4</sub>) and <sup>13</sup>C (176 MHz, methanol-*d*<sub>4</sub>) spectra data (Supplementary Table S6) with data reported in the literature<sup>13,17,37</sup>.

According to spectrophotometric analysis, the UV spectrum of compound **4** (in CH<sub>3</sub>CN: H<sub>2</sub>O: trifluoroacetic acid) revealed absorption maxima at 201.73 and 290.73 nm with HR-MS *m/z* 197.1271 (calculated for C<sub>11</sub>H<sub>16</sub>O<sub>3</sub>, [M + H]<sup>+</sup> 197.1172) (Supplementary Fig. S11). The IR spectrum of **4** revealed a conjugated carboxylic acid-carbonyl stretching-vibration absorption at 1,664 cm<sup>-1</sup>, a CH<sub>3</sub>- asymmetric stretching-vibration absorption at 1,966 cm<sup>-1</sup>, a CH<sub>3</sub>- symmetric stretching-vibration absorption at 2,874 cm<sup>-1</sup>, a CH<sub>2</sub>-asymmetric stretching-vibration absorption at 2,930 cm<sup>-1</sup>, and a carboxylic acid hydrogen bonding broad -OH stretching-vibration absorption at 3,400–2,400 cm<sup>-1</sup> (Supplementary Fig. S12). <sup>1</sup>H (700 MHz, methanol-*d*<sub>4</sub>) and <sup>13</sup>C (176 MHz, methanol-*d*<sub>4</sub>) NMR (Table 4) led to structural conclusions for compound **4**. The overall structure of **4** was established mainly based on <sup>1</sup>H-<sup>1</sup>H correlation spectroscopy (COSY) and heteronuclear multiple-bond correlation (HMBC) correlations (Fig. 5). In the <sup>1</sup>H-<sup>1</sup>H COSY spectrum, cross-peaks between δ<sub>H</sub> 1.22 ppm and 2.48 ppm, δ<sub>H</sub> 2.48 ppm and 1.57 ppm, 2.48 ppm and 1.70 ppm, 1.57 ppm and 0.91 ppm, 1.70 ppm and 0.91 ppm indicated three C-C bonds from C-9 to C-12. The correlation peak of between δ<sub>H</sub> 2.42 and 1.05 indicated C-C bond of C-6 and C-7 (Supplementary Fig. S13C). In the HMBC spectrum, observed correlated signals between δ<sub>H</sub> 5.99 (s) and three carbon (δ<sub>C</sub> 103.67, 166.88 and 167.15) indicated the skeleton of C-2, C-3, C-4, and C5 based on intensities of cross-peaks. In addition, the connection of C-5 and C-10 was confirmed based on the finding of cross-peaks showing that H-9, H-10, H-11', and H-11'' were correlated to C-5 by C-H long-range coupling. Moreover, the connection of C-2 and C-6 was confirmed from the finding of cross-peaks showing that H-6 was correlated to C-2 by C-H long-range coupling. The cross-peaks of H-6 and δ<sub>C</sub> 167.48 (C-8) indicated the bonding



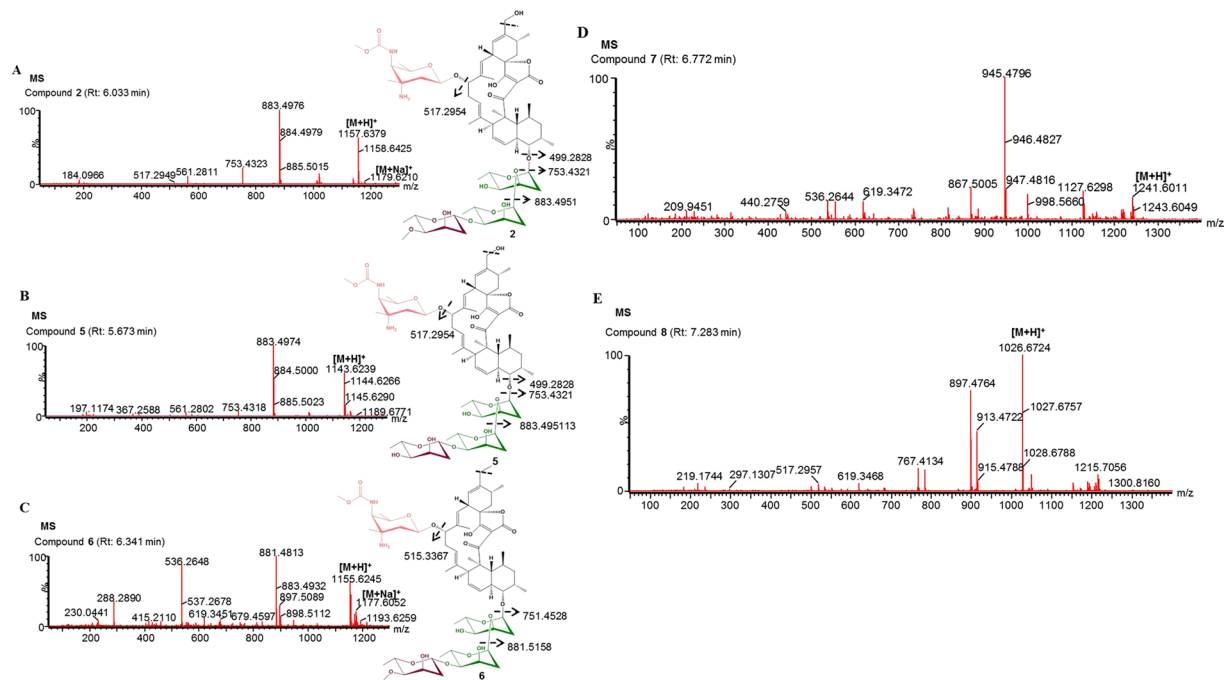


**Figure 5.** Selected key  $^1\text{H} - ^1\text{H}$  COSY (bold line) and HMBC (H ↔ C) correlations of furan type-compound 4.

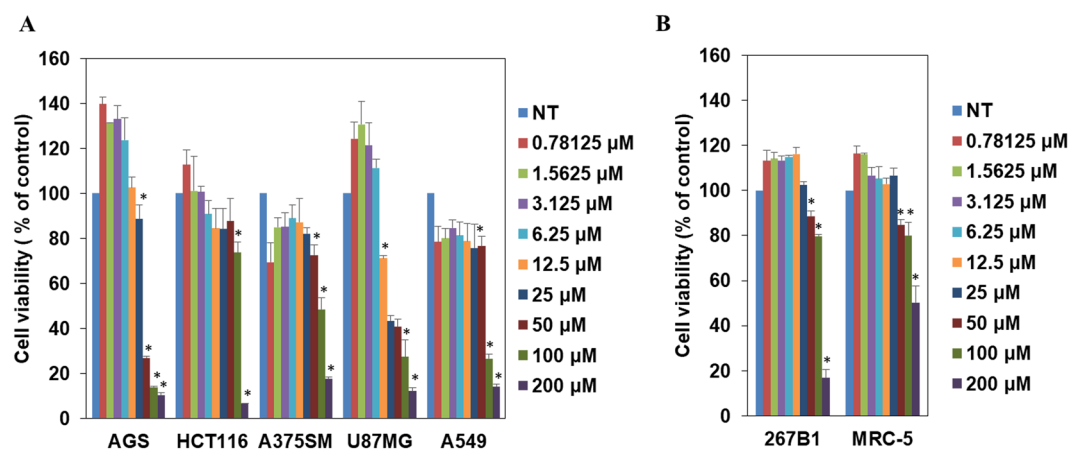
of C-3 and C-8 as the carbonyl carbon (Supplementary Fig. S13F).  $^{13}\text{C}$  NMR chemical shift values of C-5 and C-2 were around 167 ppm as the region of the carbonyl carbon of ester, amide, and carboxylic acid group. If two carbons C-5 and C-2 build the ketone form,  $^{13}\text{C}$  chemical shifts should appear at around 200 ppm. However, the  $^{13}\text{C}$  spectrum did not show chemical shifts at around 200 ppm. This observation suggests that C-5 and C-2 are carbons of the double bond with oxygen attached. The number of this oxygen should be one based on HR-MS data. This observation led to the conclusion that it has a shape of five-membered ring formations of furan. Thus, the structure of **4** was completely elucidated to be 5-(*sec*-butyl)-2-ethylfuran-3-carboxylic acid. Absolute configurations experimental CD spectra of compound **4** showed cotton effect at wavelength of 200, 205, 225, and 290 nm. The positive cotton effect in  $\pi - \pi^*$  transition of  $-\text{C}=\text{C}-$  was strong at 205 nm  $\Delta\epsilon + 15.12$  (mdeg). That of  $-\text{C}=\text{C}-\text{C}=\text{C}-$  was slightly broad at 290 nm  $\Delta\epsilon + 1.56$  (mdeg). The negative cotton effect negative in the  $\pi - \pi^*$  transition of the  $-\text{OH}$  was strong at 200 nm  $\Delta\epsilon - 19.4$  (mdeg). In the  $n - \pi^*$  transition of the  $-\text{COOH}$ , it was at 225 nm  $\Delta\epsilon - 6.7$  (mdeg) (Supplementary Fig. S14).

**Identification of lobophorin A analogs by mass spectrometry.** We further analyzed mass spectra of crude extract to determine whether predicted secondary metabolites could be detected. Mass spectra data were obtained in positive mode. From HR-MS/MS profiles, metabolic substances in crude extract samples were found to be correlated with the mass-to-charge ratio ( $m/z$ ) of molecular ions. Mass profiles of secondary metabolites existing in isolated strains were compared with genome mining data. Herein, we identified lobophorin A and its analogs in the crude extract through this technique. The proposed fragment of lobophorin A was analyzed based on the fragmentation of kijanimicin<sup>38</sup> (Supplementary Fig. S15A). In the mass profile of lobophorin A presented in Fig. 6A, fragments corresponded to  $[\text{M} + \text{H}]^+ / [\text{M} + \text{Na}]^+ 1,157.6379 / 1,179.6210$ ,  $(\text{M}-\text{B}-\text{C} + \text{H})^+ 883.4976$ ,  $(\text{M}-\text{A}-\text{B}-\text{C} + \text{H})^+ 753.4324$ , and  $(553-2\text{H}_2\text{O})^+ 517.2949$ , in full accord with the structure. Detailed fragmentation of MS/MS of lobophorin A is shown in Supplementary Fig. S15B. Compound **2** (lobophorin A, Rt: 6.033 min) and lobophorin A analogs namely compound **5** (demethylated of lobophorin A, retention time, Rt: 5.673 min), compound **6** (dehydroxylated lobophorin A, Rt: 6.341 min), compound **7** (Rt: 6.772 min) and compound **8** (Rt: 7.283 min), were detected (Fig. 6) in fraction A. Possible structures and fragmentation of compounds **5** and **6** were compared with the mass profile of lobophorin A and previously described structures<sup>36,39</sup>. These compounds displayed UV spectrum and fragmentation similar to those of lobophorin A (Supplementary Fig. S15). The first analog, compound **5** showed  $[\text{M} + \text{H}]^+$  ions at  $m/z$  1,143.6239 (calculated for  $\text{C}_{60}\text{H}_{91}\text{N}_2\text{O}_{19}$ :  $[\text{M} + \text{H}]^+ 1,143.6211$ ) (Fig. 6B). This analog has been verified and characterized by MS/MS and NMR in a previous study<sup>39</sup>. Peaks generated by MS/MS analysis showed that main fragment ions were at 883.4966, 753.4326, and 517.2950 (Supplementary Fig. S15C). The second analog, compound **6** showed  $[\text{M} + \text{H}]^+ / [\text{M} + \text{Na}]^+$  ions at  $m/z$  1,155.6245 / 1,177.6052 (calculated for  $\text{C}_{62}\text{H}_{95}\text{N}_2\text{O}_{18}$ :  $[\text{M} + \text{H}]^+ 1,155.6574$ ;  $\text{C}_{62}\text{H}_{94}\text{N}_2\text{O}_{18}\text{Na}$ :  $[\text{M} + \text{Na}]^+ 1,177.6399$ ) (Fig. 6C). Peaks generated by MS/MS analysis showed that main fragment ions were at 881.4817, 751.4147, 619.3461, and 515.2808 (Supplementary Fig. S15D). Interestingly, the third derivative, compound **7** showed  $[\text{M} + \text{H}]^+$  ions at  $m/z$  1,2041.6011 (Fig. 6D). Peaks generated by MS/MS analysis showed that main fragment ions were at 945.4792, 867.5002, 737.5001, 619.3464, 517.2946, and 499.2837 (Supplementary Fig. S15E). It was noteworthy that the fourth analog, compound **8** showed  $[\text{M} + \text{H}]^+$  ions at  $m/z$  1,026.6724 (Fig. 6E). Peaks generated by MS/MS analysis showed that main fragment ions were at 895.5777, 783.4073, 619.3467, 517.2947, and 455.3103 (Supplementary Fig. S15F). To the best of our knowledge, there is no structural reference for compound **7** or **8**. They might be new analogs of lobophorin. We also analyzed mass spectra of crude extract residue B. However, we could not correlate the data with the mass of any secondary metabolite (data not shown).

**Anticancer activity.** To assess whether compound **4** might have anticancer activity, we evaluated the inhibitory effect of compound **4** on the growth of different tumor cell lines. Compound **4** showed anti-proliferative activities against five types of cancer cell lines (Fig. 7A). This compound exhibited more sensitive growth-inhibitory activities for gastric adenocarcinoma (AGS), glioblastoma (U87MG), and lung cancer (A549) than for melanoma



**Figure 6.** HR-MS and MS/MS analyses of lobophorin analogs. (A) The component of lobophorin A generates a  $[M + H]^+$  ion at  $m/z$  1,157.6379. (B) The component of compound 5 (demethylation of lobophorin A) generates a  $[M + H]^+$  ion at  $m/z$  1,143.6239. (C) The component of compound 6 (dehydroxylation of lobophorin A) generates a  $[M + H]^+$  ion at  $m/z$  1,155.6245. (D) The component of compound 7 generates a  $[M + H]^+$  ion at  $m/z$  1,2041.6011. (E) The component of compound 8 generates a  $[M + H]^+$  ion at  $m/z$  1,026.6724.

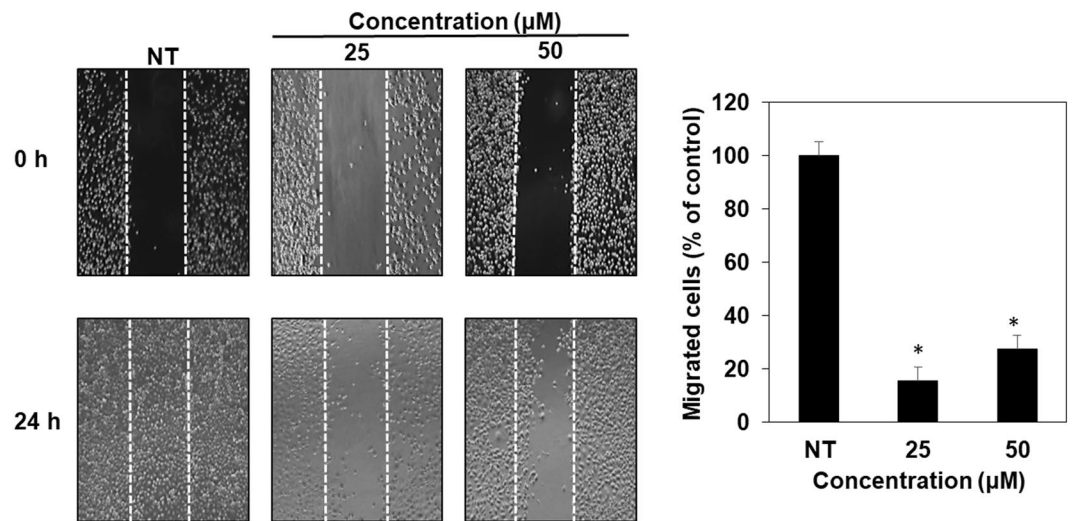


**Figure 7.** *In vitro* growth inhibitory activities of compound 4 against different cell lines. (A) cancer cell lines, (B) normal cell lines. NT: no treatment. \* $p < 0.05$  vs. control.

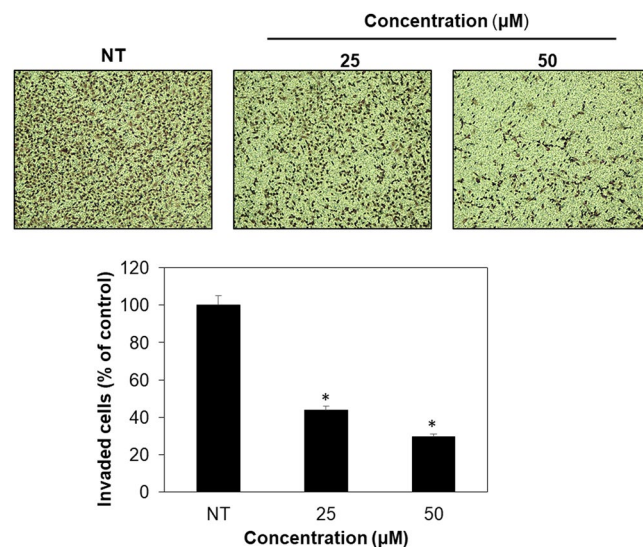
Compound 4	AGS (Gastric cancer)	HCT116 (Colon cancer)	A375SM (Melanoma)	U87MG (Glioblastoma)	A549 (Lung cancer)	267B1 (prostate epithelial)	MRC-5 (lung fibroblast)
IC <sub>50</sub> (μM)	40.5	123.7	84.67	50	58.64	157.2	>200

**Table 5.** IC<sub>50</sub> value of compound 4.

(A375SM) and colon cancer cells (HCT116) (Table 5; Fig. 7A). Compound 4 more effectively inhibited the growth of cancer cell lines than that of normal cell lines (267B1 and MRC-5) at tested doses (Table 5; Fig. 7B). Besides its inhibitory effect against cancer cell growth, the anti-metastatic activity of compound 4 against AGS gastric cancer cell line was also determined through migration and invasion assays. We found that compound 4 effectively suppressed both the migration and invasion of AGS cells after 24 h of treatment at concentrations of 25 and 50 μM



**Figure 8.** Cell migration inhibitory activity of furan-type compound against AGS cancer cells. NT: no treatment. \* $p < 0.05$  vs. control.



**Figure 9.** Cell invasion inhibitory activity of furan-type compound against AGS cancer cells. NT: no treatment. \* $p < 0.05$  vs. control.

(Figs. 8 and 9). These data suggest that compound 4 possesses an anticancer activity by inhibiting the growth and metastasis abilities of cancer cells.

## Discussion

Marine *Streptomyces* are valuable sources of various secondary metabolites for drug discovery. In this study, we found that *Streptomyces* sp. VN1 as a new strain produced many active compounds, including cinamamide, lobophorin A, cyclo-L-proline-L-tyrosine, and a furan-type compound based on HR-MS/MS and 2D-NMR studies. Besides these compounds, possible analogs of lobophorin A were found based on HR-MS/MS analysis. We found that fragments of those compounds were similar to a fragment of lobophorin A. However, in mass profiles they appeared in trace amounts. Thus, we were unable to elucidate structures of these compounds.

Cinamamide is the product of phenylalanine ammonia lyase. It has been examined in the biosynthesis of cinnamamide in cultures of *Streptomyces verticillatus* using L-phenylalanine- $carboxyl$ - $^{14}C$ <sup>40</sup>. Cinamamide is a nonlethal chemical repellent<sup>41</sup>. Recently, cinamamide derivatives have been synthesized. They show potential activities in both peripheral and central nervous systems, including antiepileptic, antidepressant, neuroprotective, analgesic, anti-inflammatory, muscle-relaxant and sedative/hypnotic properties<sup>42</sup>. Our results indicate that *Streptomyces* sp. VN1 is a promising producer of cinamamide compounds for drug discovery.

Lobophorin A and related analogs are members of the spirotreronate family. Their anti-inflammatory and antibacterial properties have been reported<sup>16,43</sup>. Genome sequence analysis results revealed that the putative

lobophorin biosynthetic gene cluster in *Streptomyces* sp. VN1 showed very high similarity with *Streptomyces* sp. FXJ7.023 and *Streptomyces* sp. SCSIO 01127 (Fig. 3; Supplementary Table S5A). Four different analogs of lobophorin (compound 5, 6, 7 and 8) were identified to be produced by *Streptomyces* sp. VN1 through mass analysis without any genetic modification. The diversity of lobophorin analogs has been previously explored. There has been a report of structural analysis of lobophorin revealing a polycyclic tetronolide core that is further modified by glycosylation and hydroxylation<sup>35</sup>. Through deletion of methyltransferase in *Streptomyces* sp. SCSIO 01127 the authors obtained two new analogs of lobophorin<sup>39</sup>. In addition, deletion of the glycosylation step in the biosynthesis pathway has achieved novel analogs of lobophorins<sup>36</sup>. The catalytic activity of core enzymes for lobophorin A biosynthesis and post-modification enzymes in *Streptomyces* sp. VN1 might be different from those in *Streptomyces* sp. SCSIO 01127. This might lead to different lobophorin A analogs. In addition, different post-modification steps might convert intermediates to different analogs. Post-modification tailoring of lobophorin A might also follow a different succession leading to accumulation of novel analogs of lobophorin A in *Streptomyces* sp. VN1. The diversity of lobophorin analogs identified in *Streptomyces* sp. VN1 might be the result of flexibility of post-modification enzymes such as glycosyltransferase, methyltransferase and hydroxylase which might have different substrate preferences among intermediates of lobophorin. The presence of an N-acetyltransferase gene in cluster 1 of *Streptomyces* sp. VN1 might have played role in adding to the diversity of lobophorin analogs. The N-acetyltransferase gene might be responsible for transfer of an acetyl group to lobophorin A to form other derivatives such as compound 7 and compound 8 (Supplementary Table S5B). A similar analog of lobophorin has been reported in a previous study where lobophorin G is detected as acetylated lobophorin A<sup>13</sup>. The diversity of lobophorin analogs identified from *Streptomyces* sp. VN1 highlights the uniqueness of lobophorin biosynthetic pathway compared to other lobophorin producer strains. Thus, *Streptomyces* sp. VN1 can be developed into a prolific producer of lobophorin and its analogs can be further diversified through genetic engineering. Biosynthetic modifications such as glycosylation<sup>44</sup>, hydroxylation<sup>45</sup> and methylation<sup>46</sup> have been successfully applied for the biosynthesis of non-natural products. These modifications could be applied for *Streptomyces* sp. VN1 to produce novel lobophorin analogs with potential medical applications.

Cyclo-L-proline-L-tyrosine was first reported to be produced by *Streptomyces* sp. strain 22-4<sup>17</sup>. It had activities against three economically important plant pathogens, *Xanthomonas axonopodis* pv. citri, *Ralstonia solanacearum* and *Clavibacter michiganensis*.

Although the structure of compound 4 has been reported previously as a synthetic by-product (<http://www.ambinter.com/search>), unfortunately, no reference NMR data of compound 4 was available. To the best of our knowledge, this is the first report of NMR spectrum and biological activity of furan-type compound 4 produced by a marine *Streptomyces*. There has been a report of moderate production of a furan-type compound by microorganism and intermediate furan-type compound in *Streptomyces* that may function as a starting unit for a secondary metabolite, such as A-factor, butyrolactone, and other analogs of methylenomycin furans (MMFs)<sup>47</sup> from *Streptomyces celicolor*.

In 2008, Jidong *et al.* reported a furan compound HS071 from *Streptomyces* sp. HS-HY-071 that showed cytotoxic activity against a cancer cell line<sup>48</sup>. However, there is no report about the biosynthesis pathway of this compound. There are some differences between structures of compound 4 and furan-compound HS071<sup>48</sup>, such as a 4-OH position in the HS071 compound where compound 4 contains a carboxyl group (Supplementary Fig. S16) probably because these two compounds are produced from two different biosynthetic pathways or because compound 4 or compound HS071 is further modified, although they come from the same biosynthesis pathway. The biosynthesis pathway of compound 4 was not explored in this study. Further studies by feeding <sup>13</sup>C in combination with genome sequence data are needed to track down the exact mechanism of compound 4 biosynthesis.

We examined the biological activity of compound 4 against five different types of cancer cell lines. Compound 4 showed growth inhibitory activity most effectively against an AGS cancer cell line with an IC<sub>50</sub> value of 40.5 μM. Although concentration ranges of compound 4 for growth inhibition of cancer cells were much higher than those of known anticancer drugs such as taxol<sup>49</sup> with effective concentrations below 10 μM, compound 4 inhibited the growth of cancer cell lines more effectively as compared to normal cell lines at tested concentrations (Fig. 7). Interestingly, compound 4 also showed anti-migration and anti-invasion activities against AGS cancer cell line after 24 hours of exposure (Figs. 8 and 9). These results confirm compound 4 possesses anticancer activity. However, as shown in cell growth assay the cell viability was high (89%) even after treatment with compound 4 at 25 μM treatment for 72 hours. Therefore, inhibitory effects of compound 4 on the migration and invasion of AGS cancer cells were not merely due to the cytotoxicity of compound 4. Such anti-invasive properties of compound 4 against cancer cell lines showcase the relevance of exploring secondary metabolites biosynthetic pathways for discovering clinically relevant compounds.

Results of *in silico* genome analysis showed that *Streptomyces* sp. VN1 encodes the putative biosynthetic gene clusters of diverse interesting secondary metabolites, such as carotenoid, frulimicin, xiamycin, divergolide, and informatipeptin. However, biosynthetic pathways for these metabolites in *Streptomyces* sp. VN1 were inactive under normal culture conditions. This study displays the biosynthetic potential of *Streptomyces* sp. VN1 through isolation and characterization of four compounds. Due to its ability to grow fast and produce various secondary metabolites, it is suitable for further metabolic exploration to produce useful metabolites. We present *Streptomyces* sp. VN1 as a good candidate for strain optimization to produce therapeutically and industrially relevant compounds. Our results reinforce the need of further exploring marine *Streptomyces* as a rich source of novel metabolites relevant for biotechnological applications.

Received: 14 March 2019; Accepted: 14 January 2020;

Published online: 04 February 2020

## References

- Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug. Discov.* **14**, 111–129, <https://doi.org/10.1038/nrd4510> (2015).
- Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335, <https://doi.org/10.1021/np200906s> (2012).
- Chaudhary, A. K., Dhakal, D. & Sohng, J. K. An Insight into the “-Omics”; Based Engineering of *Streptomyces* for Secondary Metabolite Overproduction. *Biomed. Res. Int.* **2013**, 968518, <https://doi.org/10.1155/2013/968518> (2013).
- Clardy, J., Fischbach, M. A. & Walsh, C. T. New antibiotics from bacterial natural products. *Nat. Biotechnol.* **24**, 1541–1550, <https://doi.org/10.1038/nbt1266> (2006).
- Antoraz, S., Santamaria, R. I., Diaz, M., Sanz, D. & Rodriguez, H. Toward a new focus in antibiotic and drug discovery from the *Streptomyces arsenal*. *Front. Microbiol.* **6**, 461, <https://doi.org/10.3389/fmicb.2015.00461> (2015).
- Onaka, H. Novel antibiotic screening methods to awaken silent or cryptic secondary metabolic pathways in actinomycetes. *J. Antibiot.* **70**, 865–870, <https://doi.org/10.1038/ja.2017.51> (2017).
- Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523, <https://doi.org/10.1038/nrmicro3496> (2015).
- Waksman, S. A. & Henrici, A. T. The nomenclature and classification of Actinomycetes. *J. Bacteriol.* **46**, 337–341 (1943).
- Bérdy, J. Bioactive Microbial Metabolites. *J. Antibiot.* **58**, 1–26, <https://doi.org/10.1038/ja.2005.1> (2005).
- Tanaka, Y. *et al.* Activation and products of the cryptic secondary metabolite biosynthetic gene clusters by rifampin resistance (rpoB) mutations in actinomycetes. *J. Bacteriol.* **195**, 2959–2970, <https://doi.org/10.1128/JB.00147-13> (2013).
- Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 2894, <https://doi.org/10.1038/ncomms3894> (2013).
- Devine, R., Hutchings, M. I. & Holmes, N. A. Future directions for the discovery of antibiotics from actinomycete bacteria. *Emerg. Top. Life Sci.* **1**, 1–12, <https://doi.org/10.1042/ETLS20160014> (2017).
- Chen, C. *et al.* Three antimycobacterial metabolites identified from a marine-derived *Streptomyces* sp. MS100061. *Appl. Microbiol. Biotechnol.* **97**, 3885–3892, <https://doi.org/10.1007/s00253-012-4681-0> (2013).
- Braña, A. F. *et al.* Lobophorin K, a New Natural Product with Cytotoxic Activity Produced by *Streptomyces* sp. M-207 Associated with the Deep-Sea Coral *Lophelia pertusa*. *Mar. Drugs* **15**, 144, <https://doi.org/10.3390/md15050144> (2017).
- Huang, X. L. *et al.* Streptomycinolide, an Indole Alkaloid from a Marine *Streptomyces* sp. DA22 Associated with South China Sea Sponge *Craniella australiensis*. *Helv. Chim. Act.* **94**, 1838–1842, <https://doi.org/10.1002/hlca.201100104> (2011).
- Niu, S. *et al.* Lobophorins E and F, new spiro-tetronate antibiotics from a South China Sea-derived *Streptomyces* sp. SCSIO 01127. *J. Antibiot.* **64**, 711–716, <https://doi.org/10.1038/ja.2011.78> (2011).
- Wattana-amorn, P., Charoenwongsa, W., Williams, C., Crump, M. P. & Apichaisaitaienchote, B. Antibacterial activity of cyclo (L-Pro-L-Tyr) and cyclo (D-Pro-L-Tyr) from *Streptomyces* sp. strain 22-4 against phytopathogenic bacteria. *Nat. Prod. Res.* **6419**, 1980–1983, <https://doi.org/10.1080/14786419.2015.1095747> (2015).
- Zhu, N., Zhao, P. & Shen, Y. Selective Isolation and Ansamycin-targeted Screenings of Commensal Actinomycetes from the ‘Maytansinoids-Producing’ Arboreal *Trewia nudiflora*. *Curr. Microbiol.* **58**, 87–94, <https://doi.org/10.1007/s00284-008-9284-8> (2009).
- Otoguro, M. *et al.* *Streptomyces baliensis* sp. nov., isolated from balinese soil. *Int. J. Syst. Evol. Microbiol.* **59**, 2158–2161 (2009).
- Lechevalier, M. P., De Bievre, C. & Lechevalier, H. Chemotaxonomy of aerobic Actinomycetes: Phospholipid composition. *Biochem. Syst. Ecol.* **5**, 249–260, [https://doi.org/10.1016/0305-1978\(77\)90021-7](https://doi.org/10.1016/0305-1978(77)90021-7) (1977).
- Hasegawa, T., Takizawa, M. & Tanida, S. A rapid analysis for chemical grouping of aerobic actinomycetes. *J. Gen. Appl. Microbiol.* **29**, 319–322, <https://doi.org/10.2323/jgam.29.319> (1983).
- Nybo, S. E., Shepherd, M. D., Bosserman, M. A. & Rohr, J. Genetic manipulation of *Streptomyces* species. *Curr. Protoc. Microbiol.* **19**, 1–26, <https://doi.org/10.1002/9780471729259.mc10e03s19> (2010).
- Kim, O. S. *et al.* Introducing EzTaxon-e: A prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* **62**, 716–721, [10.1099/2fijms.0.001755](https://doi.org/10.1099/2fijms.0.001755) (2012).
- Alanjary, M., Steinke, K. & Ziemert, N. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.* **47**, 276–282, <https://doi.org/10.1093/nar/gkz282> (2019).
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Göker, M. O. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinforma.* **14**, 1471–2105, <https://doi.org/10.1186/1471-2105-14-60> (2013).
- Schattner, P., Brooks, A. N. & Lowe, T. M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **33**, 686–689, <https://doi.org/10.1093/nar/gki366> (2005).
- Lagesen, K. *et al.* RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108, <https://doi.org/10.1093/nar/gkm160> (2007).
- Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000).
- Huerta-cepas, J. *et al.* eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, 286–293, <https://doi.org/10.1093/nar/gkv1248> (2016).
- Blin, K. *et al.* antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**, 36–41, <https://doi.org/10.1093/nar/gkx319> (2017).
- Miles, A. J. & Wallace, B. A. CDToolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. *Protein Sci.* **27**, 1717–1722, <https://doi.org/10.1002/pro.3474> (2018).
- Lee, S., Kwon, M. C., Jang, J. P., Sohng, J. K. & Jung, H. J. The ginsenoside metabolite compound K inhibits growth, migration and stemness of glioblastoma cells. *Int. J. Oncol.* **51**, 414–424, <https://doi.org/10.3892/ijo.2017.4054> (2017).
- Lu, Y. *et al.* Isolation and characterization of fatty acid methyl ester (FAME)-producing *Streptomyces* sp. S161 from sheep (*Ovis aries*) faeces. *Letts. Appl. Microbiol.* **57**, 200–205, <https://doi.org/10.1111/lam.12096> (2013).
- Chen, R. *et al.* Genome Mining and Activation of a Silent PKS/NRPS Gene Cluster Direct the Production of Totopotensamides. *Org. Lett.* **19**, 5697–5700, <https://doi.org/10.1021/acs.orglett.7b02878> (2017).
- Yue, C. *et al.* Cloning and identification of the lobophorin biosynthetic gene cluster from marine Cloning and identification of the lobophorin biosynthetic gene cluster from marine *Streptomyces olivaceus* strain FXJ7. 023. *Pak. J. Pharm. Sci.* **29**, 287–293 (2016).
- Li, S. *et al.* Dissecting glycosylation steps in lobophorin biosynthesis implies an iterative glycosyltransferase. *Org. Lett.* **15**, 1374–1377, <https://doi.org/10.1021/ol400342e> (2013).
- Hanai, K., Kuwae, A., Takai, T. & Senda, H. A comparative vibrational and NMR study of cis-cinnamic acid polymorphs and trans-cinnamic acid. *Spectrochim. Acta Part. A Mol. Biomol. Spectrosc.* **57**, 513–519, [https://doi.org/10.1016/S1386-1425\(00\)00401-7](https://doi.org/10.1016/S1386-1425(00)00401-7) (2001).
- Pramanik, B. N. *et al.* Special techniques of fast atom bombardment mass spectrometry for the study of oligosaccharide containing macrotetronolide antibiotic, kijanimicin. *J. Antibiot.* **37**, 818–821 (1984).
- Xiao, J. *et al.* Characterization of the sugar-O-methyltransferase LobS1 in lobophorin biosynthesis. *Appl. Microbiol. Biotechnol.* **97**, 9043–9053, <https://doi.org/10.1007/s00253-013-5083-7> (2013).
- Bezanson, G. S., Desaty, D., Emes, A. V. & Vining, L. C. Biosynthesis of cinnamamide and detection of phenylalanine ammonia-lyase in *Streptomyces verticillatus*. *Can. J. Microbiol.* **16**, 147–151 (1970).

41. Gill, E. L., Watkins, R. W., Gurney, J. E., Bishop, J. D. & Feare, C. J. Cinnamamide: A Nonlethal Chemical Repellent for Birds and Mammals. *National Wildlife Research Center Repellents Conference* **19** (1995).
42. Gunia-Krzyżak, A., Pańczyk, K., Waszkielewicz, A. M. & Marona, H. Cinnamide Derivatives for Central and Peripheral Nervous System Disorders-A Review of Structure-Activity Relationships. *Chem. Med. Chem* **10**, 1302–1325, <https://doi.org/10.1002/cmdc.201500153> (2015).
43. Jiang, Z. D., Jensen, P. R. & Fenical, W. H. Lobophorins A and B, new antiinflammatory macrolides produced by a tropical marine bacterium. *Bioorg Med. Chem. Lett.* **9**, 2003–2006 (2006).
44. Le, T. T., Pandey, R. P., Gurung, R. B., Dhakal, D. & Sohng, J. K. Efficient enzymatic systems for synthesis of novel  $\alpha$ -mangostin glycosides exhibiting antibacterial activity against Gram-positive bacteria. *Appl. Microbiol. Biotechnol.* **98**, 8527–8538, <https://doi.org/10.1007/s00253-014-5947-5> (2014).
45. Chu, L. L. *et al.* Hydroxylation of diverse flavonoids by CYP450 BM3 variants: biosynthesis of eriodictyol from naringenin in whole cells and its biological activities. *Microb. Cell Fact.* **15**, 135, <https://doi.org/10.1186/s12934-016-0533-4> (2016).
46. Pandey, R. P., Parajuli, P., Koffas, M. A. G. & Sohng, J. K. Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnol. Adv.* **34**, 634–662, <https://doi.org/10.1016/j.biotechadv.2016.02.012> (2016).
47. Niu, G., Chater, K. F., Tian, Y., Zhang, J. & Tan, H. Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol. Rev.* **40**, 554–573, <https://doi.org/10.1093/femsre/fuw012> (2018).
48. Wang, J. *et al.* HS071, A New Furan-type Cytotoxic Metabolite from *Streptomyces* sp. HS-HY-071. *J. Antibiot.* **61**, 623–626, <https://doi.org/10.1038/ja.2008.82> (2008).
49. Runowicz, C. D., Wiernik, P. H., Einzig, A. I., Goldberg, G. L. & Horwitz, S. B. Taxol in ovarian cancer. *Cancer* **71**, 1591–1596 (1993).

## Acknowledgements

This work was supported by grants (2016R1D1A1B03932956 and 2017R1A2A2A05000939) of the National Research Foundation of Korea (NRF) funded by the Korean government (MEST) and the Ministry of Science and ICT (No. 2019R1A2C1009033).

## Author contributions

H.T.N., A.R.P. and J.K.S. conceived and designed the experiments; H.T.N., A.R.P., C.T.N. and V.T.T.P. performed the experiments; H.N.L. and H.J.J. performed biological activity experiment; T.Y. analyzed the NMR data; H.T.N., D.D., T.S.K. and J.K.S. wrote and edit the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-58623-1>.

**Correspondence** and requests for materials should be addressed to J.K.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020