



Capability of a large bacterial artificial chromosome clone harboring multiple biosynthetic gene clusters for the production of diverse compounds

Kei Kudo¹ · Takehiro Nishimura^{2,6} · Miho Izumikawa³ · Ikuko Kozono³ · Junko Hashimoto³ · Manabu Fujie⁴ · Hikaru Suenaga¹ · Haruo Ikeda^{5,7} · Nori Satoh⁴ · Kazuo Shin-ya^{1,2}

Received: 30 October 2023 / Revised: 2 February 2024 / Accepted: 12 February 2024 / Published online: 4 March 2024

© The Author(s), under exclusive licence to the Japan Antibiotics Research Association 2024

Abstract

The biosynthetic gene clusters (BGCs) for the macrocyclic lactone-based polyketide compounds are extremely large-sized because the polyketide synthases that generate the polyketide chains of the basic backbone are of very high molecular weight. In developing a heterologous expression system for the large BGCs amenable to the production of such natural products, we selected concanamycin as an appropriate target. We obtained a bacterial artificial chromosome (BAC) clone with a 211-kb insert harboring the entire BGC responsible for the biosynthesis of concanamycin. Heterologous expression of this clone in a host strain, *Streptomyces avermitilis* SUKA32, permitted the production of concanamycin, as well as that of two additional aromatic polyketides. Structural elucidation identified these additional products as *ent*-gephyromycin and a novel compound that was designated JBIR-157. We describe herein sequencing and expression studies performed on these BGCs, demonstrating the utility of large BAC clones for the heterologous expression of cryptic or near-silent loci.

Deceased: Miho Izumikawa.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41429-024-00711-9>.

✉ Kazuo Shin-ya
k-shinya@aist.go.jp

¹ Department of Life Science and Biotechnology, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

² Technology Research Association for Next Generation Natural Products Chemistry, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

³ Japan Biological Informatics Consortium (JBIC), 2-4-32 Aomi, Koto-ku, Tokyo 135-0064, Japan

⁴ Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Kunigami-gun, Okinawa 904-0495, Japan

⁵ Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Minami-ku, Sagami-hara, Kanagawa 252-0373, Japan

⁶ Present address: Faculty of Pharmacy, Keio University, 1-5-30 Shibakoen, Minato-ku, Tokyo 105-8512, Japan

⁷ Present address: Technology Research Association for Next Generation Natural Products Chemistry, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

Introduction

Screening for bioactive microbial secondary metabolites has been performed for more than 70 years. Although the discovery of bioactive natural products has resulted in great benefits to humankind, the rate of discovery of skeletally-novel compounds from microorganisms has decreased significantly over time [1, 2]. To address this situation, researchers investigating natural products have made great effort to exploit the synthetic potential of microorganisms using a variety of media [3] and co-culture systems [4], and by inducing mutations in candidate microbes [5]. In contrast to these more conventional methods, recent advances in sequencing technology have revealed the presence in microorganisms of many cryptic gene clusters thought to be capable of the biosynthesis of secondary metabolites [6, 7]. This observation reminds us that we have not fully applied the abilities of microorganisms to produce various bioactive compounds. We hypothesized that heterologous expression of these cryptic biosynthetic genes will serve as a process for inducing the production of novel compounds because it has already been reported that heterologous expression can initiate even a cryptic biosynthetic gene cluster and confirm its metabolite [8]. While many studies on heterologous expression have been performed, there exist (to our knowledge) only

a limited number of reports of the heterologous production of middle-molecular-weight compounds such as macrolides, which are biosynthesized by type I polyketide synthases (PKSs). This fact presumably reflects the extremely large sizes of the corresponding genes and gene clusters, which would have led to technological challenges in cloning, transformation, and expression. To overcome these technical difficulties, we have been employing bacterial artificial chromosomes (BACs), which permit the cloning of DNA fragments up to 300 kb in size [9]. Using this technology, we have succeeded in the heterologous production of mediomycin and neomediomycin [10], JBIR-156 [11], quinolidomicin [9], and desertomycin [12], molecules generated by the products of biosynthetic gene clusters (BGCs) that are 161, 183, 137, 213, and 127 kb in length, respectively.

Concanamycin A (folimycin) (Fig. 1) [13, 14], originally isolated as an antifungal metabolite of *Streptomyces neyagawaensis* nov. sp. [15], is a potent inhibitor of the vacuolar-type proton-ATPase (V-ATPase) [16]. The concanamycin BGC, an approximately 100-kb locus consisting of six genes encoding type I PKSs, was identified by assembling the DNA sequences of three cosmid clones generated from the genome of *S. neyagawaensis* ATCC 27449 [17]. However, the heterologous expression of this BGC has not been achieved. Therefore, we targeted concanamycin(s) for the further development of our heterologous expression system for the biosynthesis of polyketide compounds derived from modular polyketide synthases. Heterologous expression of the BAC clone pKU503ccn, which harbors the concanamycin BGC, resulted in the production of concanamycins at high yield, compared to the wild-type strain, *S. neyagawaensis* IFO 13477. Interestingly, we found that a host strain carrying this

BAC also produced two other aromatic polyketides (Fig. 1) when cultured in a production medium distinct from that used to obtain concanamycin. Therefore, we also report herein a newly identified BGC for the production of *ent*-gephyromycin and a novel compound that we designate JBIR-157.

Materials and methods

Bacterial strains and growth conditions

Streptomyces neyagawaensis IFO 13477 was obtained from Institute for Fermentation, Osaka, Japan (currently, the microorganism can be distributed as NBRC 13477 from National Institute for Technology and Evaluation Biological Resource Center, Chiba, Japan). The growth of *Escherichia coli* strains DH5 α , NEB10 β (New England Biolabs, Ipswich, MA), and GM2929 *hsdS::Tn10* used Luria broth (LB), which contained (per L of deionized water) 10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl, and was adjusted to pH 7.5 following formulation. The corresponding solid medium was prepared by supplementing LB with agar at 15 g L⁻¹. For the production of secondary metabolites, *S. neyagawaensis* IFO 13477, *S. avermitilis* SUKA32 [11], or SUKA54 [18] was inoculated into a 50-mL test tube containing 15 mL of seed medium, containing (per L of deionized water) 5 g of glucose, 15 g of soya flour, and 5 g of yeast extract; the cells were cultured with reciprocal shaking at 320 rpm and 27 °C for 2 days. A 375- μ L aliquot of the resulting vegetative culture was used to inoculate a 125-mL flask containing 15 mL of either of two production media. Medium 1, used for the production of concanamycins, contained (per L of deionized water) 40 g of β -cyclodextrin, 20 g of Pharmamedia, 5 g of glycerol, 5 mg of ZnSO₄·7H₂O, 5 mg of CuSO₄·5H₂O, and 5 mg of MnCl₂·4H₂O; the pH was adjusted to 7.0 following formulation. Medium 2 (synthetic medium [19]), used for the production of *ent*-gephyromycin and JBIR-157. The fermentation cultures were grown on a rotary shaker at 180 rpm and 27 °C for 7 days. For large-scale preparations of the products, cells were grown under the same conditions in baffled 500-mL flasks, each containing 100 mL of the respective production medium.

Isolation of a concanamycin BGC-containing BAC

Genome sequencing of the *S. neyagawaensis* IFO 13477 strain was performed using PacBio RS II (Pacific Biosciences, Menlo Park, CA) and Roche 454 GS FLX Titanium chemistry (Roche, Basel, Switzerland). The sequence data were assembled using HGAP2 (Pacific Biosciences) and Newbler software package (Roche). A BAC library of *S. neyagawaensis* IFO 13477 was constructed with

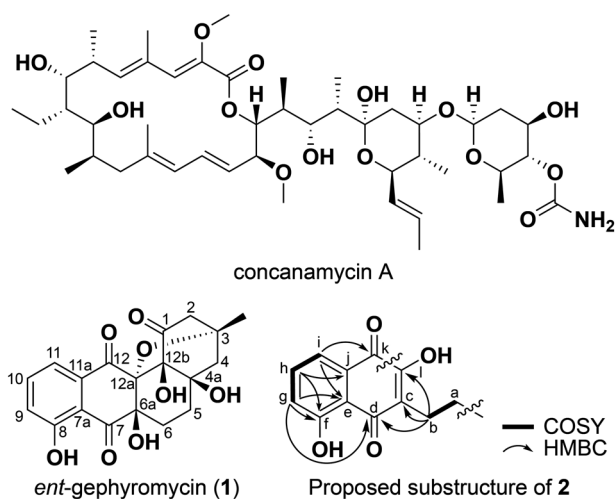


Fig. 1 Compounds identified in this study. The structure of concanamycin A (presented here as a representative concanamycin) and *ent*-gephyromycin (**1**) are shown. The proposed substructure of JBIR-157 (**2**) is shown with key correlations obtained via COSY and HMBC experiments

pKU503 [20] according to a previously reported protocol [8]. BAC clones obtained were transferred into 384-well plates containing Plusgrow II medium (Nacalai Tesque, Inc., Kyoto, Japan) supplemented with 100 µg mL⁻¹ ampicillin and 20 v/v% glycerol and their plates were stored at -80 °C. A clone carrying the concanamycin and *ent*-gephyromycin/JBIR-157 BGCs was identified by screening via PCR amplification using two pairs of primers, as follows. For the concanamycin BGC, the primer pairs con_1F (5'- AGACG CTGTA CTCCCGACCTCA -3') + con_1R (5'- GTTCT GATCGGCCTGCTCCTTCTG -3') and con_2F (5'- TCT TCTGGATGTGGCCGTTCTTCA -3') + con_2R (5'- CAG CCGGACTTCACCATCTCCTTC -3') were used to amplify the upstream and downstream regions (respectively) of that cluster; for the *ent*-gephyromycin/JBIR-157, the primer pairs gpn_1F (5'- GTACCGAGTGCCGCAGATACTCCT -3') + gpn_1R (5'- GAATCGTTACCCCAATGGTGTC -3') and gpn_2F (5'- CATCGAAGAAGATGTGTGCGGAAGA -3') + gpn_2R (5'- GACGCTTTCGAGAGATGTGTTGAG -3') were used to amplify the upstream and downstream regions (respectively) of that cluster. The identities and spans of the clones that were positive for the respective BGCs were confirmed by end-sequencing and aligning with the genome sequence obtained as described above. The inserted sequence of pKU503ccn (which includes both the concanamycin and *ent*-gephyromycin/JBIR-157 BGCs) has been deposited in the DDBJ under Accession Number LC780117.

Introduction of BACs into *S. avermitilis* SUKA strains

First, each clone was introduced into *S. lividans* TK24 $\Delta attB_{\phi K38-1}::aadA \Delta attB_{\phi BT1} \Delta attB_{\phi C31} \Delta attB_{TG1}$ harboring the SAP1 vector, a linear plasmid containing a synthetic sequence corresponding to the lysogenic bacteriophage attachment sites *attB* _{$\phi K38-1$} , *attB*_{R4}, *attB* _{$\phi BT1$} , *attB* _{$\phi C31$} , and *attB*_{TG1} of *S. avermitilis* MA-4680 [21]. After the introduction of each BAC clone, the clone was integrated into the *attB* _{$\phi C31$} site of the SAP1 vector by site-specific recombination (but not into the chromosome because it lacks *attB* sites). The SAP1 vector carrying BAC clone in *S. lividans* was transferred to *S. avermitilis* SUKA SAP1⁻ strains by simple conjugation as described previously [21]. The resulting exoconjugants were selected based on their antibiotic-resistance phenotype; the presence of a linear plasmid of the expected size was confirmed by contour-clamped homogeneous electric field (CHEF) electrophoresis [22].

Liquid chromatography/mass spectrometry (LC/MS) analysis of culture extracts

The metabolites in exoconjugants were extracted from the cultured broth with an equal volume of *n*-butanol. The organic phase was collected and evaporated to dryness under the reduced

pressure. The resulting residue was dissolved in an appropriate volume of dimethyl sulfoxide (DMSO) and analyzed by LC/MS. Specifically, analytical ultra-high performance liquid chromatography (UHPLC) and high-resolution electrospray ionization MS (HR-ESI-MS; positive mode) were performed using an ACQUITY UPLC System (Waters, Framingham, MA) in conjunction with an ethylene-bridged hybrid octadecyl-silica (BEH ODS) column (2.1 i.d. × 100 mm, 55 °C; Waters), an ACQUITY UPLC photodiode array eλ detector (Waters), and a Xevo G2 time-of-flight (TOF) system (Waters). Mobile Phase A was water + 0.1% formic acid; Mobile Phase B was acetonitrile + 0.1% formic acid. The elution program consisted of 5 – 100% B over 5 min, followed by 100% B for 1 min; the flow rate was 0.8 mL min⁻¹. A standard curve for quantification of concanamycin A was prepared by triplicate injection of 2 µL of concanamycin A samples at concentrations of 50, 25, 12.5, 6.25, 3.13, and 1.56 mg L⁻¹.

Nuclear magnetic resonance (NMR) spectroscopy

NMR spectra were recorded on an NMR 600 NB CL spectrometer (Varian, Palo Alto, CA). The coupling constants (*J*) are given in hertz units. Measurements were carried out at room temperature. Chemical shifts (δ) are reported in ppm with the residual solvent signals as the internal standards (methanol-*d*₄ at δ_H 7.26 and δ_C 77.0 ppm). The data are reported as (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, or unresolved; coupling constant(s); integration). ¹³C NMR spectra were recorded with complete ¹H decoupling.

Isolation of Compounds 1 and 2 from the fermentation culture of *S. avermitilis* SUKA32/ pKU503ccn

One liter of a fermentation culture of *S. avermitilis* SUKA32/ pKU503ccn was centrifuged to clear the cell suspension; the resulting supernatant was extracted with ethyl acetate (1 L × 3). The resulting ethyl acetate fraction was concentrated *in vacuo*, to afford 111.0 mg of crude extract. The crude extract was subjected to silica gel middle pressure-liquid chromatography (SNAP Ultra 25 g; Biotage, Uppsala, Sweden; 1 column volume (CV) = 45 mL); elution was performed using a gradient of *n*-hexane–ethyl acetate (0% ethyl acetate for 4 CV followed by a linear gradient of 0–25% ethyl acetate over 6 CV; 75 mL min⁻¹ flow) followed by a stepwise gradient of chloroform–methanol (0% methanol for 2 CV followed by a stepwise gradient of 1, 3, 5, 10, 50, and 90% methanol each for 3 CV; 75 mL min⁻¹ flow). The 5% methanol fraction (17.9 mg) was further purified by preparative reverse-phase HPLC using a CAPCELL PAK MG-II C18 column (5.0 µm, 20 i.d. × 150 mm; Shiseido, Tokyo, Japan), and eluted with a 10 mL min⁻¹ flow of 54% aqueous

methanol supplemented with 0.1% formic acid, yielding 2.1 mg of **1** (retention time 10.5 min). The 50% methanol fraction (10.5 mg) was further purified by preparative reverse-phase HPLC using a CAPCELL PAK MG-II C18 column (5.0 μm , 20 i.d. \times 150 mm; Shiseido), and eluted with 54% aqueous methanol supplemented with 0.1% formic acid, yielding 2.8 mg of **2** (17.5 min).

ent-gephyromycin (**1**):

$[\alpha]_{\text{D}}^{24} + 67.3$ (c 0.2, MeOH);

NMR: see Supplementary Table 1 and Supplementary Data 1–5;

HR-ESI-MS m/z ($[M + Na]^+$): Calcd. for $\text{C}_{19}\text{H}_{18}\text{O}_8\text{Na}$: 397.0899, Found: 397.0906

JBIR-157 (**2**):

UV λ_{max} (MeOH)nm (ϵ): 236 (15,890), 270 (14,710), 388 (2,830), 489 (2,210);

NMR: see Supplementary Table 2 and Supplementary Data 6–10;

HR-ESI-MS m/z ($[M + Na]^+$): Calcd. for $\text{C}_{19}\text{H}_{18}\text{O}_8\text{Na}$: 397.0899, Found: 397.0902

Bioinformatics

Protein database (nr) was obtained from the National Center for Biotechnology Information (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/>) and similarity searches were conducted using Basic Local Alignment Search Tool (BLAST) binary program. The binary program was made by the compilation of a source package (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.14.0/>). The Pfam database version 36.0 (September 2023, 20,795 entries, 659 clans) from the European Bioinformatics Institute (<http://ftp.ebi.ac.uk/pub/databases/Pfam/releases/Pfam36.0/>) and pfam search binary program was made by the compilation of a source code (<http://hmmer.org/>). The alignment of sequences was analyzed using Clustal Omega version 1.2.4 (<http://www.clustal.org/omega/>) [23].

Results and discussion

Isolation of a BAC clone containing the concanamycin BGC

As a first step, we sequenced the genome of *S. neyagawaensis* IFO 13477 using a PacBio RSII long-read sequencing system and a Roche 454 GS FLX short-read system. The final assembly consisted of three contigs (including a circular plasmid molecule) with a total length of 10.1 Mb. We used the results of this assembly for all sequence analyses in the present study.

Since the BGC size of concanamycin is around 100 kb, we first prepared a BAC library with 150-kb inserts, a size that was expected to encompass the complete BGC.

Screening of approximately 1500 colonies by PCR failed to yield a clone containing the entire concanamycin BGC. Therefore, we constructed a second BAC library with inserts of over 180 kb, permitting the PCR screening of an additional 1500 colonies. Among these BAC clones, we obtained two positive clones that harbored identical 211-kb fragments, which were located within the largest contig by end-sequencing the inserts of BACs. One such clone was designated pKU503ccn (Accession No. LC780117). A library of 1500 clones with fragments averaging 150 kbp in size from 10 Mbp *Streptomyces* genome could theoretically contain at least one fragment of the desired 100 kbp. In practice, however, a BAC library with much larger inserts (>180 kb) was required, and only one pair of identical clones spanning the entire concanamycin BGC was detected. The biosynthetic gene cluster for cephamycin C of *S. clavuligerus* is approximately 35 kbp, but no entire gene cluster could be obtained from partially digested fragments with restriction enzyme of *S. clavuligerus* chromosomal DNA using the cosmid vector. However, full-length clones could be obtained using fragments cut by shearing forces and cosmid vector [8]. Thus, because of the recognition sequence bias in partial digestion with restriction enzyme, it may be necessary to use larger fragments for genomic library preparation.

Analysis of secondary metabolites produced by the heterologous expression of pKU503ccn

The pKU503ccn plasmid was introduced successfully into *S. avermitilis* SUKA32, a “clean” host (lacking production of major endogenous products) that has been used previously for the heterologous production of secondary metabolites of interest [8]. Exoconjugants were cultured in several distinct media, and the resulting metabolite profiles were analyzed by UHPLC/TOF-MS, confirming the production of concanamycin A along with several other minor analogues. These metabolites were identified by comparing the retention time, UV spectra, and HR-ESI-MS profiles to those of *S. neyagawaensis* IFO 13477 and the authentic sample of concanamycin A (Fig. 2). The production yield of total concanamycins was $24.3 \pm 1.6 \text{ mg L}^{-1}$, a value twice that of the parent strain *S. neyagawaensis* IFO 13477 ($12.3 \pm 1.0 \text{ mg L}^{-1}$) grown in the same medium. Thus, as previously reported, heterologous expression provides increased, stable production of secondary metabolites [9–12].

Interestingly, a chromatogram of *n*-butanol-extract of *S. avermitilis* SUKA32/pKU503ccn cultured in Medium 2 showed two sharp, novel (compared to Medium 1) peaks that exhibited UV spectra distinct from those of concanamycins (Fig. 3a–c). Since these peaks (designated Compounds **1** and **2**) were not observed in the culture

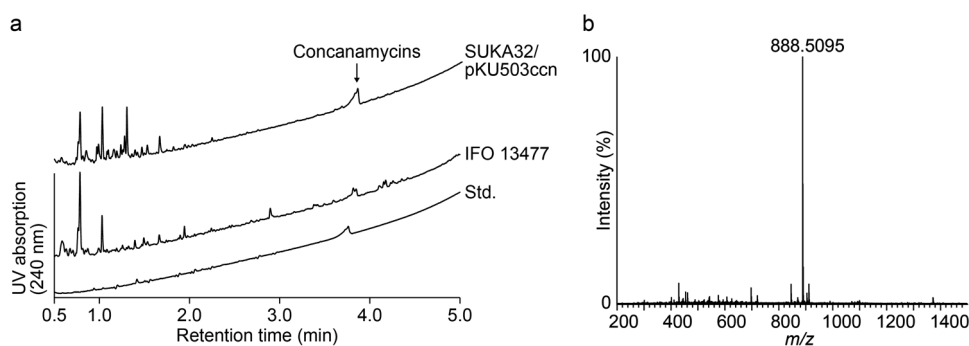


Fig. 2 Heterologous production of concanamycins. The culture extract of *S. avermitilis* SUKA32/pKU503ccn was analyzed by UHPLC/TOF-MS. **a** UV chromatograms (240 nm) of culture extracts. *S. avermitilis* SUKA32/pKU503ccn produced larger amounts of concanamycins than did *S. neyagawaensis* IFO 13477 (indicated by the trace labeled “IFO 13477”). Concanamycin A was used as the standard (indicated

by the trace labeled “Std.”). Each culture experiment was carried out in triplicate. Representative data are shown. **b** The HR-ESI-MS of concanamycin A produced by *S. avermitilis* SUKA32/pKU503ccn. Calculated m/z for $[M+Na]^+$: 888.5080; observed m/z : 888.5095; MS error +1.7 ppm

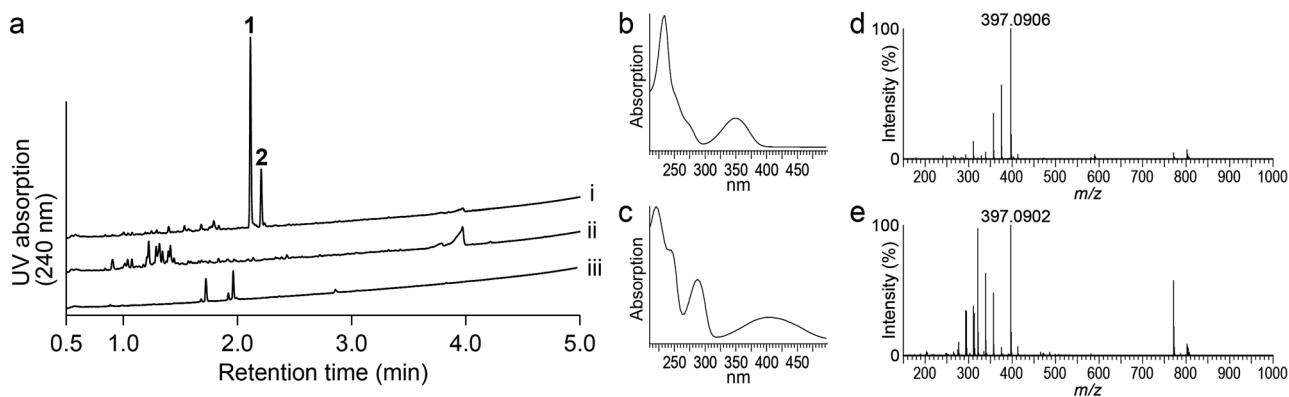


Fig. 3 Heterologous production of *ent*-gephyromycin (**1**) and JBIR-157 (**2**) by *S. avermitilis* SUKA32/pKU503ccn. The culture extracts of *S. avermitilis* SUKA32/pKU503ccn (following growth in two different media) were analyzed using UHPLC/TOF-MS. **a** UV chromatograms (240 nm) of culture extracts. *S. avermitilis* SUKA32/pKU503ccn produced **1** and **2** when grown in Medium 2 (indicated by trace “i”) but not when grown in Medium 1 (trace “ii”). *S. neyagawaensis* IFO

13477 did not produce detectable levels of **1** or **2** even when the strain was cultured in Medium 2 (trace “iii”). **b** The UV absorption spectrum of **1**. **c** The UV absorption spectrum of **2**. **d** The HR-ESI-MS of **1**. Calculated m/z for $[M + Na]^+$, 397.0899; observed m/z , 397.0906; MS error +1.8 ppm. **e** The HR-ESI-MS of **2**. Calculated m/z for $[M + Na]^+$, 397.0899; observed m/z , 397.0902; MS error +0.8 ppm

extract of *S. neyagawaensis* IFO 13477 under any of the tested culture conditions, it was expected that the cryptic gene cluster would be activated by the heterologous expression. We subsequently purified these two compounds using silica gel chromatography followed by reversed-phase chromatography. They were identified to be *ent*-gephyromycin (**1**) and a novel compound JBIR-157 (**2**), respectively. HR-ESI-MS analysis indicated that **1** and **2** had the same molecular formula, $C_{19}H_{18}O_8$ (Fig. 3d–e). Extensive analyses of 1D (1H and ^{13}C) and 2D (DQF-COSY, HSQC, and HMBC) NMR spectra revealed that **1** possessed the same planar skeleton as gephyromycin, an angucycline with a rare ether bridge [24] (Supplementary Table 1, Supplementary Data 1–5). Based on this result and the optical rotation observed with this molecule, we concluded that **1** is *ent*-gephyromycin [25]. On the other hand, for Compound **2**, the

UV and visible spectra, together with the analyses of a series of NMR spectra, suggested the existence of a 2,5-dihydroxy-1,4-naphthoquinone moiety (Fig. 1), but the 2D NMR spectra obtained in the present work was insufficient to permit determination of the remaining substructure (Supplementary Table 2, Supplementary Data 6–10). The detailed structural determination of **2** will be reported elsewhere.

Sequence analysis of the concanamycin BGC

The sequence of the concanamycin BGC determined in this study (*ccn* cluster) was compared to the previously reported sequence (*con* cluster; Genbank Accession No. DQ149987 [17]). The set of biosynthetic genes (six genes encoding PKs and 19 flanking open reading frames (ORFs) and their order and orientation were consistent between the two BGC

sequences, although several sequence differences (representing substitutions and insertion/deletions) were observed (Supplementary Fig. 1). All mismatch regions were re-confirmed by Sanger sequencing of pKU503ccn to exclude possible errors in genome sequencing or assembly. The most notable difference was a 168-bp insertion in the dehydratase-encoding region of *ccnF* in pKU503ccn compared to the previously reported sequence (*conF*). This inserted sequence corresponds to an unusual 56-amino-acid insertion between β_4 and β_5 of EryDH4 (PDB Accession No. 3EL6 [26]); to our knowledge, this insertion has not previously been discussed in the literature (Supplementary Fig. 2). Although the specific effect of this additional sequence remains undefined, the dehydratase (DH) domain of CcnF seems to function normally to give an olefin moiety at the C2 position of the growing chain. Notably, the sequence of the concanamycin BGC within the independently deposited genome sequence (NCBI Accession No. GCF_001418645.1) of *S. neyagawaensis* NRRL B-3092 (which should be identical to ATCC 27449 and IFO 13477) is consistent with that obtained in the present work. On the other hand, the corresponding PKS gene, *bfmA5*, in the BGC responsible for the biosynthesis of bafilomycin (a natural product that is structurally related to concanamycin) lacks such an insertion [27].

BGC for Compounds 1 and 2

The BGC for *ent*-gephyromycin (**1**) has not previously been elucidated. Since both **1** and **2** are aromatic polyketides, these compounds are presumed to be generated via a type II or III PKS system. Gephyromycin (enantiomer of *ent*-gephyromycin), produced by *Gephyromycinifex aptenodytis* isolated from gut of Antarctic emperor penguin, was

generated from tetraketide and geranyl pyrophosphate by type III PKS (stilbene synthase) and prenyltransferase [28]. But type III PKS system(s) and prenyltransferase were not present in the pKU503ccn and in the genome of heterologous host *S. avermitilis*. A BLAST search of sequences within the pKU503ccn insert (but outside the concanamycin BGC) revealed that these regions contained two gene clusters consisting of type II PKS system; each of these clusters included genes encoding two 3-oxoacyl-ACP synthase subunits (KS_α and KS_β) and an acyl carrier protein (ACP) (Fig. 4a). ORFs in one of these clusters showed strong sequence similarity to ORFs of conserved biosynthetic gene cluster for *Streptomyces* spore pigment (*whiE* cluster; Fig. 4b and Supplementary Table 3 [29]). Therefore, we inferred that the other set of type II PKS system-encoding ORFs and their flanking genes (hereafter referred to as the *egp* cluster) is responsible for the biosynthesis of both **1** and **2** (Fig. 4c). To prove this hypothesis, we re-screened the BAC library to obtain a BAC containing the *egp* cluster entirely lacking the concanamycin BGC or the *whiE*-like BGC. One such BAC clone (which we designated pKU503egp; Fig. 4a) was identified and introduced into *S. avermitilis* SUKA54. As expected, the exoconjugants harboring pKU503egp produced both **1** and **2**, confirming the identity of the *egp* cluster (Supplementary Fig. 3). Simultaneously, these results indicate that the type I PKS cluster (*ccn* cluster) and the type II PKS BGC (*egp* cluster) are located next to each other, and each BGC is functional for producing secondary metabolites. Also, it is rare that three BGCs (*ccn*, *egp*, and *whiE*-like cluster) are densely encoded within about 200 kb in actinomycetal genome (Fig. 4a).

We further annotated the 26 ORFs in the region of overlap between the pKU503ccn and pKU503egp inserts in

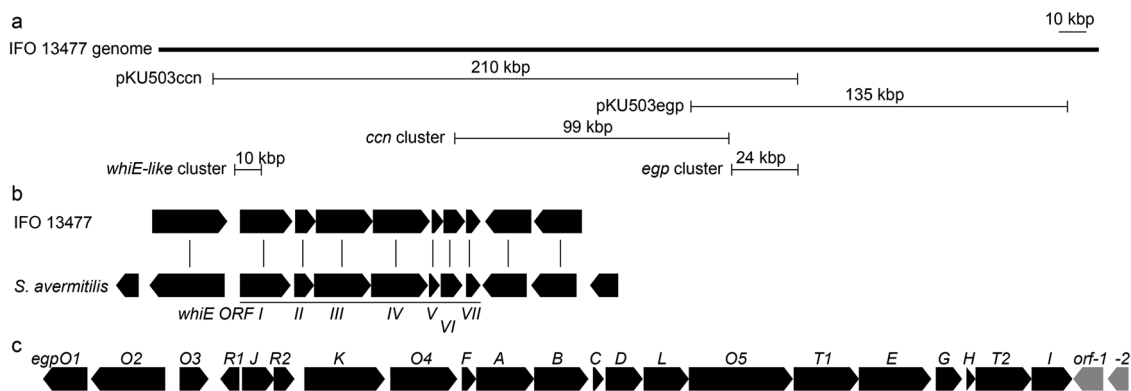


Fig. 4 Location and organization of biosynthetic gene clusters (BGCs) characterized in the present study. Schematic diagrams of BGC-containing regions. **a** The chromosomal locations of the inserts of the BACs (pKU503ccn and pKU503egp) that were isolated in the present study and of the BGCs that were characterized. A segment of the genome (bold line) is depicted. Scale bar, 10 kb. **b** Gene organization of the *whiE*-like BGCs from IFO 13477 and *S. avermitilis*. Each

pentagon represents a gene, with the pointed end indicating the orientation of the open reading frame (ORF). Vertical lines are used to indicate ORFs that are similar between the two sources (see also Supplementary Table 3). The names of ORFs similar to genes of the *whiE* cluster of *S. coelicolor* A3(2) are indicated directly below the *S. avermitilis* sequence. **c** Gene organization of the *egp* BGC

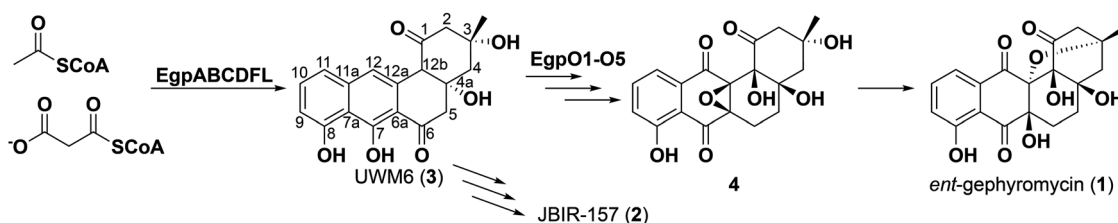


Fig. 5 Proposed biosynthetic pathway for *ent*-gephyromycin (**1**) and JBIR-157 (**2**). Based on the annotation analyses of the *egp* cluster, **1** and **2** likely are biosynthesized through a shared intermediate, UWM6 (**3**). Compound **4** is a hypothetical intermediate whose production from

3 would require multiple oxygenation/reduction steps. Definition of the detailed structure of **2** is expected to further improve the accuracy of the proposed biosynthetic pathway

an effort to predict the biosynthetic pathway for **1** and **2** (Fig. 5). Three of these ORFs (*ccn1*, *ccn2*, and *ccn3*) were identical to *con1*, *con2*, and *con3*, respectively, that had been proposed to belong to the *con* cluster, although the functions of the respective gene products remain unknown (Although unpublished data, *Con3* positively regulates the expression of *con* cluster). While the actual boundary between the *ccn* and *egp* clusters still is unclear, we tentatively have assigned 21 genes as components of the *egp* cluster (Table 1). It is noteworthy that each *egp* gene has at least one similar gene associated with a previously reported type II PKS-encoding gene cluster that has been deposited in the MIBiG (Minimum Information about a Biosynthetic Gene cluster) database [30]. In addition, the *egp* cluster was identified in the previously reported *S. neyagawaensis* genome, and also found in the *S. cyaneochromogenes* genome with the extremely high sequence similarity (Supplementary Table 4).

Proposed biosynthetic pathway for **1** and **2**

EgpA, EgpB, and EgpC show significant similarity to JadA (85% amino acid identity), JadB (78%), and JadC (73%), respectively; JadABC constitutes a minimal PKS [31]. In addition, EgpD, EgpL, and EgpF are close homologues of JadE (84% amino acid identity), JadD (75%), and JadI (81%), respectively. The *jadABCDEI* gene set is conserved across several type II PKS-encoding clusters whose products are known to be responsible for the biosynthesis of UWM6 (**3**) [32]; notably, **3** is a common biosynthetic intermediate of well-studied angucyclines such as landomycins [33], urdamycins [34], and jadomycins [35]. Accordingly, we presume that the biosynthesis of **1** and **2** employs **3** as an intermediate (Fig. 5).

The conversion from **3** to **1** requires several oxygenation/reduction steps. The hydroxylation at C-12b of angucyclines has been investigated in the context of the biosynthesis of urdamycin [36] and gaudimycin [37]. In both cases, a NADPH-dependent flavoprotein hydroxylase (UrDE or PgaE, respectively) was assigned as the responsible enzyme. In the *egp* cluster, EgpO4 has 78.1% amino acid identity to PgaE.

Additionally, the C-terminal domain of EgpO5 has 78.5% amino acid identity to the respective region of PgaM, a short-chain dehydrogenase/reductase domain that has been proposed to catalyze the formation of the C-6 hydroxy group [36]. An ethylene moiety is present on the B-ring of several angucyclines, including urdamycins I and J [38], gaudimycin A [39], and SF2315A and B [40]. However, the responsible enzyme(s) remain unidentified. The candidate participants encoded by the *egp* cluster include EgpO1 and EgpO2. EgpO1 has a 48.0% amino acid identity to ActVI-orf2, an enoyl reductase involved in actinorhodin biosynthesis [41]. On the other hand, domain analysis indicates that EgpO2 is a didomain protein. Specifically, the N-terminal domain (EgpO2-N) has a Rossmann-fold motif, while the C-terminal domain (EgpO2-C) resembles a methyltransferase. In the MIBiG database, the protein with the strongest similarity to EgpO2-N is BexH (64.0% amino acid identity), a putative sugar epimerase in the BE-7585A biosynthetic pathway [42]. However, EgpO2-N also shows some similarity (39.4% amino acid identity) to SimC7, a ketoreductase involved in simocyclinone biosynthesis [43]. The oxygenation at the C-ring to form a quinone is presumably performed by EgpO3, a protein with intermediate similarity (34.7% amino acid identity) to ActVA-orf6. ActVA-orf6 is a cofactor-independent monooxygenase that oxygenates the phenolic intermediate, 6-deoxydihydrocarafungin, into the quinone, dihydrokalafungin, during actinorhodin biosynthesis [44]. Based on these speculations, we tentatively hypothesize that **3** is converted to **1** via an intermediate consisting of **4** (Fig. 5), an enantiomer of one of two predicted precursors for gephyromycin.

Among the remaining genes, *egpR1-R2* and *egpT1-T2* may encode transcriptional regulators and transporters, respectively. Additionally, two transcriptional regulators, ORF-1 and ORF-2, are tandemly encoded in opposite direction to the putative *egpJ-I* operon. Each of them shows 52% and 47% amino acid identity, respectively, to KiqA, the positive regulator for the kinanthraquinone BGC [45]. The particular functions of ORF-1 and ORF-2 are still elusive. EgpK and EgpE show ~90% amino acid identity to the α and β subunits (respectively) of the *S. avermitilis* acetyl-CoA carboxylase.

Table 1 Deduced functions of ORFs involving *ent*-gephyromycin in *egp* cluster

ORF	aa	proposed function	BLAST hit	homology (identity/similarity)	Pfam family
Cen1	320	hypothetical protein	AAZ94385.1: putative malonyl transferase [<i>Streptomyces neyagawaensis</i>]	100%/100%	no hits
Cen2	368	hypothetical protein	AAZ94384.1: hypothetical protein [<i>Streptomyces neyagawaensis</i>]	100%/100%	no hits
Cen3	921	LuxR-family transcriptional regulator for concanamycin cluster	AAZ94383.1: putative LuxR class regulator [<i>Streptomyces neyagawaensis</i>]	100%/100%	PF13191: AAA ATPase domain PF00196: Bacterial regulatory proteins, luxR family
EgpO1	324	dehydrogenase	WP_234341498.1: zinc-binding dehydrogenase [<i>Streptomyces</i> sp. NRRL S-646]	96%/97%	PF00197: Zinc-binding dehydrogenase
EgpO2	543	(N-term.) oxidoreductase	WP_123993686.1: NAD(P)H-binding protein [unclassified <i>Streptomyces</i>]	76%/86%	PF05368: NmrA-like family
		(C-term.) methyltransferase	WP_051833649.1: class I SAM-dependent methyltransferase [<i>Streptomyces</i> sp. NRRL S-646]	95%/96%	PF08241: Methyltransferase domain PF08242: Methyltransferase domain
EgpO3	219	oxygenase	WP_208781756.1: hypothetical protein [<i>Streptomyces griseorubiginosus</i>]	43%/62%	PF12847: Methyltransferase domain PF03992: Antibiotic biosynthesis monooxygenase
EgpR1	136	regulator	WP_242760576.1: helix-turn-helix domain-containing protein [<i>Streptomyces cylindrosporus</i>]	68%/76%	PF01638: HxIR-like helix-turn-helix
EgpJ	237	oxidoreductase	WP_208781755.1: DsbA family oxidoreductase [<i>Streptomyces griseorubiginosus</i>]	60%/79%	PF01323: DSBA-like thioredoxin domain
EgpR2	150	MerR family transcriptional regulator	WP_255926062.1: redox-sensitive transcriptional activator SoxR [<i>Streptomyces rubrisolii</i>]	75%/84%	PF00376: MerR family regulatory protein PF09278: MerR, DNA binding
EgpK	592	acyl-CoA carboxylase subunit alpha	WP_189782387.1: acetyl/propionyl/methylcrotonyl-CoA carboxylase subunit alpha [<i>Streptomyces capitiformicae</i>]	92%/94%	PF00289: Biotin carboxylase, N-terminal domain PF02786: Carbamoyl-phosphate synthase L chain, ATP binding domain PF02785: Biotin carboxylase C-terminal domain
EgpO4	491	oxygenase	WP_073737890.1: FAD-dependent monooxygenase [<i>Streptomyces</i> sp. CB02488]	84%/90%	PF00364: Biotin-requiring enzyme PF01494: FAD binding domain
EgpF	108	polyketide cyclase	WP_024491216.1: TemI family type II polyketide cyclase [<i>Streptomyces</i> sp. CB02488]	83%/91%	PF04673: Polyketide synthesis cyclase
EgpA	423	3-oxoacyl-ACP synthase	APD71568.1: type II polyketide synthase 3 [<i>Streptomyces</i> sp.]	95%/97%	PF00109: Beta-ketoacyl synthase, N-terminal domain PF02801: Beta-ketoacyl synthase, C-terminal domain

Table 1 (continued)

ORF	aa	proposed function	BLAST hit	homology (identity/ similarity)	Pfam family
EgpB	403	3-oxoacyl-ACP synthase (probably chain length factor)	APD71569.1: type II polyketide synthase 4 [<i>Streptomyces</i> sp.]	91%/94%	PF00109: Beta-ketoacyl synthase, N-terminal domain PF02801: Beta-ketoacyl synthase, C-terminal domain
EgpC	85	acyl carrier protein	WP_148646228.1: acyl carrier protein [<i>Streptomyces</i> sp. CB01881]	74%/87%	PF00550: Phosphopantetheine attachment site
EgpD	273	3-oxoacyl-ACP reductase	WP_030683251.1: 3-oxoacyl-ACP reductase FabG [<i>Streptomyces cellulosae</i>]	87%/93%	PF00106: short chain dehydrogenase PF08659: KR domain
EgpL	306	aromatase	WP_190075888.1: aromatase/cyclase [<i>Streptomyces lanatus</i>]	78%/86%	PF03364: Polyketide cyclase / dehydrase and lipid transport PF10604: Polyketide cyclase / dehydrase and lipid transport
EgpO5	768	bifunctional oxygenase-reductase	WP_143941337.1: SDR family oxidoreductase [<i>Streptomyces benahoarensis</i>]	79%/86%	PF01494: FAD binding domain PF00106: short chain dehydrogenase
EgpT1	477	transmembrane efflux protein	TSB26785.1: MFS transporter [<i>Streptomyces benahoarensis</i>]	66%/74%	PF07690: Major Facilitator Superfamily
EgpE	534	methylmalonyl-CoA decarboxylase alpha subunit	WP_055713146.1: acyl-CoA carboxylase subunit beta [<i>Streptomyces torulosus</i>]	93%/96%	PF01039: Carboxyl transferase domain
EgpG	194	oxidoreductase	WP_143941340.1: NAD(P)H-dependent oxidoreductase [<i>Streptomyces benahoarensis</i>]	80%/88%	PF03358: NAFPH-deoendent FMN reductase PF02525: Flavodoxin-like fold
EgpH	65	hypothetical protein	no hits	no hits	
EgpT2	408	major facilitator transporter	WP_179081696.1: MFS transporter [<i>Streptomyces rectiverticillatus</i>]	61%/76%	PF07690: Major Facilitator Superfamily
EgpI	301	hypothetical protein	GHA00510.1: thioesterase [<i>Streptomyces echinoruber</i>]	65%/73%	PF12697: Alpha/beta hydrolase family PF12695: Alpha/beta hydrolase family
ORF-1	220	two-component system response regulator	WP_108906939.1: response regulator transcription factor [<i>Streptomyces tirandamycinicus</i>]	54%/68%	PF00486: Transcriptional regulatory protein, C terminal
ORF-2	153	hypothetical protein	WP_207315332.1: response regulator transcription factor [<i>Streptomyces laculatispora</i>]	60%/73%	PF00072: Response regulator receiver domain PF00486: Transcriptional regulatory protein, C terminal

Although the genes within *egp* cluster do not appear to encode an acetyl-CoA carboxylase ϵ subunit, the products of the *egpK* and *egpE* genes may enhance polyketide production by supplying malonyl-CoA. EgpO2-C has 49.2% amino acid identity to BexP, a methyltransferase predicted to be involved in self-resistance to the angucycline BE-7585A [42]. However, the function of EgpO2-C remains unclear, given that **1** and **2** did not show antibacterial activity (data not shown). EgpJ (which includes a DsbA-like thioredoxin domain), EgpG (a putative NADPH-dependent FMN reductase), and EgpI (a putative member of the α/β -hydrolase family) also have strong similarities to proteins encoded by type II PKS BGCs. These proteins may be responsible for protein folding, flavin mononucleotide (FMN) regeneration, and the off-loading step of polyketide synthesis, respectively, although these suggestions will need to be confirmed experimentally. Lastly, EgpH constitutes a 66-amino-acid protein without any apparent motifs; the role of this gene product in the biosynthesis of **1** and **2** remains unknown.

Concluding remarks

This work represents (to the best of our knowledge) the first report of heterologous production of concanamycins. The heterologous expression system with a suitable production titer is expected to facilitate the generation of novel concanamycins via engineering the BGC. We additionally note that our sequence analysis, performed using a long-read sequencer, refined the BGC at the nucleotide level. The data presented here should serve as a superior reference for studies requiring precise sequences, such as those involving PKS engineering.

The capacity of the BAC cloning technique to capture a large insert, along with the use of a host strain that lacks intrinsic relevant secondary metabolites, enabled us to identify not only the target compound (concanamycin) but also two type II PKS products. In combination with the results of our heterologous expression system, bioinformatic analyses provided a milestone for our understanding of the biosynthesis of **1**, a unique bridged angucycline. Further biosynthetic analyses are in progress.

Acknowledgements MI deceased on 23 December 2015. This work was supported by AMED under Grant JP19ae0101045 to KS and JSPS KAKENHI Grant Number JP23H04569 to KK.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

- Pye CR, Bertin MJ, Lokey RS, Gerwick WH, Linington RG. Retrospective analysis of natural products provides insights for future discovery trends. *Proc Natl Acad Sci USA*. 2017;114:5601–6.
- Palazzolo AME, Simons CLW, Burke MD. The natural productome. *Proc Natl Acad Sci USA*. 2017;114:5564–6.
- Singh V, Haque S, Niwas R, Srivastava A, Pasupuleti M, Tripathi CK. Strategies for fermentation medium optimization: an in-depth review. *Front Microbiol*. 2016;7:2087.
- Hoshino S, Onaka H, Abe I. Activation of silent biosynthetic pathways and discovery of novel secondary metabolites in actinomycetes by co-culture with mycolic acid-containing bacteria. *J Ind Microbiol Biotechnol*. 2019;46:363–74.
- Adrio JL, Demain AL. Genetic improvement of processes yielding microbial products. *FEMS Microbiol Rev*. 2006;30:187–214.
- Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell*. 2014;158:412–21.
- Doroghazi JR, Albright JC, Goering AW, Ju KS, Haines RR, Tchaluikov KA, et al. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol*. 2014;10:963–8.
- Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozono I, Izumikawa M, et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth Biol*. 2013;2:384–96.
- Hashimoto T, Hashimoto J, Kozono I, Amagai K, Kawahara T, Takahashi S, et al. Biosynthesis of quinolidomycin, the largest known macrolide of terrestrial origin: identification and heterologous expression of a biosynthetic gene cluster over 200 kb. *Org Lett*. 2018;20:7996–9.
- Zhang L, Hashimoto T, Qin B, Hashimoto J, Kozono I, Kawahara T, et al. Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolylol polyketides. *Angew Chem Int Ed Engl*. 2017;56:1740–5.
- Hashimoto T, Kozono I, Hashimoto J, Ueoka R, Kagaya N, Fujie M, et al. Novel macrolactam compound produced by the heterologous expression of a large cryptic biosynthetic gene cluster of *Streptomyces rochei* IFO12908. *J Antibiot*. 2020;73:171–4.
- Hashimoto T, Kozono I, Hashimoto J, Suenaga H, Fujie M, Satoh N, et al. Identification, cloning and heterologous expression of biosynthetic gene cluster for desertomycin. *J Antibiot*. 2020;73:650–4.
- Kinashi H, Someno K, Sakaguchi K, Higashijima T, Miyazawa T. Structure of concanamycin a. *Tetrahedron Lett*. 1981;22:3861–4.
- Kinashi H, Someno K, Sakaguchi K. Isolation and characterization of concanamycins A, B and C. *J Antibiot*. 1984;37:1333–43.
- Yamamoto H, Nakazawa K, Horii S, Miyake A. Studies on Agricultural Antibiotic. *J Agric Chem Soc Jpn*. 1960;34:268–72.
- Drose S, Bindseil KU, Bowman EJ, Siebers A, Zeeck A, Altendorf K. Inhibitory effect of modified bafilomycins and concanamycins on P- and V-type adenosinetriphosphatases. *Biochemistry*. 1993;32:3902–6.
- Haydock SF, Appleyard AN, Mironenko T, Lester J, Scott N, Leadlay PF. Organization of the biosynthetic gene cluster for the macrolide concanamycin A in *Streptomyces neyagawaensis* ATCC 27449. *Microbiology*. 2005;151:3161–9.
- Demachi A, Ohte S, Uchida R, Shin-Ya K, Ohshiro T, Tomoda H, et al. Discovery of prescopranone, a key intermediate in scopranone biosynthesis. *J Antibiot*. 2022;75:305–11.
- Cane DE, He X, Kobayashi S, Ōmura S, Ikeda H. Geosmin biosynthesis in *Streptomyces avermitilis*. Molecular cloning, expression, and mechanistic study of the germacradienol/geosmin synthase. *J Antibiot*. 2006;59:471–9.
- Komatsu M, Uchiyama T, Ōmura S, Cane DE, Ikeda H. Genome-minimized *Streptomyces* host for the heterologous

- expression of secondary metabolism. *Proc Natl Acad Sci USA*. 2010;107:2646–51.
21. Kim JH, Komatsu M, Shin-ya K, Ōmura S, Ikeda H. Distribution and functional analysis of the phosphopantetheinyl transferase superfamily in Actinomycetales microorganisms. *Proc Natl Acad Sci USA*. 2018;115:6828–33.
 22. Chu G, Vollrath D, Davis RW. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science*. 1986;234:1582–5.
 23. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci*. 2017;27:135–45.
 24. Bringmann G, Lang G, Maksimenka K, Hamm A, Gulder TA, Dieter A, et al. Gephyromycin, the first bridged angucyclinone, from *Streptomyces griseus* strain NTK 14. *Phytochemistry*. 2005;66:1366–73.
 25. Chang Y, Xing L, Sun C, Liang S, Liu T, Zhang X, et al. Monocyclones G–K and ent-gephyromycin A, angucycline derivatives from the marine-derived *Streptomyces* sp. HDN15129. *J Nat Prod*. 2020;83:2749–55.
 26. Keatinge-Clay A. Crystal structure of the erythromycin polyketide synthase dehydratase. *J Mol Biol*. 2008;384:941–53.
 27. Zhang W, Fortman JL, Carlson JC, Yan J, Liu Y, Bai F, et al. Characterization of the bafilomycin biosynthetic gene cluster from *Streptomyces lohii*. *ChemBioChem*. 2013;14:301–6.
 28. Zhu W-Z, Wang S-H, Gao H-M, Ge Y-M, Dai J, Zhang X-L, et al. Characterization of bioactivities and biosynthesis of angucycline/angucyclinone derivatives derived from *Gephyromycinifex aptenodytis* gen. nov., sp. nov. *Mar Drugs*. 2021;20:34.
 29. Davis NK, Chater KF. Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol Microbiol*. 1990;4:1679–91.
 30. Kautsar SA, Blin K, Shaw S, Navarro-Munoz JC, Terlouw BR, van der Hoof JJJ, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res*. 2020;48:D454–D8.
 31. McDaniel R, Ebert-Khosla S, Hopwood DA, Khosla C. Engineered biosynthesis of novel polyketides. *Science*. 1993;262:1546–50.
 32. Kulowski K, Wendt-Pienkowski E, Han L, Yang K, Vining LC, Hutchinson CR. Functional characterization of the *jadI* gene as a cyclase forming angucyclinones. *J Am Chem Soc*. 1999;121:1786–94.
 33. Westrich L, Domann S, Faust B, Bedford D, Hopwood DA, Bechthold A. Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. *FEMS Microbiol Lett*. 1999;170:381–7.
 34. Decker H, Haag S. Cloning and characterization of a polyketide synthase gene from *Streptomyces fradiae* Tu2717, which carries the genes for biosynthesis of the angucycline antibiotic urdamycin A and a gene probably involved in its oxygenation. *J Bacteriol*. 1995;177:6126–36.
 35. Han L, Yang K, Ramalingam E, Mosher RH, Vining LC. Cloning and characterization of polyketide synthase genes for jadomycin B biosynthesis in *Streptomyces venezuelae* ISP5230. *Microbiology*. 1994;140:3379–89.
 36. Patrikainen P, Kallio P, Fan K, Klika KD, Shaaban KA, Mantsala P, et al. Tailoring enzymes involved in the biosynthesis of angucyclines contain latent context-dependent catalytic activities. *Chem Biol*. 2012;19:647–55.
 37. Kallio P, Patrikainen P, Suomela JP, Mantsala P, Metsa-Ketela M, Niemi J. Flavoprotein hydroxylase PgaE catalyzes two consecutive oxygen-dependent tailoring reactions in angucycline biosynthesis. *Biochemistry*. 2011;50:5535–43.
 38. Künzel E, Faust B, Oelkers C, Weissbach U, Bearden DW, Weitnauer G, et al. Inactivation of the *urdGT2* gene, which encodes a glycosyltransferase responsible for the C-glycosyltransfer of activated d-olivose, leads to formation of the novel urdamycins I, J, and K. *J Am Chem Soc*. 1999;121:11058–62.
 39. Palmu K, Ishida K, Mantsala P, Hertweck C, Metsa-Ketela M. Artificial reconstruction of two cryptic angucycline antibiotic biosynthetic pathways. *ChemBioChem*. 2007;8:1577–84.
 40. Sasaki T, Gomi S, Sezaki M, Takeuchi Y, Kodama Y, Kawamura K. New antibiotics SF2315A and B produced by an *Excelsopora* sp. II. The structural elucidation. *J Antibiot*. 1988;41:843–8.
 41. Ishikawa K, Hashimoto M, Komatsu K, Taguchi T, Okamoto S, Ichinose K. Characterization of stereospecific enoyl reductase ActVI-ORF2 for pyran ring formation in the actinorhodin biosynthesis of *Streptomyces coelicolor* A3(2). *Bioorg Med Chem Lett*. 2022;66:128727.
 42. Sasaki E, Ogasawara Y, Liu HW. A biosynthetic pathway for BE-7585A, a 2-thiosugar-containing angucycline-type natural product. *J Am Chem Soc*. 2010;132:7405–17.
 43. Schafer M, Le TB, Hearnshaw SJ, Maxwell A, Challis GL, Wilkinson B, et al. SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone. *J Mol Biol*. 2015;427:2192–204.
 44. Sciarra G, Kendrew SG, Miele AE, Marsh NG, Federici L, Malatesta F, et al. The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. *EMBO J*. 2003;22:205–15.
 45. Takao R, Sakai K, Koshino H, Osada H, Takahashi S. Identification of the kinanthraquinone biosynthetic gene cluster by expression of an atypical response regulator. *Biosci Biotechnol Biochem*. 2021;85:714–21.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.