ORIGINAL ARTICLE

Cloning of the Gene Cluster Responsible for the Biosynthesis of Brasilicardin A, a Unique Diterpenoid

Yutaka Hayashi, Nobuyasu Matsuura, Hiroaki Toshima, Nobuya Itoh, Jun Ishikawa, Yuzuru Mikami, Tohru Dairi

Abstract Brasilicardin A (BCA), produced by *Nocardia brasiliensis* IFM 0406 (currently referred to as *N. terpenica*), has a unique structure consisting of a diterpene skeleton with L-rhamnose, *N*-acetylglucosamine, amino acid, and 3-hydroxybenzoate moieties, and exhibits potent biological activities. To understand the biosynthetic machinery of this unique compound, we have cloned the corresponding gene cluster. Firstly, we cloned a gene by PCR that encodes geranylgeranyl diphosphate synthase (GGPPS), which produces a direct precursor of diterpene compounds. We obtained four candidate genes and one of the genes was confirmed to encode a GGPPS. By sequence analysis of regions flanking the GGPPS gene, we identified eleven genes (*bra1-11*), all oriented in the same direction. We did not, however, detect any genes related to L-rhamnose and *N*-acetylglucosamine biosyntheses in the flanking regions. A gene disruption experiment did indeed show that this gene cluster was responsible for BCA biosynthesis.

Keywords *Nocardia brasiliensis*, brasilicardin A, diterpene, cloning, biosynthetic genes.

T. Dairi (Corresponding author)**, Y. Hayashi, N. Itoh:** Biotechnology Research Center, Toyama Prefectural University, Toyama 939-0398, Japan, E-mail: dairi@pu-toyama.ac.jp
N. Matsuura: Department of Life Science, Okayama University of Science 700-0005, Japan
H. Toshima: Department of Bioresource Science, College of Agriculture, Ibaraki University, Ibaraki 300-0393, Japan
J. Ishikawa: Department of Bioactive Molecules, National Institute of Infectious Diseases, Shinjuku-ku 162-8640, Japan
Y. Mikami: Medical Mycology Research Center, Chiba University, Chiba 260-8673, Japan

Introduction

Isoprenoids are the largest single family of compounds found in nature, with over 24,000 known examples, and contain industrially useful compounds, such as flavors, antibiotics, and plant hormones [1]. All isoprenoid compounds are biosynthesized from the corresponding polyprenyl diphosphate [2]. Geranyl diphosphate gives rise to monoterpenes [3], farnesyl diphosphate to sesquiterpenes [4], and geranylgeranyl diphosphate to diterpenes [5]. In eukaryotes, such as plants and fungi, which produce the vast majority of isoprenoids, the polyprenyl diphosphate is usually cyclized by an organism specific isoprenoid cyclase to make a basic skeleton, followed by successive reactions such as hydroxylation, methylation, and glycosylation, *etc*, to give many thousands of compounds [3~8].

In contrast to eukaryotes, prokaryotes are known to produce a limited number of isoprenoids. Among prokaryotes, actinomycete strains have been found to produce several isoprenoid compounds. To date, sesquiterpene cyclases, such as pentalenene synthase [9], (+)-*epi*-isozizaene synthases [10], and geosmin synthase [11], have been identified and characterized. We have also cloned several diterpene cyclase genes from actinomycetes [12~16]. All these isoprenoid cyclases were known to accept substrates that are biosynthesized from isopentenyl diphosphate, such as farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP), in a similar manner to eukaryotic enzymes. On the other hand, it was recently revealed that the structures of many isoprenoids produced by actinomycetes have mixed biosynthetic origin [16]. In many cases, isoprenoid moieties of compounds are attached to other moieties, such as a polyketide, an aromatic ring, an
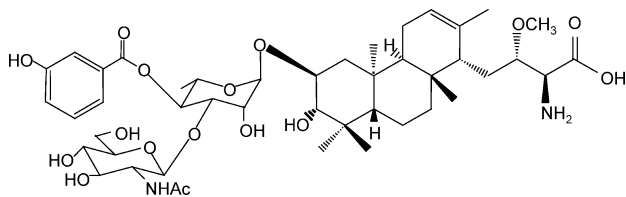
Fig. 1 Structure of brasilicardin A.

amino acid, *etc*, that have been synthesized *via* pathways that are independent of isoprenoid synthesis, to give so-called "isoprenoid-hybrid compounds" [16]. Brasilicardin A (BCA), produced by *Nocardia brasiliensis* IFM 0406, is a representative of this type of compound, which consists of a basic diterpene skeleton with L-rhamnose, *N*-acetylglucosamine, amino acid, and hydroxybenzoate moieties (Fig. 1) [17]. BCA was demonstrated to possess potent anti-tumor [18] and immunosuppressive activity [19]. Derivatives of BCA have been chemically synthesized to search for novel compounds with high immuno-suppressive activity and to examine the relationship between BCA-structures and immunosuppressive activities [20].

In this study, we cloned the BCA biosynthetic gene cluster, as a first step to understanding the biosynthetic machinery of this unique and pharmaceutically important compound. A gene disruption experiment was also performed to examine whether the gene cluster identified in this study is indeed responsible for BCA biosynthesis.

Experimental

Chemicals
[α-^{32}P]dCTP was obtained from GE Healthcare UK Ltd. (Little Chalfont, England). Isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP), FPP, and GGPP were purchased from Sigma-Aldrich (St. Louis, MO). All other chemicals were analytical grade. Authentic BCA was kindly provided by Kikkoman Co., Ltd. (Chiba, Japan).

Bacterial Strains
Nocardia brasiliensis IFM 0406 was used in the cloning experiment. Media and growth conditions for IFM 0406 were as described by Komaki *et al*. [17]. *Escherichia coli* JM110 and plasmids, pUC118 and pUC119, were used for sequence analysis. When necessary, ampicillin was added to the medium to a final concentration of 100 μg/ml.

DNA Isolation and Manipulation
Plasmids from *E. coli* were prepared using a Qiagen plasmid kit (Hilden, Germany). All restriction enzymes, T4 ligase, and calf intestinal alkaline phosphatase were obtained from Toyobo and used according to the manufacturer's protocols. Transformation of *E. coli* with plasmid DNA by electroporation was performed under standard conditions using a BTX ECM 600 electroporation system (Biotechnologies and Experimental Research, San Diego, CA). Other general procedures were performed as described by Maniatis *et al*. [21].

Cloning of Polyprenyl Diphosphate Synthase Genes by PCR
Three sets of primers were designed, based on the conserved amino acid sequences of polyprenyl diphosphate synthase genes, which had been cloned from terpentecin [22], furaquinocin A [14, 15], and BE-40644 [23] producers: PDS-1-5; 5′-(C/G/T)T(C/G)CACGACGAC-(C/G/T)T(C/G)AT(C/G)GAC(A/C/G)(A/G) (C/G)-3′, PDS-1-3; 5′-(C/G)A(C/G)GTCGTC(C/G)(A/C)(C/G)(C/G)-(A/G)(A/G)CTGGAA(C/G)GC-3′, PDS-2-5; 5′-(C/T)T-(C/G)(C/T)T(C/G)CACGACGACGT(C/G)(A/G)T(C/G)G ACG(A/C)(C/G)-3′, PDS-2-3; 5′-GTCGTC(C/G)A(C/T)-(C/G)A(A/G/T)CTGGAA(C/G)GC(C/G)(A/G)T(C/G)CC-3′, PDS-3-5; 5′-GC(C/G)ATC(C/T)T(C/G)GC(C/G)GG-(C/G)GAC(C/G/T)T(C/G)(A/T)T-3′, PDS-3-3; and 5′-(C/G)A(C/G)GTCGTC(C/G)(A/C)(C/G)(C/G)(A/G)(A/G)CTGGAA(C/G)GC-3′. These primers were used in PCR experiments to amplify polyprenyl diphosphate synthase (PDS) genes from genomic DNA of strain IFM 0406. The standard PCR amplification was performed with Takara LA PCR kit (Takara Bio Inc., Kyoto, Japan) under the following conditions: denaturation at 94°C for 1 minute, annealing at 50°C for 1 minute, elongation at 72°C for 1.5 minutes, 30 cycles. When necessary, annealing temperature was varied and other DNA polymerases, such as KOD- and Pfu-DNA polymerases (Toyobo), were used instead of Taq-derived DNA polymerase. The amplified DNA fragments were inserted into pGEM-T easy vector (Promega, Madison, WI) and were then used for sequence analysis. A cosmid library of strain IFM 0406, which was constructed by the same method as described previously [24], was screened by colony hybridization, using ^{32}P-labeled amplified fragments as probes. Positive cosmids were selected and DNA fragments that again hybridized to the probe by Southern blot analysis were subcloned into pUC119. After construction of a series of plasmids, sequencing was performed with an automatic DNA sequencer (Li-Cor, Model 4000L, Lincoln, NE).

Overproduction of the GGPPS in *E. coli*
The following primers were used to amplify the *ggpp1*

gene: 5′-CAT<u>GCATGC</u>ACCCGCTACGCCGTGCAGC-CGCCGGG-3′ (forward) and 5′-CAC<u>AAGCTT</u>TCACCA-CGTGCGCCCGATCATGCTGT-3′ (reverse). PCR was carried out under standard conditions. After sequence confirmation, a 0.9-kb *Sph*I-*Hin*dIII fragment was inserted into the same sites of pQE30. A plasmid, pQE30-ggpps, in which a recombinant protein was expressed as an *N*-terminal 6 x His-tagged fusion protein, was selected.

E. coli M15 [pREP4] harboring pQE30-ggpps was grown at 37°C in Luria-Bertani broth medium with appropriate antibiotics. Expression of the recombinant protein was induced by adding 0.1 mM of isopropyl-*β*-D-thiogalactoside when OD_{600} reached about 0.8. Cultivation was continued for an additional 12 hours at 18°C. Purification of His-tagged recombinant protein was performed according to the manufacturer's protocol (Qiagen). Purified proteins were analyzed by SDS-PAGE on 10% gels and used in an *in vitro* enzyme assay to determine the product chain length. Assay conditions and product analysis were performed as described previously [25].

Disruption of the Cyclase Gene

The bra*4* encoding a cyclase gene was disrupted by a double crossover event. A 3.0 kb *Hin*dIII/*Bgl*II DNA fragment and a 2.5 kb *Bgl*II/*Hin*dIII DNA fragment, which carried the upstream region and the downstream region of the *bra4* gene, respectively, were amplified by PCR with the following two sets of primers: 5′-ACC<u>AAGCTT</u>G-CGGAGGGTCCGGTGATGGTGAATTC-3′ and 5′-GA<u>AGATCT</u>GCCGACCTCACAGGGATATCGCCGAAT-3′; 5′-GA<u>AGATCTT</u>GGAACCGACGCCGACGCCGTG-GTGTG-3′ and 5′-ACC<u>AAGCTT</u>CTACCACGTCACC-GGAAGCCGACGAA-3′. After sequence confirmation, these two fragments and *Hin*dIII-digested pUC118 were ligated and introduced into *E. coli*. A plasmid carrying the appropriate DNA fragments was selected. A 0.8 kb *Bcl*I fragment that carried a thiostrepton resistance gene was prepared from pIJ702 [26] and was then inserted into *Bgl*II site of the plasmid, which existed between the PCR-amplified fragments. This plasmid was then digested with *Hin*dIII and the resultant fragment was inserted into the same sites of pK18mob [27] to construct pCyc-DIS. The plasmid was used for disruption of the chromosomal *bra4* gene by double crossover recombination. Introduction of the plasmid from *E. coli* into the BCA producer was performed by transconjugation and thiostrepton resistant colonies were selected [28]. Genomic DNAs were prepared from the transconjugants and subjected to PCR analysis with the following primers: 5′-GTGAGCGCGCGCA-CCCGGCGGCACTGGTGGGCGCA-3′ and 5′-AGCGT-GGCCGCGACATCGAGGAGGCGATGTGCTGA-3′, which corresponded to the *N*-terminal and *C*-terminal coding region of the *bra3* and *bra5* genes. The productivity of BCA by the disruptant was examined as follows: The disruptants were cultivated by the methods described by Komaki *et al*. [17], except for the addition of thiostrepton (50 μg/ml) to the medium. After full growth cultivation, the culture broth was centrifuged, and compounds were extracted by ethyl acetate from the supernatant. Compounds were analyzed by HPLC: column, LiChroCART 125-4.0 Superspher 100 RP-18, Kanto Chemical, Tokyo, Japan; Mobile phase, 18% CH_3CN containing TFA (0.15%) for 0 to 30 minutes, and 50% CH_3CN containing TFA (0.15%) for an additional 10 minutes; flow rate, 1.0 ml/minute; detection 220 nm.

Nucleotide Sequence Accession Number

The DNA sequences determined in this study are deposited in the DDBJ, EMBL, and GenBank databases under accession numbers AB264550 (BCA biosynthetic gene cluster) and AB268547 (a sugar biosynthetic gene cluster).

Results

Cloning of a *ggpps* Gene by PCR

We found recently that biosynthetic genes of terpentecin [12] and 3-hydroxypimeradiene (viguiepinol) [14, 15], both of which are diterpene compounds produced by actinomycetes, were located in flanking regions of the genes encoding geranylgeranyl diphosphate synthase (GGPPS), which supplies a direct precursor of diterpene compounds. Hence, we first tried to clone a GGPPS gene in order to use it as a probe to clone an entire BCA biosynthetic gene cluster. To date, many deduced amino acid sequences of PDS, including GGPPS, have been reported. For example, we could identify at least six PDS genes in both *Streptomyces coelicolor* A3(2) (SCO185, SCO565, SCO568, SCO4583, SCO5250, and SCO6763) and *S. avermitilis* (SAV1022, SAV1129, SAV1651, SAV2997, SAV3006, and SAV4864), the entire genome sequences of which have been determined [29, 30]. Furthermore, we have identified four PDS genes in *S. argenteolus*, a unique tetraterpene compound, KS-505a, producer [31]. These facts suggested that an actinomycete BCA producer has several PDS genes. Hence, we tried to clone multiple putative PDS genes by using multiple sets of primers, which were designed based on the conserved amino acid sequences of PDSs, and by varying the PCR conditions, such as annealing temperatures and Mg^{2+} concentration.

We amplified and cloned four kinds of DNA fragments (*ggpps1*~*ggpps4*), all of which are thought to encode enzymes similar to PDSs. By sequence analysis, we did not find any genes related to isoprenoid biosynthesis in the flanking regions of *ggpps3* or *ggpps4*. The *ggpps2* locus suggested that it constituted an operon, with a gene that encodes an enzyme similar to squalene-hopene cyclase, in a manner similar to that observed in *S. coelicolor* A3(2) (SCO6763 and SCO6764), *S. avermitilis* MA-4680 (SAV1650 and SAV1651), and *S. argenteolus* (a putative cyclase and *opds2* gene) [31]. In contrast, *ggpps1* constituted an operon together with a prenyl transferase
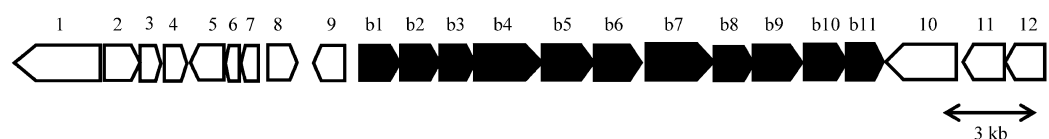
ORFs	Homologous proteins and their accession numbers
1	Hypothetical protein of *Streptomyces coelicolor* (44%), trQ9RCY2
2	Hypothetical protein of *Streptomyces avermitilis* (51%), trQ825Z0
3	Hypothetical protein of *Pseudomonas aeruginosa* (35%), trQ9I5G0
4	Hypothetical protein of *Nocardia farcinica* (35%), trQ5YX83
5	Hypothetical protein of *Nocardia farcinica* (34%), trQ5YV58
6	Putative transcriptional regulator of *Nocardia farcinica* (61%), trQ5YV57
7	RNA polymerase sigma-C factor of *Mycobacterium tuberculosis* (57% identity), spP66809
8	Hypothetical protein of *Mycobacterium tuberculosis* (35%), spQ10890
9	Dioxygenase (plaO1) for phenalinolactone biosynthesis of *Streptomyces* sp. Tü6071 (60%), trQ2I752
bra1	Putative phenylalanine aminotransferase of *Thermobifida fusca* (46% identity), spQ47KH1
bra2	Geranylgeranyl diphosphate synthase of *Methanopyrus kandleri* (36% identity), trQ8TXA1
bra3	Prenyltransferase (plaT3) for phenalinolactone biosynthesis of *Streptomyces* sp. Tü6071 (48%), trQ2I761
bra4	Diterpene cyclase (plaT2) for phenalinolactone biosynthesis of *Streptomyces* sp. Tü6071 (43%), trQ2I762
bra5	Putative epoxidase (plaT1) for phenalinolactone biosynthesis of *Streptomyces* sp. Tü6071 (47%), trQ2I763
bra6	P450 for Epothilone B hydroxylation of *Amycolatopsis orientalis* (43%), trQ156A0
bra7	Putative salicyl-AMP ligase of *Streptomyces* sp. WA46 (45%), trQ7X279
bra8	FK506 oxidase of *Streptomyces* sp. MA6548 (49%), trQ9ZGA5
bra9	Hypothetical protein of *Nocardia farcinica* (49%), tr|Q5YZT5
bra10	Rhamnosyltransferase of *Burkholderia cenocepacia* HI2424 (37%), trQ4LYP6
bra11	*O*-methyltransferase of *Solibacter usitatus* (46%), trQ445Z3
10	AfsR of *Streptomyces coelicolor* (44%), spP25941
11	LysR-family transcriptional regulator of *Streptomyces coelicolor* (52%), trQ9L030
12	Putative carboxyphosphonoenolpyruvate phosphonomutase of uncultured crenarchaeote (37%), trQ1ERB4

Fig. 2 Geranylgeranyl diphosphate synthase gene and its flanking regions in *Nocardia brasiliensis*.

ORFs deduced by sequence analysis are shown. Filled arrows and open arrows represent the putative BCA biosynthetic genes (*bra* genes) and the BCA-unrelated genes, respectively. sp, Swissplot; tr, UniProt (Universal Protein Resource).

homologue (*bra3*) and a putative isoprenoid cyclase (*bra4*) (Fig. 2), both of which possibly participate in the biosynthesis of BCA. Hence, we first examined whether *ggpps1* encodes a GGPPS by producing a recombinant enzyme.

A recombinant *ggpps1* product was expressed in *E. coli* as an *N*-terminal His-tagged fusion protein, and an *in vitro* assay was performed. Soluble protein extracts from *E. coli* harboring pQE-GGPPS and pQE30 (no insert) were analyzed by SDS-PAGE. The His-tagged *ggpps1* product was expressed at a high level. The enzyme was then purified (Fig. 3) and used in the enzyme assay. When the assay was carried out using DMAPP and [1-¹⁴C] IPP as substrates, geranylgeraniol (C20) was clearly detected as a major product by TLC analysis (Fig. 3), suggesting that the *ggpps1* gene encodes a GGPPS, as expected.

Cloning and Identification of the BCA Biosynthetic Gene Cluster

Considering that antibiotic biosynthesis genes cloned from actinomycetes are usually clustered in the genome, all genetic information for the biosynthesis of BCA was expected to reside in the flanking region of the *ggpps1*
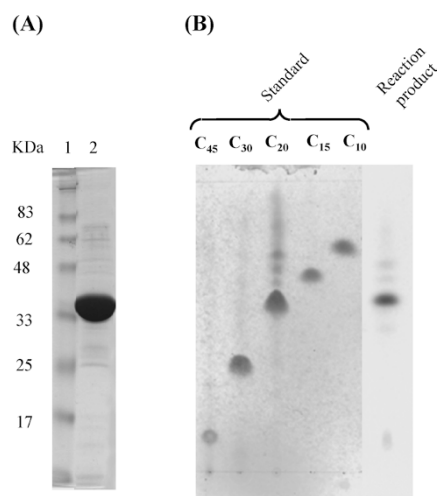

Fig. 3 Electrophoresis of the overproduced and purified *bra2* product (A) and TLC autoradiography of the alcohols obtained by hydrolysis of the products (B).

A, Molecular mass marker (lane 1) and purified His-tagged *bra2* product (lane 2) were analyzed on SDS-PAGE (12.5%). Proteins were stained with Coomassie brilliant blue R-250. B, The sample obtained by incubation of [1-¹⁴C]IPP and DMAPP with the recombinant enzyme was analyzed by TLC. Spots of authentic standard alcohols (C₁₀ to C₄₅) are also shown.
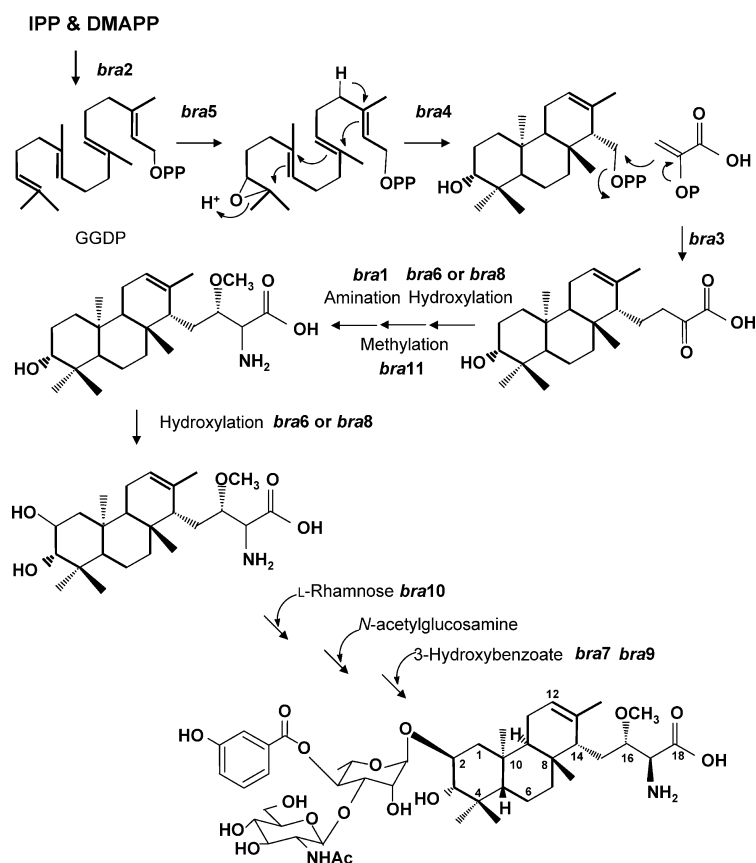

Fig. 4 Proposed biosynthetic pathway of BCA.

gene. Hence, the nucleotide sequences of both the region upstream of *ggpps1* (*bra2*) and of the region downstream of the putative isoprenoid cyclase genes (*bra4*) were determined. We identified 11 genes, all oriented in the same direction. To understand the functions of individual ORFs determined by DNA sequencing, we searched the public databases with FASTA and BLAST programs [32, 33]. The results are summarized in Fig. 2. In brief, each of the gene products had a significant similarity to one of: NikT-like amino-transferase responsible for nikkomycin biosynthsis (*bra1*), GGPPS (*bra2*, *ggpps1*), prenyltransferase (*bra3*), isoprenoid cyclase (*bra4*), FAD-epoxidase (*bra5*), P450 (*bra6*), 2,3-dihydroxybenzoate-AMP ligase (*bra7*), oxidase (*bra8*), hypothetical protein (*bra9*), dTDP-rhamnose transferase (*bra10*), methyl transferase (*bra11*). Among them, *bra3*, *bra4*, and *bra5* products had high amino acid identities to *plaT3*, *plaT2*, and *plaT1* products, respectively, all of which were recently cloned from *Streptomyces* sp. Tü6071 and confirmed to participate in the early step of biosynthesis of phenalinolactone, a diterpene compound [34]. Since the early step of the proposed biosynthetic pathway of phenalinolactone was thought to be identical of those of BCA (Fig. 4), high identities of amino acid sequences between the Bra3-5 and the PlaT3-1 are reasonable. This observation also suggested that the *bra* genes are involved in the BCA biosynthesis though we did not detect any genes related to L-rhamnose and *N*-acetylglucosamine biosyntheses in the flanking regions of the bra genes.

Disruption of Cyclase Gene (*bra4*)

We constructed a mutant, in which the cyclase gene (*bra4*) was replaced with thiostrepton resistance gene, to confirm that the gene cluster was responsible for BCA biosynthesis. A disruption plasmid, pCyc-DIS (Fig. 5A), was constructed and introduced into the BCA producer by transconjugation. Three thiostrepton-resistant colonies were randomly selected and their genomic DNAs were prepared, and used as templates for PCR analysis to study gene disruption. As shown in Fig. 5A, replacement of the *bra4* gene with the thiostrepton-resistance gene by a double crossover was clearly confirmed. The disruptants and the wild type strain
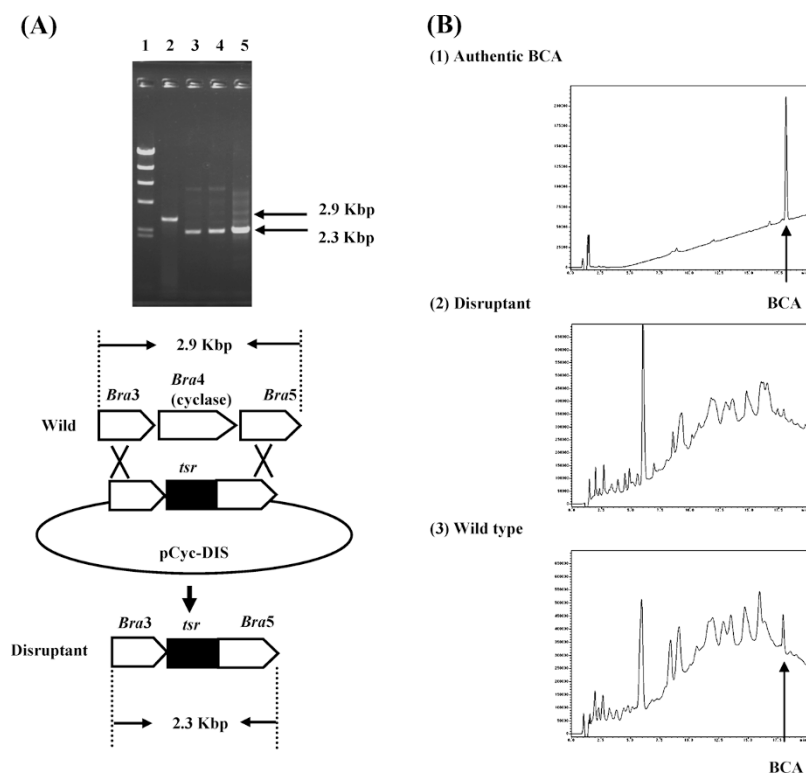

Fig. 5 PCR Analysis of the genomic DNA of the *bra4* (cyclase gene)-inactivated strains (A) and BCA productivity of the disruptants (B).

(A), Molecular marker (lane 1), PCR products from genomic DNA of wild-type strain (lane 2), and of disruptant (lanes 3~5). The amplified fragments are shown schematically. Arrows indicate the primers used in PCR analysis; they correspond to the *N*-terminal and *C*-terminal coding region of the *bra3* and *bra5* genes. Filled boxes indicate the thiostrepton resistant gene (0.9 Kb). (B), Authentic BCA (1), products of disruptant (2), and products of wild-type strain (3) were analyzed by HPLC.

were cultivated in the production medium, and their productivities of BCA were examined. None of the disruptants produced BCA, in contrast to the wild type strain, which produced significant amounts of BCA (Fig. 5B). This showed that the cluster, including the putative cyclase gene (*bra4*), was indeed involved in the biosynthesis of BCA. We next searched for an intermediate compound in the culture broth of the *bra4*-disrupted strain, which would be a substrate for Bra4 (cyclase). However, we did not detect a compound specifically accumulated in the culture broth of the disruptant.

Discussion

We found recently that most actinomycete strains are equipped with only the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway for the formation of IPP, a common precursor of isoprenoids. In addition to this pathway, we also found that a part of actinomycete strains possess both the mevalonate (MV) and MEP pathways [16]. In these strains, isoprenoid biosynthetic genes, such as terpentecin [12], BE-40644 [14, 15], viguiepinol [15] and furaquinocin A [15] were located just upstream and/or downstream of the MV pathway gene cluster. In this study, we did not exploit this observation to clone a BCA biosynthetic gene cluster because the BCA producer has only the MEP pathway and BCA was demonstrated to be synthesized *via* the MEP pathway [35]. Hence, we first cloned a gene encoding GGPPS, which supplies a direct precursor of diterpene compounds, and we then analyzed the flanking regions of the GGPPS gene. We expected that GGPPS genes and BCA biosynthetic genes would be clustered. We succeeded in cloning *bra1* to *bra11*, which might participate in BCA biosynthesis, though we did not find any genes related to L-rhamnose and *N*-acetylglucosamine biosyntheses in the flanking regions of the bra genes.

An NDP-glucose-4,6-dehydratase is known to participate in the biosynthesis of L-rhamnose [36]. Hence, we tried to clone an NDP-glucose-4,6-dehydratase gene from the BCA producer by PCR and one candidate gene was successfully cloned. In the flanking region of the gene, we also identified a putative dTDP-glucose synthase gene and a dTDP-4-keto-L-rhamnose 3,5-epimerase gene, both of which are involved in the biosynthesis of L-rhamnose [36] (accession number AB268547). By gene walking analysis, however, the putative L-rhamnose biosynthetic genes and the BCA biosynthetic genes were judged to be located more than 100 kb apart. Considering that L-rhamnose and *N*-acetylglucosamine are used for lipopolysaccharide

biosynthesis and peptidoglycan biosynthesis, respectively [36], it is reasonable to assume that the L-rhamnose biosynthetic gene cluster is located at a different locus relative to the BCA biosynthetic gene cluster. In fact, Gullon *et al*. have cloned a biosynthetic gene cluster of steffimycin, an anthracycline compound with an L-rhamnose moiety, from *Streptomyces steffisburgensis* [37]. In this case, they also detected no genes related to L-rhamnose biosynthesis in the flanking regions of the steffimycin biosynthetic gene cluster.

Of the eleven *bra* genes, we were able to estimate the function of seven genes (*bra1*, *2*, *3*, *4*, *5*, *10*, and *11*) by homology searches and by heterologous expression experiments. Although we have not postulated the detailed functions of the other genes, probable functions were estimated by intensive homology searches. The *bra9* gene product showed 34% identity to PapA5 of *Mycobacterium tuberculosis*, which was recently identified to be a new type of acyltransfrase that catalyzes a condensation of acyl-CoA thioesters with medium chain alcohols to give ester formation [38]. The *bra7* gene product was similar to enzymes, which catalyze adenylation of specific aromatic compound, such as 3-hydroxypicolinic acid (*snbA*) [39], salicylate (*sdgA*) [40], and 2,3-dihydroxybenzoic acid (*mxcE*) [41]. This activates the aromatic compounds to enable condensation with other molecules. Among them, *snbA* was predicted to catalyze the transfer of 3-hydroxypicolinic acid to the thiol acceptor of CoA *via* an adenylation step [39]. Taken together, these facts suggest that the *bra7* gene product might convert 3-hydroxybenzoate into 3-hydroxybenzoyl-CoA *via* adenylation, followed by attachment to L-rhamnose by the *bra9* gene product. *bra6* and *bra8* had significant similarity to a P450 and to a pteridine-dependent oxidase, respectively. Considering the structure of BCA, two positions, the 2 and 16 positions, might be hydroxylated by these enzymes, although we did not investigate the exact position specificity. In conclusion, the proposed biosynthetic pathway is summarized in Fig. 4.

We have been interested in diterpene cyclases found in actinomycetes. To date, we have cloned three diterpene cyclase genes, terpentedienol diphosphate synthase (TDPS) gene [12, 13], tuberculosinol diphosphate synthase (TBPS) gene [42], and *ent*-copalyl diphosphate synthase (CDPS) gene [14]. All these genes possess a DXDD motif that mediates substrate binding, *via* chelation of a divalent metal ion, and catalyze cyclization reactions of GGPP by protonation at the terminal-double bond of the substrate. Moreover, we have succeeded in expression of these recombinant enzymes and characterized their properties. Hence, the *bra4* gene product, which has significant

```
Bra     ---MTTRTIDTALETGAEAVLRHRAADGLFDYDSGVGATLATAGATLVLHLHDPVGSRAYVTGGVAHLSRTQNGD
phena   MCSDVDSRVDSSVARAAQALFAIQRPEGSWPNRRPT-AVLGTAGAVVALHLADPERSRDLIERGAQWLVAAQNAD
eCDP    ----MNVTSFAALRAAAQDIVDEMIAD---PYGLTSPSVYETARMVVSAPWLEGHRQR------VEFLLAQQHED
TDP     -------------------MKDRAAD---PVTKFSPSPYETGQFLRISERADVGTPQ------IDYLLATQRPD
3377c   ------------METFRTLLAKAALG-----NGISSTAYDTAWVAKLGQLDDELSDLA-----LNWLCERQLPD

Bra     GGWGSVPGQPTEAVTTAIAAAVLHLVDADGSGDRVRAGRAALDRLGGLSGIADPALSAVCHRFHAMAG-WAPPAP
phena   GGWGGVPRAASQLVPTVVTASALTLTAPRTAGEPARRALDLIRSLGGVESLADPGMVHMATTFLALAGLRSLHGS
eCDP    GTWGGP--AAYGLLPTLSAVDALLSVAGTQDAR------------RVAGAVESGLAALAGRFPRNVELPDTIAV
TDP     GLWGSVG---FELVPTLGAVAGLSSRPEYADRAG----------VTDAVARACEKLWELALGEGGLPKLPDTVAS
3377c   GSWGAEFPFCYEDRLLSTLAAMISLTSNKHRRRR-------------AAQVEKGLLALKNLTSGAFEGPQLDIK

Bra     VRVPLWVFALPR--LRRTRMSFRSGVLAATALAQARAEPGGLLRRLIDHAATPAALRVLYGIHRDEGGTGAFG--
phena   RRIPLELLLLPRR-LWRPRLSFRAAPFVAMAFIQAHDAAPKGVGGLLHRLARPTALRLLQEMERGENDRGGYG--
eCDP    ELLVPWLIEQVD--QRLSRMDDRGDLPGRLDLQADTGTLSGIRELLRQNTGIP---------EKTWHSLEALG--
TDP     EIIVPSLIDLLSEVLQRHRPAVGGKAGQEQEFPSPPGANAELWRQLSDRIARGQAIP-----KTAWHTLEAFHPL
3377c   DATVGFELIAPTLMAEAARLGLAICHEESILGELVGVREQKLRKLGGSKINKHITAA----FSVELAGQDGVG--

Bra     ADAWAAGMTGLGLVRSGAAPDLTEAIADFLRRSVFADGSWTIVE-IRLTYTGFAATGLCDAGYGADPR-----LA
phena   GDNWLAAVVCIGLCRTGAPAHMIAATVDYLRSNVRPDGSWHIMQGLDLIGGSYVARGLADAGYRDDPR-----LA
eCDP    APAVRSGTVTPMGGAVGASPAATSAWLGDPPHTDAAKACLAYLHQTQARHGGPVS-GITSISYFELAW-----VV
TDP     PKQFAATVTPAADGAVTCSPSSTAAWLSAVG-TDAGASTRAYLDEAQSRYGGAIP-MGSSMPYFEVLW-----VL
3377c   --MLDVDNLQETNGSVKYSPSASAYFALHVK--PGDKRALAYISSIIQAGDGGAP-AFYQAEIFEIVWSLWNLSR

Bra     VTADGVRAAQLTRPFTMLDCPPGGWSYSGPDGWP---VTL**ESAE**LLSALAKMPGADRDPHVARGARWLAGRQDTR
phena   RARQWLRGCQQDQAFPLYDAPPGGWGWEGPHGWP---NFL**DSAN**VLAALTPASREEPAEHLRTGLDWLVSRQDGR
eCDP    TALSGS-GLDVDIPAQVPDILRTALGANGLSAGPG--LPA**DSDD**TSAALHALDLLGKPESVDCLWEYDTG-----
TDP     NLVLKY-FPDVPIPREIIEEIAAGFSDSGIGGGPG--LPP**DGDD**TAYANLAGDKLGAPTHPEILMKFWAED----
3377c   TDIDLSDPEIVRTYLPYLDHVEQHWVRGRGVGWTGNSTLE**DCDT**TSVAYDVLSKFGRSPDIGAVLQFEDAD----

Bra     GSWSLWVRDTALANDGPCPYLTAQAVDALLDAGIPPEDNGIRRAARWLAGVQRPD-GSFESLWYRGHTAGTAQVI
phena   GSWGTFVPDTTLANDGPCPYTTAQCVESLLDGGLSRGDPRIAKALDWLLAAQRPD-GTYEALWYRGLTAATATAL
eCDP    LYFTCFPKERTPSTSTNAHILVALADR--RGQGDTRYDHAAERVGGWLVEQQQPD-GRWMDKWHASPYYATACGA
TDP     -HFVSYPGEQTPSETVNAHALEYLNHLR-MRRGITEFGAVEDACAEWVISQQTED-GCWYDKWNVSPYYSTAACV
3377c   -WFRTYFHEVGPSISTNVHVLGALKQA-----GYDKCHPRVRKVLEFIRSSKEPGRFCWRDKWHRSAYYTTAHLI

Bra     GALARIT--DPATDPTVRGGVRWLLATQHSDGSWGPGDG-------RAGSVEETAWALRALLEAGIAPG-TAEID
phena   VALARG---GVGDHPAARRAREALLAGQLADGSWGPGDTGTPGDDPSRGTVEETAWALRGLLAAGLPAQ-DDRLR
eCDP    AAMARLD--GPRTSAALDDAIRWVLDTQHADGSWGRWEG----------TGEETAYALQVLNHRAAPDRPALEAI
TDP     EALLDARKQDEPQLDSLRRAREWLLRHQTDSGGWGMAEP----------SPEETAYAVLALDLFASRGGEGAEEC
3377c   CAASNYD------DALCSDAIGWILNTQRPDGSWGFFDG--------QATAEETAYCIQALAHWQRHSG----TS

Bra     RAADWLLARQQPDGHWPEAPVSAYIRRCVHYPNGAITAGLALRALAAYRRGLGR----KDDDG
phena   RAAEWIVSAQGDDGLWDANPVCLHIRRLAYYADGLIGNGLVLRALGAYRSALAAGPVARREAS
eCDP    RAGRAFLSGHVEDDRRNPP---LWHDKDLYTPVRVIRAEILGTLAATQRLAEAE----KEARA
TDP     AAAISRAKEFFTDESRENPP--LWMGKDLYTPFRIVDVTVMCGRAVVGRY------------
3377c   LSAQISRAGGWLSQHCEPPYAPLWIAKTLYCSATVVKAAILSALRLVDESNQ----------
```

Fig. 6 Alignment of eubacterial diterpene cyclases.

Multiple alignment of the *bra4* gene product (Bra) with PlaT2 (phena), *ent*-copalyl diphosphate synthase (eCDP), terpentedienol diphosphate synthase (TDP), and tuberculosinol diphosphate synthase (Rv3377c) is shown. Aspartate rich regions that are known to be responsible for substrate binding are underlined.

similarity to diterpene cyclases, was also expressed as a recombinant enzyme and used in an *in vitro* assay, using GGPP as a substrate. However, we detected no reaction products under various conditions. Very recently, a biosynthetic gene cluster of phenalinolactone, a diterpene compound, was cloned from *Streptomyces* sp. Tü6071 by searching for an NDP-glucose-4,6-dehydratase gene that participates in the formation of L-amicetose attached to an aglycon of phenalinolactone [34]. In the cluster, the *plaT2* gene has significant similarity to isoprenoid cyclases and is annotated to be a diterpene cyclase responsible for the biosynthesis of phenalinolactone. They suggested that the

plaT2 product uses epoxy-GGPP as a substrate based on the following observations: The *plaT1* gene, which is located just downstream of the *plaT2* gene, had significant similarity to eukaryotic squalene epoxidase genes and the *plaT2* product lacked a DXDD motif, in a similar manner to those of oxidosqualene cyclases that use epoxy-squalene as a substrate. To predict whether the *bra4* gene product would also function in a similar manner, a multiple alignment of TDPS, TBPS, CDPS, PlaT2 and Bra4 was constructed using the ClustalW program [43]. As shown in Fig. 6, both PlaT2 and Bra4 had no typical DXDD motif, in contrast to TDPS, TBPS, and CDPS. Moreover, the

bra5 product had significant identity to PlaT1. Taking these results together, Bra4 might use epoxy-GGPP as a substrate. To examine this possibility, an *in vitro* experiment is essential. This study is now in progress and will be reported in the near future.

References

1. Connolly JD, Hill RA. Dictionary of terpenoids, Chapman and Hall, NY (1992)

2. Dewick PM. The biosynthesis of C5~C25 terpenoid compounds. Nat Prod Rep 19: 181–222 (2002)

3. Wise ML, Croteau R. Comprehensive natural products chemistry no. 2, isoprenoids, including carotenoids and steroids. pp. 97–153, Elsevier Science, Oxford (1999)

4. Cane DE. Comprehensive natural products chemistry no. 2, isoprenoids, including carotenoids and steroids. pp. 155–200, Elsevier Science, Oxford (1999)

5. MacMillan J, Beale MH. Comprehensive natural products chemistry no. 2, isoprenoids, including carotenoids and steroids. pp. 217–243, Elsevier Science, Oxford (1999)

6. Bohlmann J, Meyer-Gauen G, Croteau R. Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc Natl Acad Sci USA 95: 4126–4133 (1998)

7. Trapps S, Croteau R. Defensive resin biosynthesis in conifers. Annu Rev Plant Physiol Plant Mol Biol 52: 689–724 (2001)

8. Keeling CI, Bohlmann J. Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defense of conifers against insects and pathogens. New Phytol 170: 657–675 (2006)

9. Cane DE, Sohng JK, Lamberson CR, Rudnicki SM, Wu Z, Lloyd MD, Oliver JS, Hubbard BR. Pentalenene synthase: purification, molecular cloning, sequencing, and high-level expression in *Escherichia coli* of a terpenoid cyclase from *Streptomyces* UC5319. Biochemistry 33: 5846–5857 (1994)

10. Lin X, Hopson R, Cane DE. Genome mining in *Streptomyces coelicolor*: molecular cloning and characterization of a new sesquiterpene synthase. J Am Chem Soc 128: 6022–6023 (2006)

11. Jiang J, He X, Cane DE. Geosmin biosynthesis. *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. J Am Chem Soc 128: 8128–8129 (2006)

12. Dairi T, Hamano Y, Kuzuyama T, Itoh N, Furihata K, Seto H. Eubacterial diterpene cyclase genes essential for production of the isoprenoid antibiotic terpentecin. J Bacteriol 183: 6085–6094 (2001)

13. Hamano Y, Kuzuyama T, Itoh N, Furihata K, Seto H, Dairi T. Functional analysis of eubacterial diterpene cyclases responsible for biosynthesis of a diterpene antibiotic, terpentecin. J Biol Chem 277: 37098–37104 (2002)

14. Kawasaki T, Kuzuyama T, Kuwamori Y, Matsuura N, Itoh N, Furihata K, Seto H, Dairi T. Presence of copalyl diphosphate synthase gene in an actinomycete possessing the mevalonate pathway. J Antibiot 57: 739–747 (2004)

15. Kawasaki T, Hayashi Y, Kuzuyama T, Furihata K, Itoh N, Seto H, Dairi T. Biosynthesis of a natural polyketide-isoprenoid hybrid compound, furaquinocin A: identification and heterologous expression of the gene cluster. J Bacteriol 188: 1236–1244 (2006)

16. Dairi T. Studies on biosynthetic genes and enzymes of isoprenoids produced by actinomycetes. J Antibiot 58: 227–243 (2005)

17. Komaki H, Nemoto A, Tanaka Y, Takagi H, Yazawa K, Mikami Y, Shigemori H, Kobayashi J, Ando A, Nagata Y. Brasilicardin A, a new terpenoid antibiotic from pathogenic *Nocardia brasiliensis*: fermentation, isolation and biological activity. J Antibiot 52: 13–19 (1999)

18. Komaki H, Tanaka Y, Yazawa K, Takagi H, Ando A, Nagata Y, Mikami Y. Antitumor activity of brasilicardin A, a novel terpenoid antibiotic from *Nocardia brasiliensis*. J Antibiot 53: 75–77 (2000)

19. Shigemori H, Komaki H, Yazawa K, Mikami Y, Nemoto A, Tanaka Y, Sasaki T, In Y, Ishida T, Kobayashi J. Brasilicardin A. A novel tricyclic metabolite with potent immunosuppressive activity from actinomycete *Nocardia brasiliensis*. J Org Chem 163: 6900–6904 (1998)

20. Komatsu K, Tsuda M, Tanaka Y, Mikami Y, Kobayashi, J. SAR studies of brasilicardin A for immunosuppressive and cytotoxic activities. Bioorg Med Chem 13: 1507–1513 (2005)

21. Maniatis T, Fritsch EF, Sambrook J. Molecular cloning. A

laboratory manual, Cold Spring Harbor Laboratory, NY (1982)

22. Hamano Y, Dairi T, Yamamoto M, Kawasaki T, Kaneda K, Kuzuyama T, Itoh N, Seto H. Cloning of a gene cluster encoding enzymes responsible for the mevalonate pathway from a terpenoid-antibiotic-producing *Streptomyces* strain. Biosci Biotechol Biochem 65: 1627–1635 (2001)

23. Kawasaki T, Kuzuyama T, Furihata K, Itoh N, Seto H, Dairi T. A relationship between the mevalonate pathway and isoprenoid production in actinomycetes. J Antibiot 56: 957–966 (2003)

24. Dairi T, Ohta T, Hashimoto E, Hasegawa M. Organization and nature of fortimicin A (astromicin) biosynthetic genes studied using a cosmid library of *Micromonospora olivasterospora* DNA. Mol Gen Genet 236: 39–48 (1992)

25. Kawasaki T, Hamano Y, Kuzuyama T, Itoh N, Seto H, Dairi T. Interconversion of the product specificity of type I eubacterial farnesyl diphosphate synthase and geranylgeranyl diphosphate synthase through one amino acid substitution. J Biochem 133: 83–91 (2003)

26. Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton CJ, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schremp H. Gene Manipulation of *Streptomyces*. A laboratory manual, The John Innes Foundation, Norwich, UK (1985)

27. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Puhler A. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145: 69–73 (1994)

28. Onaka H, Taniguchi S, Igarashi Y, Furumai T. Characterization of the biosynthetic gene cluster of rebeccamycin from *Lechevalieria aerocolonigenes* ATCC 39243. Biosci Biotechol Biochem 67: 127–138 (2003)

29. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417: 141–147 (2002)

30. Omura S, Ikeda H, Ishikawa J, Hanamoto A, Takahashi C, Shinose M, Takahashi Y, Horikawa H, Nakazawa H, Osonoe T, Kikuchi H, Shiba T, Sakaki Y, Hattori M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. Proc Natl Acad Sci USA 98: 12215–12220 (2001)

31. Hayashi Y, Onaka H, Itoh N, Seto H, Dairi T. Cloning of the gene cluster responsible for biosynthesis of KS-505a (longestin), a unique tetraterpenoid. Biosci Biotechol Biochem 71: 3072–3081 (2007)

32. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 215: 403–410 (1990)

33. Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444–2448 (1988)

34. Durr C, Schnell HJ, Luzhetskyy A, Murillo R, Weber M, Welzel K, Vente A, Bechthold A. Biosynthesis of the terpene phenalinolactone in *Streptomyces* sp. Tü6071: Analysis of the gene cluster and generation of derivatives. Chem Biol 13: 365–377 (2006)

35. Shigemori H, Komaki H, Yazawa K, Mikami Y, Nemoto A, Tanaka Y, Kobayashi J. Biosynthesis of diterpene moiety of brasilicardin A *via* non-mevalonate pathway in *Nocardia brasiliensis*. Tetraherdon Lett 40: 4353–4354 (1999)

36. Trefzer A, Salas JA, Bechthold A, Genes and enzymes involved in deoxysugar biosynthesis in bacteria. Nat Prod Rep 16: 283–399 (1999)

37. Gullon S, Olano C, Abdelfattah MS, Brana AF, Rohr J, Mendez C, Salas JA. Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin. Appl Environ Microbiol 72: 4172–4183 (2006)

38. Onwueme KC, Ferreras JA, Buglino J, Lima CD, Quadri LEN. Mycobacterial polyketide-associated proteins are acyltransferases: Proof of principle with *Mycobacterium tuberculosis* PapA5. Proc. Natl. Acad. Sci. USA 101: 4608–4613 (2004)

39. Crecy-Lagard V, Blanc V, Gil P, Naudin L, Lorenzon S, Famechon A, Bamas-Jacques M, Crouzet J, Thibaut D. Pristinamycin I biosynthesis in *Streptomyces pristinaespiralis*: molecular characterization of the first two structural peptide synthetase genes. J Bacteriol 179: 705–713 (1997)

40. Ishiyama D, Vujaklija D, Davies J. Novel pathway of salicylate degradation by *Streptomyces* sp. Strain WA46.

Appl Environ Microbiol 70: 1297–1306 (2004)

41. Silakowski B, Kunze B, Nordsiek G., Blöcker H, Höfle G, Muller R. The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sg a15. Eur J Biochem 267: 6476–6485 (2000)

42. Nakano C, Okamura T, Sato T, Dairi T, Hoshino T. *Mycobacterium tuberculosis* H37Rv 3377c encodes the diterpene cyclase for producing halimane skeleton. Chem Commun Issue 8: 1016–1018 (2005)

43. Thompson JD, Higgins DG, Gibson TJ. CLUSTALW. Nucleic Acids Res 22: 4673–4680 (1994)