J. Microbiol. Biotechnol. (2017), 27(3), 439–449
https://doi.org/10.4014/jmb.1606.06056

jmb

The Stress-Responsive and Host-Oriented Role of Nonribosomal Peptide Synthetases in an Entomopathogenic Fungus, *Beauveria bassiana* S

Hang Liu[1†], Linan Xie[1†], Jing Wang[1], Qiannan Guo[1], Shengnan Yang[1], Pei Liang[2], Chengshu Wang[3], Min Lin[1], Yuquan Xu[1], and Liwen Zhang[1*]

[1]*Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R. China*
[2]*Department of Entomology, China Agricultural University, Beijing 100083, P.R. China*
[3]*Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, P.R. China*

Received: June 27, 2016
Revised: October 18, 2016
Accepted: November 13, 2016

First published online
November 14, 2016

**Corresponding author*
Phone: +86-10-82105498;
Fax: +86-10-82106106;
E-mail: zhangliwen@caas.cn

†These authors contributed equally to this work.

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872

Beauveria bassiana infects a number of pest species and is known to produce insecticidal substances, such as the nonribosomal peptides (NRPs) beauvericin and bassianolide. However, most NRPs and their biological roles in *B. bassiana* remain undiscovered. To identify NRPs that potentially contribute to pathogenesis, the 21 predicted NRP synthetases (NRPSs) or NRPS-like proteins of *B. bassiana* ARSEF 2860 were primarily ranked into three functional groups: basic metabolism (7 NRPSs), pathogenicity (12 NRPSs), and unknown function (2 NRPSs). Based on the transcript levels during in vivo growth on diamondback moth (*Plutella xylostella* (Linnaeus)), half of the Group II NRPSs were likely to be involved in infection. Given that the metabolites biosynthesized by these NRPSs remain to be determined, our result underlines the importance of the NRPSome in fungal pathogenesis, and will serve as a guide for future genomic mining projects to discover functionally essential and structurally diverse NRPs in fungal genomes.

Keywords: Nonribosomal peptide synthetase, secondary metabolites, entomopathogen

Introduction

Fungi are well known for their production of a wide range of secondary metabolites (SMs) [1]. Although often not required for life or propagation, SMs have various biological activities, such as antibiotics (*e.g.*, penicillin), lowering cholesterol (*e.g.*, lovastatin), and mediators of pathogenesis (*e.g.*, beauvericin and oosporein [2, 3]). Fungi mainly produce the nonribosomal peptides (NRPs) and polyketides (PKs), in contrast to the terpenes and alkaloids of plants [1]. Fungal genomes often harbor a number of core genes encoding polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPSs), and the hybrid PKS-NRPS. These core enzymes contain multiple domains and are typically responsible for synthesizing the first intermediate in the biosynthetic pathways. In addition, the SM biosynthesis often requires several tailoring enzymes to mature the intermediate. The encoding genes of core and tailoring enzymes are usually co-regulated and located at the same locus in the genome, which often defines a gene cluster [1, 4].

The highly conserved domain organizations enable the fast identification of NRPS and PKS genes in a given genome [5]. The existence of co-regulated gene clusters also adds to the confidence of retrieving a biosynthetic pathway in silico. In recent years, fast expansions of available fungal genomes and genomic technologies have successfully linked many previously known or new products to their genetic bases, such as in the well-studied *Aspergillus* spp. [6]. However, it has become increasingly apparent that the vast number of predicted biosynthesis genes of microorganisms is not reflected by the metabolic profile observed under laboratory fermentation conditions. The opportunistic human

pathogen *Aspergillus fumigatus*, one of the best known fungi, has 8 out of 14 NRPSs already characterized [7], and other well-studied fungi such as *Metarhizium* spp., *Beauveria* spp., *Magnaporthe oryzae*, *etc*., only have 2–4 characterized NRP products. One important obstacle is that the SM gene loci remain silent in the absence of a particular (in most cases unknown) trigger. These cryptic genes may code for the biosynthesis of important virulence factors, toxins, or antibiotics that are even more efficient than the generally expressed ones when against specific stresses or host. Therefore, investigation of the functions of these silent genes is important to the applications of this largely untapped reservoir of potentially bioactive compounds, and at the same time to understanding the panorama of secondary metabolism in the success of fungal survival.

Pathogenic fungi are well-known carriers of SM synthetic machineries. They use these SMs to assist their pathogenic lifestyles; for example, as cytotoxins to host cells, suppressant of host immune systems, antibiotics to defend the food sources from other competitors, or mediators towards abiotic stresses [8, 9]. Species of *Cordyceps* are mostly entomopathogenic fungi that parasitize the larvae or pupae of insects. *Beauveria* (teleomorph: *Cordyceps*, Hypocreales) *bassiana* is a broad-host-range facultative entomopathogen that has been considered a model organism in the study of entomopathogenic fungi [10], and developed as biocontrol agents against different insect pests. The *B. bassiana* genome

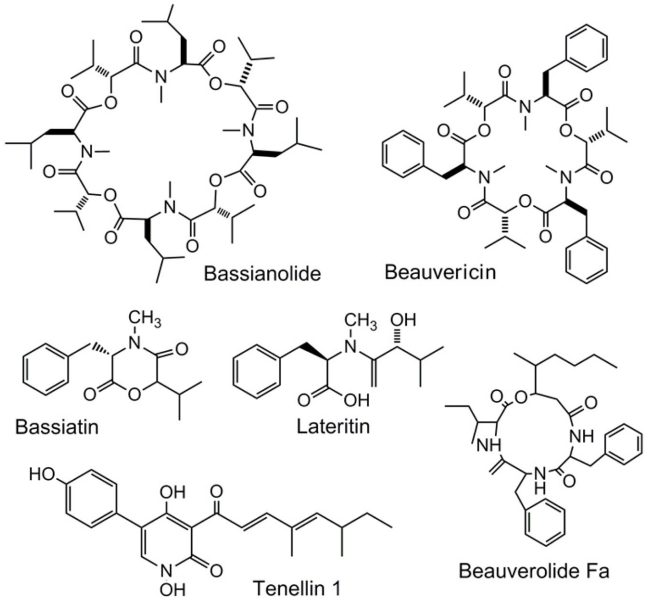

Fig. 1. Representative secondary metabolites of *Beauveria bassiana* [3, 47, 54, 55].

possesses 21 NRPSs (including NRPS-like proteins), providing a relatively replete pool to thoroughly study the roles of the NRPSome in the fungal lifecycle. Through this, the function and/or product of each identified NRPS or PKS-NRPS were primarily inferred to screen the NRPSs potentially associated with the pathogenicity pattern under regular laboratory condition or during infection. Thus, in this study, all the NRPSs and PKS-NRPSs in the currently available genome sequence of *B. bassiana* ARSEF 2860 [11] were identified and analyzed as the first step to overview the significance of its NRPSome in entomopathogenesis.

Materials and Methods

Beauveria bassiana Strains, Chemicals, and Culture Conditions

B. bassiana ARSEF 2860 was maintained on potato dextrose agar (PDA, Becton, Dickinson and Company, USA). Aerial conidia were collected in Triton X-100 (0.05%) from *B. bassiana* strains grown on PDA plates at 28°C for 14 days. Conidial concentrations were determined using serial dilutions of conidia growth on PDA plates. Chemicals (purchased from Sigma–Aldrich, Germany) were of the highest purities available.

Phylogenetic Tree Construction and Identification of Putative Gene Clusters

The phylogenetic relationships among NRPSs were constructed based on the alignments of amino acid sequences of the adenylation and/or condensation domains for NRPS. Adenylation and/or condensation domain-containing proteins from the *B. bassiana* genome were identified using the AMP-binding (PF00501) and condensation domain (PF00668) models from PFAM (http://pfam.sanger.ac.uk/) and the program hmmsearch. Fragments corresponding to the A and C domains from the NRPSs were aligned with MUSCLE (ver. 3.6) [12]. The phylogenetic relationship was constructed using the maximum likelihood and maximum parsimony methods by MEGA (ver. 5.2) [13]. Statistical support for branches within the phylogenetic trees was generated by bootstrap analysis with 1,000 pseudoreplicates. The compared NRPSs with known function or that have been studied were selected from the literature [4, 6, 8, 14–18]. Identification of potential gene clusters and conserved gene synteny was conducted using the on-line service antiSMASH 2.0.2 [19] and the NCBI BLASTp. The first program allowed for the comparison of *B. bassiana* contigs with a database of fungal and bacterial genomes maintained by the service curators.

Gene Expression Analysis

Gene expression of the NRPSs on PDA media and during infection of diamondback moth (*Plutella xylostella* (Linnaeus)) was analyzed by PCR and by quantitative RT-PCR. Total DNA and total RNA were extracted using the Trizol extraction method

(Invitrogen, USA) as instructed by the manufacturer manual. cDNA was then reverse transcribed using a Fast Quant RT Kit (with gDNase) (Tiangen, China). The PCRs were done with Easy Taq DNA polymerase (Transgene) using 30 amplification cycles with the temperature profile 94°C for 30 sec, 65°C for 30 sec, and 72°C for 60 sec. The primer sequences are listed in Table S3 and were used at 0.4 μmol/l each (final concentration). DNase-treated total RNA without reverse transcription and RNase-free DI-H_2O were both used as negative control templates. Genomic DNA was used as the template during the PCR for the positive control. All primers can amplify the genome DNA with a single band.

Results

The Reservoir of NRPSs in the Genome of *Beauveria bassiana*

The numbers of the NRPS genes in the *B. bassiana* genome are moderate compared with other insect or plant parasitic fungi [11]. The *B. bassiana* genome encodes 21 NRPS or NRPS-like and three PKS-NRPS hybrid proteins (Fig. 2, Table S1). The predicted proteins are composed by the various combinations of adenylation (A), condensation (C),

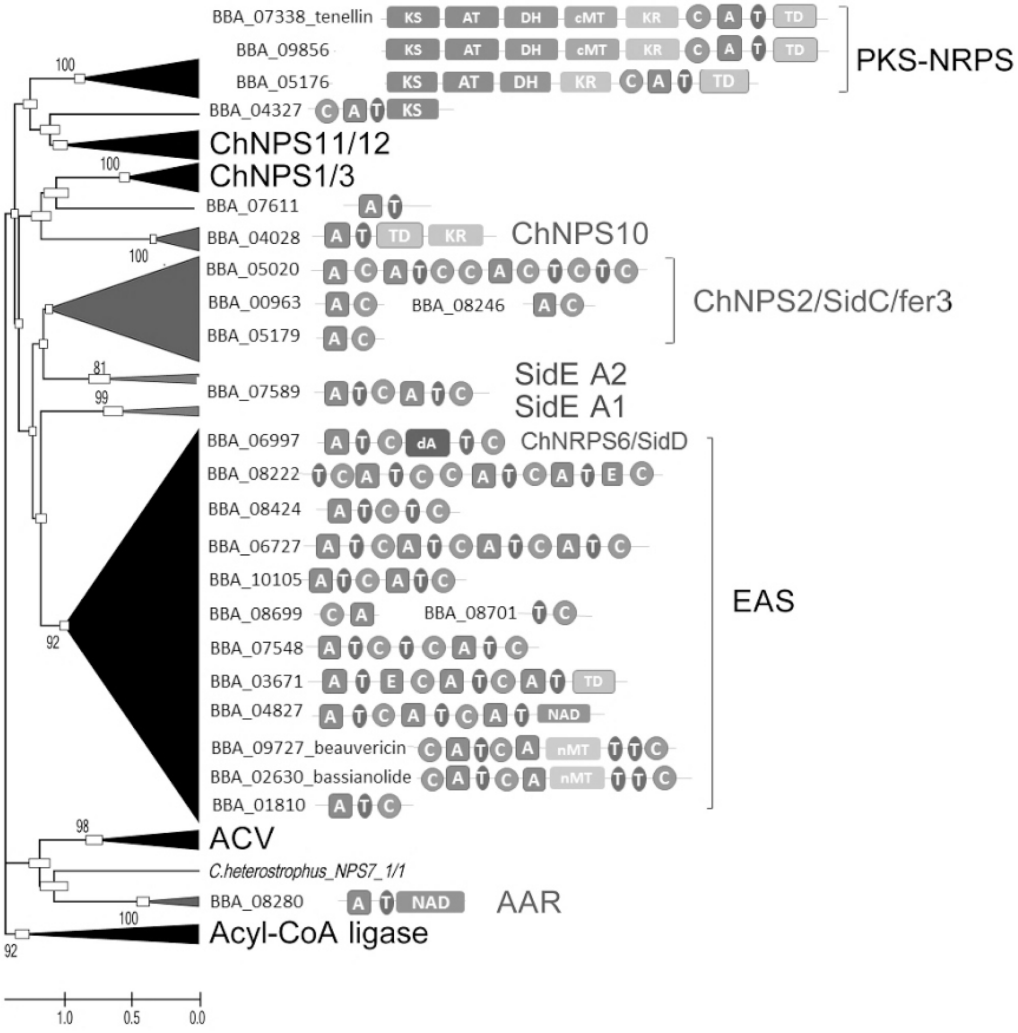

Fig. 2. Domain organization and phylogenetic relationships of nonribosomal protein synthetase (NRPSs) in *Beauveria bassianna*. The phylogeny of the *B. bassiana* NRPSs was inferred by maximum likelihood analysis of deduced amino acid sequences of the adenylation (A) domains. The full tree is included in the supporting information (Fig. S1). Major clades grouped by the *Cochliobolus heterostrophus* and other known NRPSs according to previous studies [8, 18] are marked by brackets to the right of the tree. Numbers above branches indicate the percentage bootstrap support (when >60) for each branch point based on 1,000 pseudoreplicates. The Acyl-CoA ligases from *B. bassiana* are used as the outgroup. Designations contain species name, NCBI index number, and the order of A domains in that NRPS out of the total number. Abbreviations of domains are used as follow: A = adenylation; C = condensation; T = thiolation; NAD = short-chain dehydrogenases/reductase; TD = thioester reductase; KS = ketoacyl synthase; AT = acyl transferase; E = epimerization; DH = dehydratase; KR = ketoreductase; ER = enolylreductase; MT = methyltransferase.

thiolation (T), and other modification domains. The number of A domains varies from 1 to 4; no very long NRPS was found as in the peptaibol-producing *Trichoderma* species [20]. Nine NRPSs end with typical terminal C domains (Fig. S1C), while two and four end with terminal reductase (R) domains and thioester reductase (TD) domains, respectively, and these domains may serve as the releasing and cyclization mechanisms of the peptide products [21–23]. One protein, BBA_06661, although annotated as NRPS in the NCBI database, does not have the A domain or functional enzymes nearby and was thus not included here.

The phylogenetic analysis can help predict the possible functions of the encoding NRPSs in the fungal lifecycle, especially when both the core enzyme and the gene cluster are conserved [4, 24, 25]. Besides this, resolving their orthologous relationships can reveal the evolution track of these highly variable enzymes. In this study, the phylogenetic tree of the NRPSs was constructed using the deduced amino acid sequences of the A domains as described in the literature [26]. The resultant tree has similar topology to the previous report (Figs. S1 and 2) [8, 26], and based on the classification by Bushley *et al*. [25], the *B. bassiana* NRPSs distribute in two main groups: mono/bimodular bacterial-origin NRPSs and multimodular eukaryotic NRPSs. Using the phylogenomic background, the *B. bassiana* NRPSs can be divided into three functional groups, as discussed in the following section.

The phylogeny of the C domains is similar to but more obscure than that of the A domains (Fig. S1C). Terminal C domains comprise a unique and exclusive group. Those conserved NRPSs (*i.e.*, siderophore clades *ChNPS6* and *ChNPS2*, PKS-NRPS, and *SidE*) still form well-supported C-domain clades.

To further characterize the function of the *B. bassiana* NRPSome, transcript levels were measured under in vitro growth condition on PDA for 6 and 9 days, and after infection of diamondback moth (*Plutella xylostella* (Linnaeus)) in vivo for 6 days. Previous studies have shown that the *B. bassiana* NRPS genes are dynamically expressed under different treatments and during different infection periods [11, 27]. Using the NRPSs producing beauvericin and bassianolide as examples, although peaked at different periods, these two genes expressed all through the infection processes despite of treatment methods (*i.e.*, insects treated

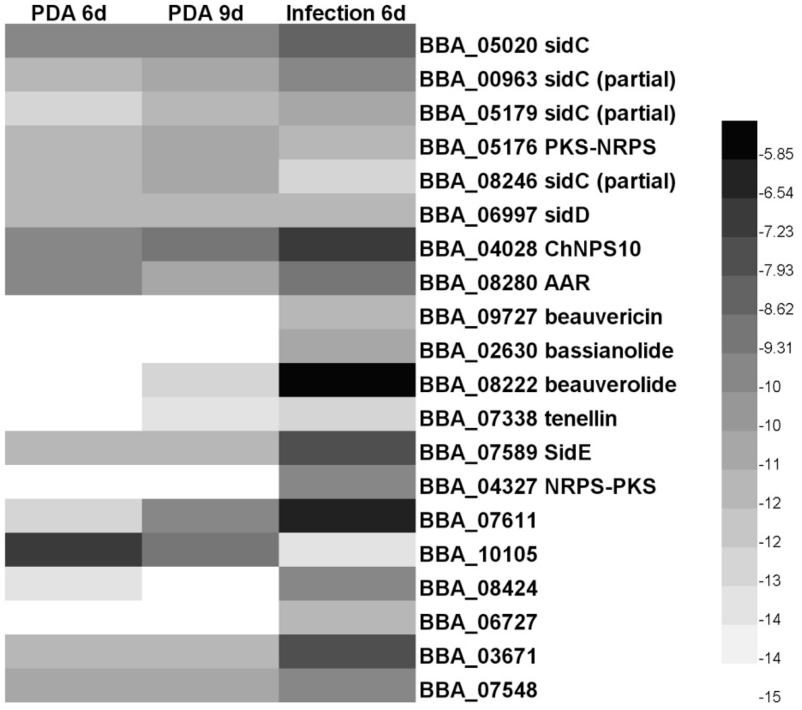

Fig. 3. Expression levels of the *Beauveria bassiana* nonribosomal protein synthetase (NRPS) genes using quantitative reverse transcription real-time PCR (qRT-PCR) in three conditions: 6 or 9 days on potato dextrose agar (PDA) media and 6 days after infection of diamondback moth.

18S rDNA was used as the internal standard. The data values were averaged by triplicates and indicate the difference between the expression levels of the targeted genes and 18S rDNA; that is, less negative represents higher expression.

by conidia suspension or injected with blastospore) [29]. Therefore, the transcript levels of NRPS genes during infection of diamondback moth can primarily screen the NRPSs universally active or specific to the certain type of host, but cannot reflect the time-course changes. PCR of the reverse-transcribed cDNA was used to qualitatively test whether each NRPS gene was expressed or not (Fig. S4), while quantitative reverse transcription real-time PCR (qRT-PCR) was used to quantify the differential transcription level (Figs. 3 and S4). A gene is described as "expressed" when (i) its transcript level is significantly higher than the negative control (*i.e.*, the RNA extracts without reverse transcription), and (ii) the PCR band of its reverse-transcribed cDNA is observable.

Conserved NRPSs and Their Predicted Functions

Compared with the multimodular NRPSs, the mono/bimodular NRPSs are relatively conserved with domain organization suggestive of functional conservation [27]. The functions of these conserved NRPSs can be speculated according to the evolutionary background deduced by their phylogenetic relationship. These NRPSs are often observed to be involved in sexual development, reproduction, iron metabolism, conidial development, and cytotoxicity [28–30], and are widely spread in plant and animal pathogens as well as saprophytes. In addition, the transcription of the conserved NRPS genes tends to be steady among different conditions and/or growth periods in this study and the literature [11].

Group I: Basic Metabolism

The NRPSs of Group I comprise a large portion of the *B. bassiana* NRPSome and are relatively indispensable during both saprophytic and pathogenic lifestyles. Consequently, they are equally transcribed during all tested conditions at a level generally higher than other NRPS genes (Figs. 3 and S4), and therefore Group I NRPSs were ascribed as "basic metabolism." Four *B. bassiana* NRPSs group within the subfamily of *SidC/fer3* synthesizing ferricrocin and ferrichrome intracellular siderophores; but only one,

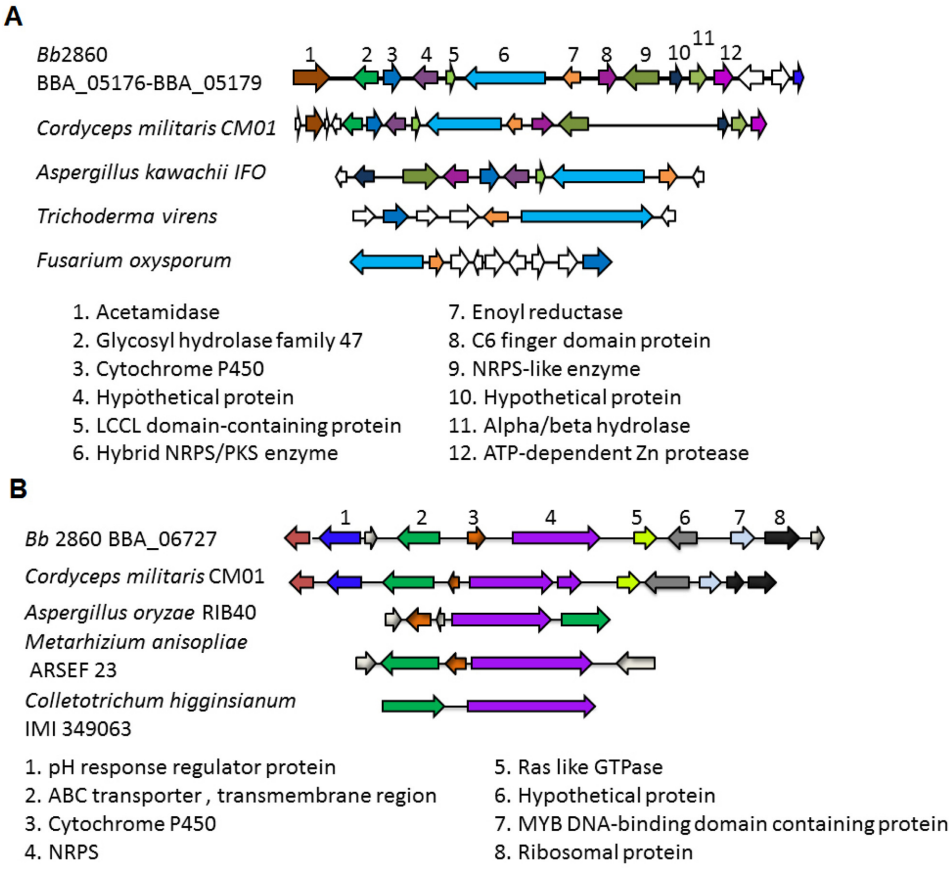

Fig. 4. Synteny and rearrangement of the conserved nonribosomal protein synthetase (NRPS) gene clusters in *Beauveria bassiana*. The organization of gene clusters in *B. bassiana* was compared with the previously described ones included in the antiSMASH database.

BBA_05020, possesses canonical domain organization as the well-documented *Sid2* of the basidiomyces *Ustilago maydis* [31], *Sib1* of the ascomycetes *Schizosaccharomyces pombe* [32], and *SidC* of *A. fumigatus* [33]. BBA_05020 also has the highest transcript level in this study, and the previous transcriptomic measurement suggested its product is a major iron-chelator (Fig. 3) [11]. The other three *SidC*-like NRPSs (BBA_00963, BBA_05179, and BBA_08246) only contain one A and one C domain, similar to the first module of *SidC/fer3* (Fig. 2). Interestingly, BBA_00963 and BBA_05179 have a PKS (BBA_00967) or PKS-NRPS hybrid (BBA_05176) gene located nearby, respectively, together with a separate enoylreductase (ER). Similar cluster organization can be found in *Cordyceps militaris* CM01 and *Aspergillus kawachii* IFO 4308 (Fig. 4A). The *B. bassiana* NRPSome also includes a homolog of the extracellular siderophore synthetase *ChNPS6/SidD* (BBA_06997), a proved virulence determinant in various pathogens [28, 34]. Different to the *SidC*-like NRPSs, BBA_06997 is a fungal-specific siderophore synthetase. Two other genes in this group, BBA_04028 and BBA_08280, are respectively homologous to *ChNPS10*, an NRPS-like protein involved in morphological development [25], and *AAR* (amino-adipate reductase), which is responsible for reducing α-amino-adipic acid in the fungal lysine biosynthetic pathway and is widely present in the fungi kingdom [26]. The domain architectures as well as gene clusters of these two NRPSs are both conserved. More description about this group of NRPSs is in the supplementary results.

Group II: Pathogenicity

Most of the NRPSs that are possibly involved in pathogenicity are not conserved at the gene and gene cluster levels, except for the *sidE*-like NRPS, BBA_07589. Both A and C domains of BBA_07589 are homologous to those of the *SidE* NRPS from the human pathogen *A. fumigatus* (>99 bootstrap value). BBA_07589 and *SidE* share the same domain architecture, with terminal C domains as the releasing mechanism, and a C6 transcription factor and a MFS transporter in their vicinity. *SidE* is phylogenetically close to the *SidC/fer3* siderophore synthetase and is originally assumed to be the same sort. However, instead of siderophore, *SidE* is shown to produce fumarylalanine, which is structurally related to the established pharmaceuticals exerting immunomodulatory activity [7]. Thus, BBA_07589 is very likely to produce a similar product, and play roles in evasion of the host immune system.

The domain organization of BBA_04327 is similar to the *Magnaporthe oryzae* NRPS *TAS1*, which synthesizes tenuazonic acid (TeA) using isoleucine and acetoacetyl-CoA [35]. TeA is a well-known mycotoxin produced by various plant pathogenic fungi. *TAS1* is a unique NRPS-PKS hybrid enzyme that begins with an NRPS module and is followed by a single KS domain, in contrast to other PKS-NRPS hybrid enzymes. This KS domain conducts the final cyclization step for TeA release instead of extension of the ketone chain [35]. BBA_04327 shares 36% identity and 98% coverage with *TAS1*, and thus may convey a similar synthetic capability. Its cluster, however, is very different from that of MGG_07803, including a transcription factor and a drug resistance transporter protein, suggesting a different product function. This combination is also presented in the genomes of *Metarhizium* spp. and *Fusarium* spp.

Fast-Evolving NRPSs Possible for Pathogenicity-Related Functions

Twelve *B. bassiana* NRPSs are classified into Group II that are likely to be involved in pathogenicity. They are mostly multimodular NRPSs unique to euascomycetes. Multimodular subfamilies of NRPSs are restricted to fungi, show less-stable domain architectures, and biosynthesize metabolites that perform more niche-specific functions than mono/bimodular NRPS products [26]. The phylogenetic assignment of their A domains is more complicated, with either a very low bootstrap value (<50) or discrepancies between two computing algorithms (*i.e.*, maximum likelihood or maximum parsimony), indicating a fast-changing nature. These NRPSs often do not have a conserved gene cluster, even within the closely related species. The expression levels of pathogenicity-related NRPSs vary significantly under different conditions and with different hosts; for example on PDA media, per os in diamondback moth, on locust wing, in cotton bollworm blood, or with root extract (Figs. 3 and S4) [11]. Consequently, their functions are difficult to predict. However, this fast-changing nature is beneficial for the fungi to adapt to the changing environment. It was proposed that genes involved in stress responses were more likely to undergo duplication and loss than growth related genes [19]. Therefore, pathogenicity-related NRPSs are the most interesting objects to study. If functional, they are more likely to undertake niche-specific functions to gain advantages in a certain environment, as in the host organism for pathogenic fungi. As examples, the synthetases producing the well-known insecticidal or cytotoxic substances beauvericin, serinocyclin, and ergot alkaloid belong to this group.

Possible Beauverolide Synthetase

The NRPS BBA_08222 and PKS BBA_08219 grouped with the emericellamide synthases *EasA* (NRPS) and *EasB* (PKS) in the respective phylogenetic trees. The *EasA*-like NRPS, BBA_08222, contains 13 domains divided into three modules corresponding to three A domains and possibly the consequent substrate amino acid incorporated (Fig. 2). These three A domains are counterparts to the first three of *EasA* in *A. nidulans* FGSC A4, which has five modules as $(T\text{-}C)_1$ $(A\text{-}T\text{-}C)_{2\text{-}4}$ $(A\text{-}T\text{-}C)_5$ (Fig. 3) [24]. The highly reducing iterative type I polyketide synthase BBA_08219 and the genes located nearby are also similar to the emericellamide synthetic gene cluster, although the order has been slightly shuffled (Fig. S3). The detailed description of this cluster is included in the supplementary results. Therefore, the biosynthesis process of this cluster, if active, may follow the same procedure with different chain length, number of monomers, and/or modifications, which may initiate from the linear polyketide intermediate by BBA_08219. After release, the intermediate is converted to a CoA thioester by the 4-coumarate-CoA ligase, and then shuttled by the acetyltransferase to the first T domain of the NRPS. The product chain will be further extended by the NRPS and released by the terminal C domain. The substrate selectivity of A domains was analyzed using five publicly available bioinformatic software programs, antiSMASH [36], NRPSpredictor2 [37], NRPSsp [38], SEQL-NRPS [39], and PKS/NRPS Analysis [40] (Table S1). Unfortunately, no consensus result was achieved. All three substrates are likely to be hydrophobic amino acids like alanine, valine, or leucine. The correctness of these bioinformatic software tools on predicting the substrate selectivity of fungal NRPS may be hindered by the fact that a majority of the characterized NRPS used to develop these programs are bacterial in origin and understanding of fungal NRPS is still in the early stages. To determine the product structure requires chemical analysis and identification of more fungal NRPs with proper gene sequencing. In summary, the product of the BBA_08222 gene cluster should contain a polyketide-derivative followed by three amino acids, the last of which is right-handed owing to the presence of the epimerization (E) domain. Given this predicted synthetic process, the BBA_08222 and its surrounding gene cluster are likely to be responsible for the biosynthesis of the known products, beauverolides, which are cyclodepsipeptides consisting of two L-amino acids, one D-amino acids, and one hydroxyl acid such as 3-hydroxy-4-methyloctanoic acid [41, 42], in accordance with the domain organization of BBA_08222.

NPS8-Like NRPSs

Among the highly variable NRPS clades, the *ChNPS8* clade is relatively conserved (Fig. S1). The NRPS (BBA_07548) has A domains somewhat similar to those of *ChNPS8* (Figs. 2 and S1). It ends with a terminal C domain, and its genomic locus lacks enzymes likely to provide substrate or perform post-modification. Therefore, BBA_07548 may independently biosynthesize the final metabolite. This gene was transcribed under both PDA and infection of diamondback moth conditions with levels similar to the siderophore synthetases (Fig. 4), suggesting a relatively large demand of its product under these conditions. However, its expression was very weak on locust wing or with root extract, and was not detected in cotton bollworm blood [51]. Near the *ChNPS8* clade, three *B. bassiana* NRPSs (BBA_06727, BBA_10105, and BBA_08699) form an individual clade (Fig. S1A). The transcription of BBA_06727 can only be activated by infection of diamondback moth among the tested conditions (Fig. 3) [11]. BBA_06727 also has a terminal C domain. The first and second A domains (A1 and A2) are highly similar to A3 and A4, respectively; the C1, C2, and C3, but not C4, domains also group together, which suggests an ancient duplication events occurred before the evolving of the terminal C domain. A cytochrome P450 and an ABC transporter are located nearby, and the cluster organization can also be found in *C. militaris*, *M. robertsii*, and *A. oryzae* (Fig. 4b). BBA_08699 only contains one A domain and is separated with BBA_08701 (comprising single T and terminal C domains) by a lipase, and showed no or negligible transcript levels under the tested conditions (Fig. 3) [11]. BBA_10105 has two highly similar A domains (identity = 61%). It was moderately transcribed with root extract and in cotton bollworm blood [11], but downregulated during infection of diamondback moth (Fig. 3). A transposase-like protein is located nearby, suggesting a possible exogenous origin of this gene.

***PerA*- or *DtxS1*-Like NRPS**

Another NRPS in this group is BBA_08424 that only showed transcription during infection of diamondback moth (Fig. 3). Its sole A domain groups with the *Epichloe festucae* peramine synthetase (PerA), and *Metarhizium* spp. destruxin synthetase (DtxS1). This NRPS has one A domain and two C domains, probably providing a homodimer as an intermediate. An aminotransferase located in the same locus may provide its substrate by catalyzing the transformation between an amino acid and an α-keto acid. A cytochrome P450 and a hydroxylase may exert further modification and/or cyclization. This process resembles

the biosynthesis of destruxins, which requires an aldo-keto reductase (DtxS3) to provide substrate for the NRPS (DtxS1), a cytochrome P450 (DtxS2), and an aspartic acid decarboxylase (DtxS4) for post-NRPS modification, although the domain architecture of *DtxS1* is very different ((A-T-C)$_{1-3}$(A-nMT-T-C)$_{4-5}$(A-T-C)$_6$) [43]. Destruxins are cyclohexadepsipeptides that exhibit a wide variety of biological activities, best known as insecticidal and phytotoxic substances [43, 44]. Destruxins also can inhibit the immunity of host insects like diamondback moth, *Bombyx mori*, and *Drosophila melanogaster* [45]. Peramine is an insect deterrent that mediates symbiotic interactions of *E. festuca* with its grass host [46]. Similarly, the product of the BBA_08424 cluster is possibly involved in the interaction with the host insect as a virulence factor.

NRPSs with A Domains Dispersed in Far Apart Subfamilies

A few NRPSs have A domains distributing in phylogenetically far apart subfamilies; that is, BBA_03671, BBA_04827, and the NRPSs synthesizing the known insecticidal mycotoxins beauvericin (*BbbeaS*, BBA_09727) and bassianolide (*BbbslS*, BBA_02630). Beauvericin and bassianolide are significant but nonessential virulence factors of *B. bassiana*, which are both formed by recursive condensations of the dipeptidol monomers [3, 47]. The similarity between *BbbeaS* and *BbbslS* sequences indicates a possible same origin. The expression of *BbbeaS* and *BbbslS* was induced by the presence of the host, which agrees with the previous results measured during the infection of *Triantoma infestans*, either by immersing the insects in conidial suspensions or by injecting them with blastospores [27]. However, both of the genes were not express during growth on locust hind wings or in cotton bollworm blood [11], suggesting a transcription highly dependent on conditions. Similarly, BBA_03671 was upregulated during infection of diamondback moth, but remained silent under other tested conditions (Fig. 3) [11]. BBA_04827 did not show expression (Fig. S4) [11]. Both BBA_04827 and BBA_03671 end with a reductase domain (Fig. 2). BLAST results of BBA_03671, if not considering *Beauveria* species, are from the class of Eurotiomycetes instead of Sordariomycetes, suggesting a possible origin from horizontal transfer.

Group III: Unknown Function

NRPSs with unresolved phylogeny had low bootstrap support and discrepancies between the algorisms used, and thus we were not able to infer their possible products or function. In this sort of NRPSs, BBA_07611 is an unusual NRPS with single A and T domains followed by a terminal domain with unknown function. This domain is found exclusively in NRPSs (more likely with bacteria origin) and always as the final domain (supplementary results) [48]. Another NRPS, BBA_01810, has the domain architecture of A-T-C. Its A domain is somewhat similar to the second and third ones of HC-toxin synthase in *T. marneffei*, and is grouped close to the serinocyclin synthase.

Discussion

Phylogenetics paired with comparative genomics is a useful tool to predict information of the *Beauveria bassiana* NRPSome, on what they might produce and what their functions may be. Transcript levels under different conditions help primarily to screen genes participating in certain biological processes. The *B. bassiana* genome harbors 21 NRPS or NRPS-like genes, three of which collaborate with PKS or PKS-NRPS hybrid genes. Of these 21 clusters, the products of two have been previously characterized (bassianolide and beauvericin). Functions of the NRPSome can be divided into three groups. Group I include five clusters involved in iron metabolism (BBA_05020, BBA_00963-BBA_00967, BBA_05176-BBA_05179, BBA_08246, and BBA_06997), one amino-adipate reductase (AAR) (BBA_08280), and one involved in morphological development (BBA_04028). In addition, BBA_07589 is highly similar to the *SidE* synthesizing fumarylalanine, whereas BBA_04327 resembles the *TAS1* NRPS-PKS hybrid enzyme synthesizing TeA [35]. Based on the phylogenetic relationship and the structure of the gene cluster, it is also predicted that BBA_08222 was responsible for the biosynthesis of beauverolide. These NRPSs have a relatively certain phylogenetic relationship, and their expression levels are in accordance with their predicted functions. However, if both the NRPS core gene and the surrounding genes are not similar to any known NRPS, it is difficult to infer its function based only on bioinformatics. Some NRPSs even do not have homologs in closely related species. The A domain of the rest of the NRPSs always group with the most fast-changing clades, suggesting vivid evolution events sculpting their structure to cope with the specific environments of the fungi, such as against certain hosts, competitors, or stress. In the case of entomopathogenic fungi such as *B. bassiana*, their "environmental pressure" mainly comes from the host, and thus the majority of these fast-evolving NRPSs are associated with pathogenesis. As examples, the two NRPSs that synthesize beauvericin and bassianolide (BBA_09727 and BBA_02630) both belong to this group.

The fast-evolving NRPSs may also be under more restricted regulation control and are only activated upon infection. Their expression tends to be highly dependent on host type and growth conditions [11]. Consequently, they are usually the less known NRPSs and require more attention. Genes involved in stress response are usually under restricted control, and their expression requires a certain condition to trigger as demonstrated in this and previous study [49]. In recent years, the molecular characterization of stress response in pathogenic fungi revealed that general signaling pathways involved in stress response are also associated with virulence, especially GPI-anchored proteins and G-protein-coupled membrane receptors, mitogen-activated protein kinase (MAPK) pathways; for example, (i) pheromone/nutrient sensing (Fus3/Kss1), (ii) cell wall integrity (Mpk1), and (iii) high osmolarity (Hog1), and the PKA/adenyl cyclase pathway, and various downstream transcription factors, such as Msn2, CreA, and Pac1 [49, 50]. For example, a homolog of the *msn2/seb1* transcription factor identified in *B. bassiana* participated in the resistance to Calcofluor White, H_2O_2, and Congo Red. This gene also showed pH-dependent deregulation of oosporein, a secondary metabolite produced by PKS, and the null mutant *ΔBbmsn2* was impaired in virulence in both topical and intrahemocoel injection bioassays against *Galleria mellonella* [51]. The TeA biosynthetic gene, *TAS1*, was regulated by Hog1-related MAPK involved in the response to hyperosmotic stress in *M. oryzae* [35]. In addition, SMs such as NRPs and PKs can be regulated by transcriptional factors located in the same locus with the core enzymes, while the global regulator *Lae* can regulate the expression of multiple SM gene clusters in *Fusarium verticillioides* and *Aspergillus* spp. [52, 53]. Despite of the increasing discoveries, most of the regulation networks and the specific trigger for the SMs in entomopathogenic fungi are as yet unknown. Our study provides NRPS genes selectively expressed during infection, which can be model systems to investigate the regulation of SM production.

Our procedure is useful in genome mining studies to screen SM genes of certain interest, such as pathogenesis in this study. Our results also help demonstrate the function of the NRPSome in an entomopathogenic fungus. The specific biological roles as well as biosynthetic machineries of each functional gene are interesting topics for further investigations.

Acknowledgments

We thank the Research Facility Center of the Biotechnology Research Institute for providing equipment. This work was supported by the National Basic Research Program of China (2015CB755700 to M.L. and Y.X.), Natural Science Foundation of China (31570093 to Y.X.), Natural Science Foundation of China (31500079 to L.Z.), and the "948" Project of the Ministry of Agriculture of China (2016-X43 to Y.X.).

References

1. Keller NP, Turner G, Bennett JW. 2005. Fungal secondary metabolism – from biochemistry to genomics. *Nat. Rev. Microbiol.* **3:** 937-947.
2. Feng P, Shang Y, Cen K, Wang C. 2015. Fungal biosynthesis of the bibenzoquinone oosporein to evade insect immunity. *Proc. Natl. Acad. Sci. USA* **112:** 11365-11370.
3. Xu YQ, Orozco R, Wijeratne EMK, Gunatilaka AAL, Stock SP, Molnar I. 2008. Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus *Beauveria bassiana*. *Chem. Biol.* **15:** 898-907.
4. Collemare J, Griffiths S, Iida Y, Karimi Jashni M, Battaglia E, Cox RJ, de Wit PJ. 2014. Secondary metabolism and biotrophic lifestyle in the tomato pathogen *Cladosporium fulvum*. *PLoS One* **9:** e85877.
5. de Wit PJGM, van der Burgt A, Ökmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, *et al.* 2012. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genet.* **8:** e1003088.
6. Sanchez JF, Somoza AD, Keller NP, Wang CCC. 2012. Advances in *Aspergillus* secondary metabolite research in the post-genomic era. *Nat. Prod. Rep.* **29:** 351-371.
7. Steinchen W, Lackner G, Yasmin S, Schrettl M, Dahse H-M, Haas H, Hoffmeister D. 2013. Bimodular peptide synthetase SidE produces fumarylalanine in the human pathogen *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* **79:** 6670-6676.
8. Gibson DM, Donzelli BGG, Krasnoff SB, Keyhani NO. 2014. Discovering the secondary metabolite potential encoded within entomopathogenic fungi. *Nat. Prod. Rep.* **31:** 1287-1305.
9. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stuber K, *et al.* 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330:** 1543-1546.
10. Rohlfs M, Churchill ACL. 2011. Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genet. Biol.* **48:** 23-34.
11. Xiao G, Ying SH, Zheng P, Wang ZL, Zhang S, Xie XQ, *et al.* 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci. Rep.* **2:** 483.
12. Edgar R. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

13. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28:** 2731-2739.
14. Bolker M, Basse CW, Schirawski J. 2008. *Ustilago maydis* secondary metabolism – from genomics to biochemistry. *Fungal Genet. Biol.* **45:** S88-S93.
15. Guo CJ, Wang CC. 2014. Recent advances in genome mining of secondary metabolites in *Aspergillus terreus*. *Front. Microbiol.* **5:** 717.
16. Inglis DO, Binkley J, Skrzypek MS, Arnaud MB, Cerqueira GC, Shah P, *et al.* 2013. Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. *BMC Microbiol.* **13:** 91-91.
17. Frisvad JC, Smedsgaard J, Larsen TO, Samson RA. 2004. Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. *Stud. Mycol.* 201-241.
18. Boettger D, Hertweck C. 2013. Molecular diversity sculpted by fungal PKS-NRPS hybrids. *Chembiochem* **14:** 28-42.
19. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, *et al.* 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39:** W339-W346.
20. Keswani C, Mishra S, Sarma BK, Singh SP, Singh HB. 2014. Unraveling the efficient applications of secondary metabolites of various *Trichoderma* spp. *Appl. Microbiol. Biotechnol.* **98:** 533-544.
21. Kopp F, Marahiel MA. 2007. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* **24:** 735-749.
22. Gao X, Haynes SW, Ames BD, Wang P, Vien LP, Walsh CT, Tang Y. 2012. Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat. Chem. Biol.* **8:** 823-830.
23. Du LC, Lou LL. 2010. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **27:** 255-278.
24. Chiang Y-M, Szewczyk E, Nayak T, Davidson AD, Sanchez JF, Lo H-C, *et al.* 2008. Molecular genetic mining of the *Aspergillus* secondary metabolome: discovery of the emericellamide biosynthetic pathway. *Chem. Biol.* **15:** 527-532.
25. Bushley KE, Raja R, Jaiswal P, Cumbie JS, Nonogaki M, Boyd AE, *et al.* 2013. The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. *PLoS Genet.* **9:** e1003496.
26. Bushley KE, Turgeon BG. 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. *BMC Evol. Biol.* **10:** 1.
27. Lobo LS, Luz C, Fernandes EKK, Juarez MP, Pedrini N. 2015. Assessing gene expression during pathogenesis: use of qRT-PCR to follow toxin production in the entomopathogenic fungus *Beauveria bassiana* during infection and immune response of the insect host *Triatoma infestans*. *J. Invertebr. Pathol.* **128:** 14-21.
28. Oide S, Moeder W, Krasnoff S, Gibson D, Haas H, Yoshioka K, Turgeon BG. 2006. NPS6, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. *Plant Cell* **18:** 2836-2853.
29. Oide S, Krasnoff SB, Gibson DM, Turgeon BG. 2007. Intracellular siderophores are essential for ascomycete sexual development in heterothallic *Cochliobolus heterostrophus* and homothallic *Gibberella zeae*. *Eukaryot. Cell* **6:** 1339-1353.
30. Atanasova L, Knox BP, Kubicek CP, Druzhinina IS, Baker SE. 2013. The polyketide synthase gene *pks4* of *Trichoderma reesei* provides pigmentation and stress resistance. *Eukaryot. Cell* **12:** 1499-1508.
31. Winterberg B, Uhlmann S, Linne U, Lessing F, Marahiel MA, Eichhorn H, *et al.* 2010. Elucidation of the complete ferrichrome A biosynthetic pathway in *Ustilago maydis*. *Mol. Microbiol.* **75:** 1260-1271.
32. Schwecke T, Gottling K, Durek P, Duenas I, Kaufer NF, Zock-Emmenthal S, *et al.* 2006. Nonribosomal peptide synthesis in *Schizosaccharomyces pombe* and the architectures of ferrichrome-type siderophore synthetases in fungi. *Chembiochem* **7:** 612-622.
33. Wallner A, Blatzer M, Schrettl M, Sarg B, Lindner H, Haas H. 2009. Ferricrocin, a siderophore involved in intra- and transcellular iron distribution in *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* **75:** 4194-4196.
34. Stack D, Neville C, Doyle S. 2007. Nonribosomal peptide synthesis in *Aspergillus fumigatus* and other fungi. *Microbiology* **153:** 1297-1306.
35. Yun C-S, Motoyama T, Osada H. 2015. Biosynthesis of the mycotoxin tenuazonic acid by a fungal NRPS–PKS hybrid enzyme. *Nat. Commun.* **6:** 8758.
36. Minowa Y, Araki M, Kanehisa M. 2007. Comprehensive analysis of distinctive polyketide and nonribosomal peptide structural motifs encoded in microbial genomes. *J. Mol. Biol.* **368:** 1500-1517.
37. Rottig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O. 2011. NRPSpredictor2 – a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39:** W362-W367.
38. Khayatt BI, Overmars L, Siezen RJ, Francke C. 2013. Classification of the adenylation and acyl-transferase activity of NRPS and PKS systems using ensembles of substrate specific hidden Markov models. *PLoS One* **8:** 10.
39. Knudsen M, Sondergaard D, Tofting-Olesen C, Hansen FT, Brodersen DE, Pedersen CN. 2016. Computational discovery of specificity-conferring sites in non-ribosomal peptide synthetases. *Bioinformatics* **32:** 325-329.
40. Bachmann BO, Ravel J. 2009. Methods for in silico prediction of microbial secondary metabolic pathways from DNA sequence data. *Methods Enzymol.* **458:** 181-217.

41. Jegorov A, Haiduch M, Sulc M, Havlicek V. 2006. Nonribosomal cyclic peptides: specific markers of fungal infections. *J. Mass Spectrom.* **41:** 563-576.
42. Jegorov A, Paizs B, Kuzma M, Zabka M, Landa Z, Sulc M, *et al.* 2004. Extraribosomal cyclic tetradepsipeptides beauverolides: profiling and modeling the fragmentation pathways. *J. Mass Spectrom.* **39:** 949-960.
43. Wang B, Kang QJ, Lu YZ, Bai LQ, Wang CS. 2012. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. *Proc. Natl. Acad. Sci. USA* **109:** 1287-1292.
44. Pedras MSC, Irina Zaharia L, Ward DE. 2002. The destruxins: synthesis, biosynthesis, biotransformation, and biological activity. *Phytochemistry* **59:** 579-596.
45. Pal S, St. Leger RJ, Wu LP. 2007. Fungal peptide destruxin A plays a specific role in suppressing the innate immune response in *Drosophila melanogaster*. *J. Biol. Chem.* **282:** 8969-8977.
46. Tanaka A, Tapper BA, Popay A, Parker EJ, Scott B. 2005. A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. *Mol. Biol.* **57:** 1036-1050.
47. Xu YQ, Rozco R, Wijeratne EMK, Espinosa-Artiles P, Gunatilaka AAL, Stock SP, Molnar I. 2009. Biosynthesis of the cyclooligomer depsipeptide bassianolide, an insecticidal virulence factor of *Beauveria bassiana*. *Fungal Genet. Biol.* **46:** 353-364.
48. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, *et al.* 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43:** D222-D226.
49. Ortiz-Urquiza A, Keyhani NO. 2015. Stress response signaling and virulence: insights from entomopathogenic fungi. *Curr. Genet.* **61:** 239-249.
50. Abad A, Fernandez-Molina JV, Bikandi J, Ramirez A, Margareto J, Sendino J, *et al.* 2010. What makes *Aspergillus fumigatus* a successful pathogen? Genes and molecules involved in invasive aspergillosis. *Rev. Iberoam. Micol.* **27:** 155-182.
51. Luo ZB, Li YJ, Mousa J, Bruner S, Zhang YJ, Pei Y, Keyhani NO. 2015. *Bbmsn2* acts as a pH-dependent negative regulator of secondary metabolite production in the entomopathogenic fungus *Beauveria bassiana*. *Environ. Microbiol.* **17:** 1189-1202.
52. Butchko RAE, Brown DW, Busman M, Tudzynski B, Wiemann P. 2012. Lae1 regulates expression of multiple secondary metabolite gene clusters in *Fusarium verticillioides*. *Fungal Genet. Biol.* **49:** 602-612.
53. Bayram O, Braus GH. 2012. Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. *FEMS Microbiol. Rev.* **36:** 1-24.
54. Xu YQ, Wijeratne EMK, Espinosa-Artiles P, Gunatilaka AAL, Molnar I. 2009. Combinatorial mutasynthesis of scrambled beauvericins, cyclooligomer depsipeptide cell migration inhibitors from *Beauveria bassiana*. *Chembiochem* **10:** 345-354.
55. Halo LM, Heneghan MN, Yakasai AA, Song Z, Williams K, Bailey AM, *et al.* 2008. Late stage oxidations during the biosynthesis of the 2-pyridone tenellin in the entomopathogenic fungus *Beauveria bassiana*. *J. Am. Chem. Soc.* **130:** 17988-17996.