J Ind Microbiol Biotechnol (2014) 41:233–250
DOI 10.1007/s10295-013-1327-x

REVIEW

SIMB
Society for Industrial Microbiology and Biotechnology

Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters

Haruo Ikeda · Kazuo Shin-ya · Satoshi Omura

Received: 28 June 2013 / Accepted: 4 August 2013 / Published online: 29 August 2013

Abstract To date, several actinomycete genomes have been completed and annotated. Among them, *Streptomyces* microorganisms are of major pharmaceutical interest because they are a rich source of numerous secondary metabolites. *S. avermitilis* is an industrial microorganism used for the production of an anthelmintic agent, avermectin, which is a commercially important antiparasitic agent in human and veterinary medicine, and agricultural pesticides. Genome analysis of *S. avermitilis* provides significant information for not only industrial applications but also understanding the features of this genus. On genome mining of *S. avermitilis*, the microorganism has been found to harbor at least 38 secondary metabolic gene clusters and 46 insertion sequence (IS)-like sequences on the genome, which have not been searched so far. A significant use of the genome data of *Streptomyces* microorganisms is the construction of a versatile host for heterologous expression of exogenous biosynthetic gene clusters by genetic engineering. Since *S. avermitilis* is used as an industrial microorganism, the microorganism is already optimized for the efficient supply of primary metabolic precursors and biochemical energy to support multistep biosynthesis. The feasibility of large-deletion mutants of *S. avermitilis* has been confirmed by heterologous expression of more than 20 exogenous biosynthetic gene clusters.

Keywords Genome mining · *Streptomyces avermitilis* · IS · Heterologous expression · Secondary metabolite · Versatile host

Dedicated to Prof. Sir David A. Hopwood on the occasion of his 80th birthday and in recognition of his exceptional contributions to *Streptomyces* genetics and molecular biology, and the honor of his friendship.

Special issue: Genome Mining for Natural Products Discovery.

Electronic supplementary material The online version of this article (doi:10.1007/s10295-013-1327-x) contains supplementary material, which is available to authorized users.

H. Ikeda (✉)
Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 252-0373, Japan
e-mail: ikeda@ls.kitasato-u.ac.jp

K. Shin-ya
National Institute of Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

S. Omura
Kitasato Institute for Life Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan

Introduction

Streptomyces avermitilis is a Gram-positive bacterium in the genus *Streptomyces*, which belongs to the family *Streptomycetaceae* and class Actinobacteria. *Streptomyces* microorganisms belonging to the family *Streptomycetaceae* are unique among soil bacteria because they form filamentous mycelia, aerial hyphae, and conidial spores during their life cycle. This process, unique among Gram-positive bacteria, requires the specialized coordination of metabolism and is more complex than in other Gram-positive bacteria. The most interesting property of this genus is its ability to produce secondary metabolites, including antibiotics and bioactive compounds of value in human and veterinary medicine, agriculture, and as unique biochemical tools. Structural diversity is observed in these secondary metabolites that encompass not only antibacterial, antifungal, antiviral, and antitumor compounds, but also

metabolites with immunosuppressant and antihypertensive properties. The genus *Streptomyces* is a rich source of secondary metabolites in which common intermediates in the living cell are condensed into more complex structures by defined biochemical pathways. *Streptomyces avermitilis* (MA-4680 = ATCC 31267, NRRL 8165, NCBIM 12804, JCM 5070) was isolated from a soil sample collected in Shizuoka Prefecture, Japan in 1978 by one of the screening programs for the discovery of antiparasitic metabolites [4]. The microorganism produces a potent anthelmintic agent, avermectins (**1**), which are a series of 16-membered macrocyclic lactones. They are pentacyclic polyketide-derived compounds linked to a disaccharide of the methylated deoxysugar. Related anthelmintic polyketide-derived compounds, milbemycins [43] and nemadectins [8], have been discovered; however, avermectins retain the only glycosylated compounds to date. Their semisynthetic derivatives, particularly ivermectins, have been used as antiparasitic agents since 1981 and as agricultural pesticides since 1985. Ivermectin has been used for livestock farming and in health care of companion animals. The efficacy of ivermectin in human onchocerciasis has made it a promising candidate for the control of one of the most insidious and intractable of all tropical diseases. Furthermore, ivermectin has been found to be effective against the human disease strongyloidiasis and as a miticidal agent. Thus, *S. avermitilis* is the most important industrial microorganism for the production of anthelmintic agents and antiparasites.

At the beginning of this century, the genome project of *S. avermitilis* was started in our laboratory. We subsequently reported the sequence analysis of this microorganism, covering 99 % of its genome, in 2001 [35] and the results allowed a first genome-level glimpse into the extensive secondary metabolic dexterity of microorganisms of the genus *Streptomyces*. The genome analysis was completed in 2003 [17]. To date, several *Streptomyces* genomes have been completed and are publicly available [1–3, 32, 34, 37, 47]. These studies revealed that *Streptomyces* have large (ca. 6–10 Mb), linear chromosomes containing well over 20 biosynthetic gene clusters for secondary metabolites, which encode the biosynthesis of polyketides by polyketide synthases (PKSs), peptides by non-ribosomal peptide synthetases (NRPSs) or ribosomal synthesis, bacteriocins, terpenoids, shikimate-derived metabolites, aminoglycosides, and other natural products [33]. The accumulation of information about genes involved in the biosynthesis of secondary metabolites allows the identification of biosynthetic genes by genome mining in other draft genome sequence data of other *Streptomycetaceae* microorganisms.

The identification and characterization of biosynthetic gene clusters have proved to be invaluable tools for the elucidation of the biosynthesis of secondary metabolites as well as a potentially rich source of information on cryptic metabolites encoded by silent biosynthetic pathways. In some *Streptomyces* microorganisms carrying interesting gene clusters of secondary metabolites, systematic analysis of interesting metabolic pathways cannot be applied because the productivity of the metabolites is too low, introduction of DNA for the characterization of each gene involving biosynthesis cannot be applied, or because of the genetic instability of the producing microorganism. There has been considerable recent interest in the development of engineered bacterial strains for the efficient heterologous production of secondary metabolites [12, 31, 36]. *S. avermitilis* has already proven to be a highly efficient producer of secondary metabolites. Because this strain is already optimized for the efficient supply of primary metabolic precursors and biochemical energy to support multistep biosynthesis, it is therefore an attractive host for the heterologous production of secondary metabolites.

In this review we describe genome mining of *S. avermitilis* and the construction of a versatile host for the efficient production of natural products by controlled minimization of the genome of *S. avermitilis*.

Future of *S. avermitilis* genome

The chromosome of *S. avermitilis* contains 9,025,608 bases with 70.7 % average GC content and encodes at least 7,574 protein-coding genes. The microorganism harbors two linear plasmids, SAP1 and SAP2, one of which, SAP2 (ca. 200 kb), was easily eliminated from the microorganism during cultivation. The other, SAP1, was extremely stably maintained during cultivation and its genome contains 94,287 bases (69.2 % average GC content) with 96 potential protein-coding genes [17]. The replication origin (*oriC*) of the linear chromosomes of other *Streptomyces* microorganisms, *S. coelicolor* A3(2), *S. scabie* 87.22, and *S. griseus* IFO 13350, is located in the middle of the chromosomes. In contrast, *oriC* on the *S. avermitilis* chromosome is shifted about 800 kb away from the center. Comparative analysis of *S. avermitilis* and other *Streptomyces* chromosomes revealed a 6–6.5-Mb, highly conserved internal region where most essential genes are located, with a similar order and direction among other *Streptomyces* microorganisms [17, 34]. This region also shows structural similarity to the other actinomycete chromosomes [2]. This finding implies that the 6–6.5-Mb internal region is the underlying backbone of *Streptomyces* chromosomes. On the other hand, there are variable and less conserved subtelomeric regions near both ends of the chromosome (telomeres). About half (or more) of the genes related to secondary metabolism were found in the

subtelomeric regions, where no known essential gene was found, in *Streptomyces* chromosomes [17, 34]. The length of terminal inverted repeats (TIRs; telomeres) at both ends of the chromosome varies from 49 to 132,910 bp in *Streptomyces* species. The TIRs of *S. avermitilis* are the shortest of the genome-sequenced *Streptomycetaceae* microorganisms.

In addition, the subtelomeric regions in *S. avermitilis* contained most of the mobile elements in the genome. No reports have described the distribution of mobile elements on the genomes of *Streptomyces* microorganisms in detail. Only a functional insertion sequence (IS), IS*493* [40], and transposon, Tn*4556* [10], have been reported. From the genome analysis of *S. avermitilis*, 111 genes encoding transposase are found in the genome. Among them, 87 and 13 genes are located in the left and right subtelomeric regions, respectively (Fig. 1). Since the remaining 11 genes are located in the 6.5-Mb internal region, 78.4 % of putative transposase genes lie in both subtelomeric regions. This finding suggests that frequent horizontal transfer may have preferentially occurred in the subtelomeric regions of this microorganism. The bacterial IS consists of genes encoding transposase, perfect or imperfect inverted repeats (IRs) flanking the transposase genes, and direct repeated target duplication short sequences outside of both IRs. Forty-six IS-like structures are found in the *S. avermitilis* genome and 67.4 % of IS-like sequences are located in left (31 ISs) and right (9 ISs) subtelomeric regions, respectively, and 6 ISs are in the internal region (Fig. 1). They contain 10–37-bp perfect or imperfect IRs and 2–9-bp target duplication sequences flanking both IRs in some IS-like structures (Fig. 2 and Table S1). Although we have not confirmed their in vivo function (transposition from the replicon to replicon), some IS-like sequences appear to be fully functional as IS elements from their structures. Target duplication sequences containing –CTAG– are frequently found in some IS-like sequences (Table S1). The IS recognizes the target sequence in the replicon, and inserts and duplicates the target sequence flanking IRs. Since this target duplication sequences contain the translational stop codon, TAG, in the 3' region, the protein-coding gene containing the target sequence for IS will not be disrupted by the insertion of IS.

Almost all of the IS-like sequences are found as single copies in the chromosome, but six types of IS-like sequence exist as multiple copies in the chromosome (Fig. 2). Two copies of IS*Sav1* (*A* and *B*), which resemble the IS*4* family and contain a gene encoding IS*Fsp6*-like transposase, are located in the left subtelomeric region. These IS-like sequences contain relatively long inverted repeats (36/37 bp matches) and the predicted target sequence is 9 bp. Three IS*Sav4* (*A*, *B*, and *C*), which are the IS*701* family and contain a gene encoding IS*Azvi8*-like transposase, consist of 16-bp perfect inverted repeats and all copies are located in the left subtelomeric region. Three IS*Sav5* (*A, B*, and *C*) are located in both left (one copy) and right (two copies) subtelomeric regions. These predicted IS-like sequences consist of relatively long perfect inverted repeats (30/30 bp) and an 8-bp target sequence flanking both IRs. Each IS-like sequence contains a gene encoding IS*L3* family IS*Fsp1*-like transposase. Thus, these three types of IS-like sequence are identical. The remaining three types of IS-like sequence (IS*Sav6, 7*, and *8*) are similar to those of each group. One of the IS-like sequences, IS*Sav4C*, inserts into a gene encoding nucleotide-binding protein, and the protein-coding gene was truncated and inactivated. A gene encoding nucleotide-binding protein is found in an operon for volatile monoterpenoid alcohol, 2-methylisoborneol (**47**), synthesis in which three genes encoding nucleotide-binding protein, 2-methylisoborneol synthase, and geranyl diphosphate methyltransferase, respectively, are located. Four protein-coding genes, including putative transposase genes, lie downstream of IS*Sav4C*, and a truncated gene encoding 2-methylisoborneol synthase is located downstream of the four genes. This truncation is induced by a deletion mutation, and the deduced polypeptide of truncated gene product lacked the second conserved metal-binding motif in ordinary terpene

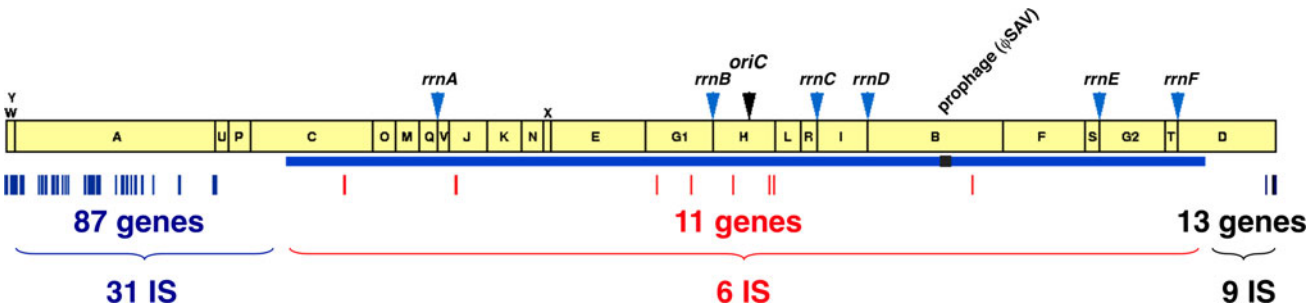


Fig. 1 *Ase*I physical map of *S. avermitilis* chromosome and distribution of genes encoding transposase and IS-like sequences. *rrnA–rrnF* and *oriC* indicate rRNA operon (5S, 16S, and 23S rRNA) and replication origin, respectively. *Thick blue bar* is the core region of *Streptomyces*. *Vertical bars* are putative transposase genes

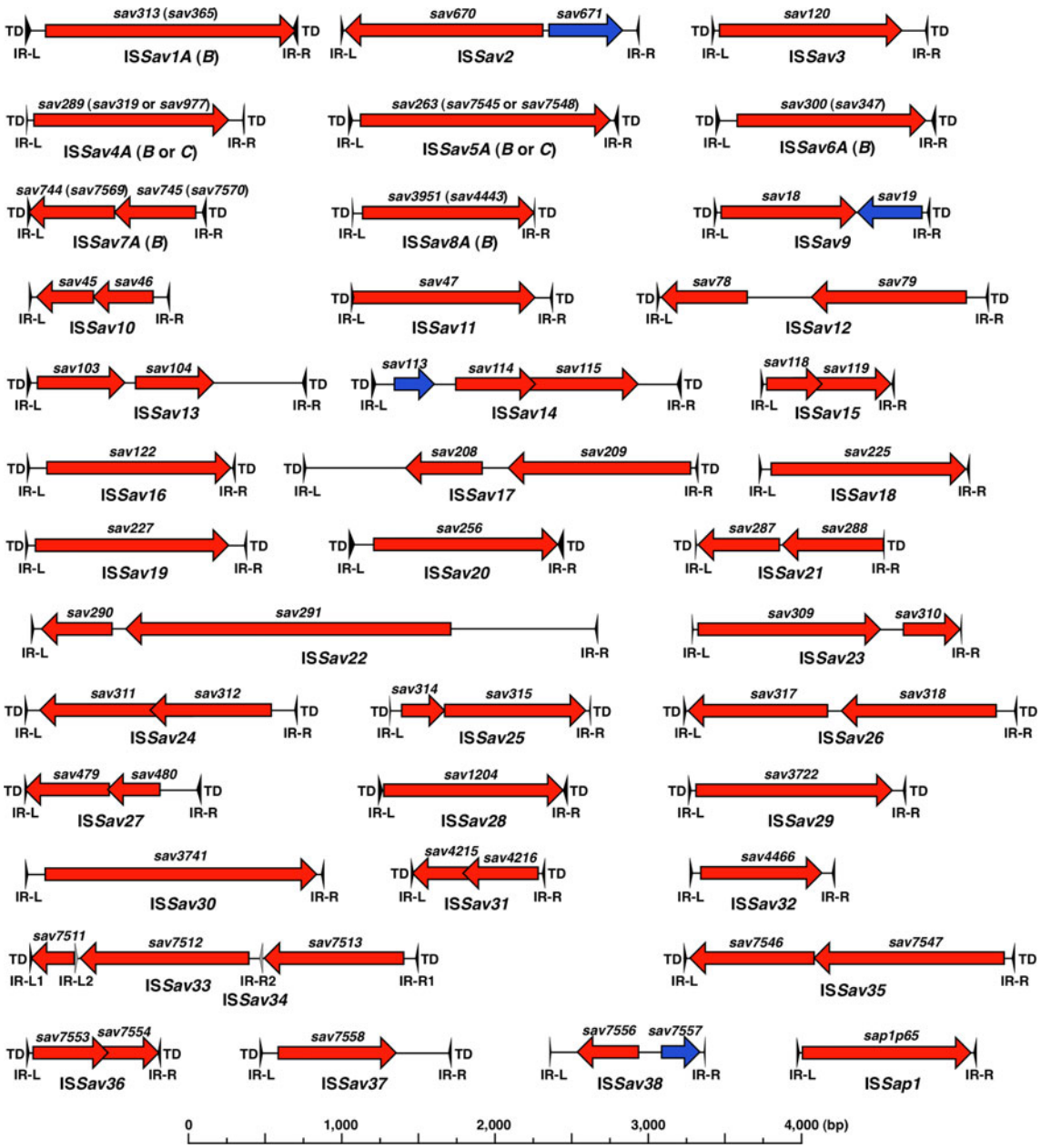


Fig. 2 Structures of IS-like sequences in *S. avermitilis* genome. *Red arrows* indicate genes encoding transposase and *blue arrows* are genes encoding hypothetical protein. *Filled triangles* indicate inverted repeat (IR) sequences. *IR-L* left side of IR, *IR-R* right site of IR, *TD* target duplication sequence. The information about each IS-like sequence is summarized in Table S1 and http://avermitilis.ls.kitasato-u.ac.jp

synthases. Furthermore, an intact gene (*sav983*) encoding geranyl diphosphate 2-methyltransferase is located downstream of the truncated 2-methylisoborneol synthase gene. Thus, *S. avermitilis* is capable of producing volatile monoterpenoid alcohol, 2-methylisoborneol, but deletion of the region encoding the second conserved metal-binding motif in the monoterpene synthase gene prevented the production of 2-methylisoborneol in this microorganism [22].

Genome mining of *S. avermitilis*

Complete sequencing and annotation of the *S. avermitilis* genome allowed for its rational mining to identify many of the orphan pathway products. Prior to the availability of the genome sequence of *S. avermitilis*, the structure and biosynthesis of avermectin [16] and that of another polyketide macrocyclic lactone, oligomycins (**2**), had already been established. A further 11 polyketide synthase (PKS) gene

Table 1 Biosynthetic gene clusters for secondary metabolites in *S. avermitilis*

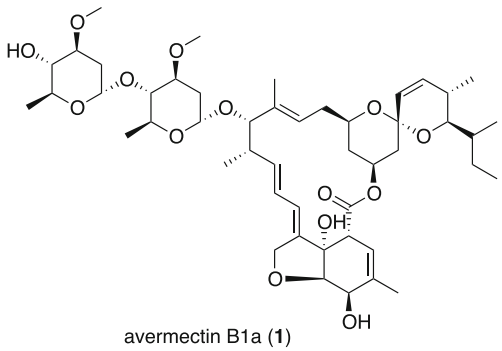
No.	Genes	Location (nt)	Actual or predicted product[a]
1	*sav76 (ams)*	86,073–87,080	**Avermitilol** (**15**), **avermitilone** (**16**)
2	*sav100–101*	113,361–118,594	Polyketide
3	*sav257–259*	299,873–303,052	Microcin
4	*sav407–419 (pte)*	486,648–567,017	**Filipin** (**3**)
5	*sav603–609*	754,376–763,277	Non-ribosomal peptide (siderophore)
6	*sav837–869*	991,134–1,042,269	Non-ribosomal peptide
7	*sav935–953 (ave)*	1,132,045–1,212,960	**Avermectin** (**1**)
8	*sav1019–1025 (crt)*	1,285,187–1,293,904	**Isorenieratene** (**7**)
9	*sav1136–1137 (melC)*	1,424,869–1,426,085	**Melanin**
10	*sav1249–1251*	1,549,424–1,554,224	Polyketide–non-ribosomal peptide hybrid
11	*sav1550–1552*	1,893,266–1,912,282	Polyketide
12	*sav1650–1654 (hop)*	2,020,191–2,026,846	**Squalene** (**8**)
13	*sav2163 (geo)*	2,635,583–2,637,760	**Geosmin** (**5**), **germacradienol** (**6**)
14	*sav2267–2269*	2,765,027–2,768,005	γ-Butyrolactone ?
15	*sav2277–2282*	2,775,228–2,784,841	Polyketide
16	*sav2367–2369*	2,878,682–2,894,413	Polyketide
17	*sav2372–2388*	2,896,543–2,914,291	Aromatic polyketide
18	*sav2465–2467*	3,007,876–3,012,729	Siderophore
19	*sav2835–2842 (spp)*	3,480,598–3,487,905	**Spore pigment**
20	*sav2890–2903 (olm)*	3,534,525–3,634,592	**Oligomycin** (**2**)
21	*sav2989–2999 (ptl)*	3,744,875–3,757,141	**Neopentalenoketolactone** (**10**)
22	*sav3031–3032 (ezs)*	3,788,761–3,791,219	**Albaflavenol** (**12**), **albaflavenone** (**13**)
23	*sav3155–3164*	3,930,088–3,942,062	Non-ribosomal peptide
24	*sav3193–3202*	3,977,231–3,994,940	Non-ribosomal peptide
25	*sav3636–3651*	4,494,250–4,526,990	Non-ribosomal peptide
26	*sav3653–3667*	4,527,901–4,541,568	Aromatic polyketide
27	*sav3704, sav3706*	4,583,057–4,587,741	**Avenolide** (**22**; **autoregulator**)
28	*sav5149 (hpd)*	6,253,610–6,254,755	**Ochronotic pigment**
29	*sav5269–5274 (sid)*	6,383,181–6,386,774	**Nocardamine** (**18**), **deferrioxamine B** (**19**)
30	*sav5361–5362*	6,500,109–6,501,345	Melanin
31	*sav5686–5689*	6,877,533–6,881,873	Microcin
32	*sav6395–6398 (ect)*	7,670,236–7,673,431	**Ectoine** (**20**), **5-hydroxyectoine** (**21**)
33	*sav6632–6633*	7,931,869–7,937,201	Non-ribosomal peptide
34	*sav7130–7131*	8,490,207–8,492,484	**Tetrahydroxynaphthalene** (**4**)
35	*sav7161–7165*	8,522,760–8,530,697	Non-ribosomal peptide
36	*sav7184*	8,553,602–8,558,155	Polyketide
37	*sav7320–7323*	8,730,975–8,737,039	Vibrioferrin-like siderophore
38	*sav7360–7362*	8,777,769–8,789,766	Polyketide

[a] Products observed are indicated in bold

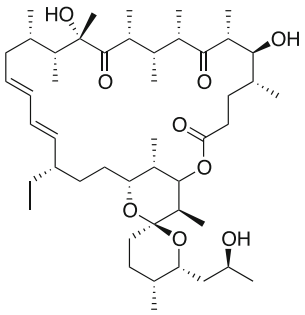
clusters were deduced from the genome sequence, leading to the subsequent isolation of a third group of macrocyclic lactones, filipins (**3**), which belong to the polyene family of antifungal compounds previously isolated from other *Streptomyces* microorganisms. In addition to the enormous diversity of PKS pathways in *S. avermitilis*, the genome harbors secondary metabolic gene clusters for at least eight non-ribosomal peptides, six terpenoids, and assorted pigments, osmolytes, siderophores, and bacteriocins, for a total of 38 secondary metabolic gene clusters (Table 1). As shown in Fig. 3, about 20 metabolites were detected by general cultivation or genetic manipulation of the biosynthetic genes.

Biosynthetic gene clusters for polyketide compounds

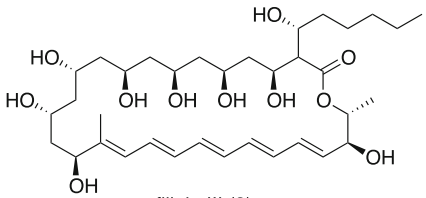
The gene cluster for oligomycin (**2**) biosynthesis at 3,534,525–3,634,592 nt of the genome contains 11 protein-coding genes, including seven genes (*olmA1* to *olmA7*) encoding multifunctional polypeptides of the type-I PKSs. These clustered PKS genes responsible for oligomycin biosynthesis together encode 16 homologous sets of



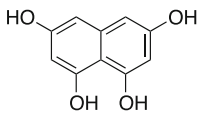
avermectin B1a (**1**)



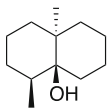
oligomycin A (**2**)



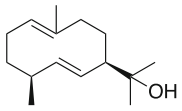
filipin III (**3**)



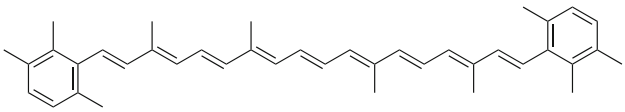
tetrahydroxynaphthalene (**4**)



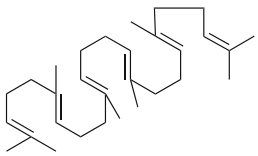
geosmin (**5**)



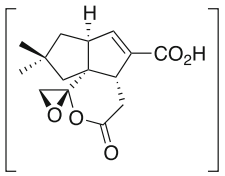
germacradienol (**6**)



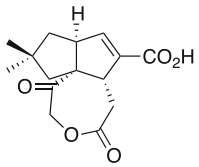
isorenieratene (**7**)



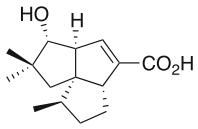
squalene (**8**)



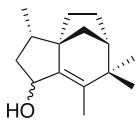
neopentalenolactone F (**9**)



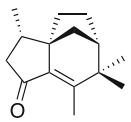
neopetalenoketolactone (**10**)



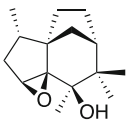
pentalenic acid (**11**)



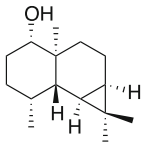
albaflavenols (**12**)



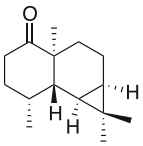
albaflavenone (**13**)



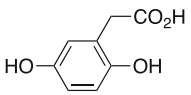
4β,5β-epoxy-2-*epi*-zizaan-6β-ol (**14**)



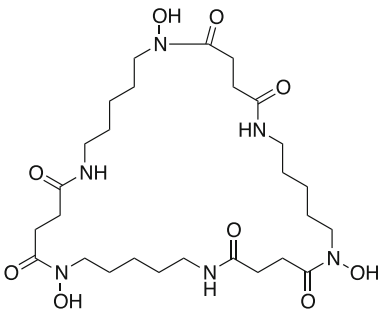
avermitilol (**15**)



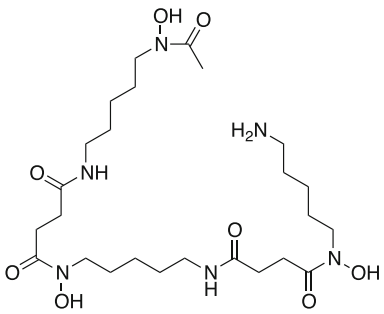
avermitilone (**16**)



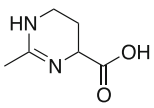
homogentisic acid (**17**)



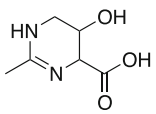
nocardamine (**18**)



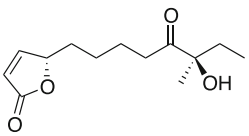
deferrioxamine B (**19**)



ectoine (**20**)



5-hydroxyectoine (**21**)



avenolide (**22**)

◀ **Fig. 3** Structures of secondary metabolites from *S. avermitilis*. Albaflavenols (**12**), albaflavenone (**13**), 4β,5β-epoxy-2-*epi*-zizaan-6β-ol (**14**), avermitilol (**15**), and avermitilone (**16**) are produced using an alternative promoter. Tetrahydroxynaphthalene was produced by the expression in the multi-copy plasmid

modules, each catalyzing a specific round of polyketide chain elongation (Fig. 4). The total of 79 constituent active sites, in which six are enzymatically nonfunctional, makes this the most complex multifunctional enzyme system identified to date. The clustered genes encoding PKS are organized as two sets of five and two genes, *olmA1–olmA2–olmA3–olmA6–olmA7* (*sav2899–2895*) and *olmA4-olmA5* (*sav2892–2893*), which are convergently transcribed. Between the two sets of PKS genes lie one gene, *olmB* (*sav2894*), which encodes cytochrome P450 monooxygenase (CYP107W1 [26]) catalyzing hydroxylation at C12 of oligomycin C in post-polyketide modification. Two open reading frames, *olmRI* and *olmRII*, encoding a regulatory function to activate the transcription of oligomycin biosynthetic genes are located to the left of the PKS genes. An adjacent gene encodes a monofunctional thioesterase that may be involved in the removal of an aberrant intermediate arising from incorrect functioning of chain extension processes. On the right of PKS genes, one open reading frame encoding crotonyl-CoA reductase is involved in the supply of an unusual polyketide extender unit (ethylmalonyl-CoA). We isolated many mutants by transposon mutagenesis previously [15] and a mutant failed to produce oligomycin caused by insertion of a transposon, Tn*4560*, in the *olmA5* region.

The biosynthetic gene cluster for filipin (**3**) at 486,648–567,017 nt of the genome contains 13 protein-coding genes. Five of these genes (*pteA1–pteA5*), encoding multifunctional polypeptides of type-I PKSs, were unidirectionally transcribed. These PKS genes responsible for biosynthesis of a polyene compound, filipin, together encode 13 homologous sets of modules. The 62 constituent active sites, of which two are enzymatically nonfunctional, are found in deduced polypeptides translated from five PKS genes (Fig. 5). Two genes, *pteC* and *pteD*, are located downstream of PKS genes. The gene products of *pteC* and *pteD* were expressed in *Escherichia coli,* characterized and defined as cytochrome P450 monooxygenase catalyzing hydroxylation at C26 and C1′, respectively, in the post-polyketide modification [26, 48, 49]. Many other type-I PKS genes are composed of multifunctional polypeptides; however, corresponding polyketide compounds are not found under any culture conditions.

During the sporulation of *S. avermitilis* on solid medium, the mature spores changed in color to deep brown by the accumulation of spore pigment associated with the *whiE*-homologous type-II PKS cluster *spp* (*sav2835–2842*). Disruption of the *spp* cluster led to albino spores and its transformants carrying a set of genes encoding type-III PKS and cytochrome P450 (*sav7130–7131*) on the multi-copy plasmid began to produce a diffusible brown pigment, which is another polyketide melanin that will be generated by spontaneous polymerization of tetrahydroxynaphthalene (**4**), suggesting that a gene encoding type-III PKS is an unusual role in *S. avermitilis* and the production of the brown pigment would cause an increase in the expression level of these genes by the desired gene-dosage effect.

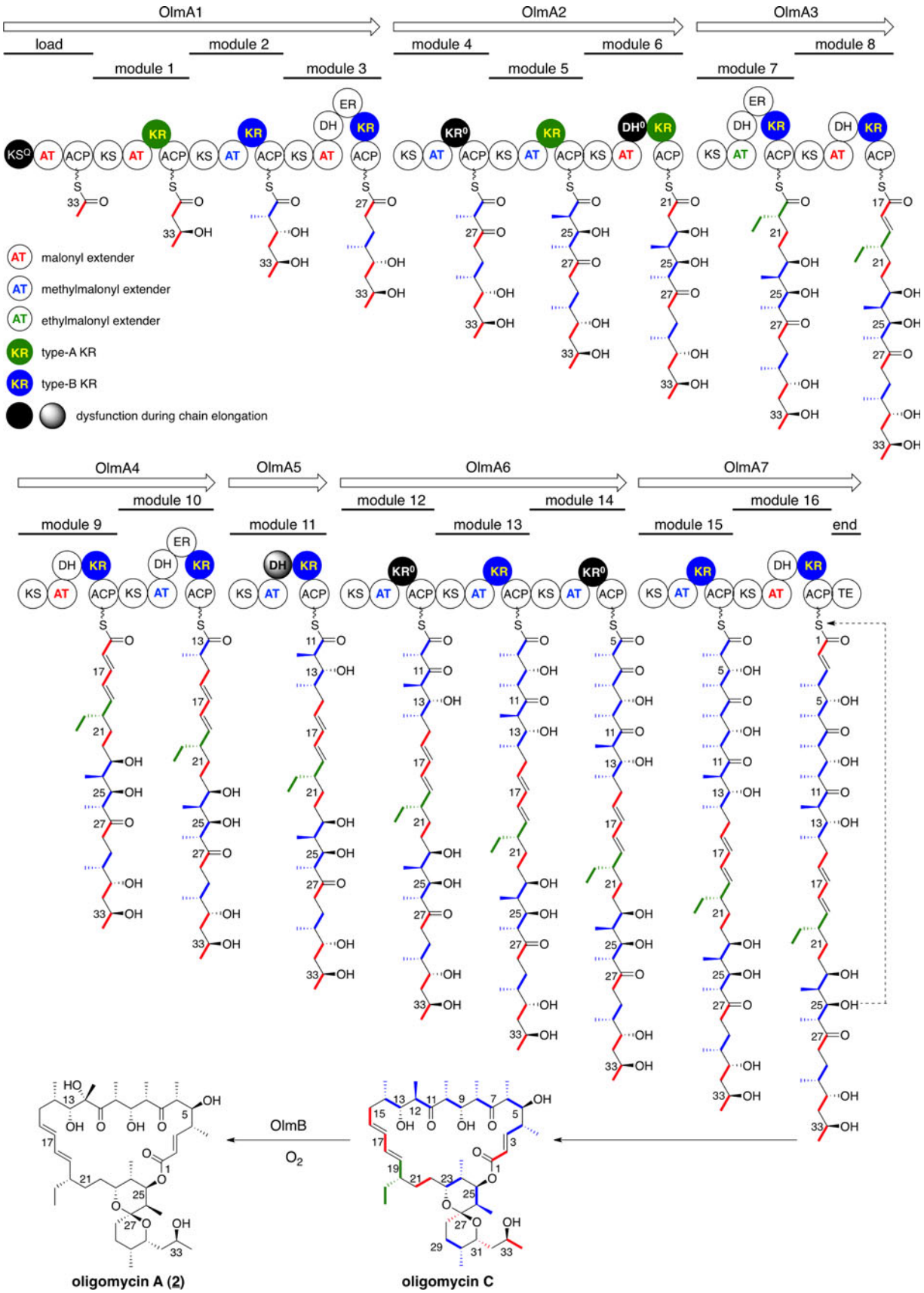
Biosynthetic gene clusters for peptide compounds

As *S. avermitilis* contains enormously diverse PKS pathways, the microorganism also has eight gene clusters for non-ribosomal peptide biosynthesis. However, all NRPS gene clusters are silent in the microorganism because peptide compounds synthesized by NRPSs have not been produced from *S. avermitilis* under any culture conditions. The reasons for this might be concerned with the expression of genes encoding NRPS or post-translational modification involving the phosphopantetheinylation of apo-NRPS to holo-NRPS. However, this prediction has been rejected by the heterologous expression of the biosynthetic gene cluster for cephamycin C (**40**) in *S. avermitilis*, which is synthesized by one NRPS containing three modules [23].

Biosynthetic gene clusters for terpene compounds

Many *Actinomycetales* microorganisms produce volatile terpenoid alcohols, 2-methylisoborneol (**47**), geosmin (**5**), and/or germacradienol (**6**). *S. avermitilis* has at least six gene clusters for terpenoid metabolites. Among them, two gene clusters are involved in the biosynthesis of a caroteinoid metabolite, isorenieratene (**7**), and squalene (**8**). Neither metabolite was found in the liquid culture, but the production of isorenieratene was promoted by light irradiation of the solid medium. On the other hand, squalene (hopanoid metabolites were also produced but the structures were not elucidated) was accumulated in mature spores on the solid medium.

Four other genes were involved in the cyclization of acyclic allylic diphosphate substrate to generate terpenoid hydrocarbon or alcohol. These four genes were predicted to be sesquiterpene synthases by clustering analysis of each polypeptide sequence. It is known that many *Actinomycetales* microorganisms produce volatile sesquiterpenoid alcohol, germacradienol (**6**) and its degraded product, geosmin (**5**). *S. avermitilis* also produces a large amount of these volatile metabolites on solid medium. A gene encoding germacradienol/geosmin synthase of *S. avermitilis* was cloned and the recombinant protein catalyzed the reaction of farnesyl-diphosphate to germacradienol and geosmin [5].


OlmA1
load
module 1
module 2
module 3
OlmA2
module 4
module 5
module 6
OlmA3
module 7
module 8
AT malonyl extender
AT methylmalonyl extender
AT ethylmalonyl extender
KR type-A KR
KR type-B KR
dysfunction during chain elongation
OlmA4
module 9
module 10
OlmA5
module 11
OlmA6
module 12
module 13
module 14
OlmA7
module 15
module 16
end
OlmB
O2
oligomycin A (2)
oligomycin C

◀ **Fig. 4** Model of oligomycin C (C12-deoxy oligomycin A) formation and predicted structure of the oligomycin PKSs. Putative enzymatic activity abbreviations: *ACP* acyl carrier protein, *AT* acyltransferase, *DH* dehydratase, *KR* β-oxoacyl-ACP reductase, *KS* β-oxoacyl-ACP synthase, *TE* thioesterase

The second gene encoding sesquiterpene synthase was found in a 13.4-kb *ptl* gene cluster encoding 13 predicted protein-coding sequences that we initially thought might represent the set of biosynthetic genes for pentalenolactone itself. The deduced amino acid sequence of the gene product of *ptlA* (*sav2998*) in the *ptl* cluster has high homology to the well-characterized pentalenene synthase of *S. exfoliatus* UC5319, which produces a sesquiterpenoid metabolite, pentalenolactone (**46**) [44]. By systematic expression of each of the individual protein-coding sequences of the *ptl* gene cluster, we identified the function of each gene product [6, 20, 38, 39, 50–52]. Furthermore, these results were supported by analysis of the metabolites produced by mutants carrying in-frame deletion of each gene in the cluster [20]. Unexpectedly, PtlE (SAV_2994) was identified as flavin-dependent Baeyer–Villiger monooxygenase and generated the previously unknown metabolite neopentalenolactone D, an isomer of the expected product pentalenolactone D of *S. exfoliatus* UC5319 [20]. The *ptl* gene cluster of *S. avermitilis* is likely responsible for the biosynthesis of a previously unknown metabolite, predicted to be neopentalenolactone F (**9**). This speculation has been supported by the large-deletion derivatives of *S. avermitilis* producing neopentalenoketolactone (**10**), most likely resulting from rearrangement of neopentalenolactone F (**9**) [20]. The pentalenolactone-producing *Streptomyces* microorganisms produced a shunt product, pentalenic acid (**11**), together with pentalenolactone. Similarly, wild-type *S. avermitilis* mainly produced pentalenic acid (**11**). This shunt pathway from 1-deoxypentalinic acid is generated by the cyclization of farnesyl diphosphate to pentalenene by PtlA (SAV_2998), and the generated pentalenene, oxidized to 1-deoxypentalenic acid by cytochrome P450 monooxygenase, PtlI (SAV_2999), and then to pentalenic acid, has been identified to be catalyzed by cytochrome P450, CYP105D7 (SAV_7469), the gene of which is located outside the *ptl* gene cluster [41].

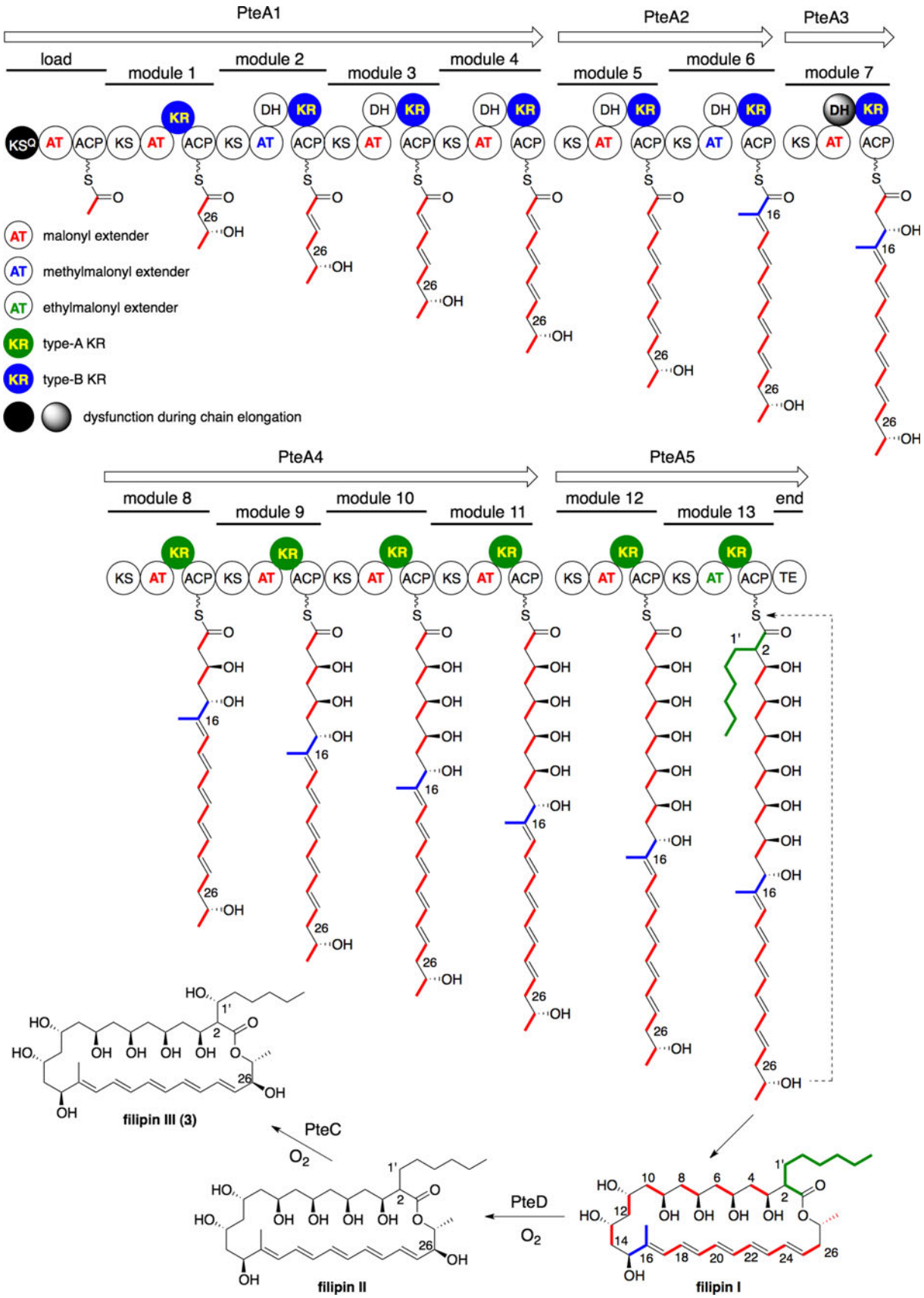
Two other genes encoding sesquiterpene synthase are silent in *S. avermitilis* because the microorganism produces no sesquiterpenoid metabolites, except for geosmin, under any culture condition. Previously, a gene product of *sco5222* in *S. coelicolor* A3(2) has been characterized as *epi*-isozizaene synthase [28]. The orthologous SAV_3032 protein of *S. avermitilis* has high homology to SCO5222 and also generated *epi*-isozizaene, and the overexpression of *sav3032* using an alternative promoter in the large-deletion mutant of *S. avermitilis* produced abundant *epi*-isozizaene in the mycelium. The genes encoding *epi*-isozizaene synthase of *S. coelicolor* A3(2) and *S. avermitilis* are translationally coupled with downstream genes encoding cytochrome P450, CYP170A1 (SCO5223) and CYP170A2 (SAV_3031), respectively. Co-expression of *sav3031* with *sav3032* using an alternative promoter in a large-deletion mutant of *S. avermitilis* generated both (4*R*)- and (4*S*)-albaflavenols (**12**) and albaflavenone (**13**), as well as a large amount of a novel epoxy alcohol, 4β,5β-epoxy-2-*epi*-zizaan-6β-ol (**14**) [42], a previously unknown trace metabolite in *S. coelicolor* A3(2).

Another silent sesquiterpene synthase gene, *ams* (*sav76*), is located in the left subtelomeric region of the *S. avermitilis* chromosome. Incubation of purified recombinant SAV_76 protein with farnesyl diphosphate in the presence of magnesium ions gave a mixture consisting of a novel sesquiterpenoid alcohol, avermitilol (**15**), accompanied by the known isomer viridiflorol, as well as germacrenes A and B. In vivo overexpression of *sav76* using an alternative promoter in a large-deletion mutant of *S. avermitilis* also produced the aforementioned metabolites and a relatively large amount of an oxidized derivative of avermitilol, avermitilone (**16**) [9].

Biosynthetic gene clusters for other compounds

No biosynthetic gene clusters for aminoglycoside and shikimate-derived compounds were located in the *S. avermitilis* genome, but several other gene clusters for assorted pigments, osmolyte compounds, siderophores, and bacteriocins were found. The microorganism produces spore pigment and another diffusible deep brown pigment (ochronotic pigment) in liquid medium and brown pigment production was enhanced by the addition of L-tyrosine to the medium. Transformants of *E. coli* harboring cloned DNA containing *sav5149* of *S. avermitilis* produced brown pigment. The *sav5149* was annotated to a gene encoding 4-hydroxyphenylpyruvate dioxygenase. SAV_5149 protein catalyzes the conversion of 4-hydroxyphenylpyruvate, which is generated by transamination of L-tyrosine, to homogentisic acid (**17**) [11]. A brown pigment that is characteristically similar to diffusible pigment of *S. avermitilis* is observed in the urine of human alkaptonuric patients. Alkaptonuria is a genetic disease (deficiency in the gene encoding homogentisic acid 1,2-dioxygenase) that causes the inability to metabolize homogentisic acid (**17**) and increases its concentration in urine, followed by its oxidation and polymerization to a brown ochronotic pigment.

Many microorganisms require iron as an essential element for many metabolic and cellular signal pathways. Since iron exists as an insoluble oxidized form under aerobic conditions, microorganisms secrete iron-chelating compounds, siderophores, which solubilize insoluble iron,


PteA1
PteA2
PteA3
load
module 1
module 2
module 3
module 4
module 5
module 6
module 7
PteA4
PteA5
module 8
module 9
module 10
module 11
module 12
module 13
end
malonyl extender
methylmalonyl extender
ethylmalonyl extender
type-A KR
type-B KR
dysfunction during chain elongation
filipin III (3)
filipin II
filipin I
PteC
PteD
O2

◀ **Fig. 5** Model of filipin I formation and predicted structure of the filipin PKSs. Putative enzymatic activity abbreviations: *ACP* acyl carrier protein, *AT* acyltransferase, *DH* dehydratase, *KR* β-oxoacyl-ACP reductase, *KS* β-oxoacyl-ACP synthase, *TE* thioesterase

and incorporate iron–siderophore complexes into the cells. In the genome sequence of *S. avermitilis*, it is expected that some gene clusters, which are classified into three groups, non-ribosomal peptide (*sav603–609*), hydroxamate (*sav5269–5274*), and catecholate (*sav7320–7323*) siderophores, are involved in the biosynthesis of siderophore. Wild-type *S. avermitilis* produces deferrioxamine-related hydroxamate siderophores, nocardamine (**18**) and deferrioxamine B (**19**), in a culture under iron-limited conditions [46]. However, unlike *S. coelicolor* A3(2) [27], *S. avermitilis* does not produce additional siderophores, non-ribosomal peptide, and catecholate compounds, under the same conditions, although the microorganism possesses the genetic capacity. Disruption of the *sav5272* gene encoding lysine decarboxylase abolishes the ability to produce nocardamine, suggesting that the gene cluster is involved in nocardamine biosynthesis. The gene cluster consists of four biosynthetic genes (*sav5269–5272*) and two transporter genes (*sav5273–5274*), the gene products of which may act to incorporate nocardamine/iron complex. The characteristic IdeR-binding motif was found in upstream regions of both *sav5272* and *sav5274*. The *ideR* encodes iron-dependent repressor protein and it is suspected that IdeR regulates the expression of the gene cluster for nocardamine biosynthesis. Mutants with disrupted *ideR* produce more nocardamine and deferrioxamine B than the wild-type strain. Nocardamine and deferrioxamine B production was completely suppressed at 5 μM Fe^{3+} in the wild-type strain, but *ideR*-disrupted mutants were not suppressed under the same conditions [46].

It is known that osmolytes are water-soluble compounds affecting osmosis and play a role in maintaining cell volume and fluid balance. The evolutionary advantage of the osmolyte systems is compatibility with the macromolecular structure and functioning at high and variable osmolyte concentrations without modifying cellular proteins to function in concentrated intracellular solutions. These organic osmolytes include polyhydric alcohols, amino acids, and others. A halophilic bacterium [7], which can grow in the presence of a high concentration of salts, accumulates a tetrahydropyrimidine derivative, ectoine (**20**), as an osmolyte in cells. A single operon with four protein-coding genes (*sav6395–6398*) resembles the biosynthetic gene cluster for ectoine of other bacteria, but ectoine-producing *Halomonas elongata* [7], *Marinococcus halophilus* [29], and *Bacillus pasteurii* [25] have three genes. On the other hand, *S. clavuligerus*, *S. griseus*, *S. peucetius* [30], and *S. antibioticus* [18] accumulate not only ectoine (**20**) but also 5-hydroxyectoine (**21**) in high ionic strength medium, suggesting that the fourth gene is involved in the ectoine hydroxylation step (**20**). As shown in Fig. 6a and b, wild-type *S. avermitilis* also accumulates both ectoine (**20**) and 5-hydroxyectoine (**21**) under hypertonic conditions and the disruption of *sav6395–6398* resulted in the failed accumulation of ectoine (**20**) and 5-hydroxyectoine (**21**). Furthermore, mutants lacking the fourth gene (*ectD*; *sav6395*) accumulated ectoine but not 5-hydroxyectoine (Fig. 6c).

Small extracellular signaling molecules, so-called autoregulators, are able to trigger secondary metabolite production in the genus *Streptomyces* [13, 21]. Many γ-butyrolactone autoregulators, including A-factor of *S. griseus*, virginiae butanolide of *S. virginiae*, and SCB1 of *S. coelicolor* A3(2), have since been identified in several *Streptomyces* species. Although an orthologous gene (*sav2269*) of *S. avermitilis* was found in *afsA*, *barX*, and *scbA*, which encode γ-butyrolactone biosynthesis enzyme, no corresponding γ-butyrolactone autoregulators were detected in the culture broth. However, a novel type of autoregulator molecule named avenolide (**22**), which is essential for eliciting avermectin production, has been discovered from the culture of wild-type *S. avermitilis* using an assay for specific binding to the receptor protein (AvaR1; SAV_3705) and restoration of avermectin production in an *aco* (*sav3706*)-disrupted mutant, which fails to produce avenolide [21]. Structure elucidation revealed that avenolide (**22**) is a butenolide but not a γ-butyrolactone autoregulator [21, 45]. The detailed biosynthetic pathway to avenolide has not been elucidated yet; however, two genes, *sav3706* (*aco*) and *sav3704*, encoding acyl-CoA oxidase and cytochrome P450 (CYP154B2), respectively, are required for avenolide biosynthesis.

Recently, antimycobacterial peptides with a lasso structure, named lariatins, were discovered from an *Actinomycetales* microorganism, *Rhodococcus jostii* K01-B0171 [19]. After cloning the lariatin biosynthetic gene cluster for lariatins, it was revealed that the peptide backbone is synthesized by ribosomal peptide synthesis but not non-ribosomal peptide synthesis. These genes form a single operon and encode a precursor peptide, an unknown function protein, a ring formation enzyme, processing peptidase, and an ABC transporter for the excretion of metabolites. Two sets of gene clusters, *sav257–259* and *sav5686–5689*, are similar to the lariatin gene cluster and each cluster contains five genes, but the former gene cluster is incomplete because the fifth gene encoding the ABC transporter is truncated by the insertion of IS*Sav20*. The latter gene cluster was cloned, and the recombinant plasmids carrying the intact gene cluster and the gene cluster controlled by the alternative promoter were introduced into the large-deletion mutant of *S. avermitilis*. As neither

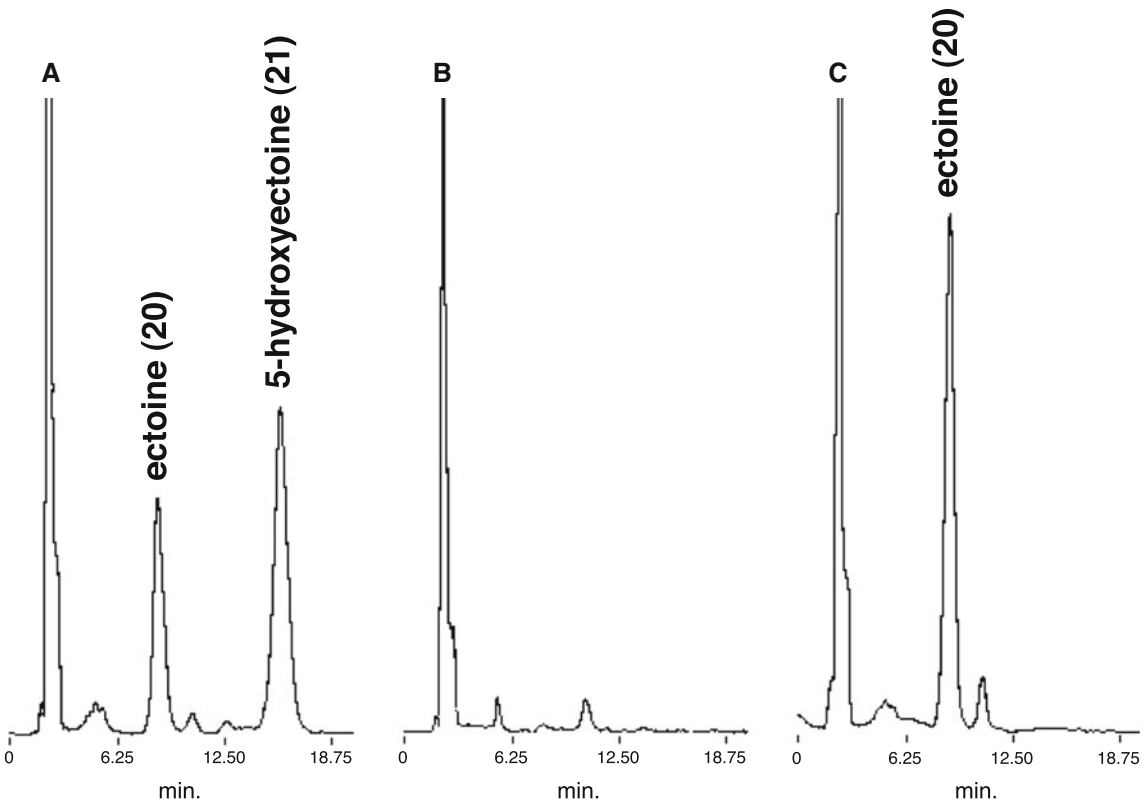


Fig. 6 HPLC of mycelial extracts of *A* wild type, *B* Δ*ectA-D* mutant, and *C* Δ*ectD* mutant. Each strain was cultured at 28 °C for 3 days in hypertonic medium containing 10 g glycerol, 2 g L-asparagine·monohydrate, 0.6 g $MgSO_4 \cdot 7H_2O$, 3.5 g K_2HPO_4, 6.8 g KH_2PO_4, and 14.6 g NaCl per liter adjusted to pH 7.0. Mycelium from 10 ml culture was sedimented by centrifugation and the mycelium was extracted with 0.4 ml $CHCl_3$/methanol/water (5:10:4) for 5 min. Then 0.13 ml $CHCl_3$ and 0.13 ml water were added and agitated. Phases were separated and the aqueous layer was subjected immediately to HPLC (column, Shodex Asahipak NH_2P504D (4.6ϕ × 150 mm); mobile phase, 80 % aqueous acetonitrile; flow rate, 0.6 ml/min; detection at 210 nm). Authentic ectoine (**20**) and 5-hydroectoine (**21**) were used as standards

transformant produced peptide compounds, the gene cluster appears to be dysfunctional.

Construction of versatile host for heterologous expression of biosynthetic gene clusters

Genome analysis of *Streptomyces* species has revealed that these microorganisms each have large linear chromosomes that contain over 20 secondary metabolic gene clusters encoding the biosynthesis of polyketides by PKSs, including type-I, -II and -III PKSs, peptides by NRPSs or ribosomal peptide synthesis, terpenoids, shikimate-derived metabolites, aminoglycosides, and other natural products. Although the origins of these secondary metabolic gene clusters are not clear, they might be most likely acquired through horizontal gene transfer or the duplication of related genes. These microorganisms gained many genes during evolution and a large amount of coding capacity is devoted to the biosynthesis of secondary metabolites at ca. 10 % of the genome. However, not all gene clusters for secondary metabolite biosynthesis are expressed and the metabolites generated by these gene products (biosynthetic enzymes) are produced under ordinary culture conditions. Many biosynthetic gene clusters are still unknown in these microorganisms. Before the completion of the genome analysis of *Streptomyces* species, biosynthetic gene clusters for secondary metabolites were typically identified by targeting the proposed biosynthetic or self-resistance genes, or by various in vivo strategies such as transposon mutagenesis or by complementing biosynthetically blocked mutants.

Genome mining offers an alternative method that is direct and reliable, especially for those gene sets associated with secondary metabolites without specific biosynthetic reactions to probe. Thus, the search for proposed biosynthetic genes or clusters for secondary metabolites has been improved. Detailed characterization of biosynthetic gene clusters for secondary metabolites has proved to be an invaluable tool for the elucidation of the biosynthetic mechanism and the regulation mechanism of the expression of biosynthetic genes. Furthermore, strategic genetic engineering of biosynthetic gene clusters for secondary metabolites and their regulatory networks allows improvement of the productivity of the metabolites and the

production of analogues by combinatorial biosynthesis. Genetic instability has been observed in many *Streptomyces* species. Genetic instability of the producing microorganism prevents the genetic analysis and engineering of biosynthetic genes. The introduction of DNA into secondary metabolite-producing microorganisms of interest by transformation or conjugation is a fundamental experimental procedure for genetic analysis, although many *Actinomycetales* microorganisms are unable to accept DNA thus impairing such genetic experiments. To overcome such a requirement, appropriate prerequisites are the availability of the relevant biosynthetic gene clusters controlling the production of the secondary metabolite of interest and suitable genetic systems for in vivo manipulation of the corresponding genes in heterologous hosts. Although the expression of biosynthetic genes is indispensable for the production of secondary metabolites, efficient production of the desired secondary metabolites requires an optimum relationship of timing and flux between primary and secondary metabolism, because all secondary metabolites are ultimately derived from primary metabolic building blocks and require an adequate source of energy and reducing equivalents derived from primary metabolism.

As described above, *S. avermitilis*, which is used for industrial production of the anthelmintic agent avermectin, has already proven to be a highly efficient microorganism producing secondary metabolites. This microorganism has been optimized for efficient supply of primary metabolic precursors, with equivalent energy and reducing ability to support the multistep biosynthesis of the secondary metabolites. Genetic instability of the producing microorganisms is a serious problem in the industrial production of secondary metabolites. Interestingly, *S. avermitilis* is stable in comparison with other *Streptomyces* microorganisms, which might depend on the length of TIRs on the linear chromosome [24]. Genetic manipulation systems, including genetic recombination [14] by mating, transposon mutagenesis using *Streptomyces* transposon [15], gene knockout and integration by transformation or conjugation, are fairly well developed in *S. avermitilis* [15, 16, 23]. Therefore, it is an attractive host for heterologous production of secondary metabolites. We therefore constructed a versatile host for the efficient production of secondary metabolites by controlled minimization of the genome of *S. avermitilis* [22, 23].

A series of large-deletion mutants of *S. avermitilis* SUKA2 to 22, which consist of a 7,509,588- to 7,352,064-bp (deletion of 1,516,020–1,673,544 bp; corresponding to 83.20–81.46 % of the wild-type genome) linear chromosome by general homologous recombination or site-specific recombination using Cre-*loxP* [23], could grow on minimum medium without any supplements, because no essential genes were deleted from the chromosome [24]. On the other hand, large-deletion mutants did not produce any endogenous metabolites, including major metabolites, avermectin (**1**), oligomycin (**2**), filipin (**3**), and terpenoid metabolites. Since the large deletions that we introduced into *S. avermitilis* also eliminate 78 % of putative transposase genes and 67 % of IS-like sequences, this would be expected to further improve genetic stability in these mutants. Enhancement of the already intrinsic genetically stable characteristics of *S. avermitilis* therefore make it especially suitable as a source of endogenous secondary metabolites and as a host for the heterologous production of secondary metabolites derived from exogenous biosynthetic gene clusters.

Heterologous expression of exogenous biosynthetic gene clusters

We have examined the heterologous expression of more than 20 exogenous biosynthetic gene clusters in the large-deletion mutants of *S. avermitilis* and these large-deletion mutants have been shown to be versatile and effective hosts for the expression of heterologous gene clusters governing the production of a variety of secondary metabolites, including aminoglycosides, nucleosides, ribosomal and non-ribosomal peptides, shikimate-derived metabolites, and terpenes. Almost all expression of heterologous biosynthetic gene clusters in the large-deletion mutants was effectively observed and the metabolites detected are shown in Fig. 7. In some cases, production was not observed. Pladienolides (**31**), metabolites of *S. platensis* Mer-11107, are antitumor macrocyclic polyketides with a unique mode of action and the entire gene cluster of pladienolide biosynthesis contains eight protein-coding genes, including a gene encoding a pathway-specific regulator, spanning a distance of 65 kb. Transformants harboring BAC clone (75 kb) carrying the entire biosynthetic gene cluster for pladienolide did not produce desired metabolites. Transcription analysis revealed that a gene encoding the transcriptional activator for the pladienolide biosynthetic gene, *pldR*, was not expressed. The result indicates that the appropriate regulator to activate *pldR* expression might not be present in the heterologous *S. avermitilis* host, but pladienolide production was accomplished by the expression of *pldR* using an alternative promoter (*ermE* promoter of *Saccharopolyspora erythraea* NRRL 2338) [23]. Although understanding the higher level of regulatory systems that control the expression of pathway-specific regulatory genes for individual biosynthetic gene clusters is essential to control metabolite production, only a few cases have been studied to date. Heterologous expression of the biosynthetic gene cluster for pladienolide could be also accomplished by introducing the specific regulator underlying the expression of *pldR* in *S. platensis*. The second

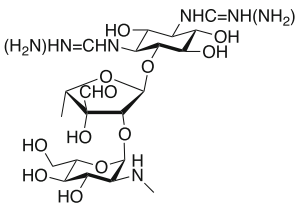
case is the heterologous expression of a biosynthetic gene cluster specific to a proteasome inhibitor, lactacystin (**41**), of *S. lactacystinaeus* OM-6519. The biosynthetic cluster for lactacystin contains five protein-coding genes, including genes encoding PKS-NRPS hybrid and NRPS, but no pathway-specific regulatory gene(s). The transformants carrying the entire gene cluster did not produce lactacystin. In general, transcription of biosynthetic genes is regulated by pathway-specific regulator(s), of which gene(s) lies in the gene cluster. The gene cluster might be transcribed by a specific sigma factor of RNA polymerase or the transcription might be regulated by a specific regulator, the gene of which is located outside of the gene cluster. *S. avermitilis* does not seem to possess the requisite orthologous gene(s). Since all five genes in the gene cluster are transcribed in the same direction and these genes form an operon, an alternative promoter sequence was introduced upstream of the front of the gene in this operon. Desired transformants produced abundant lactacystin in the culture broth [24].

In many cases, the productivity of exogenous metabolites by heterologous expression of the intact biosynthetic gene cluster in large-deletion mutants of *S. avermitilis* was improved in comparison with those of the original producing microorganisms. Chloramphenicol (**42**) productivity in transformants of the large-deletion mutants of *S. avermitilis* carrying the entire biosynthetic gene cluster for chloramphenicol of *S. venezuelae* ATCC 10712 was tenfold or more higher than that of the original producer, *S. venezuelae* [24]. Interestingly, *S. avermitilis* does not possess the biosynthetic gene cluster for shikimate-derived metabolites such as chloramphenicol. The results indicate that *S. avermitilis* has naturally optimized primary metabolism of the shikimate pathway and/or the regulator protein of *S. avermitilis* to activate the pathway-specific regulator of biosynthetic genes for chloramphenicol and is more efficient than the native host *S. venezuelae*.

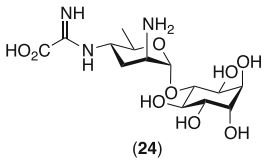
Although *Streptomyces* microorganisms possess many biosynthetic gene clusters for secondary metabolites, many of them are silent in the microorganism. Therefore, it is not clear what the function of many gene products is nor the identity of the respective metabolite biosynthesis by these silent genes. *S. clavuligerus* ATCC 27065 is known as an industrial microorganism producing compounds of clinical importance such as an cephamycin C (**40**), an antibacterial β-lactam antibiotic, and clavulanic acid (**39**), a potent β-lactamase inhibitor. Similarly to *S. avermitilis*, genomic analysis of *S. clavuligerus* ATCC 27065 suggests that this organism harbors several silent gene clusters [32]. Among these gene clusters, the size and organization of genes in a 26-kb region of mega-plasmid pSCL4 were quite similar to those of a 30-kb region containing product of the biosynthetic gene cluster for a phosphoglycolipid antibiotic, moenomycin, of *S. ghanaensis* ATCC 14772. However, no antimicrobial phosphoglycolipid compounds have been detected from the culture of *S. clavuligerus* to date. This observation indicates that the biosynthetic gene cluster, spanning a distance of 26 kb, is unknown in this microorganism. Similarly, *S. clavuligerus* ATCC 27065 also possesses the biosynthetic gene cluster for holomycin (**38**), but the wild-type strain hardly produces holomycin and its related metabolites. Holomycin production was observed in the mutants with blocked cephamycin C and/or clavulanic biosynthesis. The entire biosynthetic gene cluster for the phosphoglycolipid compound was introduced into the large-deletion mutants of *S. avermitilis* and the resultant transformants began to produce an antimicrobial compound. Ultimately, the compound was purified from mycelium and was identified as a moenomycin-related antibiotic, pholipomycin (**27**). Although wild-type *S. clavuligerus* produced a trace amount of holomycin, the large-deletion mutants of *S. avermitilis* carrying the entire gene cluster for holomycin biosynthesis produced sufficient holomycin (more than 40-fold higher than the original producer). The results suggest that a silent biosynthetic gene cluster of *S. clavuligerus* is activated in *S. avermitilis* [24]. This is the first observation of the activation of biosynthetic genes by the heterologous expression of an exogenous biosynthetic gene cluster.

The feasibility of using large-deletion mutants of *S. avermitilis* as a heterologous host has been shown by the effective expression of more than 20 biosynthetic gene clusters (Table 2). Furthermore, because large-deletion mutants of *S. avermitilis* no longer produced major endogenous metabolites, primary metabolism seemed to be efficiently exploited to generate precursors of exogenous biosynthetic gene clusters. Transformation of large-deletion mutants of *S. avermitilis* using less than ca. 45-kb recombinant plasmids, including cosmid clones, was efficiently accomplished, but the introduction of BAC clone with a length of ca. 100 kb was inefficient. Efficient introduction of DNA of such a size by transformation and/or conjugation from *E. coli* using RP4-based transfer function will need to be improved in order to take full advantage of *S. avermitilis* as a heterologous host.

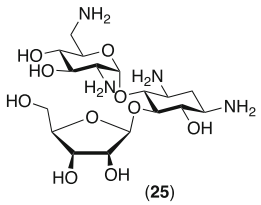
Fig. 7 Structures of metabolites produced by heterologous expression of exogenous biosynthetic gene clusters in the large-deletion mutants of *S. avermitilis*. Streptomycin (**23**), kasugamycin (**24**), ribostamycin (**25**), sangivamycin (**26**), pholipomycin (**27**), nosokomycin B (**28**), erythromycin A (**29**), bafilomycin B1 (**30**), pladienole B (**31**), aureothin (**32**), nemedectin α (**33**), leptomycin (**34**), asukamycin (**35**), oxytetracycline (**36**), resistomycin (**37**), holomycin (**38**), clavulanic acid (**39**), cephamycin C (**40**), lactacystin (**41**), chloramphenicol (**42**), thiostrepton (**43**), rebeccamycin (**44**), novobiocin (**45**), pentalenolactone (**46**), and 2-methylisoborneol (**47**). Details of transformants carrying the entire gene clusters are summarized in Table 2



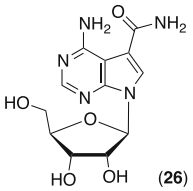
(23)



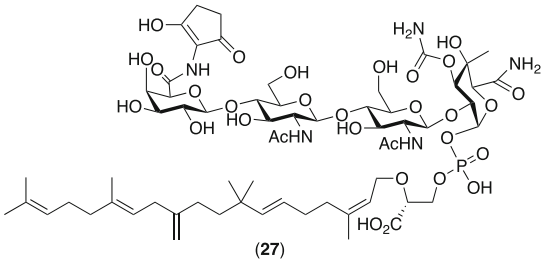
(24)



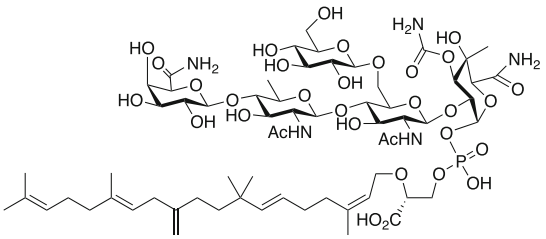
(25)



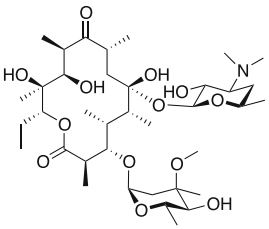
(26)



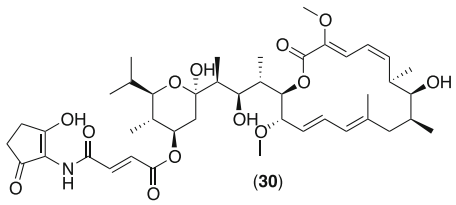
(27)



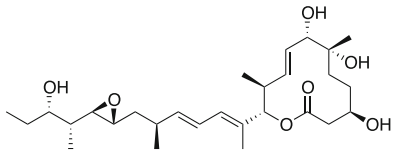
(28)



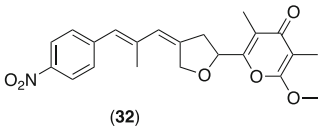
(29)



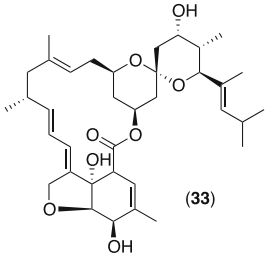
(30)



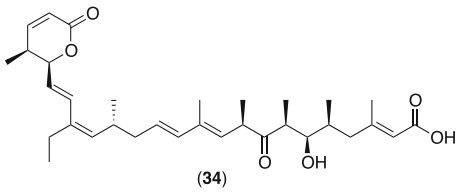
(31)



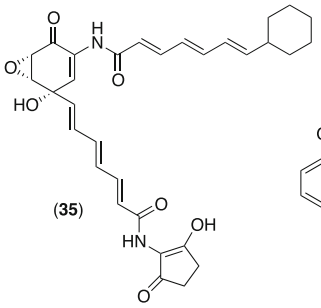
(32)



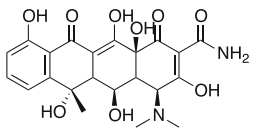
(33)



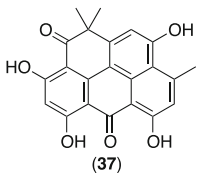
(34)



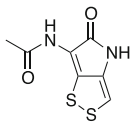
(35)



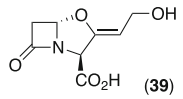
(36)



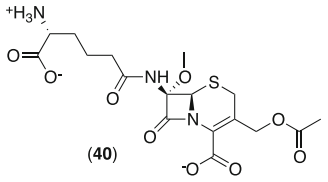
(37)



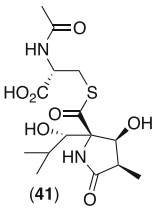
(38)



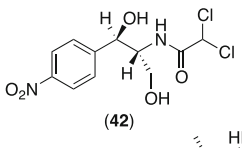
(39)



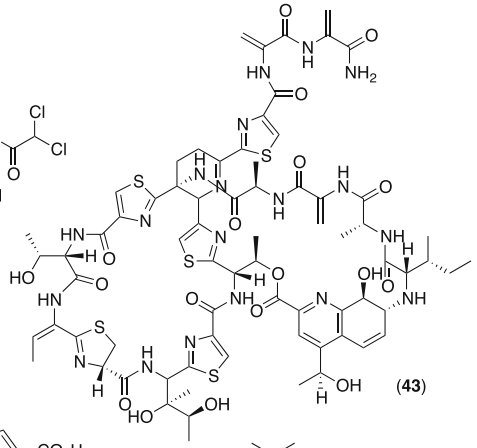
(40)



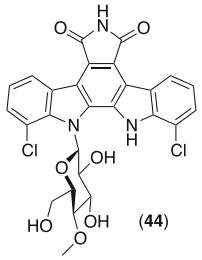
(41)



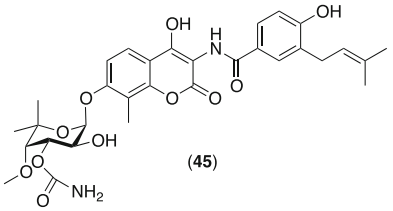
(42)



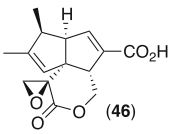
(43)



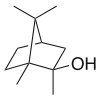
(44)



(45)



(46)



(47)

Table 2 Heterologous expression of exogenous biosynthetic gene clusters in large-deletion mutants of *S. avermitilis*

Gene cluster	Type of metabolite	Source of gene cluster [a]	Productivity [b]			
			(mg/l)	Expression[e]	Host[l]	Reference
Streptomycin	Aminoglycoside	*Streptomyces griseus* IFO 13350	176	Native	SUKA4, 5, 17	[23]
Kasugamycin	Aminoglycoside	*S. kasugaensis* MB273	7	Native[f]	SUKA17	[24]
Ribostamycin	Aminoglycoside	*S. ribosidificus* ATCC 21294	8	Native	SUKA17	[24]
Sangivamycin	Nucleoside	*S. rimosus* subsp. *rimosus* ATCC14673	56	Native	SUKA22	
Pholipomycin	Phosphoglycolipid	*S. clavuligerus* ATCC 27065	20	Native	SUKA22	[24]
Nosokomycin	Phosphoglycolipid	*Streptomyces* sp. K04-0144	2	Native	SUKA22	
Erythromycin	Polyketide (type-I PKS)	*Saccharopolyspora erythraea* NRRL 2338	4	Native	SUKA22	[24]
Bafilomycin	Polyketide (type-I PKS)	*Kitasatospora setae* KM-6054	16	Native	SUKA22	[24]
Pladienolide	Polyketide (type-I PKS)	*S. platensis* Mer-11107	250	*ermEp*[g]	SUKA5	[23]
Aureothin	Polyketide (type-I PKS)	*Streptomyces* sp. MM23	116	Native	SUKA17, 22	[24]
Nemadectin	Polyketide (type-I PKS)	*S. cyaneogriseus* subsp. *noncyanogenus* NRRL 15774	104[c]	Native	SUKA22	[24]
Leptomycin	Polyketide (type-I PKS)	*Streptomyces* sp. EM52	5	Native	SUKA22	[24]
Asukamycin	Polyketide (type-II PKS)	*S. nodosus* subsp. *asukaensis* ATCC 29757	153	Native	SUKA22	
Oxytetracycline	Polyketide (type-II PKS)	*S. rimosus* subsp. *rimosus* NRRL 2234	20	Native	SUKA17, 22	[24]
Resistomycin	Polyketide (type-II PKS)	*Streptomyces* sp. NA97	360	Native	SUKA22	
Clavulanic acid	Peptide	*S. clavuligerus* ATCC 27065	16	*ccaR*[h]	SUKA22	[24]
Holomycin	Peptide (NRPS)	*S. clavuligerus* ATCC 27065	8	Native	SUKA22	[24]
Cephamycin C	Peptide (NRPS)	*S. clavuligerus* ATCC 27065	85	Native	SUKA17	[23]
Lactacystin	Peptide (NRPS/PKS)	*S. lactacystinaeus* OM-6519	30	*rpsJp*[i]	SUKA22	[24]
Thiostrepton	Peptide (ribosomal)	*S. laurentii* ATCC 31255	92	*tsr*[j]	SUKA22	
Chloramphenicol	Shikimate	*S. venezuelae* ATCC 10712	262	Native	SUKA22	[24]
Rebeccamycin	Shikimate	*Lechevalieria aerocolonigenes* IFO 13195	7	Native	SUKA17, 22	[24]
Novobiocin	Shikimate	*S. anulatus* 3533-SV4 GM95	1	Native	SUKA22	[24]
Pentalenolactone	Terpene	*S. cyaneogriseus* subsp. *noncyanogenus* NRRL 15774	12	Native	SUKA22	
		S. arenae TÜ469/*S. avermitilis* MA-4680	23	*ermEp*[k]	SUKA16	[52]
2-Methylisoborneol	Terpene	*Sa. erythraea* NRRL 2338	ND[d]	Native	SUKA16	[22]

[a] Each entire biosynthetic gene cluster was introduced into large-deletion mutant of *S. avermitilis* (SUKA4, 5, 17, or 22) using bacteriophage ϕC31- or TG1-based integrating vector. No different expression of the gene cluster in transformants was observed between *attB-ϕC31* and *attB-TG1* on the genome

[b] Conditions for vegetative culture and fermentation culture were described previously [23]. In almost all cases, semisynthetic medium [23] was used for the production and another yeast extract/malts extract-based medium [24] was used when the productivity was low in the semisynthetic medium

[c] Newly isolated transformants produced more nemadectins than that of previously isolated clones (1.6 mg/l)

[d] *ND* not determined quantitative analysis

Since metabolite is volatile compound, the productivity was confirmed by GC–MS analysis using authentic sample

[e] Almost all biosynthetic gene clusters were expressed by their native promoters ("native")

[f] *kasT* gene encoding pathway-specific regulator in the biosynthetic gene cluster for kasugamycin of *S. kasugaensis* MB273 was expressed in the multi-copy plasmid vector that gave the desired gene-dosage effect

[g] An extra-copy of *pldR* gene encoding pathway-specific regulator, which was controlled by *ermE* promoter (*ermEp*) of *Saccharopolyspora erythraea* NRRL 2338, was introduced into transformants carrying the entire biosynthetic gene cluster for pladienolide

[h] *ccaR* gene encoding pathway-specific regulator in the gene cluster for cephamycin C biosynthesis of *S. clavuligerus* ATCC 27065 was co-expressed with the entire biosynthetic gene cluster for clavulanic acid

[i] *rpsJ* (*sav4925*; encoding ribosomal protein S10) promoter of *S. avermitilis* was introduced upstream of the front of the gene in the gene cluster

[j] The entire biosynthetic gene cluster for thiostrepton was co-expressed with thiostrepton resistance gene, *tsr,* of *S. azureus* ATCC 14921 because the gene cluster lacks self-resistance gene

[k] Pentalenolactone-producing transformants of *S. avermitilis* SUKA16 (Δ*ptl*-cluster; *sav2990–sav3002*) were constructed by introduction of three segments, *ptl*-cluster Δ*ptlD* Δ*ptlE* of *S. avermitilis*, *ermEp*-controlled *pntE-pntD* and *ermEp*-controlled *pntM* of *S. arenae* TÜ469, using bacteriophages TG1-, ϕBT1-, and ϕC31-based integrating vectors, respectively

[l] All hosts derived from *S. avermitilis* for heterologous expression of biosynthetic gene cluster for secondary metabolite were described previously [23]

Acknowledgments We thank M. Komatsu, K. Komatsu, H. Koiwai, S. Takamatsu, I. Kozone, M. Izumikawa, and J. Hashimoto for the collaborative experiments. We also thank Y. Shiga for valuable advice about IS elements. We gratefully acknowledge support from a research Grant-in-Aid for Scientific Research on Priority Areas (to H.I.), a research Grant-in-Aid for Scientific Research on Innovative Areas (to H.I.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, a research grant from the Institute for Fermentation, Osaka, Japan (to H.I.), and a research Grant-in-Aid for Scientific Research from the

New Energy and Industrial Technology Development Organization (NEDO; to K.S. and H.I.).

References

1. Barbe V, Bouzon M, Mangenot S, Badet B, Poulain J, Segurens B, Vallenet D, Marlière P, Weissenbach J (2011) Complete genome sequence of *Streptomyces cattleya* NRRL 8057, a producer of antibiotics and fluorometabolites. J Bacteriol 193:5055–5056
2. 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 (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417:141–147
3. Bignell DRD, Seipke RF, Huguet-Tapia JC, Chambers AH, Parry RJ, Loria R (2010) *Streptomyces scabies* 87-22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant–microbe interactions. Mol Plant Microbe Interact 23:161–175
4. Burg RW, Miller BM, Baker EE, Birnbaum J, Currie SA, Hartman R, Kong YL, Monaghan RL, Olson G, Putter I, Tunac JB, Wallick H, Stapley EO, Oiwa R, Omura S (1979) Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. Antimicrob Agents Chemother 15:361–367
5. Cane DE, He X, Kobayashi S, Omura S, Ikeda H (2006) Geosmin biosynthesis in *Streptomyces avermitilis*. J Antibiot 59:471–479
6. Cane DE, You Z, Omura S, Ikeda H (2007) Pentalenolactone biosynthesis. Molecular cloning and assignment of biochemical function to PtlF, a short-chain dehydrogenase from *Streptomyces avermitilis*, and identification of a new biosynthetic intermediate. Arch Biochem Biophys 459:233–240
7. Cánovas D, Vargas C, Calderon MI, Ventosa A, Nieto JJ (1998) Characterization of the genes for the biosynthesis of the compatible solute ectoine in the moderately halophilic bacterium *Halomonas elongata* DSM 3043. Syst Appl Microbiol 21:487–497
8. Carter GT, Nietsche JA, Hertz MR, Williams DR, Siegel MM, Morton GO, James JC, Borders DB (1988) LL-F28249 antibiotic complex: a new family of antiparasitic macrocyclic lactones: isolation, characterization and structures of LL-F28249 α, β, γ, λ. J Antibiot 41:519–529
9. Chou WKW, Fanizza I, Uchiyama T, Komatsu M, Ikeda H, Cane DE (2010) Genome mining in *Streptomyces avermitilis*: cloning and characterization of SAV_76, the synthase for a new sesquiterpene, avermitilol. J Am Chem Soc 132:8850–8851
10. Chung ST (1987) Tn*4556*, a 6.8-kilobase-pair transposable element of *Streptomyces fradiae*. J Bacteriol 169:4436–4441
11. Denoya CD, Skinner DD, Morgenstern MR (1994) A *Streptomyces avermitilis* gene encoding a 4-hydroxyphenylpyruvic acid dioxygenase-like protein that directs the production of homogentisic acid and an ochronotic pigment in *Escherichia coli*. J Bacteriol 176:5312–5319
12. Feng Z, Wang L, Rajski SR, Xu Z, Coeffet-LeGal MF, Shen B (2009) Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts. Bioorg Med Chem 17:2147–2153
13. Hsiao NH, Nakayama S, Merlo ME, Vries M, Bunet R, Kitani S, Nihira T, Takano T (2009) Analysis of two additional signaling molecules in *Streptomyces coelicolor* and the development of a butyrolactone-specific reporter system. Chem Biol 16:951–960
14. Ikeda H, Kotaki H, Omura S (1987) Genetic studies of avermectin biosynthesis in *Streptomyces avermitilis*. J Bacteriol 169:5615–5621
15. Ikeda H, Takada Y, Pang CH, Tanaka H, Omura S (1993) Transposon mutagenesis by Tn*4560* and applications with avermectin-producing *Streptomyces avermitilis*. J Bacteriol 175:2077–2082
16. Ikeda H, Nonomiya T, Usami M, Ohta T, Omura S (1999) Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. Proc Natl Acad Sci U S A 96:9509–9514
17. Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol 21:526–531
18. Inbar L, Lapidot A (1988) The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. J Biol Chem 263:16014–16022
19. Inokoshi J, Matsuhama M, Miyake M, Ikeda H, Tomoda H (2012) Molecular cloning of the gene cluster for lariatin biosynthesis of *Rhodococcus jostii* K01-B0171. Appl Microb Biotechnol 95:451–460
20. Jiang J, Tetzlaff CN, Takamatsu S, Iwatsuki M, Komatsu M, Ikeda H, Cane DE (2009) Genome mining in *Streptomyces avermitilis*: a biochemical Baeyer-Villiger reaction and discovery of a new branch of the pentalenolactone family tree. Biochemistry 48:6431–6440
21. Kitani S, Miyamoto KT, Takamatsu S, Herawati E, Iguchi H, Nishitomi K, Uchida M, Nagamitsu T, Omura S, Ikeda H, Nihira T (2011) Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. Proc Natl Acad Sci U S A 108:16410–16415
22. Komatsu M, Tsuda M, Omura S, Oikawa H, Ikeda H (2008) Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. Proc Natl Acad Sci U S A 105:7422–7427
23. Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H (2010) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. Proc Natl Acad Sci U S A 107:2646–2651
24. Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozone I, Izumikawa M, Hashimoto J, Takagi M, Omura S, Shin-ya K, Cane DE, Ikeda H (2013) Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. ACS Syn Biol 2:384–396
25. Kuhlmann AU, Bremer E (2002) Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. Appl Env Microbiol 68:772–783
26. Lamb DC, Ikeda Nelson DR, Ishikawa J, Skaug T, Jackson C, Omura S, Waterman MR, Kelly SL (2003) Cytochrome P450 complement of the avermectin-producer *Streptomyces avermitilis* and comparison to that of *Streptomyces coelicolor* A3(2). Biochem Biophys Res Commun 307:610–619
27. Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat Chem Biol 1:265–269
28. Lin X, Cane DE (2009) Biosynthesis of sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor*. Mechanism and stereochemistry of the enzymatic formation of *epi*-isozizaene. J Am Chem Soc 131:6332–6333
29. Louis P, Galinski EA (1997) Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. Microbiology 143:1141–1149

30. Malin G, Lapidot A (1996) Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. J Bactriol 178:385–395
31. Martinez A, Kolvek SJ, Yip CLT, Hopke J, Brown KA, MacNeil IA, Osburne MS (2004) Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. Appl Environ Microbiol 70:2452–2463
32. Medema MH, Trefzer A, Kovalchuk A, Berg M, Müller U, Heijne W, Wu L, Alam MT, Ronning CM, Nierman WC, Bovenberg RAL, Breitling R, Takano E (2010) The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol Evol 2:212–224
33. Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26:1326–1384
34. Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, Yamashita A, Hattori M, Horinouchi S (2008) Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. J Bacteriol 190:4050–4060
35. 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 (2001) Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc Natl Acad Sci U S A 98:12215–12220
36. Pfeifer B, Hu Z, Licari P, Khosla C (2002) Process and metabolic strategies for improved production of *Escherichia coli*-derived 6-deoxyerythronolide B. Appl Environ Microbiol 68:3287–3292
37. Pullan ST, Chandra G, Bibb MJ, Merrick M (2011) Genome-wide analysis of the role of GlnR in *Streptomyces venezuelae* provides new insights into global nitrogen regulation in actinomycetes. BMC Genomics 12:175
38. Quaderer R, Omura S, Ikeda H, Cane DE (2006) Pentalenolactone biosynthesis. Molecular cloning and assignment of biochemical function to PtlI, a cytochrome P450 of *Streptomyces avermitilis*. J Am Chem Soc 128:13036–13037
39. Seo MJ, Zhu D, Endo S, Ikeda H, Cane DE (2011) Genome mining in *Streptomyces*. Elucidation of the role of Baeyer-Villiger monooxygenases and non-heme iron-dependent dehydrogenase/oxygenases in the final steps of the biosynthesis of pentalenolactone and neopentalenolactone. Biochemistry 50:1739–1754
40. Solenberg PJ, Burgett SG (1989) Method for selection of transposable DNA and characterization of a new insertion sequence, IS*493*, from *Streptomyces lividans*. J Bacteriol 171:4807–4813
41. Takamatsu S, Xu LH, Fushinobu S, Shoun H, Komastu M, Cane DE, Ikeda H (2011) Pentalenic acid is a shunt metabolite in the biosynthesis of the pentalenolactone family of metabolites: hydroxylation of 1-deoxypentalenic acid mediated by CYP105D7 (SAV_7469) of *Streptomyces avermitilis*. J. Antibiot 64:65–71
42. Takamatsu S, Lin X, Nara A, Komatsu M, Cane DE, Ikeda H (2011) Characterization of a silent sesquiterpenoid biosynthetic pathway in *Streptomyces avermitilis* controlling *epi*-isozizaene albaflavenone biosynthesis and isolation of a new oxidized *epi*-isozizaene metabolite. Microb Biotechnol 4:184–191
43. Takiguchi Y, Mishima H, Okuda M, Terao M, Aoki A, Fukuda R (1980) Milbemycins, a new family of macrolide antibiotics: fermentation, isolation and physico-chemical properties. J Antibiot 33:1120–1127
44. Tetzlaff CN, You Z, Cane DE, Takamatsu S, Omura S, Ikeda H (2006) A gene cluster for biosynthesis of the sesquiterpenoid antibiotic pentalenolactone in *Streptomyces avermitilis*. Biochemistry 45:6179–6186
45. Uchida M, Takamatsu S, Arima S, Miyamoto KT, Kitani S, Nihira T, Ikeda H, Nagamitsu T (2011) Total synthesis and absolute configuration of avenolide, extracelluar factor in *Streptomyces avermitilis*. J Antibiot 64:781–787
46. Ueki M, Suzuki R, Takamatsu S, Takagi H, Uramoto M, Ikeda H, Osada H (2009) Nocardamin production by *Streptomyces avermitilis*. Actinomycetologica 23:34–39
47. Wang XJ, Yan YJ, Zhang B, An J, Wang JJ, Tian J, Jiang L, Chen YH, Huang SX, Yin M, Zhang J, Gao AL, Liu CX, Zhu ZX, Xiang WS (2010) Genome sequence of the milbemycin-producing bacterium *Streptomyces bingchenggensis*. J Bacteriol 192:4526–4527
48. Xu L, Fushinobu S, Ikeda H, Wakagi T, Shoun H (2009) Crystal structures of cytochrome P450 105P1 from *Streptomyces avermitilis*: conformational flexibility and histidine-liganded state. J Bacteriol 191:1211–1219
49. Xu L, Fushinobu S, Takamatsu S, Wakagi T, Ikeda H, Shoun H (2010) Regio- and stereospecificity of filipin hydroxylation sites revealed by crystal structures of cytochrome P450 105P1 and 105D6 from *Streptomyces avermitilis*. J Biol Chem 285:16844–16853
50. You Z, Omura S, Ikeda H, Cane DE (2006) Pentalenolactone biosynthesis. Moleclar cloning and assignment of biochemical function to PtlH, a non-heme iron dioxygenase of *Streptomyces avermitilis*. J Am Chem Soc 128:6566–6567
51. You Z, Omura S, Ikeda H, Cane DE, Jogl G (2007) Crystal structure of the non-heme iron dioxygenase PtlH in pentalenolactone biosynthesis. J Biol Chem 282:36552–36560
52. Zhu D, Seo MJ, Ikeda H, Cane DE (2011) Genome mining in *Streptomyces*. Discovery of an unprecedented P450-catalyzed oxidative rearrangement that is the final step in the biosynthesis of pentalenolactone. J Am Chem Soc 133:2128–2131