SCIENCE CHINA
Life Sciences

SPECIAL TOPIC: Biosynthesis and regulation of secondary metabolites in microorganisms
• RESEARCH PAPER •

July 2013 Vol.56 No.7: 628–637
doi: 10.1007/s11427-013-4505-1

Cloning and characterization of the gene cluster required for beauvericin biosynthesis in *Fusarium proliferatum*

ZHANG Tao[1,2†], ZHUO Ying[1†], JIA XiaoPeng[1,2], LIU JinTao[1,2], GAO Hong[1], SONG FuHang[1], LIU Mei[1*] & ZHANG LiXin[1*]

[1]*CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

Received April 24, 2013; accepted May 28, 2013

Beauvericin, a cyclohexadepsipeptide-possessing natural product with synergistic antifungal, insecticidal, and cytotoxic activities. We isolated and characterized the *fpBeas* gene cluster, devoted to beauvericin biosynthesis, from the filamentous fungus *Fusarium proliferatum* LF061. Targeted inactivation of the *F. proliferatum* genomic copy of *fpBeas* abolished the production of beauvericin. Comparative sequence analysis of the FpBEAS showed 74% similarity with the BbBEAS that synthesizes the cyclic trimeric ester beauvericin in *Beauveria bassiana*, which assembles *N*-methyl-dipeptidol monomer intermediates by the programmed iterative use of the nonribosomal peptide synthetase modules. Differences between the organization of the beauvericin loci in *F. proliferaturm* and *B. bassiana* revealed the mechanism for high production of beauvericin in *F. proliferatum*. Our work provides new insights into beauvericin biosynthesis, and may lead to beauvericin overproduction and creation of new analogs via synthetic biology approaches.

beauvericin, *Fusarium proliferatum*, biosynthesis gene cluster, synthetic biology

Citation: Zhang T, Zhuo Y, Jia X P, et al. Cloning and characterization of the gene cluster required for beauvericin biosynthesis in *Fusarium proliferatum*. Sci China Life Sci, 2013, 56: 628−637, doi: 10.1007/s11427-013-4505-1

The incidence of invasive fungal infections has dramatically increased in the past two decades, especially in patients who are severely immunocompromised because of cancer chemotherapy, organ or bone marrow transplantation, or through immunodeficient viral infections [1,2]. Resistance to existing classes of drugs is also on the rise. In our previous work we showed that beauvericin (BEA) has dramatic synergistic antifungal activity with low dosages of ketoconazole (KTC) against diverse fungal pathogens both *in vitro* and *in vivo* [1]. Beauvericin belongs to the cyclic nonribosomal hexadepsipeptide family of natural products, originally isolated from entomopathogenic fungi, including *Beauveria bassiana* [3]. Its molecular structure consists of an alternating sequence of three *N*-methyl-L-phenylalanine and three D-hydroxyisovaleric acids. It has diverse biological functions, including insecticidal, synergistic antifungal [1], and cytotoxic activities [4].

Nonribosomal peptide synthesis occurs via a process of stepwise condensations by nonribosomal peptide synthetases (NRPSs) [5]. These are modularly organized multienzyme complexes in which each module is responsible for the elongation of proteinogenic and non-protein amino acids, as well as carboxyl and hydroxyacids. Each module of the NRPS system is composed of domains which are distinct catalytic folds with highly conserved core motifs, im-

†Contributed equally to this work
*Corresponding author (email: zhanglixin@im.ac.cn; lium@im.ac.cn)

portant for their catalytic activities [6]. A minimum of three domains are necessary for one NRPS elongation module: an adenylation (A)-domain for substrate recognition and activation; a peptidyl carrier protein (PCP)-domain that tethers the growing peptide chain and the incoming aminoacyl unit; and a condensation (C)-domain to catalyze peptide bond formation. Optional domains include epimerization (E), heterocyclization (Cy), and oxidation (Ox) domains, which may edit the enzyme-bound precursors or intermediates. The full-length NRPS product is normally released by a thioesterase (TE) domain giving rise to free acids, lactones, or lactams [7]. NRPSs that synthesize cyclooligomer peptides assemble oligopeptide monomer intermediates by the programmed iterative use of their modules; this differs from the classical NRPS paradigm described in bacteria [6], and these monomers are also then used in a recursive oligomerization and cyclization process [7].

As previously reported from the Zocher lab, the gene, *esyn1*, corresponding to the enniatin synthetase from *Fusarium equiseti* (*F. scirpi*), encoding an ORF of 9393 bp (3131 amino acids) has been isolated and characterized [8,9]. Beauvericin biosynthesis is catalyzed by the beauvericin synthetase via a nonribosomal mechanism. The cloning and enzymatic characterization of the beauvericin synthetase from *B. bassiana* was described by Peeters et al. [10] and Xu et al. [11]. The beauvericin biosynthesis genes are clustered with the ketoisovalerate reductase (encoded by the *kivr* gene), which is responsible for catalyzing the NADPH-specific reduction of ketoisovaleric acid to hydroxyisovalerate [11,12]. In addition to *Beauveria*, the genera *Isaria* and *Fusarium* constitute a large group also possessing beauvericin-production capabilities. Those in *Fusarium* are greatly superior in beauvericin production. There is an impressive series of publications on the fermentation and higher productivity of beauvericin derived from *Fusarium* species. It has been reported that the titer of beauvericin produced by *F. denticulatum* could be 3000 μg g^{-1} of dry culture [13]. Even in liquid *Fusarium*-defined medium, *F. oxysprum* KFCC 11363P could still produce beauvericin at a level of 420 mg L^{-1}; much higher than production in *B. bassiana* [14,15]. However, the beauvericin synthetase and corresponding encoding gene cluster in *Fusarium* has not been characterized. To gain greater insight into the unique and intriguing features of beauvericin biosynthesis in *Fusarium*, we undertook both genetic and biochemical studies. We describe here the identification and functional analysis of the beauvericin biosynthetic gene cluster in *F. proliferatum* LF061. Sequence analysis revealed the presence of a nonribosomal peptide synthetase and a pair of auxiliary pathogenicity genes upstream of beauvericin synthetase. We also described biochemical experiments supporting the proposed functions of these encoded biosynthetic enzymes.

1 Materials and methods

1.1 Strains, media, and genetic manipulation

All *Escherichia coli* strains were grown in Luria-Bertani (LB) medium. *Fusarium proliferatum* LF061 was maintained on potato dextrose agar (PDA). Isolation of *Fusarium* genomic DNA was carried out using the modified protocol from Moller et al. [16]. The genomic library of *F. proliferatum* was constructed according to the manufacturer's instructions (Epicentre Technologies Corp., Wisconsin, USA, Cat. No. CCFOS059). *Agrobacterium tumefaciens* AGL-1 with the plasmid pRF-HU2 [17] was used for *Fusarium* transformation. Antibiotic concentrations routinely used were ampicillin 100 μg mL^{-1}, chloramphenicol 12.5 μg mL^{-1}, kanamycin 50 μg mL^{-1}, and hygromycin 50 μg mL^{-1}.

1.2 Library screening, sequencing, and bioinformatics analysis

Two sets of degenerate primers were used to amplify A-domain homologous sequences using *F. proliferatum* total DNA as the template. The primer pairs EA1f (5′-GCYTAYGTYATYTTYACYWSYGG-3′) and EA1r (5′-TTYGAYGCYWSYGTYTGG-3′), and EA2f (5′-CCKC-GRATYTTRACYTGRSWRTCCAT-3′) and EA2r (5′-CARATRGTRGTYTCRGTDGGRCCRTA-3′) (S=C, G; Y=C, T; B=C, G, T; H=A, C, T; R=A, G; W=A, T; D=A, G, T), were based on the A3, A8, A4, and A5 motifs of typical NRPS adenylation domains. This was confirmed by sequencing and BLAST searching. To screen the positive clones, colony PCR was performed using a pair of specific primers (Table 1).

DNA sequencing was conducted by SinoGenoMax Co., Ltd (Beijing). Sequencher 4.9 (Gene Codes Corporation, Michigan, USA) and Vector NTI suite 10.0 (Invitrogen, San Diego, USA) were used for sequence assembly and analysis. Gene products of the individual ORFs were deduced and analyzed using the programs AUGUSTUS [18,19] and FGENESH (http://linux1.softberry.com/berry.phtml?topic=fgenes_plus&group=programs&subgroup=gfs). Gene models were manually curated using exon/intron boundary predictions from SPLICEVIEW (http://zeus2.itb.cnr.it/~webgene/wwwspliceview_ex.html) and multiple sequence alignments. The UMA algorithm was used for the prediction of domain boundaries in NRPSs [20], and A domain specificity signatures were derived by NRPS predictor [21]. TMHMM (http://www.cbs.dtu.dk/services/TMHMM) was used for secondary structure analysis. Deduced amino acid sequences of 17 NRPSs were subjected to protein phylogenetic analysis. A phylogenetic tree was generated using the neighbor joining method of Saitou and Nei [22] with MEGA 4.0 software [23]. The length of each branch pair represents the evolutionary distance between the sequences.

Table 1 Primers used in this study

Primer	Sequence 5′–3′	Description	Fragment size
EA-QF	CAACTGGTCGACCTA AGGGTGTCATGG	Fosmid library screening primers	
EA-QR	ATGAGGCCGTCGCCTACTCTGTACC	Fosmid library screening primers	750 bp
EA-SF	GCTTGGACCTCAGCAGTTCG	Fosmid library screening primers	
EA-SR	TCAACTTTCATTCCAAACGTCG	Fosmid library screening primers	515 bp
B-HRS1F	GGTCTTAAUCGTCGTCTCGTCTCACCT	*fpBeas* gene disruption primers	
B-HRS 1R	GGCATTAAUCCGCCTTGTCAATCTCAC	*fpBeas* gene disruption primers	1562 bp
B-HRS2F	GGACTTAAUGCCACGAAACTCATCTCA	*fpBeas* gene disruption primers	
B-HRS 1R	GGGTTTAAUTGCTGTTCCACCTCACTCT	*fpBeas* gene disruption primers	1463 bp

1.3 Gene disruption of *fpBeas* in *F. proliferatum* LF061

The disruption cassette contained the *hph* selectable marker cloned between two PCR-amplified DNA fragments covering the 5′- and 3′- regions of the *fpBeas* gene. The target region (1.4 kb) is located in the middle of the *fpBeas* encoding A_2 and partial AdoMet domain. Amplification of the two homologous recombination sequences (HRS) with primers that contain 5′-deoxyuridine extensions was performed with PfuTurbo®Cx Hotstart DNA polymerase (Stratagene, Cedar Creek, TX, US) for the USER (uracil-specific excision reagent) friendly cloning reactions [17]. pRF-HU2 contains two regions including a nicking enzyme (*Nt. Bbv*C I) and *Pac* I restriction enzyme sites, in which the two amplicons, B-HRS1 (1.4 kb) and B-HRS2 (1.5 kb), were designed for single step directional cloning. Thus, a B-HRS1-*hph*-*TtrpC*-B-HRS2 cassette was constructed, in which the B-HRS1 and B-HRS2 arms were functioning in homologous recombination. The *hph* encodes a hygromycin phosphotransferase and acts as a selection marker in *Fusarium*, and *TtrpC* is a tryptophan terminator from *Aspergillus nidulans*. This resulted in the construction of a vector containing the hygromycin-resistance cassette flanked by a 1.4 kb region (corresponding to the C_2 domain) and a 1.5 kb region (corresponding to the AdoMet domain and partial T_2 domain) of *fpBeas*. Colonies resistant to hygromycin at high concentrations were then grown in solid medium in Erlenmeyer flasks for 10 d. *F. proliferatum* colonies with hygromycin resistance were chosen for PCR identification and fermentation. Genomic DNA isolated from *F. proliferatum* transformants with hygromycin resistance was used for PCR identification. Two pairs of primers were used to identify the transformants: HygU and HygL for the hygromycin phosphotransferase gene; and geneT3-RF1 for the left border of integration locus.

1.4 Fermentation, extraction, and analysis of beauvericin

The ability of different strains to produce beauvericin was measured following culturing on maize grain. Erlenmeyer flasks (250 mL) were filled with 20 g of cracked maize and 11 mL of distilled water, and then autoclaved twice for 30 min at 121°C. After the flasks were autoclaved, an additional 11 mL of sterilized distilled water was added to each flask. After cooling, the flasks were inoculated with 1 mL of conidial suspension from the PDA culture containing 1×10^7 spores mL^{-1}. Cultures were incubated at 25°C in the dark for two weeks with shaking once a day. Beauvericin was extracted from *F. proliferatum* fermentations and analyzed by HPLC. The fermented cultures were extracted with 86% acetonitrile in water. The clarified supernatants were collected, as these contained more than 90% of the beauvericins (data not shown). The fraction was analyzed by an Agilent 1100 RP-HPLC system (ZORBAX SB C18 column, 4.6 mm×150 mm, pore size, 5 μm). The mobile phase was composed of 80% methanol and 20% water. The column was eluted with the mobile phase at a flow rate of 1 mL min^{-1} at room temperature and the elution was monitored at 210 nm.

1.5 Nucleotide sequence accession number

The nucleotide sequence of the beauvericin biosynthetic gene cluster (*bea*) of *F. proliferatum* LF061 has been deposited with GenBank under accession number JF826561.

2 Results

2.1 Cloning and sequence analysis of beauvericin biosynthetic locus

To isolate the beauvericin biosynthesis genes, we screened a *Fusarium proliferatum* LF061 genomic fosmid library with primers amplifying a BEAS gene. Two sets of degenerate oligonucleotide primers were designed: corresponding to conserved amino acid regions within adenylation domains of known NRPS synthetase (see Materials and methods). Further selectivity can be gained by amplifying fragments of unique functional domains from the NRPS genes. Thus, genomic DNA from *F. proliferatum* LF061 was used as the template in PCRs with A domain-selective PCR primers (EA1F and EA1R). These reactions afforded a 750 bp product, which was cloned into an *E. coli* vector and sequenced. The translated PCR product proved to be homologous with other fungal adenylation domains, and particularly resembled that of the enniatin synthetase from

Fusarium scirpi and *Gibberella pulicaris* (accession numbers ADB27871 and ACR78148, data not shown). A genomic fosmid library of strain LF061 was constructed with more than 50000 recombinant clones; each clone harbored about 40 kb of genomic DNA from *F. proliferatum* LF061. To clone a larger portion of the beauvericin biosynthetic gene cluster, nine fosmids were selected and the specific fragment (750 bp) was successfully amplified, showing that they might carry the beauvericin synthetase gene from strain LF061. Among these, three representative overlapping fosmids, fpBeas1#, fpBeas5#, and fpBeas8#, were selected for DNA sequencing using a shotgun method, and a ~43 kb contig centered on the NRPS-encoding gene was assembled and analyzed. The arrangement of the beauvericin biosynthetic gene cluster (*bea*) is shown in Figure 1A. The contig contains 15 ORFs and the overall GC content of the sequenced region was 49.16%. The sequence was analyzed by comparison with database sequences, and functional assignments of individual ORFs are deposited in GenBank under accession number JF826561 and summarized in Table 2. Consistent with the structure of beauvericin, those ORFs identified within the *bea* cluster included a 9413 bp gene (*fpBeas*) encoding a putative cyclooligomer depsipeptide synthetases (CODS), and a *kivr* gene that encodes a novel NADPH-dependent 2-ketoisovalerate reductase (KIVR), responsible for pyruvate metabolism to D-Hiv [12]. Sequence analysis suggested that the *fpBeas* gene has two exons (5575 and 3703 bp) and is interrupted by a 134 bp intron. The contig contains a further thirteen open reading frames at the left border of the sequenced region. BLAST analysis revealed that *orf1* codes a hypothetical thioesterase, and *orf3*, *orf4*, *orf5*, *orf6* encode putative secreted triacylglycerol lipase, chitinase, zinc-dependent metalloprotease, and putative muramidase. The translation of predicted *orf8* yielded a short chain type dehydrogenase, with function unclear. Immediately upstream of the KIVR-encoding gene, *orf10* codes for a 1301 aa polypeptide interrupted by ten introns; it displays high degree of similarity to transporters of the ABC-type multidrug transport system family [24], and thus might be involved in ATP-dependent efflux of

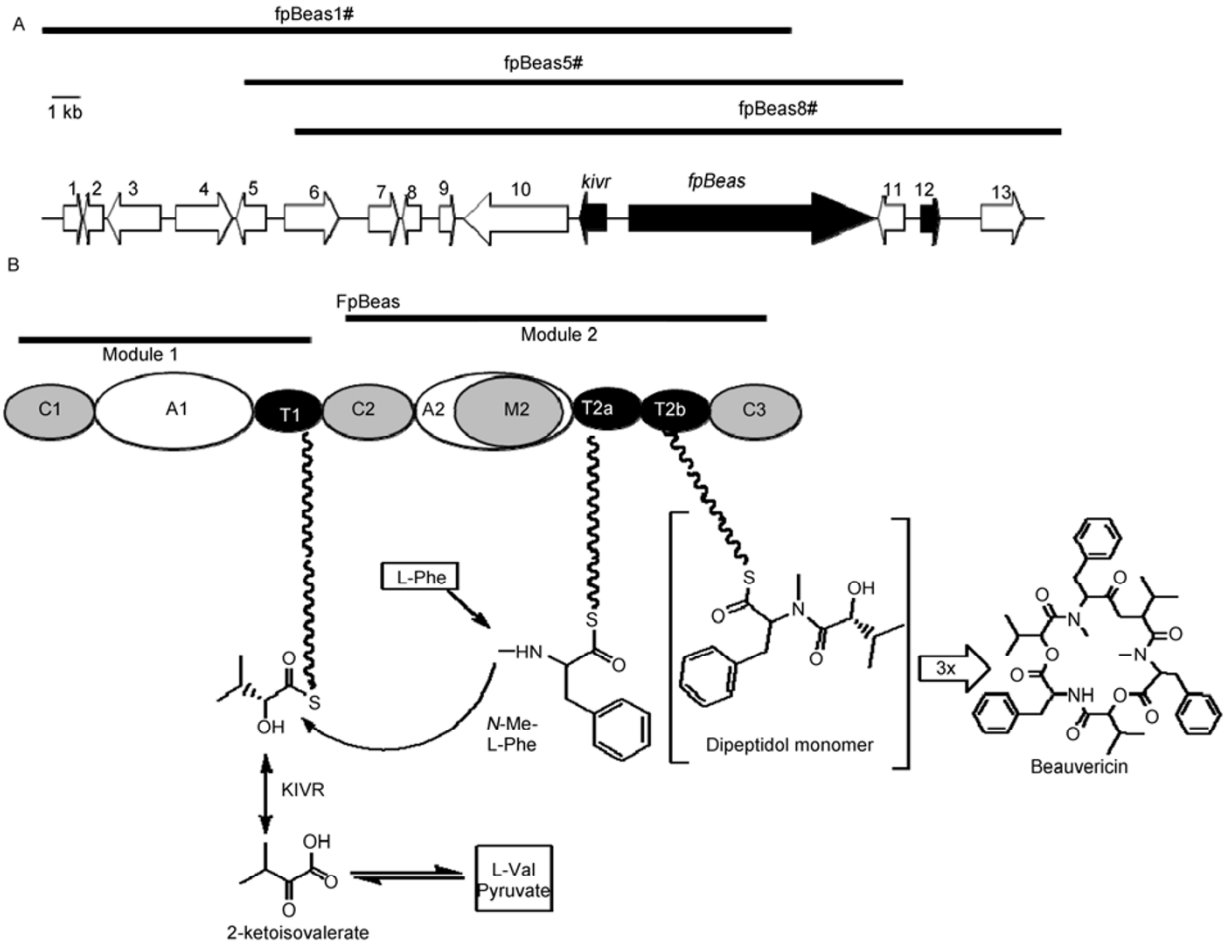

Figure 1 Diagrammatic representation of the *bea* biosynthetic locus and the proposed biosynthesis of beauvericin. A, ORF arrangement of the *bea* gene cluster and flanking regions. The black arrow represents *fpBeas*, *kivr*, and the regulator gene; genes flanking the cluster are in white, and the three sequenced fosmids are indicated as bold lines. B, Module organization and model of the biosynthesis of beauvericin from primary metabolites on the FpBEAS beauvericin synthetase (domains not to scale).

Table 2 Summary of the deduced proteins from the beauvericin locus

Gene modules (start-end, bp)	Protein length (aa)	Closest GenBank homolog, organism[a], function of the homolog	Identity/simiarity (%)	Conserved domain	*E*-value[b]
Orf1 (544–1272)	224	XP_001798883, *P. nodorum*, hypothetical protein	57/71	COG0824, FcbC, predicted thioesterase	2×10^{-67}
Orf2[c] (1396–2088)	230	EEU43407, *N. haematococca*, hypothetical protein	68/81	COG1670, RimL, putative acetyltransferase	1×10^{-85}
Orf3[c] (2449–4533)	655	EEU37338, *N. haematococca*, hypothetical protein	58/71	cd01823, SEST_like. A family of secreted SGNH-hydrolases, triacyl-glycerol lipase	1×10^{-161}
Orf4 (5226–7165)	544	EEU37337, *C. globosum*, hypothetical protein	65/79	COG3325, ChiA, Chitinase, ChtBD, Chitin binding domain	0.0
Orf5[c] (7517–8632)	371	EEU37337, *N. haematococca*, hypothetical protein	38/52	smart00235, Zinc-dependent metallo-protease	6×10^{-59}
Orf6 (10169–13032)	686	EEU37422, *N. haematococca*, hypothetical protein	59/74	LysM domain-containing protein, putative muramidase	2×10^{-91}
Orf7 (15201–16378)	374	EEU37421, *N. haematococca*, hypothetical protein	38/53	ND	3×10^{-53}
Orf8[c] (16467–17272)	235	EEU34292, *N. haematococca*, hypothetical protein	68/80	short chain type dehydrogenase	2×10^{-90}
Orf9 (18293–18847)	184	XP_385076, *G. zeae* PH-1 , hypothetical protein	47/61	ND	2×10^{-21}
Orf10[c] (19172–23623)	1301	XP_387057, *G. zeae*, hypothetical protein	57/74	COG1132, ABC-type multidrug transport system	0.0
kivr[c] (24722–25950)	390	ACI30654, *B. bassiana* ketoisovalerate reductase	60/75	COG1893, Ketopantoate reductase, putative 2-dehydropantoate 2-reductase	7×10^{-132}
fpBeas (27092–36504)	3092	ADB27871, *F. oxysporum*, cyclic peptide synthetase	66/80	cyclic peptide synthetase, Peptide synthetase, or enniatin synthase	0.0
Orf11[c] (36749–37828)	342	EEU34450, *N. haematococca,* hypothetical protein	59/72	Predicted membrane protein	7×10^{-114}
Orf12 (39528–40190)	220	EEU43647, *N. haematococca,* hypothetical protein	51/67	putative C6 zinc transcription factor (PRO1)	2×10^{-26}
Orf13 (42373–443979)	479	EEU43611, *N. haematococca,* hypothetical protein	72/82	flavin containing amine oxidase	0.0

a) Genus abbreviations: *P*, *Phaeosphaeria*; *N*, *Nectria*; *C*, *Chaetomium*; *G*, *Gibberella*; *B*, *Beauveria*; *F*, *Fusarium*. b) *E*-value for the match against the conserved domain. c) Reverse complement. ND, not detected.

beauvericin. Secondary structure analysis revealed that ORF10 contained eleven membrane-spanning domains. At the right border of the sequenced region, the deduced gene product of *orf13* resembles flavin-containing amine oxidase. *orf12* encodes a putative Gal4-like transcriptional regulator with a $Zn(II)_2Cys_6$ binuclear cluster DNA-binding domain, and might thus be associated with a regulatory role. Two conserved hypothetical proteins (ORF7 and ORF9) and one deduced protein that might be involved in membrane transport (ORF11) are also encoded in the sequenced locus.

2.2 Inactivation of the *fpBeas* gene in *F. proliferatum* LF061

To further identify the function of the deduced FpBeas, an internal fragment of the *fpBeas* gene was replaced in the genome of *F. proliferatum* LF061 with the *hph* as a selectable marker gene encoding resistance to hygromycin. Hygromycin-resistant *F. proliferatum* colonies (≈130) were picked out and subcloned onto solid medium containing increasing concentrations of hygromycin. Mutants that had undergone double homologous recombination events were identified by PCR (Figure 2A and B). The ~600 bp fragment can be amplified from the genomic DNA of both mutants (m1 and m2), but it did not appear in that of the wild type (W). Furthermore, in mutant 1 (m1), geneT3-RF1 primers gave a 1.9 kb product, while neither mutant 2 (m2) nor the wild type could be amplified. The result indicated that *fpBeas* was inactivated in the mutant 1, while mutant 2 is an ectopic integrant with the *hph* gene insertion at an unusual locus. The ectopic integration of the *hph* gene in some other place in the genome may disrupt the regulatory genes or some other related genes, and this may to some extent affect beauvericin biosynthesis. We have attempted to rescue the *fpBeas* disruption construct in m2 and in several other ectopic integrant strains, but the integration loci in the genomes of those strains were not consistent (data not shown). Both mutants were further fermented under beauvericin-producing conditions, and the fermentation extracts were analyzed by HPLC (as shown in Figure 2C). When compared with the wild type strain LF061 (top), production of beauvericin was abolished in m1 strain (middle). In contrast, m2 strain can produce low levels of beauvericin (bottom). The data coincide with the molecular identification results, which show that the *fpBeas* gene is responsible for the biosynthesis of beauvericin. As expected, serial passages of the *fpBeas* gene knockout isolates on nonselective media show that the characters of beauvericin production are stable (results not shown).

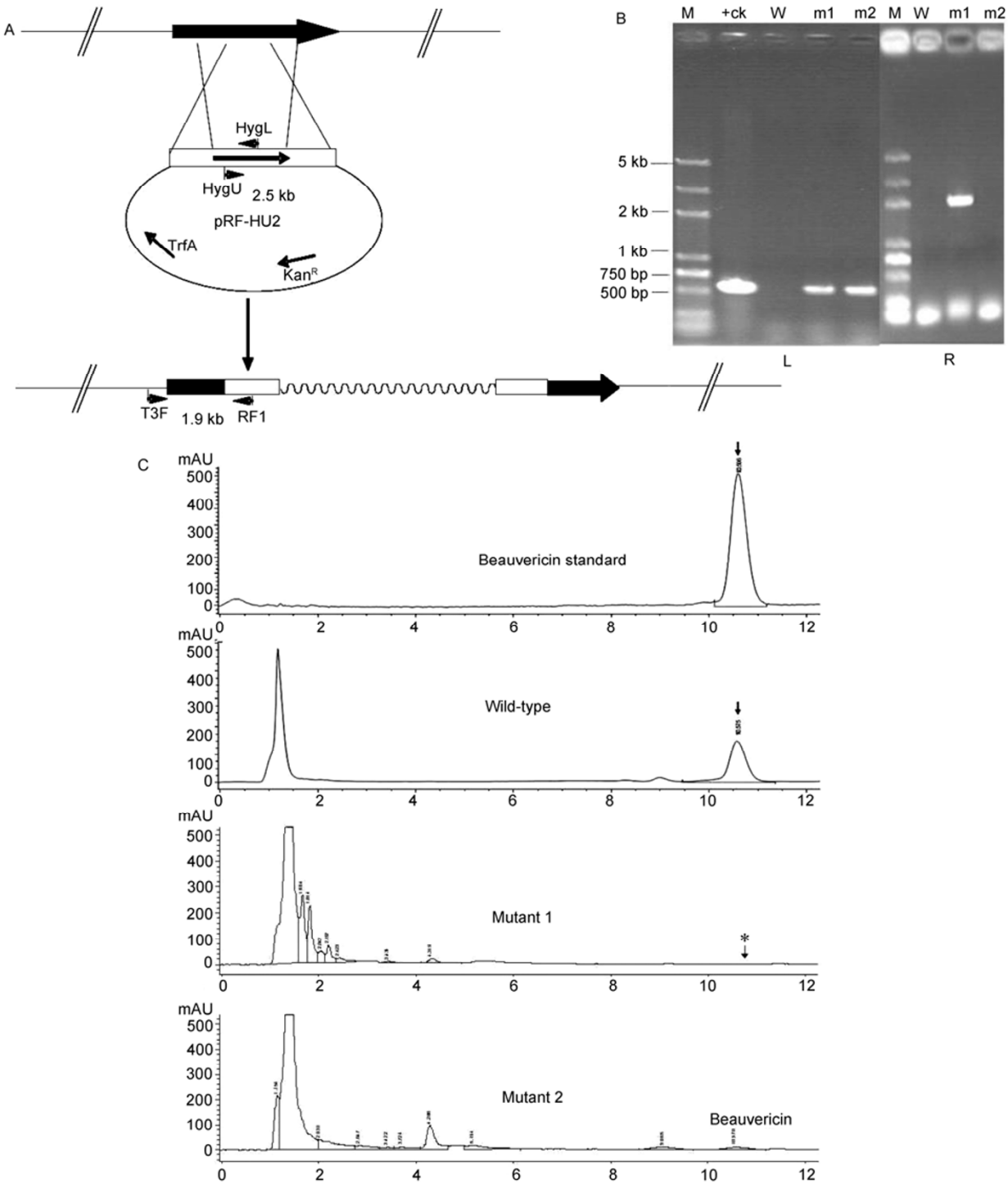

Figure 2 Inactivation of FpBEAS in *F. proliferatum* LF061, PCR identification, and HPLC analysis of the mutant strain. A, Double crossover gene replacement to inactivate the *fpBeas* gene (black arrow). The top line shows the genes in the wild type strain. The bottom line shows the mutants, in which the *fpBeas* gene was inactivated by the hygromycin cassette. The angled arrows indicate the PCR primers that were used to amplify the DNA fragments shown in (B). B, Ethidium bromide-stained agarose gels showing the PCR analysis of the mutants. Lanes in gel: M, marker DL2000 plus (cas. No. BM111, Transgen); +ck, *Agrobacterium tumefaciens* AGL-1 mutant which has the pRF-HU2 plasmid; W, genomic DNA of *Fusarium proliferatum* LF061; m1, mutant 1; m2 mutant 2. Left side shows amplification of the hygromycin-resistance gene in mutants using primers HygU and HygL (588 bp); right side shows amplification of the left border in mutants using primers gene-T3 and RF1 (1.9 kb). C, HPLC analysis of beauvericin production in *F. proliferatum* LF061 (wild-type) and the mutants. Mutant 1, a representative isolate of the *fpBeas* knockout mutant; mutant 2, a ectopic integration mutant. *, the expected position of the beauvericin peak.

2.3 Bioinformatic analysis of the beauvericin synthetase

Our sequence analysis showed that FpBEAS belongs to the growing family of nonribosomal peptide synthetases. The coding sequence of the peptide synthetases spans 9276 bp corresponding to 3092 amino acids. The deduced molecular mass of FpBEAS (343181 Da) is almost the same as in earlier reports [8,9,14]. FpBEAS shows 64% identity and 79% similarity to ESYN (GenBank accession No. CAA79245), and close homology to the nonribosomal cyclodepsipeptide synthetases of *B. bassiana* (ACR78148, 61% identity and 76% similarity; ACI30655, 60% identity and 74% similarity). A similar level of identity is also noted with the NRPS-homologous part of the bassianolide synthetase of *Xylaria* sp. BCC 1067 (GenBank accession No. ABR28366). Phylogenetic analysis (Figure 3A) revealed that FpBEAS and other fungal NRPSs constitute a different branch from the bacterial group in the tree, and the group includes ESYN [8], BbBEAS [14], BbBSLS [25], and NRPSXY [26]. The overall sequence similarity between the fungal and prokaryotic groups is in the range of 17%–25%. This implies that the two branches evolved from different ancestors.

All the characterized fungal cyclooligomer depsipeptide synthetases (CODS), such as beauvericin, bassianolide, and enniatin synthetases, feature identical module and domain organization: $C_1A_1T_1$-$C_2A_2M_2T_{2a}T_{2b}$-C_3 (Figure 1B). The D-Hiv is activated by the A_1 domains within the first modules of these CODSs and installed as a thioester onto the T_1 domains of the same modules. The L-Phe is specifically activated by the A2 domain of FpBEAS and BbBEAS beauvericin synthetase, while branched-chain amino acids are activated by the A_2 domains of enniatin and bassianolide synthase, and loaded onto the twin T_2 domains within module 2. There is an integrated *N*-methyltransferase domain in beauvericin synthetases, which is responsible for the editing function. It has also been found in the enniatin synthetases [27]. The *N*-methyltransferase domains (M_2) of the CODSs display a high level of conservation among *N*-methyl peptide synthetases from prokaryotic and eukaryotic origins [5], and are inserted into a flexible loop region between core motifs 8 and 9 of the A_2 domains (Figure 1B). This "stuffer" region often accommodates different "editing" domains in diverse NRPSs; in the editing process, S-adenosyl-L-methionine (AdoMet) acts as the methyl group donor. The condensation of D-Hiv and the *N*-Me-amino acid to form dipeptidol is catalyzed by domain C_2 in CODSs (Figure 1B). An N-terminal C_1 and a C-terminal C_3 domain are found in the CODSs, which might play a role in the successive cyclooligomerization process [9,28]. The C_1 and C_3 domains show overall similarity to condensation domains, but their core motifs are substantially divergent from the canonical forms [5,29]. However, they retain a modified form of the active His-motif signature, SHALVDT in C_1 and SHALYDG in C_3 domains versus HHXXXDG for the canonical motif, in which the second His-residue and the conserved Asp appears essential for the catalytic function of the domain [30].

2.4 Comparative analysis of the beauvericin biosynthetic locus in *F. proliferatum* and *B. bassiana*

The organization of the *F. proliferatum bea* gene cluster (Table 3) shows a striking similarity to that of the *B. bassiana bbBeas* gene cluster. Although other organisms also tend to group these biosynthetic clusters, *F. proliferatum* and *B. bassiana* are the only known organisms to group these genes into a single operon. Hence, we compared in detail the gene organization of the *fpBeas* gene cluster with that of the *bbBeas* (Figure 3B). Presumably, the *bea* operon is similarly arranged, although only *kivr* and *fpBeas* have been identified to date, demonstrating that the arrangement of these genes is well conserved between these two fungal species. On average they exhibited 60% identity between the *kivr* gene of *B. bassiana* and that of *Fusarium proliferatum*, and 61% identity between *bbBeas* and *fpB*eas. There are still some differences between the organization of the beauvericin locus in *F. proliferaturm* and *B. bassiana*. First, the *fpBeas* gene cluster shows a more integrative and complicated function than the *bbBeas* biosynthetic cluster. The detailed functions of the *orf1*, *orf3*, *orf4*, *orf5*, and *orf6* gene products have not yet been characterized, however, they show similarity to thioesterase, lipase, chitinase, zinc-dependent metalloprotease, and putative muramidase, respectively, suggesting these lytic enzymes might provide entry into the host by degradation of the cuticle [31]. These auxiliary pathogenicity factor-encoding genes are clustered together with beauvericin biosynthetic genes in *Fusarium*. Second, in the left border of the *bbBeas* cluster, *orf1* encodes a putative Gal4-like transcriptional regulator with a $Zn(II)_2Cys_6$ binuclear cluster DNA-binding domain. However, a putative Gal4-like transcriptional regulator that is smaller in size (220 aa) encoding the *orf12* gene exists in the right border of the *fpBeas* biosynthetic cluster, downstream of the nonribosomal peptide synthetase FpBEAS-encoding gene. Similar regulators are often encoded in secondary metabolic biosynthetic gene clusters of fungi [14,32]. The regulation of secondary metabolic gene clusters in fungi occurs at several levels, some of which are specific to the respective biosynthetic pathway, and others that function as global regulators [33]. Zinc cluster transcriptional regulators function in a wide range of processes, including regulation of genes involved in stress responses, pleiotropic drug resistance, amino acid and vitamin synthesis, and carbon and nitrogen metabolism [34]. This indicates that engineering of the Gal4-like transcriptional regulator could enhance the production of beauvericin in *F. proliferaturm* LF061 if synthetic biology approaches are used. Third, in the region upstream of *kivr*, *orf10* encodes an ABC-type multidrug transporter, which could retain immunotolerance to beau-

Table 3 Specificity-determining signatures of NRPS adenylation domains[a)]

Synthetase[b)]	A Domain	Specificity[c)]	235[d)]	236	239	278	299	301	322	330	331	517
Fungal D-Hiv signature[e)]			G	A	L	x	I/V	V	G	S/T	I	K
FpBEAS	A_1	D-Hiv	G	A	L	M	V	V	G	S	I	K
BbBEAS	A_1	D-Hiv	G	A	L	M	I	V	G	S	I	K
FeESYN	A_1	D-Hiv	G	A	L	H	V	V	G	S	I	K
BbBSLS	A_1	D-Hiv	G	A	L	M	V	V	G	S	I	K
XsBSLS	A_1	D-Hiv	G	A	L	L	V	V	G	T	I	K
PFSYN	A_1	D-(Ph)Lac	G	A	V	I	V	V	A	T	M	K
StVlml	A_1	Kiv	A	A	L	W	I	A	V	S	G	K
BcCesB	A_1	Kiv	V	G	F	W	V	A	V	S	D	K
KsKtzG	A_1	Kiv	V	T	Y	F	N	G	P	S	G	K
Fungal Phe signature			D	A/G	W	I/L	I/L	A/G	A/G	I/V	C	K
FpBEAS	A_2	Phe	D	G	Y	C	M	A	G	A	L	K
BbBEAS	A_2	Phe	D	G	Y	I	M	A	A	V	M	K
AfPSES	A_1	Phe	D	A	Y	T	M	A	A	V	G	K
CpEALS	A_2	Phe	D	L	V	G	M	A	A	V	G	K
AfGliP	A_2	Phe	D	Y	N	T	Y	T	A	I	C	K
BbBSLS	A_2	Leu	D	G	Y	I	I	G	G	V	F	K
XsBSLS	A_2	Leu	D	A	W	L	V	G	A	V	M	K
FeESYN	A_2	V(LI)	D	G	W	F	I	G	I	I	I	K
FsESYN	A_2	I(LV)	D	G	W	F	A	G	V	M	I	K

a) Amino acids identical to those in the A-domain signatures of FpBEAS are shown in black font on a grey background. Amino acids similar (V=I=L, A=G, S=T, W=Y=F) to those in the A-domain signatures of FpBEAS are shown in bold type. b) NRPS abbreviations (and GenBank accession numbers): *B. bassiana* beauvericin synthetase (EU886196); FeESYN, *F. equiseti* enniatin synthetase (CAA79245); BbBSLS, *B. bassiana* bassianolide synthetase(FJ439897); XsBSLS, *Xylaria* sp. bassianolide synthetase (ABR28366); PFSYN, *Mycelia sterilia* PF1022 synthetase (BD013055); StVlm1, *Streptomyces tsusimaensis* valinomycin synthetase subunit 1 (ABF61888); BcCesB, *Bacillus cereus* cereulide synthetase B subunit (ABK00751); KsKtzG, *Kutzneria* sp. kutzneride synthetase G subunit (ABV56585); AfPSES, *A. fumigatus* pseurotin A synthetase (ABS87601); CpEALS, *Claviceps purpurea* ergot alkaloid synthetase (O94205); AfGliP, *A. fumigatus* gliotoxin synthetase (ABE60889); FsESYN, *F. sambucinum* enniatin synthetase (CAA88634). c) A-domain specificities: D-Hiv, D-2-hydroxyisovalerate; Kiv, 2-ketoisovalerate; V(LI), valine preferred, leucine and isoleucine accepted; I(LV), isoleucine preferred, leucine and valine accepted. d) Numbering according to PheA. e) Specificity signatures were derived by NRPSpredictor.

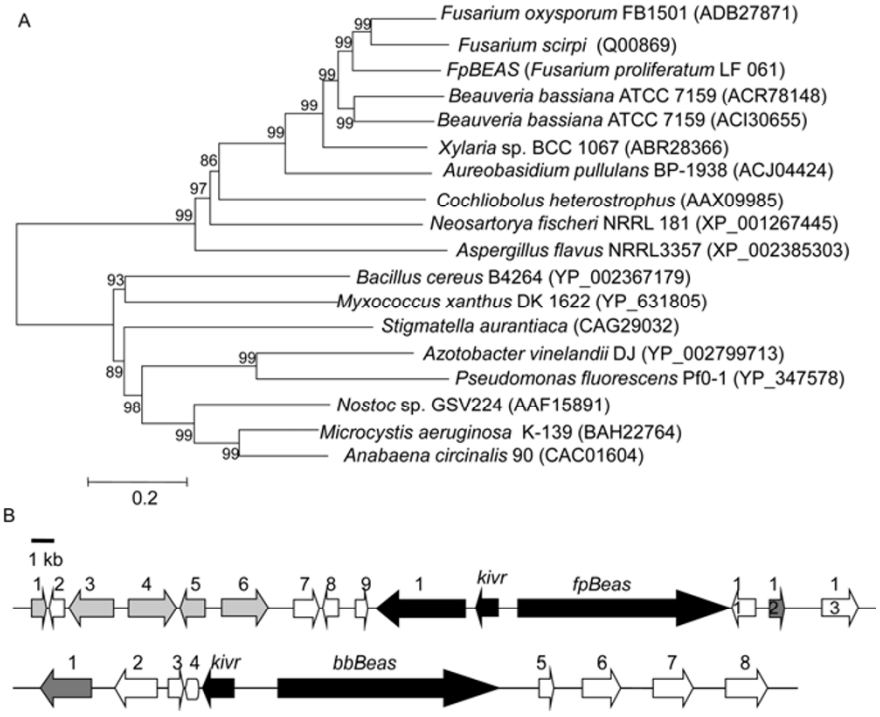

Figure 3 Comparative sequence analysis of the deduced FpBEAS beauvericin synthetase. A, A phylogenetic tree was created based on amino acid sequences of nonribosomal peptide synthetase (NRPS) showing the phylogenetic position of FpBEAS and its counterparts from other origins. Sequence accession numbers and strain names are shown. Bootstrap values (expressed as percentages of 1000 replications) >50% are shown at branch points. B, Comparison of the *fpBeas* biosynthetic cluster of *F. proliferatum* LF061 with the *bbBeas* locus of *Beauveria bassiana*. Black and charcoal-grey arrows indicate homologous regions (domains not to scale).

vericin. Furthermore, the ABC-type multidrug transporter-containing cluster gave more than ten-fold-higher beauvericin titers than the BEA biosynthetic cluster in *B. bassiana*. In contrast, there is no counterpart in the *bbBeas* gene cluster. This might be the cause of the higher productivity of beauvericin seen in the genus *Fusarium*. Manipulating the transporter gene may trigger further enhancement of beauvericin production.

3 Discussion

In an effort to generate a more detailed understanding of the mechanisms of cyclooligomer depsipeptide biosynthesis in *Fusarium*, a ~43 kb fragment of marine fungus *F. proliferaturm* LF061 genomic DNA was cloned and shown to be involved in beauvericin biosynthesis. The synthesis of cyclooligomer depsipeptides by both filamentous fungi and bacteria significantly diverges from the collinear and processive logic of canonical NRPSs [5,14]. Therefore, CODSs possess a parallel synthetic scheme that involves the programmed iterative mechanism of dipeptidol monomer intermediates from amino acid and D-2-hydroxycarboxylic acid precursors. It has been demonstrated that sequence alignments with two structurally characterized A domains can predict the amino acids that line the active site cavities of A domains in NRPSs [5,35]. We analyzed the 10-amino acid signatures of the fungal CODSs (Table 3). Accordingly, the first position of the specificity signatures clearly distinguished the D-Hiv-specific A_1 domains of these NRPSs from the amino acid-specific A_2 domains (Gly in A_1 domains, Asp in A_2 domains). Further, this is in accordance with the fungal A-domain specificity signature, GALx(I/V)VG(S/T)IK, for the 2-hydroxycarboxylate precursor D-Hiv. A similar comparison of the A_2 domain specificity signature with the fungal A domains was also unproductive. We propose a somewhat degenerate Phe specificity signature for fungi supported by the A_2 domain of *B. bassiana* beauvericin synthetase, A_1 domain of *A. fumigatus* pseurotin A synthetase, A_2 domain from *Claviceps purpurea* ergot alkaloid synthetase, and *A. fumigatus* gliotoxin synthetase, all of which activate Phe. Whereas the signature for the Phe-activating A domain of BbBEAS and AfPSES are rather similar to that of FpBEAS, the signature of AfGliP and CpEALS are highly different. This discrepancy might reflect the existence of synonymous nonribosomal "codons" for Phe activation in fungal A domains: similar divergence has also been noted in bacterial NRPSs [5,36]. The resulting convergent evolution would lead to a higher degeneracy of the code and might confound *a priori* predictions. Alteration of substrate specificity by such "keyhole surgery" on the A domain has been demonstrated in the lab [6].

The physiological functions of secondary metabolites in the producing organism are often obscure or unknown [32]. The pathogenicity of *Fusarium* isolates does not appear to be related to the origin of the organism as beauvericin-producing strains have been found in both terrestrial and aquatic habitats [37]. Conidia of the necrotrophic fungus *Botrytis cinerea* are able to adhere to plant surfaces by secreting an extracellular matrix, allowing them to germinate and penetrate cuticles [38]. Investigations into the pathogenicity factors of the insect pathogen *B. bassiana* have focused on lytic enzymes that provide entry into the insect host by degrading the cuticle [31]. In addition to these enzymes, the *bea* cluster also encodes chitinase, zinc-dependent metalloprotease, and putative muramidase. Thus these lytic enzymes are believed to loosen the tight structure of the host and strengthen pathogen persistence, although we have no unequivocal experimental evidence. Further investigation would unveil the specific purpose of these lytic enzymes. We propose herein that beauvericin is a significantly contributing, if not indispensable, factor for the pathogenicity of *F. proliferatum* LF061; supporting endophytic and saprophytic lifestyles that may demand a secondary metabolome with diverse biotic effectors and localized stress regulator activities [39–41].

We are very grateful for the technical support provided by Dr. Rasmus JN Frandsen, in our Fusarium-disruption experiment as well as providing the pRF-HU2 plasmid. This work was supported by the National Basic Research Program of China (2013CB734000, 2012CB725200), the CAS Pillar Program (XDA04074000), and Ministry of Science and Technology of China (2011ZX11102-011-11). ZHANG LiXin is an Awardee for the National Distinguished Young Scholar Program in China.

1 Zhang L, Yan K, Zhang Y, et al. High-throughput synergy screening identifies microbial metabolites as combination agents for the treatment of fungal infections. Proc Natl Acad Sci USA, 2007, 104: 4606–4611
2 Hoffman H L, Ernst E J, Klepser M E. Novel triazole antifungal agents. Exp Opin Invest Drugs, 2000, 9: 593–605
3 Hamill R L, Higgens C E, Boaz H E, et al. The structure of beauvericin, a new depsipeptide antibiotic toxic to *Artemia salina.* Tetrahedron Lett, 1969, 49: 4255–4258
4 Roeske R W, Isaac S, Steinrau L, et al. Synthesis and ion-transport properties of peptide antibiotic beauvericin. Fed Proc, 1971, 30: 1282–1288
5 Schwarzer D, Finking R, Marahiel M A. Nonribosomal peptides: from genes to products. Nat Prod Rep, 2003, 20: 275–287
6 Finking R, Marahiel M A. Biosynthesis of nonribosomal peptides. Ann Rev Microbiol, 2004, 58: 453–488
7 Kopp F, Marahiel M A. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. Nat Prod Rep, 2007, 24: 735–749
8 Pieper R, Haese A, Schroder W, et al. Arrangement of catalytic sites in the multifunctional enzyme enniatin synthetase. Eur J Biochem, 1995, 230: 119–126
9 Haese A, Schubert M, Herrmann M, et al. Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalyzing n-methyldepsipeptide formation in fusarium-scirpi. Mol Microbiol, 1993, 7: 905–914
10 Peeters H, Zocher R, Kleinkauf H. Synthesis of beauvericin by a multifunctional enzyme. J Antibiotics, 1988, 41: 352–359
11 Xu Y, Orozco R, Wijeratne E M, et al. Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus beauveria bassiana. Chem Biol, 2008, 15: 898–907

12 Zhang T, Jia X, Zhuo Y, et al. Cloning and characterization of a novel 2-ketoisovalerate reductase from the beauvericin producer fusarium proliferatum lf061. BMC Biotechnol, 2012, 12: 55
13 Moretti A, Mule G, Ritieni A, et al. Further data on the production of beauvericin, enniatins and fusaproliferin and toxicity to artemia salina by fusarium species of gibberella fujikuroi species complex. Int J Food Microbiol, 2007, 118: 158–163
14 Xu Y Q, Orozco R, Wijeratne E M K, et al. Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus beauveria bassiana. Chem Biol, 2008, 15: 898–907
15 Lee H S, Song H H, Ahn J H, et al. Statistical optimization of growth medium for the production of the entomopathogenic and phytotoxic cyclic depsipeptide beauvericin from fusarium oxysporum kfcc11363p. J Microbiol Biotechnol, 2008, 18: 138–144
16 Moller E M, Bahnweg G, Sandermann H, et al. A simple and efficient protocol for isolation of high-molecular-weight DNA from filamentous fungi, fruit bodies, and infected-plant tissues. Nucleic Acids Res, 1992, 20: 6115–6116
17 Frandsen R J N, Andersson J A, Kristensen M B, et al. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. BMC Mol Biol, 2008, 9: 70
18 Stanke M, Morgenstern B. Augustus: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res, 2005, 33: W465–W467
19 Stanke M, Steinkamp R, Waack S, et al. Augustus: a web server for gene finding in eukaryotes. Nucleic Acids Res, 2004, 32: W309–W312
20 Udwary D W, Merski M, Townsend C A. A method for prediction of the locations of linker regions within large multifunctional proteins, and application to a type I polyketide synthase. J Mol Biol, 2002, 323: 585–598
21 Rausch C, Weber T, Kohlbacher O, et al. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res, 2005, 33: 5799–5808
22 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol, 1987, 4: 406–425
23 Tamura K, Dudley J, Nei M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol, 2007, 24: 1596–1599
24 Sharom F J. Abc multidrug transporters: structure, function and role in chemoresistance. Pharmacogenomics, 2008, 9: 105–127
25 Xu Y Q, Rozco R, Wijeratne E M K, et al. Biosynthesis of the cyclooligomer depsipeptide bassianolide, an insecticidal virulence factor of beauveria bassiana. Fungal Genet Biol, 2009, 46: 353–364
26 Jirakkakul J, Punya J, Pongpattanakitshote S, et al. Identification of the nonribosomal peptide synthetase gene responsible for bassianolide synthesis in wood-decaying fungus *xylaria* sp. BCC1067. Microbiol-SGM, 2008, 154: 995–1006
27 Billich A, Zocher R. N-methyltransferase function of the multifunctional enzyme enniatin synthetase. Biochemistry, 1987, 26: 8417–8423
28 Glinski M, Urbanke C, Hornbogen T, et al. Enniatin synthetase is a monomer with extended structure: evidence for an intramolecular reaction mechanism. Arch Microbiol, 2002, 178: 267–273
29 Rausch C, Hoof I, Weber T, et al. Phylogenetic analysis of condensation domains in nrps sheds light on their functional evolution. BMC Evol Biol, 2007, 7: 78
30 Sieber S A, Marahiel M A. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem Rev, 2005, 105: 715–738
31 Fan Y H, Fang W G, Guo S J, et al. Increased insect virulence in beauveria bassiana strains overexpressing an engineered chitinase. Appl Environ Microbiol, 2007, 73: 295–302
32 Yu J H, Keller N. Regulation of secondary metabolism in filamentous fungi. Ann Rev Phytopathol, 2005, 43: 437–458
33 Brakhage A A. Regulation of fungal secondary metabolism. Nat Rev Microbiol, 2013, 11: 21–32
34 MacPherson S, Larochelle M, Turcotte B. A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol Mol Biol Rev MMBR, 2006, 70: 583–604
35 Challis G L, Ravel J, Townsend C A. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol, 2000, 7: 211–224
36 vonDohren H, Keller U, Vater J, et al. Multifunctional peptide synthetases. Chem Rev, 1997, 97: 2675–2705
37 Palmero D, Iglesias C, de Cara M, et al. Species of fusarium isolated from river and sea water of southeastern spain and pathogenicity on four plant species. Plant Dis, 2009, 93: 377–385
38 Doss R P, Potter S W, Chastagner G A, et al. Adhesion of nongerminated botrytis-cinerea conidia to several substrata. Appl Environ Microbiol, 1993, 59: 1786–1791
39 Turgeon B G, Oide S, Bushley K. Creating and screening cochliobolus heterostrophus non-ribosomal peptide synthetase mutants. Mycol Res, 2008, 112: 200–206
40 Wang L, Tian X, Wang J, et al. Autoregulation of antibiotic biosynthesis by binding of the end product to an atypical response regulator. Proc Natl Acad Sci USA, 2009, 106: 8617–8622
41 Xu G, Wang J, Wang L, et al. Pseudo gamma-butyrolactone receptors respond to antibiotic signals to coordinate antibiotics biosynthesis. J Biol Chem, 2010, 285: 27440–27448