



Published in final edited form as:

Chem Biol. 2010 May 28; 17(5): 483–494. doi:10.1016/j.chembiol.2010.03.015.

Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from *Penicillium aethiopicum*

Yit-Heng Chooi, Ralph Cacho, and Yi Tang*

Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095

SUMMARY

Penicillium aethiopicum produces two structurally interesting and biologically active polyketides: the tetracycline-like viridicatumtoxin **1** and the classic antifungal agent griseofulvin **2**. Here, we report the concurrent discovery of the two corresponding biosynthetic gene clusters (*vrt* and *gsf*) by 454 shotgun sequencing. Gene deletions confirmed two nonreducing PKSs (NRPKS), *vrtA* and *gsfA*, are required for the biosynthesis of **1** and **2**, respectively. Both PKSs share similar domain architectures and lack a C-terminal thioesterase domain. We identified *gsfI* as the chlorinase involved in the biosynthesis of **2**, as deletion of *gsfI* resulted in the accumulation of dechlorogriseofulvin **3**. Comparative analysis with the *P. chrysogenum* genome revealed that both clusters are embedded within conserved syntenic regions of *P. aethiopicum* chromosomes. Discovery of the *vrt* and *gsf* clusters provided the basis for genetic and biochemical studies of the pathways.

INTRODUCTION

Fungal polyketides are an important class of secondary metabolites, which include blockbuster drugs, mycotoxins and pigments. There has been a growing interest in the engineering and heterologous expression of fungal polyketide synthase (PKS) genes and pathways to enable rational production of novel compounds (Schümann and Hertweck, 2006). Understanding the molecular genetic basis of the biosynthesis of these structurally diverse compounds is the underlying key to achieve this goal. Over the past decade, there has been a steady growth of knowledge on the mechanisms of fungal polyketide biosynthetic pathways (Cox, 2007). Numerous fungal polyketides have been linked to their PKS genes and gene clusters (Hoffmeister and Keller, 2007), including some from the cryptic metabolic pathways uncovered via genome mining (Bergmann et al., 2007; Bok et al., 2006; Chiang et al., 2008). Nevertheless, the biosynthetic pathways of many previously known fungal metabolites, some of which have important pharmaceutical applications, still remain to be elucidated.

The filamentous fungus *Penicillium aethiopicum* Frisvad is known to produce a number of interesting secondary metabolites, including viridicatumtoxin (**1**), griseofulvin (**2**), and tryptoquialanine (**4**) (Figure 1), which are used as chemotaxonomic markers for the species (Frisvad and Samson, 2004). Tryptoquialanine is a tetrapeptide structurally similar to tryptoquivaline, which is a known tremorgen (Ariza et al., 2002; Clardy et al., 1975). Viridicatumtoxin is a hybrid polyketide-isoprenoid compound and is a rare example of

*Correspondence: yitang@ucla.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

tetracycline-like compounds produced by fungi. Compound **1** shares a common tetracyclic carboxamide core with the well-known tetracycline intermediate anhydrotetracycline. **1** has been reported to cause nephrotoxicity (Hutchison et al., 1973) as well as exhibiting modest antitumor activity (Raju et al., 2004). An epoxide derivative of the compound, viridicatumtoxin B, was isolated from *Penicillium* sp. FR11 along with **1**; both were shown to inhibit the growth of methicilin- and quinolone-resistant *Staphylococcus aureus* at 8–64 times higher activity than tetracycline (Zheng et al., 2008). Another known fungal compound with a tetracyclic carboxamide core is anthrotainin (TAN-1652) (Ishimaru et al., 1993; Wong et al., 1993), which is an inhibitor of neuropeptide substance P binding and is a potential non-steroidal anti-inflammatory agent. The biosynthetic convergence of tetracycline-like compounds across the bacterial and fungal kingdoms, and their various bioactivities, further suggest that compounds of this structural type contain an “evolutionarily privileged scaffold”, which is capable of interacting with a variety of biological targets. The co-occurrence of tetracyclic carboxamide core in both the tetracyclines from bacteria and **1** from fungi provides an excellent example to study the convergence of the polyketide pathway. Previous isotope feeding study (de Jesus et al., 1982) showed that **1** is produced by a mechanism significantly different from the bacterial tetracyclines (Thomas, 2001). The formation and attachment of the spirobicyclic ring of isoprenoid origin on a tetracycline scaffold also warrants detailed biosynthetic investigation.

Griseofulvin, which affects the function of mitotic spindle microtubules in mitosis, is an antifungal drug and has been in use for many years in medical and veterinary applications (Finkelstein et al., 1996). Although the use of **2** is mostly superseded by newer and more effective synthetic antifungal agents, it remains useful for the treatment of some dermatophytes such as *Tinea capitis* (ringworm of the scalp) and *Tinea pedis* (athlete’s foot). Recently, there is a renewed interest in **2** owing to its specific antiproliferative and antimitotic activities towards cancer cells (Panda et al., 2005; Rebacz et al., 2007). The recent studies suggest that **2** acts by inhibiting centrosomal clustering in tumor cells with supernumerary centrosomes, causing multipolar mitoses, and subsequently, apoptosis (Rebacz et al., 2007). Interestingly, **2** has also been shown to suppress hepatitis C virus replication *in vitro* (Jin et al., 2008). The discovery of the new potentials of **2** has led to new synthetic efforts to search for superior analogs (Rønneest et al., 2009). Structurally, **2** has an interesting cyclization pattern, where the polyketide backbone is folded to allow formation of orcinol and phloroglucinol ring structures through a Claisen and an aldol reactions, respectively. A stereospecific oxidative coupling reaction was proposed for the formation of the grisan structure of **2** (Barton and Cohen, 1957). The biosynthetic pathway of **2** has been extensively studied using isotopic incorporation (Harris et al., 1976; Lane et al., 1982; Rhodes et al., 1963; Simpson and Holker, 1977), but the genes and enzymes involved in the biosynthesis remain unknown.

The significant biological properties and unusual structural features of **1** and **2** have motivated us to perform a genome scanning of *P. aethiopicum* with DNA pyrosequencing technology and search for the two corresponding gene clusters. A bioinformatic search for PKS genes and comparative analysis with *Penicillium chrysogenum* genome revealed the putative biosynthetic gene clusters for both compounds, which were confirmed by targeted gene deletions and RNA-silencing. The two gene clusters provide an important step towards understanding the enzymatic basis of biosynthesis of **1** and **2**.

RESULTS AND DISCUSSION

Genome-wide analysis of *P. aethiopicum* PKS genes revealed two putative gene clusters for viridicatumtoxin and griseofulvin biosynthesis

The 454 shotgun sequencing of *P. aethiopicum* IBT 5753 with the GS FLX Titanium series generated a total of ~572 million bases with an average sequencing read length of 367.5 bases. Assembly of the unpaired shotgun sequence reads resulted in 1522 contigs, which consist of

28,925,551 non-redundant bases, with an N50 of 149.2 kilobases. Based on the phylogenetic analysis of β -tubulin sequences, *P. aethiopicum* is closely related to *P. chrysogenum*, and both species are grouped under *Penicillium* subgenus *Penicillium*, section *Chrysogena* (Samson et al., 2004). Assuming that the size of *P. aethiopicum* genome is close to the sequenced *P. chrysogenum* genome (32.19 Mb), the total non-redundant bases were estimated to cover about 90% of the fungal genome.

Using a local BLASTP program queried against a database consists of all the contigs with an arbitrary KS domain sequences, a total of 30 putative, complete PKS genes were found in the *P. aethiopicum* genome. Out of them, there are a total of 7 non-reducing PKSs (NRPKSs), 17 highly-reducing PKSs (HRPKS), 3 partially-reducing PKSs (PRPKSs) and 3 HRPKS-nonribosomal peptide synthetases (NRPSs) hybrids. For simplicity, the PKSs are designated as PaPKS[contig number] (Table S1).

Compared to *P. chrysogenum*, which has a total of 21 putative PKS genes, the *P. aethiopicum* genome contains more putative PKS genes. Based on a chemotaxonomic study, the two closely related species are known to produce distinct sets of secondary metabolites (Frisvad and Samson, 2004). Since *P. chrysogenum* is not known to produce **1** and **2**, we reasoned that the orthologous PKSs are not involved in the biosynthesis of the two compounds. Sequence alignments and phylogenetic analysis implied that out of the 30 putative PKSs genes in *P. aethiopicum*, 9 are orthologous to *P. chrysogenum* (Table S1 and Figure S1).

We reasoned that NRPKSs are likely involved in the biosynthesis of the carbon skeletons of **1** and **2**, since these compounds are derived from aromatic polyketide precursors. Among the *P. aethiopicum* NRPKSs that are non-orthologous to those found in *P. chrysogenum*, PaPKS0274 and PaPKS0880 fall into the same clade as the recently identified *Aspergillus nidulans* asperthecin synthase (AptA) (Szewczyk et al., 2008) and *Aspergillus terreus* atrochrysonic acid synthase (ACAS) (Awakawa et al., 2009), both of which produce fused-ring aromatic compounds (Figure S1). Both PaPKS0274 and PaPKS0880 consist of starter unit:ACP transacylase (SAT) (Crawford et al., 2006), ketosynthase (KS), malonyl-CoA:ACP transacylase (MAT), product template (PT) (Crawford et al., 2009) and acyl carrier protein (ACP) domains, but lack a thioesterase/Claisen cyclase (TE/CLC) domain (Fujii et al., 2001) at the C-terminal for chain release. Since the ACAS product, atrochrysonic acid, has a similar structure to the BCD rings of **1**, we predicted that one of these two PKSs may be involved in the biosynthesis of **1**.

Given that genes involved in fungal secondary metabolic pathways are often clustered together (Hoffmeister and Keller, 2007), we searched the genes surrounding the two putative PKS genes for clues about the final polyketide products. Preliminary bioinformatic analysis of contig 00274 identified a pair of genes related to isoprenoid pathway downstream of PaPKS0274 gene. The presence of both genes may be associated with the isoprenoid portion of the spirobicyclic ring in **1**. Other genes encoding putative oxygenases, aminotransferase and *O*-methyltransferase were also found in the vicinity of the PaPKS0274 gene, which respectively match the structural features present in **1**. On contig 00880, the presence of three putative *O*-methyltransferase genes and a halogenase gene flanking the PaPKS0880 gene immediately points to a potential correlation to the structure of **2**.

VrtA and GsfA are two PKSs essential for the biosynthesis of viridicatumtoxin and griseofulvin respectively

Liquid chromatography-mass spectrometry (LC-MS) analysis of the ethyl acetate extract from *P. aethiopicum* wild type (*WT*) stationary culture grown in yeast malt extract glucose (YMEG) medium showed six distinct peaks at 283 nm. Two of the peaks have the retention time, UV spectra and m/z value matching the authentic standards of **1** (RT = 30 min, m/z 548 [M+H-

H₂O)⁺ and **2** (RT = 24 min, *m/z* 353 [M+H]⁺) (Figure 2). Another peak at RT = 22 min showed a UV spectrum that is similar to **2** and *m/z* 319 [M+H]⁺, which matches with the molecular weight of dechlorogriseofulvin **3** (Figure 1). The compound was purified from a large culture of *P. aethiopicum* and the structure was confirmed to be **3** by comparing the ¹H NMR and ¹³C NMR spectra to the previously published spectra for **2** (Simpson and Holker, 1977) and epidechlorogriseofulvin (Jarvis et al., 1996) (Table S3). Two additional peaks (RT = 27 and 28 min) with UV spectra similar to previously reported for tryptoquivalines and tryptoquialanines were also detected (Ariza et al., 2002; Clardy et al., 1975). Both peaks have a mass of [M+H]⁺ *m/z* = 519, which matches with the molecular weight of **4**. The additional peak could be an isomer of **4**, analogous to tryptoquivaline and isotryptoquivaline (Yamazaki et al., 1976). The identity of the remaining peak at RT = 19.5 min is undetermined.

To verify the proposed associations between the putative gene clusters and the metabolites **1** and **2**, we developed a transformation system for *P. aethiopicum* based on existing methods for *A. nidulans* using *bar* resistant marker (Chooi et al., 2008; Nayak et al., 2006). Using double-homologous deletion cassettes with a *bar* resistant marker, the PaPKS0274 and PaPKS0880 genes were deleted by double crossover recombination (Figure 3). Approximately 100 glufosinate-resistant transformants were picked and screened with PCR using a *bar* gene primer and primers outside of the deletion cassette (Table S2). A total of 6 and 5 potential recombinants were identified for PaPKS0274 and PaPKS880 respectively. Three PCR positive clones from each knockout experiments were chosen and confirmed by Southern hybridization (Figure 3). All the PCR positive clones (Δ PKS0274-I16, Δ PKS0274-I18, Δ PKS0274-I23, Δ PKS880-I7, Δ PKS880-I12 and Δ PKS880-I9) showed the expected band size for correct insertion of deletion cassette by double crossover recombination into the corresponding genomic copy of the two PKS genes. Analysis of the PaPKS0274 disruptants showed that the production of **1** was totally abolished (Figure 2), while the other metabolites were unaffected. For the PaPKS0880 disruptants, only the production of **2** and **3** were abolished (Figure 2). These results confirmed that PaPKS0274 is essential for the biosynthesis of **1**, while PaPKS0880 is required for the biosynthesis of **2** and **3**.

The two PKS gene clusters for **1** and **2** were designated as *vrt* and *gsf* respectively, and the two NRPKS genes were designated as *vrtA* and *gsfA*, respectively (Table 2 and 3). Interestingly, both PKSs have the same domain architecture and lack a fused TE/CLC domain. The functional assignment of VrtA and GsfA, which produce polyketides with different chain lengths and backbone folding patterns in **1** and **2**, suggested that this family of “TE-less” PKSs (assigned as clade IV NR-PKSs here, Figure S1) may account for a greater structural diversity among fungal polyketides.

We also tested whether the production of **1** and **2** in *P. aethiopicum* can be attenuated or abolished by using RNA silencing. This strategy has been previously employed in the identification of the chaetoglobosin gene cluster from *P. expansum* (Schumann and Hertweck, 2007), and in the study of tailoring oxidations during tenellin biosynthesis in *Beauveria bassiana* (Halo et al., 2008). Gene fragments of PaPKS0274 (3.8 kB) and PaPKS0880 (2.5kB) were separately cloned into the pBARGPE1 plasmid (Pall and Brunelli, 1993) in the antisense orientation under the control of a *gpdA* promoter. The resultant plasmids, pBAR274si and pBAR880si, were randomly integrated into the *P. aethiopicum* genome. We recovered 9 clones for PaPKS0274 and 6 clones for PaPKS0880 that contain the full length *gpdA* promoter with contiguous PKS gene fragments. For the *vrtA* silencing, 3 out of the 9 PCR positive clones screened did not produce **1**, while the other 6 clones showed various degree of reduced viridicatumtoxin production compared to wild type (Figure S2). In the case for *gsfA*, all 6 clones continued to produce **2** and **3**, but at various lower titers compared to wild type (Figure S3). The level of silencing of the two PKS genes is likely linked to the copy number of the silencing constructs integrated in the transformants, the expression level of the antisense gene driven by

the *gpdA* promoter and the endogenous transcript levels of the two corresponding PKS genes. The inability to silence *gsfA* completely may be due to the high level of endogenous gene expression as reflected in the relatively large amount of **2** produced in the culture.

Comparative genomics analysis revealed the *vrt* and *gsf* cluster embedded within conserved syntenic regions of *P. aethiopicum* genome

A more detailed bioinformatic analysis of the two biosynthetic loci revealed adjacent genes that are highly similar and syntenic to those found in *P. chrysogenum* genome (80–99% identity). Except the regions surrounding *vrtA* and *gsfA*, the conserved syntenic blocks span the whole contig 0274 and 0880, and they correspond to a specific locus in *P. chrysogenum* contig Pc00c21 and Pc00c06, respectively (Figure 4). Under the assumption that secondary metabolic genes could be unique to individual species while housekeeping and primary metabolic genes are usually conserved among closely related organisms, we used those syntenic sets of orthologous genes to tentatively assign the boundaries of the two putative clusters (Figure 4).

In the case of the *vrt* gene cluster, 14 putative genes (designated as *vrtA-L*, *vrtR1* and *vrtR2*) were found between *orf3* and *orf4*, which share 93% and 92% identity to Pc21g16710 and Pc21g16780, respectively (Figure 4A and Table 1). Most of the genes upstream of *orf3* and downstream of *orf4* also share synteny and high similarities with genes in the corresponding locus on the *P. chrysogenum* genome, except that several putative *P. chrysogenum* pseudogenes in the locus appear to have been replaced with the *vrt* cluster. The *gsf* cluster consists of 13 putative genes (designated as *gsfA-K*, *gsfR1* and *gsfR2*) and was found inserted between *orf4'* and *orf5'*, which are highly similar to Pc06g01020 and Pc06g01000 (92% and 94% identity respectively). The *gsf* cluster replaces Pc06g01010 in the *P. chrysogenum* genome (Figure 4B and Table 2). The genes immediately flanking the *gsf* cluster also remained syntenic with the corresponding locus in the *P. chrysogenum* genome. These observations of *vrt* and *gsf* clusters embedded within conserved syntenic regions led to speculation that *P. aethiopicum* may have acquired the two gene clusters via horizontal gene transfer (Walton, 2000). However, careful examinations of the Pc00c21 and Pc00c06 contigs revealed the presence of small conserved regions (some annotated as pseudogenes), which appeared to be fragments of *vrt* and *gsf* gene clusters, in the corresponding loci in *P. chrysogenum* (Figure 4). This suggests that *P. chrysogenum* may have possessed the *vrt* and *gsf* gene clusters in the past but have lost them during evolution.

Both clusters have two transcription factors (*vrtR1* and *vrtR2*, *gsfR1* and *gsfR2*) that contain Zn(II)₂Cys₆ DNA-binding domains, which are similar to pathway specific regulators such as *aflR* and *ctnA* for aflatoxin and citrinin biosynthesis respectively (Ehrlich et al., 1999; Shimizu et al., 2007). It is unclear if both or only one of the transcription factors is involved in the regulation of the corresponding cluster.

The putative biosynthetic pathways for the biosynthesis of viridicatumtoxin

Previous isotope incorporation study showed that the biosynthesis of the tetracyclic carboxamide core of **1** is significantly different from tetracyclines in bacteria (de Jesus et al., 1982), including: 1) differences in the cyclization regioselectivity of the carbon backbone (Thomas, 2001); 2) the non-acetate origin of C3; and 3) the retention of the oxygen at C4a of **1** from [1-¹³C¹⁸O]-acetate. These differences exclude the participation of a fully aromatic tetracene intermediate, such as the pretetramid intermediate in oxytetracycline biosynthesis (Thomas and Williams, 1983; Zhang et al., 2007). Based on the acetate labeling pattern and the new gene cluster data, a putative pathway for biosynthesis of **1** can be envisioned (Figure 5). The PT domain of a NRPKS has recently been shown to mediate aromatic polyketide cyclization (Crawford et al., 2009). Since the PT domain of VrtA shares high similarity to those

from AptA and ACAS (61% and 50% identity), which were shown to produce polyketide products with C6–C11 first-ring cyclization regioselectivity, we reasoned that the cyclization of the polyketide backbone of **1** by VrtA may also proceed via a similar route. We hypothesized that VrtA may utilize a malonamoyl-CoA starter unit, which is generated by VrtB, followed by sequential condensation of eight malonyl-CoA units to form the polyketide backbone. The cyclizations of the BCD rings were assumed to follow the pattern observed in atrochryson biosynthesis (Awakawa et al., 2009), in which the oxygen at C4a is retained from an acetate unit (Figure 5). Cyclization of the last ring (ring A) along with offloading of the intermediate **5** could be mediated by VrtG, which has protein sequence similarity to the recently discovered β -lactamase-type TE (AptB and ACTE, 60% and 49% protein identity) (Awakawa et al., 2009; Szewczyk et al., 2008).

The hypothesis of a malonamoyl-CoA starter unit produced by VrtB is supported by the high sequence similarity of VrtB to acetoacetyl-CoA synthetases (AACSSs) (44% protein identity with the rat homolog, AACSS_RAT), and the structural similarity of the AACSS substrate, acetoacetate, to malonamate. AACSS activates acetoacetate as an acyl-AMP intermediate followed by ligation to the free thiol of CoA to form acetoacetyl-CoA (Fukui et al., 1982). Similarly, the formation of malonamoyl-CoA from malonamate could be catalyzed by VrtB. Furthermore, *P. aethiopicum* has an additional copy of putative AACSS gene (PaAACSS1086) on contig1086, which is 96% identical to the only AACSS homolog (Pc13g10810) in *P. chrysogenum*, while VrtB shares only 50% protein identity with the putative PaAACSS (Pc13g10810). This implies that VrtB is most likely dedicated to the *vrt* pathway, while PaAACSS1086 is involved in the primary metabolism.

VrtJ has a pyridoxal 5'-phosphate (PLP) binding site and a conserved domain similar to those found in threonine aldolase (TA), aspartate aminotransferase and β -eliminating lyase (Table 1). The closest characterized homolog of VrtJ is Gly1 from *Candida albicans* (39% protein identity), which is a TA that catalyzes degradation of L-threonine to glycine and acetaldehyde (McNeil et al., 2000). The presence of an additional copy of TA-like gene (PaTA0310) in *P. aethiopicum*, which shares higher identity (92%) to the only homolog in *P. chrysogenum* (Pc12g01020) than VrtJ (54%), again implies that PaTA0310 is most likely to be the primary TA involved in glycine metabolism. VrtJ is therefore likely to be dedicated to the *vrt* pathway and may be involved in the synthesis of the malonamate substrate for VrtB. A malonamoyl-CoA starter unit is also likely involved in oxytetracycline biosynthesis, however, the malonate-derived carboxamide of oxytetracycline is in contrast to the acetate-derived carboxamide of **1** (Thomas and Williams, 1983). One possibility could be a malonamoyl-CoA starter derived from asparagine, presumably synthesized by VrtJ and VrtB collaboratively (Figure 5). Such a pathway, however, does not match the previous acetate labeling study as incorporation of an asparagine-derived malonamoyl-CoA would result in the C3 of **1** being labeled by the C2 of [1,2- ^{13}C]-acetate. This is due to asparagine is originating from oxaloacetate, and the formation of oxaloacetate from [1,2- $^{13}\text{C}_2$]-acetate proceeds via a symmetrical succinate intermediate in the tricarboxylic acid cycle (Ogasawara and Liu, 2009). Further gene targeting along with biochemical characterization of the four gene products, VrtA/B/G/J, are likely to shed light to the enzymatic reactions involved in the biosynthesis of the tetracyclic carboxamide core.

The proposed post-PKS tailoring steps following the formation of intermediate **5** are two hydroxylations and an *O*-methylation (Figure 5). The hydroxylations at C5 and C12a are predicted to be catalyzed by one or two of the oxygenases (VrtH, VrtI, VrtE, and VrtK), while the methoxy group is likely to be formed by the *O*-methyltransferase (VrtF). The pathway that leads to formation of the spirobicyclic ring of **1** is proposed to involve at least three gene products in the cluster (Figure 5). VrtD, which is similar to *trans*-isoprenyl diphosphate synthases (Table 1) such as the *Neurospora crassa* farnesyl pyrophosphate synthase (Homann et al., 1996), is predicted to catalyze the formation of geranylpyrophosphate. Based on the

shared homology of VrtC with aromatic prenyltransferases, such as *A. fumigatus* dimethylallyltryptophan synthase FgaPT2 (Metzger et al., 2009) (Table 1), the protein is predicted to catalyze the attachment of the geranyl moiety to the C ring of **1** to yield **6** (Figure 5). Most aromatic prenyltransferases identified in fungi transfer the prenyl group to a tryptophan or indole moiety (Heide, 2009). Intriguingly, VrtC is proposed to catalyze the transfer of a geranyl group to the aromatic C ring of the tetracyclic polyketide intermediate of **1**. Prenylation at the same C6 position is also observed in a similar fungal tetracyclic compound, hypomycesin, which has an acetyl group in place of the carboxamide in **1** (Breinholt et al., 1997). Thus, VrtC and the corresponding prenyltransferase in hypomycesin pathway can be considered as potential enzymes that can modify tetracycline compounds at C6.

We propose that the cyclization of the geranyl moiety of **6** can be initiated by a cytochrome P450 enzyme-catalyzed hydroxylation (VrtE or VrtK) at the allylic C17, which upon protonation of the hydroxyl can lead to a C17 carbocation (Figure 5). Attack of C20 on the bridging C15 forms the C15–C20 bond and transfers the positive charge to C19. A hydride shift from C15 to C19 is followed by C7 attack on C15 to form the spirobicyclic ring (Figure 5). The 1,3-hydride shift was previously proposed to rationalize the incorporation pattern of ²H-labeled mevalonate (Horak et al., 1988). The cyclization of the geranyl side chain to the spirobicyclic ring in **1** is most likely enzyme-mediated; however, no terpene cyclase gene is found in the vicinity of the *vrt* cluster. Thus, formation of the spirobicyclic ring in **1** could be mediated by the prenyltransferase (VrtC) as similarly proposed in paspaline biosynthesis (Saikia et al., 2006) or facilitated by the P450 enzyme.

The putative biosynthetic pathway for the biosynthesis of griseofulvin

The biosynthesis of **2**, as implied by the previously established pathway (Harris et al., 1976), required at least seven biosynthetic enzymes (Figure 6). Formation of the heptaketide backbone by GsfA is initiated by priming with acetyl-CoA, followed by sequential condensations of six malonyl-CoA units (Simpson and Holker, 1977). Since neither a fused TE/CLC nor a stand-alone TE is present, we proposed that the C1–C6 Claisen cyclization of the polyketide backbone could be mediated by the PT domain of GsfA, while the C8–C13 aldol cyclization could occur spontaneously to afford the benzophenone intermediate of **2** (Figure 6). Alternatively, the PT domain may mediate the cyclization of both aromatic rings. As PT domains of fungal PKSs are so far known to only catalyze aldol cyclizations (Crawford et al., 2009), the mechanism of folding and cyclization of the polyketide backbone by GsfA requires further investigation.

Isotope feeding experiment by Harris et al. (1976) established that two of the *O*-methylation steps at C6 and C12 hydroxyl groups occur immediately after the formation of the benzophenone intermediate to yield griseophenone C, while the methylation of the hydroxyl group at C10 occurs at a later stage (Figure 6). Three *O*-methyltransferases (GsfB/C/D) are encoded in the gene cluster and can therefore catalyze the required methylation steps.

An interesting feature in the biosynthesis of **2** is the formation of the heterocyclic ring by a stereospecific phenol oxidative coupling, which leads to the unique spiro structure of **2**. Geodin is another grisan compound similar to **2**, which also undergoes a similar stereospecific oxidative coupling. It has been established that the *A. terreus* dihydrogeodin oxidase (DHGO), which converts the benzophenone dihydrogeodin to (+)-geodin, is a multicopper blue protein similar to laccases (Fujii et al., 1987; Huang et al., 1995). The structural similarity of **2** to geodin suggests that a similar enzyme may be responsible for the oxidative coupling reaction, but no such enzyme is encoded in the *gsf* cluster. A P450 oxygenase (GsfF) was found in the *gsf* cluster, although no oxidation is required en route to **2**. We proposed that GsfF may catalyze the stereospecific oxidative coupling reaction of griseophenone C to form the grisan core (Figure 6). Oxidative coupling reactions catalyzed by cytochrome P450 enzymes are not unprecedented. A well-known example is OxyB involved in the phenol coupling reaction

during the biosynthesis of vancomycin (Zerbe et al., 2004). GsfK, a putative NAD(P)-dependent oxidoreductase, is likely to catalyze the stereospecific reduction step at the C-ring resulting in the *R* conformation at the C2, forming the second chiral center in **2** and **3** (Figure 6).

GsfI is a halogenase that shares high similarity (60% identity) to the recently discovered flavin-dependent chlorinase (RadH) in the biosynthesis of radicicol (Wang et al., 2008). In this current study, **3** is a major product in the *P. aethiopicum* culture when grown on stationary YMEG liquid culture, albeit often produced at a slightly lower ratio to **2**. To confirm the function of GsfI, the gene was disrupted in the same manner as the two PKSs. Two positive *gsfI* disruptants, Δ *gsfI-IV5* and Δ *gsfI-IV11*, were found after screening approximately 100 glufosinate-resistant transformants. LC-MS analysis showed that disruption of *gsfI* completely abolished the production of **2**, while production of **3** remained (Figure 2). This confirmed the role of GsfI as a halogenase responsible for the regiospecific chlorination of **3** at the C13 position to form **2**. The result also suggests that chlorination does not affect the formation of the grisan ring and could therefore be one of the last steps in the **2** biosynthetic pathway in *P. aethiopicum* (Figure 6).

In conclusion, we have successfully associated two polyketide products to their corresponding gene clusters by bioinformatics and comparative analysis of the *P. aethiopicum* genome, and subsequently confirmed the functions of three of the biosynthetic genes by targeted gene replacement. As opposed to the previously reported low gene targeting efficiency in other *Penicillium* spp. (Casqueiro et al., 1999; Schumann and Hertweck, 2007), our results showed that gene targeting by double homologous recombination in *P. aethiopicum* is relatively feasible. The *P. aethiopicum* transformation system developed in this study will facilitate future mechanistic studies of the two pathways, as well as the numerous other PKSs found in the genome of this fungus.

SIGNIFICANCE

The present study unveiled the biosynthetic gene clusters of two interesting polyketide spiro compounds, **1** and **2**, **1** is a rare example of fungal compound that is structurally similar to the tetracycline antibiotics from bacteria, while **2** is an important antifungal drug that has been in use for a long time for treating dermatophyte infections. Localization of the two gene clusters within conserved syntenic regions of the *P. aethiopicum* genome raises interesting questions regarding the evolution of clustering of secondary metabolite genes in fungi and may provide insights to the underlying genetic basis of *Penicillium* chemotaxonomy.

The *vrt* cluster is an interesting example of a biosynthetic pathway that encodes for a polyketide-isoprenoid hybrid compound in fungi. Understanding the exact mechanism of which the anhydrotetracycline-like core of **1** is formed by further targeted gene deletion and biochemical characterization of the gene products may provide new chemical insights. The unusual folding and cyclization of the isoprenoid moiety to form the spirobicyclic ring also deserves further investigation. More importantly, unveiling the clustered genes for biosynthesis of **1** has opened up the possibilities to generate novel tetracycline analogs using combinatorial biosynthetic strategies by integrating bacterial and fungal genes that can act on the tetracycline scaffolds.

There has been a renewed interest in **2** owing to its newly discovered anti-cancer and anti-viral properties. It is significant that this is the first report of the genetic basis of biosynthesis of **2**. Further functional characterization of the gene cluster will advance our understanding of biosynthesis of **2** and could lead to isolation of intermediates that may have important implications in producing useful structural analogs either by enzymatic or chemical modifications.

EXPERIMENTAL PROCEDURES

Strains and culture conditions

P. aethiopicum, IBT 5753, was obtained from the IBT culture collection (Kgs. Lyngby, Denmark) and maintained on YMEG agar (4 g/l yeast extract, 10g/l malt extract, 16 g/l agar) or glucose minimal medium (GMM) (Cove, 1966) at 28°C.

454 sequencing and bioinformatic analysis

The genomic DNA used for sequencing was prepared as described (Gauch et al., 1998) from mycelium grown in stationary liquid culture. The shotgun sequencing was performed at the GenoSeq (UCLA Genotyping and Sequencing Core) with the GS FLX Titanium system (Roche).

The 454 sequencing reads were assembled into contigs with the GS De Novo Assembler software (Roche). The contigs were converted into BLAST database format for local BLAST search using stand-alone BLAST software (ver. 2.2.18) downloaded from the NCBI website. Gene predictions were carried out using the FGENESH program (Softberry) and manually checked by comparing with homologous gene/proteins in the GenBank database. Functional domains in the translated protein sequences were predicted using Conserved Domain Search (NCBI) or InterproScan (EBI). Phylogenetic analysis was performed with MEGA4 software (Tamura et al., 2007).

Fungal transformation and gene disruption in *P. aethiopicum*

Polyethylene glycol-mediated transformation of *P. aethiopicum* was done essentially as described previously for *A. nidulans* (Andrianopoulos and Hynes, 1988; Chooi et al., 2008), except that the protoplasts were prepared with 3 mg/ml lysing enzymes (Sigma-Aldrich) and 2 mg/ml Yatalase (Takara Bio). Construction of fusion PCR knockout cassettes containing the *bar* gene were carried out as described (Szewczyk et al., 2007), except that the homologous regions flanking the resistant marker were increased to 2 kb. Fusion PCR products were gel purified and sequenced before using for transformation. The *bar* gene with the *trpC* promoter was amplified from the plasmid pBARKS1 (Pall and Brunelli, 1993), which was obtained from the Fungal Genetics Stock Center (FGSC). RNA-silencing plasmids were constructed from pBARGPE1 (Pall and Brunelli, 1993) obtained from FGSC. Glufosinate used for the selection of *bar* transformants was prepared by extracting twice with equal volume of 1-butanol from commercial herbicide Finale (Bayer), which contains 11.33 % (w/v) glufosinate-ammonium (Hays and Selker, 2000). The resulting aqueous phase was filter-sterilized and used directly at 40 µl per ml of GMM with 1.2M sorbitol and 10 mM ammonium tartrate as the sole nitrogen source. Miniprep genomic DNA from *P. aethiopicum* transformants was used for PCR screening of gene deletants and was prepared as described previously for *A. nidulans* (Chooi et al., 2008). Primers used for amplification of fusion PCR products and screening of transformants, were listed in Table S2. Southern hybridizations were performed with DIG-High Prime DNA labeling and detection starter kit II (Roche Applied Science) following manufacturer's protocol.

Chemical analysis and compound isolation

For small scale analysis, the *P. aethiopicum* wild type and transformants were grown in YMEG liquid medium (20mL) for 7 days at 28°C without shaking. The cultures were extracted with equal volume of ethyl acetate and evaporated to dryness. The dried extracts were dissolved in methanol for LC-MS analysis. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization and a Phenomenex Luna 5µ 2.0 × 100 mm C18 reverse-phase column. Samples were

separated at a flow rate of 0.1 ml/min on a linear gradient of 5 to 95% solvent B in 30 min followed by isocratic 95% solvent B for another 15 mins (solvent A: 0.1% (v/v) formic acid, solvent B: CH₃CN with 0.1% (v/v) formic acid). The identity of **1** and **2** were confirmed by comparing the UV spectra, retention time and *m/z* value to the authentic standards: **1** (a gift from Dr. Wongon Kim, Korea Research Institute of Bioscience & Biotechnology) and **2** (Sigma-Aldrich).

Dechlorogriseofulvin was isolated using a solvent-solvent partitioning scheme as described previously (Hutchison et al., 1973) followed by crystallization in MeOH. Briefly, the ethyl acetate extract from a two liters stationary liquid culture of a Δ *gsfI* mutant was evaporated to dryness and partition between CHCl₃/H₂O. The CHCl₃ layer was further partitioned with hexane/MeOH (90%). The 90% MeOH layer was dried completely and redissolved in a small volume of MeOH (1 ml). Formation of white crystals, which was later confirmed to be **3**, was observed when the MeOH extract was left undisturbed at room temperature (further crystallized at 4°C). The crystals (28 mg) were collected by filtration, washed in cold methanol and dried. Purity of the compound was checked by LC-MS and the structure was confirmed by NMR with reference to published spectra (Table S3).

HIGHLIGHTS

- 454 sequencing unveiled the gene clusters for viridicatumtoxin **1** and griseofulvin **2**.
- *vrt* and *gsf* clusters lie within conserved syntenic regions of *P. aethiopicum* genome.
- VrtA and GsfA are nonreducing PKSs required for biosynthesis of **1** and **2**.
- *gsfI* encodes chlorinase involved in the biosynthesis of **2**.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to W.G. Kim (Korea Research Institute of Bioscience & Biotechnology) for providing a standard of **1**. We thank Dr. W. Zhang for initial identification of **1** from *P. aethiopicum*.; X. Xie and W. Xu for assistance with NMR of **3**; and Prof. John Vederas and Prof. Neil Garg for helpful discussions. We also like to acknowledge the suggestions provided by a reviewer on the biosynthetic pathway of **1**. This project was supported by a David and Lucile Packard Fellowship and NIH (1R01GM085128).

References

- Andrianopoulos A, Hynes MJ. Cloning and analysis of the positively acting regulatory gene *amdR* from *Aspergillus nidulans*. *Mol Cell Biol* 1988;8:3532–3541. [PubMed: 3062382]
- Ariza MR, Larsen TO, Duus JO, Barrero AF. *Penicillium digitatum* metabolites on synthetic media and citrus fruits. *J Agric Food Chem* 2002;50:6361–6365. [PubMed: 12381117]
- Awakawa T, Yokota K, Funa N, Doi F, Mori N, Watanabe H, Horinouchi S. Physically discrete β -lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem Biol* 2009;16:613–623. [PubMed: 19549600]
- Barton, D.; Cohen, T. *Festschrift Prof Dr Arthur Stoll*. Birkhäuser Verlag; Basel: 1957. Some biogenetic aspects of phenol oxidation; p. 117
- Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat Chem Biol* 2007;3:213–217. [PubMed: 17369821]

- Bok JW, Hoffmeister D, Maggio-Hall LA, Murillo R, Glasner JD, Keller NP. Genomic mining for *Aspergillus* natural products. *Chem Biol* 2006;13:31–37. [PubMed: 16426969]
- Breinholt J, Jensen GW, Kjær A, Olsen CE, Rosendahl CN. Hypomycesin - an antifungal, tetracyclic metabolite from *Hypomyces aurantius*: production, structure and biosynthesis. *Acta Chem Scand* 1997;51:855–860.
- Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW. Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. *J Biol Chem* 2000;275:25533–25539. [PubMed: 10825154]
- Casqueiro J, Gutierrez S, Banuelos O, Hijarrubia MJ, Martin JF. Gene targeting in *Penicillium chrysogenum*: disruption of the *lys2* gene leads to penicillin overproduction. *J Bacteriol* 1999;181:1181–1188. [PubMed: 9973344]
- Chiang YM, Szewczyk E, Nayak T, Davidson AD, Sanchez JF, Lo HC, Ho WY, Simityan H, Kuo E, Praseuth A, et al. Molecular genetic mining of the *Aspergillus* secondary metabolome: discovery of the emericellamide biosynthetic pathway. *Chem Biol* 2008;15:527–532. [PubMed: 18559263]
- Chooi YH, Stalker DM, Davis MA, Fujii I, Elix JA, Louwhoff SHJJ, Lawrie AC. Cloning and sequence characterization of a non-reducing polyketide synthase gene from the lichen *Xanthoparmelia semiviridis*. *Mycol Res* 2008;112:147–161. [PubMed: 18280724]
- Clardy J, Springer JP, Buechi G, Matsuo K, Wightman R. Tryptoquivaline and tryptoquivalone, two tremorgenic metabolites of *Aspergillus clavatus*. *J Am Chem Soc* 1975;97:663–665. [PubMed: 1133368]
- Cove DJ. The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. *Biochim Biophys Acta* 1966;113:51–56. [PubMed: 5940632]
- Cox R. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org Biomol Chem* 2007;5:2010–2026. [PubMed: 17581644]
- Crawford JM, Dancy BCR, Hill EA, Udway DW, Townsend CA. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc Natl Acad Sci USA* 2006;103:16728–16733. [PubMed: 17071746]
- Crawford JM, Korman TP, Labonte JW, Vagstad AL, Hill EA, Kamari-Bidkorpeh O, Tsai SC, Townsend CA. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* 2009;461:1139–1143. [PubMed: 19847268]
- de Jesus A, Hull W, Steyn P, van Heerden F, Vlegaar R. Biosynthesis of viridicatumtoxin, a mycotoxin from *Penicillium expansum*. *J Chem Soc, Chem Commun* 1982:902–904.
- Ehrlich KC, Montalbano BG, Cary JW. Binding of the C6-zinc cluster protein, AFLR, to the promoters of aflatoxin pathway biosynthesis genes in *Aspergillus parasiticus*. *Gene* 1999;230:249–257. [PubMed: 10216264]
- Finkelstein E, Amichai B, Grunwald MH. Griseofulvin and its uses. *Int J Antimicrob Agents* 1996;6:189–194. [PubMed: 18611708]
- Frisvad, JC.; Samson, RA. Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*. A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. In: Samson, RA.; Frisvad, JC., editors. *Studies in Mycolgy 49: Penicillium subgenus Penicillium: new taxonomic schemes, mycotoxins and other extrolites*. Utrecht, The Netherlands: Centraalbureau voor Schimmelcultures; 2004. p. 1-174.
- Fujii I, Iijima H, Tsukita S, Ebizuka Y, Sankawa U. Purification and properties of dihydrogeodin oxidase from *Aspergillus terreus*. *J Biochem* 1987;101:11–18. [PubMed: 3032923]
- Fujii I, Watanabe A, Sankawa U, Ebizuka Y. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*. *Chem Biol* 2001;8:189–197. [PubMed: 11251292]
- Fukui T, Ito M, Tomita K. Purification and characterization of acetoacetyl-CoA synthetase from *Zoogloea ramigera* I-16-M. *Eur. J Biochem* 1982;127:423–428.
- Gauch, S.; Hermann, R.; Feuser, P.; Oelmuller, U.; Bastian, H. Isolation of nucleic acids using anion-exchange chromatography: Qiagen-tip based methods. In: Karp, A.; Ingram, DS.; Isaac, PG., editors. *Molecular Tools for Screening Biodiversity*. London: Chapman and Hall; 1998. p. 54-59.

- Halo LM, Heneghan MN, Yakasai AA, Song Z, Williams K, Bailey AM, Cox RJ, Lazarus CM, Simpson TJ. Late stage oxidations during the biosynthesis of the 2-pyridone tenellin in the entomopathogenic fungus *Beauveria bassiana*. *J Am Chem Soc* 2008;130:17988–17996. [PubMed: 19067514]
- Harris C, Roberson J, Harris T. Biosynthesis of griseofulvin. *J Am Chem Soc* 1976;98:5380–5386. [PubMed: 956563]
- Hays S, Selker E. Making the selectable marker *bar* tighter and more economical. *Fungal Genet Newsl* 2000;47:107.
- Heide L. Prenyl transfer to aromatic substrates: genetics and enzymology. *Curr Opin Chem Biol* 2009;13:171–179. [PubMed: 19299193]
- Hoffmeister D, Keller N. Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep* 2007;24:393–416. [PubMed: 17390002]
- Homann V, Mende K, Arntz C, Ilardi V, Macino G, Morelli G, Böse G, Tudzynski B. The isoprenoid pathway: cloning and characterization of fungal FPPS genes. *Curr Genet* 1996;30:232–239. [PubMed: 8753652]
- Horak R, Maharaj V, Marais S, Heerden F, Vlegaar R. Stereochemical studies on the biosynthesis of viridicatumtoxin: evidence for a 1,3-hydride shift in the formation of the spirobicyclic ring system. *J Chem Soc, Chem Commun* 1988;1988:1562–1564.
- Huang, K-x; Fujii, I.; Ebizuka, Y.; Gomi, K.; Sankawa, U. Molecular cloning and heterologous expression of the gene encoding dihydrogeodin oxidase, a multicopper blue enzyme from *Aspergillus terreus*. *J Biol Chem* 1995;270:21495–21502. [PubMed: 7665560]
- Hutchison RD, Steyn PS, van Rensburg SJ. Viridicatumtoxin, a new mycotoxin from *Penicillium viridicatum* Westling. *Toxicol Appl Pharmacol* 1973;24:507–509. [PubMed: 4122267]
- Ishimaru, T.; Tsuboya, S.; Saijo, T. Tetracyclic compounds, its production and use thereof. *JP*. 06–040995. 1993.
- Jarvis B, Zhou Y, Jiang J, Wang S, Sorenson W, Hintikka E, Nikulin M, Parikka P, Etzel R, Dearborn D. Toxicogenic molds in water-damaged buildings: dechlorogriseofulvins from *Memnoniella echinata*. *J Nat Prod* 1996;59:553–554. [PubMed: 8786360]
- Jin H, Yamashita A, Maekawa S, Yang P, He L, Takayanagi S, Wakita T, Sakamoto N, Enomoto N, Ito M. Griseofulvin, an oral antifungal agent, suppresses hepatitis C virus replication in vitro. *Hepatol Res* 2008;38:909–918. [PubMed: 18624717]
- Lane M, Nakashima T, Vederas J. Biosynthetic source of oxygens in griseofulvin. Spin-echo resolution of oxygen-18 isotope shifts in carbon-13 NMR spectroscopy. *J Am Chem Soc* 1982;104:913–915.
- McNeil J, Flynn J, Tsao N, Monschau N, Stahmann K, Haynes R, McIntosh E, Pearlman R. Glycine metabolism in *Candida albicans*: characterization of the serine hydroxymethyltransferase (SHM1, SHM2) and threonine aldolase (GLY1) genes. *Yeast* 2000;16:167. [PubMed: 10641038]
- Metzger U, Schall C, Zocher G, Unsöld I, Stec E, Li S, Heide L, Stehle T. The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. *Proc Natl Acad Sci USA* 2009;106:14309. [PubMed: 19706516]
- Nayak T, Szewczyk E, Oakley CE, Osmani A, Ukil L, Murray SL, Hynes MJ, Osmani SA, Oakley BR. A versatile and efficient gene targeting system for *Aspergillus nidulans*. *Genetics* 2006;172:1557–1566. [PubMed: 16387870]
- Ogasawara Y, Liu HW. Biosynthetic studies of aziridine formation in azicemicins. *J Am Chem Soc* 2009;131:18066–18068. [PubMed: 19928906]
- Pall M, Brunelli J. A series of six compact fungal transformation vectors containing polylinkers with multiple unique restriction sites. *Fungal Genet Newsl* 1993;40:59.
- Panda D, Rathinasamy K, Santra M, Wilson L. Kinetic suppression of microtubule dynamic instability by griseofulvin: implications for its possible use in the treatment of cancer. *Proc Natl Acad Sci USA* 2005;102:9878–9883. [PubMed: 15985553]
- Raju MS, Wu GS, Gard A, Rosazza JP. Microbial transformations of natural antitumor agents. 20 Glucosylation of viridicatumtoxin. *J Nat Prod* 2004;45:321–327.
- Rebacz B, Larsen T, Clausen M, Ronnest M, Löffler H, Ho A, Kramer A. Identification of griseofulvin as an inhibitor of centrosomal clustering in a phenotype-based screen. *Cancer Res* 2007;67:6342. [PubMed: 17616693]

- Rhodes A, Somerfield GA, Mcgonagle MP. Biosynthesis of griseofulvin. Observations on the incorporation of (¹⁴C)griseophenone C and (³⁶C)griseophenones B and A. *Biochem J* 1963;88:349–357. [PubMed: 14063874]
- Rønneest M, Rebacz B, Markworth L, Terp A, Larsen T, Kraemer A, Clausen M. Synthesis and structure-activity relationship of griseofulvin analogues as inhibitors of centrosomal clustering in cancer cells. *J Med Chem* 2009;52:3342–3347. [PubMed: 19402668]
- Saikia S, Parker E, Koulman A, Scott B. Four gene products are required for the fungal synthesis of the indole-diterpene, paspaline. *FEBS letters* 2006;580:1625–1630. [PubMed: 16494875]
- Samson, RA.; Seifert, KA.; Frisvad, JC. Phylogenetic analysis of *Penicillium* subgenus *Penicillium* using partial α -tubulin sequences. In: Samson, RA.; Frisvad, JC., editors. *Studies in Mycology 49: Penicillium subgenus Penicillium: new taxonomic schemes, mycotoxins and other extrolites*. Utrecht, The Netherlands: Centraalbureau voor Schimmelcultures; 2004. p. 175-200.
- Schumann J, Hertweck C. Molecular basis of cytochalasan biosynthesis in fungi: gene cluster analysis and evidence for the involvement of a PKS-NRPS hybrid synthase by RNA silencing. *J Am Chem Soc* 2007;129:9564–9565. [PubMed: 17636916]
- Schumann J, Hertweck C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J Biotechnol* 2006;124:690–703. [PubMed: 16716432]
- Shimizu T, Kinoshita H, Nihira T. Identification and *in vivo* functional analysis by gene disruption of *ctnA*, an activator gene involved in citrinin biosynthesis in *Monascus purpureus*. *Appl Environ Microbiol* 2007;73:5097–5103. [PubMed: 17586673]
- Simpson TJ, Holker JSE. ¹³C-NMR studies on griseofulvin biosynthesis and acetate metabolism in *Penicillium patulum*. *Phytochemistry* 1977;16:229–233.
- Szewczyk E, Chiang Y, Oakley C, Davidson A, Wang C, Oakley B. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microbiol* 2008;74:7607. [PubMed: 18978088]
- Szewczyk E, Nayak T, Oakley CE, Edgerton H, Xiong Y, Taheri-Talesh N, Osmani SA, Oakley BR. Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nat Protocols* 2007;1:3111–3120.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–1599. [PubMed: 17488738]
- Thomas R. A biosynthetic classification of fungal and streptomycete fused-ring aromatic polyketides. *ChemBioChem* 2001;2:612–627. [PubMed: 11828498]
- Thomas R, Williams D. Oxytetracycline biosynthesis: mode of incorporation of [1-¹³C]- and [1, 2-¹³C₂]-acetate. *J Chem Soc, Chem Commun* 1983;1983:128–130.
- Walton JD. Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fung Genet Biol* 2000;30:167–171.
- Wang S, Xu Y, Maine E, Wijeratne E, Espinosa-Artiles P, Gunatilaka A, Molnár I. Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. *Chem Biol* 2008;15:1328–1338. [PubMed: 19101477]
- Wong S, Kullnig R, Dedinas J, Appell K, Kydd G, Gillum A, Cooper R, Moore R. Anthrotainin, an inhibitor of substance P binding produced by *Gliocladium catenulatum*. *J Antibiot* 1993;46:214–221. [PubMed: 7682212]
- Yamazaki M, Fujimoto H, Okuyama E. Structure determination of six tryptoquivaline-related metabolites from *Aspergillus fumigatus*. *Tetrahedron Lett* 1976;17:2861–2864.
- Zerbe K, Woithe K, Li DB, Vitali F, Bigler L, Robinson JA. An oxidative phenol coupling reaction catalyzed by OxyB, a cytochrome P450 from the vancomycin-producing microorganism. *Angew Chem Int Ed* 2004;43:6709–6713.
- Zhang W, Watanabe K, Wang CCC, Tang Y. Investigation of early tailoring reactions in the oxytetracycline biosynthetic pathway. *J Biol Chem* 2007;282:25717–25725. [PubMed: 17631493]
- Zheng CJ, Yu HE, Kim EH, Kim WG. Viridicatumtoxin B, a new anti-MRSA agent from *Penicillium* sp FR11. *J Antibiot* 2008;61:633–637. [PubMed: 19168978]

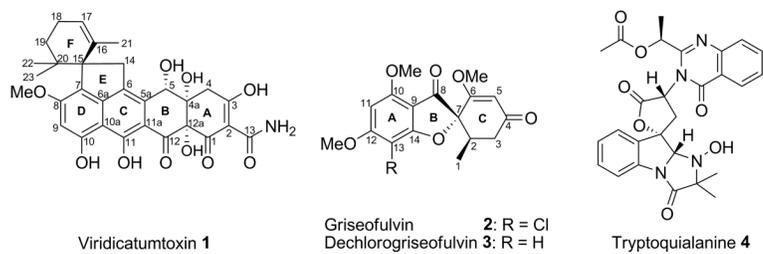


Figure 1. Structures of the major secondary metabolites produced by *P. aethiopicum*. The numbering schemes for the carbon atoms for **1–3** are based on previous studies (de Jesus et al., 1982; Simpson and Holker, 1977).

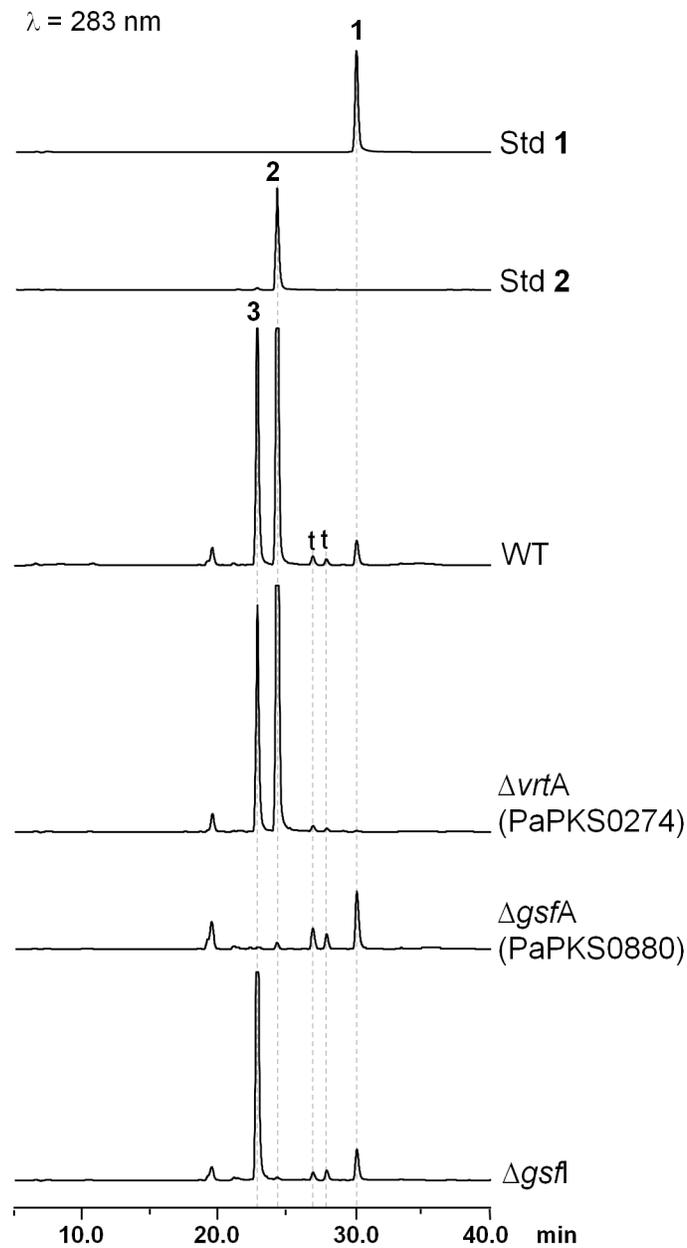


Figure 2. HPLC analysis (283 nm) of metabolites produced by wild type and mutant *P. aethiopicum* strains. The mutant $\Delta vrtA$, $\Delta gsfA$ or $\Delta gsfI$ no longer produced **1**, **2** or **3**, respectively. Standards of compounds **1** and **2** are shown. The identity of **3** was established by NMR characterization. Peaks labeled with **t** have UV spectra and m/z values matching to **4** and can be isomers.

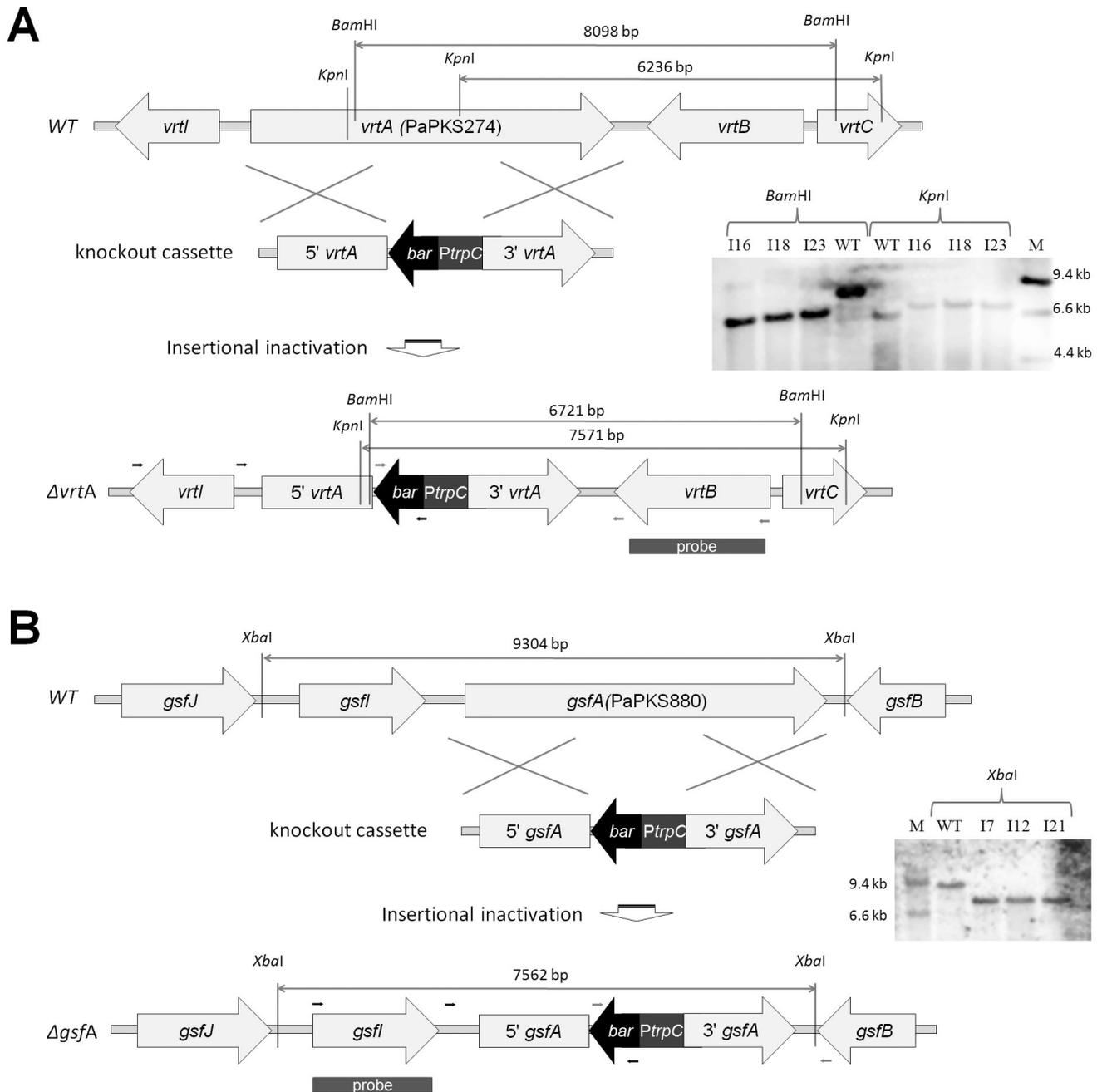
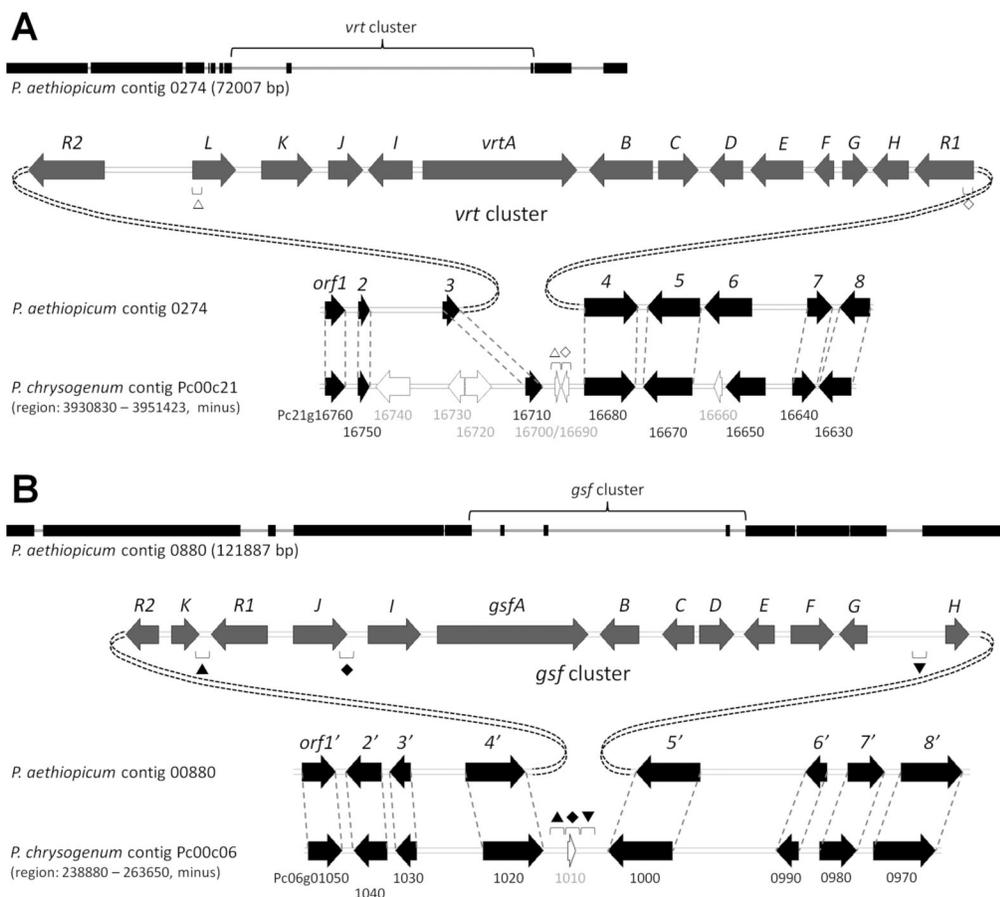


Figure 3. Homologous recombination schemes for deletion of *vrtA* and *gsfA* genes in *P. aethiopicum*. The knockout cassettes for (A) *vrtA* (PaPKS0274) (B) *gsfA* (PaPKS880) are shown with marked restriction sites. The corresponding Southern blot results for the wild type and the mutants are displayed on the right. Small arrows indicate approximate binding sites of the primers used in PCR screening.

**Figure 4.**

Putative *vrt* and *gsf* clusters embedded within conserved syntenic regions of *P. aethiopicum* genome. (A) *vrt* cluster on contig 0274 and the corresponding locus on *P. chrysogenum* contig Pc00c21. (B) *gsf* cluster on contig 0880 and the corresponding locus on *P. chrysogenum* contig Pc00c06. Blocks on the thin lines indicate conserved syntenic blocks on the contigs. The symbols (Δ ◇ ▲ ◆ ▼) indicate short conserved regions within the *vrt* and *gsf* clusters that are remained at the corresponding loci in *P. chrysogenum*. Block arrows: black, genes predicted not to be involved in the biosynthesis of **1** and **2**; white, putative pseudogenes.

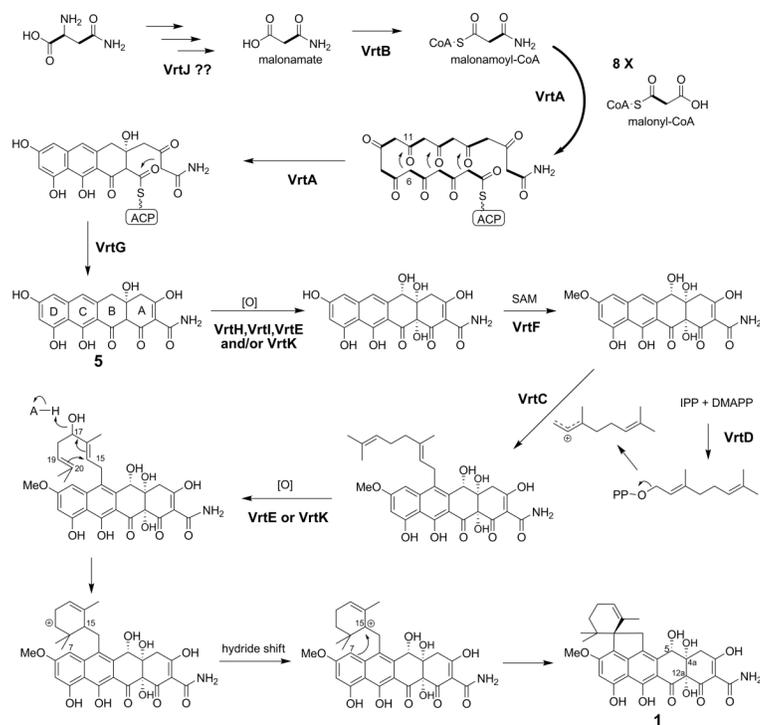


Figure 5. Proposed pathway for the biosynthesis of viridicatumtoxin **1**. DMAPP – dimethylallyl pyrophosphate; IPP – isopentenyl pyrophosphate.

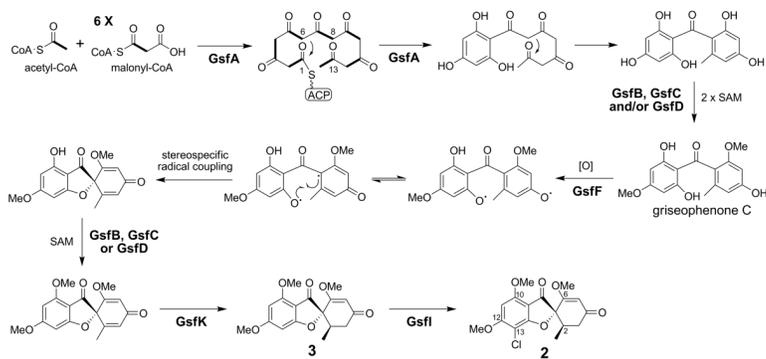


Figure 6. Proposed biosynthesis of griseofulvin **2**. The function of the halogenase GsfI has been confirmed using knockout as shown in Figure 2.

Putative genes within and flanking the *vrr* cluster.

Table 1

Gene	Size (bp/aa)	BLASTP homolog	Identity/Similarity (%)	Conserved Domain	E value
<i>vrrA</i>	5578/1824	<i>A. nidulans</i> , AN6000.3 (AptA)	60/75	SAT-KS-MAT-PT-ACP	< 1.0e-180
<i>vrrB</i>	2312/723	<i>Pyrenophora tritici-repentis</i> , PTRG_06131 <i>Rattus norvegicus</i> , AACR_RAT	51/64 40/56	PRK03584, acetoacetyl-CoA synthetase	
<i>vrrC</i>	1452/483	<i>Microsporium canis</i> , MCYG_03599 <i>A. fumigatus</i> , FgaPT2 [314X]	45/59 20/36	TIGR03429, aromatic prenyltransferase	2e-73
<i>vrrD</i>	1206/349	<i>Neurospora crassa</i> , FACPS_NEUCR [Q92250]	54/68	cd00685, trans-isoprenyl diphosphate synthases	7e-46
<i>vrrE</i>	1898/520	<i>A. terreus</i> , ATEG_04107	33/52	pfam00067, cytochrome P450	5e-24
<i>vrrF</i>	717/238	<i>A. niger</i> , An11g07340	55/65	pfam08242, methyltransferase domain	1e-04
<i>vrrG</i>	924/307	<i>Polyangium cellulosum</i> , JetF	35/49	type 12	
<i>vrrH</i>	1305/413	<i>A. nidulans</i> , AN6001.3 (AptB)	60/74	smart00849, metallo- β -lactamase superfamily	2e-24
<i>vrrI</i>	1597/413	<i>A. nidulans</i> , AN6002.3 (AptC) <i>A. clavatus</i> , ACLA_098360	50/69 39/50	COG0654, UbiH and related FAD-dependent oxidoreductases	4e-11
<i>vrrJ</i>	1249/388	<i>Nicotiana tabacum</i> , Ntc12 gibberillin 20-oxidase	18/35	COG3491, PbcC, IPNS and related dioxygenases	2e27
<i>vrrK</i>	1847/534	<i>Candida albicans</i> , Gly1 [GLY1_CANAL]6	47/61	pfam03171, 2OG-Fe(II) oxygenase	2e-11
<i>vrrL</i>	1570/503	<i>A. flavus</i> , <i>verB</i> desaturase	37/53	cd06502, low-specificity threonine aldolase (TA), PLP-dependent aspartate aminotransferase superfamily (fold I)	1e-77
<i>vrrM</i>	1847/534	<i>A. nidulans</i> , <i>stcL</i> desaturase	37/54	pfam01212, Beta-eliminating lyase	4e-70
<i>vrrN</i>	1570/503	<i>A. niger</i> , An13g00720	31/47	pfam00067, p450, Cytochrome P450	3e-41
<i>vrrO</i>	2148/715	<i>Neosartorya fischeri</i> , NFIA_025860	80/89	cd06174, MFS, The Major Facilitator Superfamily transporters	1e-31
<i>vrrP</i>	2757/836	<i>A. niger</i> , An11g07350	70/81	pfam04082, fungal specific transcription factor domain	2e-09
<i>orf1</i>	789/262	<i>P. chrysogenum</i> , Pc21g16760	34/48	smart00066, GAL4-like Zn(II)2Cys6 binuclear cluster DNA-binding domain	2e-05
<i>orf2</i>	444/147	<i>P. chrysogenum</i> , Pc21g16750	91/95	COG0400, Predicted esterase	8e-12
<i>orf3</i>	675/224	<i>P. chrysogenum</i> , Pc21g16710	99/100	No conserved domain detected	
<i>orf4</i>	1974/657	<i>P. chrysogenum</i> , Pc21g16680	93/95	COG4297, uncharacterized protein with double-stranded beta helix domain	3e-14
<i>orf5</i>	1499/462	<i>P. chrysogenum</i> , Pc21g16670	92/94	No putative conserved domains detected	
<i>orf6</i>	1722/573	<i>Uncinocarpus reesii</i> , UREG_05338	85/92	cd05120, aminoglycoside 3'-phosphotransferase (APH) and choline kinase (ChoK) family.	1e-04
			26/40	cd00204, ANK, ankyrin repeats; mediate protein-protein interactions	0.004

Gene	Size (bp/aa)	BLASTP homolog	Identity/Similarity (%)	Conserved Domain	E value
<i>orf7</i>	888/263	<i>P. chrysogenum</i> , Pc21g16640	97/99	pfam09177, Syntaxin 6, N-terminal	2e-21
<i>orf8</i>	1112/319	<i>P. chrysogenum</i> , Pc21g16630	95/98	pfam05739, SNARE domain pfam00956, NAP, Nucleosome assembly protein	2e-08 1e-09

Genus abbreviation: *A. -Aspergillus*, *P. - Penicillium*.

Sequence of the *vrf* gene cluster was submitted to GenBank under the accession GU574477.

Table 2

Putative genes within and flanking the *gsf* cluster.

Gene	Size (bp/aa)	BLASTP homolog	Identity/Similarity (%)	Conserved Domain	E value
<i>gsfA</i>	5672/1790	<i>B. fuckeliana</i> , PKS14	57/71	SAT-KS-MAT-PT-ACP	
		<i>A. terreus</i> , ATEG_08451 (ACAS)	42/60		
<i>gsfB</i>	1442/419	<i>P. marneffei</i> , PMAA_079120	41/61	pfam00891, <i>O</i> -methyltransferase	7e-33
		<i>G. fujikuroi</i> , Bik3 bikaverin <i>O</i> -methyltransferase ³	29/45	COG2226, UbiE, Methylase involved in ubiquinone/menaquinone biosynthesis	9e-07
<i>gsfC</i>	1191/396	<i>N. haematococca</i> , NECHADRAFT_81295	32/51	pfam00891, <i>O</i> -methyltransferase	2e-20
		<i>A. flavus</i> , <i>omitB</i> aflatoxin <i>O</i> -methyltransferase ⁴	24/44		
<i>gsfD</i>	378/1301	<i>P. chrysogenum</i> , Pc12g06140	26/41	pfam00891, <i>O</i> -methyltransferase	1e-21
		<i>A. flavus</i> , <i>omitB</i> aflatoxin <i>O</i> -methyltransferase ⁴	27/45		
<i>gsfE</i>	1134/377	<i>A. nidulans</i> , AN9028.2	61/74	COG0451, WcaG, Nucleoside-diphosphate-sugar epimerases/dehydratases	1e-09
<i>gsfF</i>	1603/462	<i>N. haematococca</i> , NECHADRAFT_3047	37/53	pfam00067, p450, Cytochrome P450	2e-37
<i>gsfG</i>	1033/299	<i>C. globosum</i> , CHGG_02201	24/37	cd00204, ANK, ankyrin repeats; ankyrin repeats mediate protein-protein interactions	7e-37
<i>gsfH</i>	889/237	<i>A. oryzae</i> , AO090001000095	81/89	cd01012, YcaC_related, YcaC related amidohydrolases	1e-31
<i>gsfI</i>	1976/533	<i>C. chiversii</i> , RadH flavin-dependent halogenase	61/75	pfam04820, tryptophan halogenase	3e-15
				COG0644, FixC, dehydrogenases (flavoproteins)	3e-17
<i>gsfJ</i>	2031/554	<i>A. nidulans</i> , AN8459.2	47/61	TIGR00711, efflux_EmrB, drug resistance transporter, EmrB/QacA subfamily	1e-24
<i>gsfK</i>	1028/251	<i>P. chrysogenum</i> , Pc12g16460	40/59	PRK06953, short chain dehydrogenase	4e-24
<i>gsfR1</i>	2118/688	<i>A. nidulans</i> , AN8460.2	52/68	pfam04082, fungal specific transcription factor domain	2e-05
<i>gsfR2</i>	1248/415	<i>Sclerotinia sclerotiorum</i> , SS1G_05579	53/67	smart00066, GAL4-like Zn(II)2Cys6 binuclear cluster DNA-binding domain	8e-09
<i>orf1'</i>	1246/351	<i>P. chrysogenum</i> , Pc06g01050	86/90	No conserved domain detected	
<i>orf2'</i>	1310/366	<i>P. chrysogenum</i> , Pc06g01040	96/99	cd03445, Thioesterase II repeat2	7e-27
				cd03444, Thioesterase II repeat1	2e-20
<i>orf3'</i>	765/211	<i>P. chrysogenum</i> , Pc06g01030	80/86 [‡]	No conserved domain detected	
<i>orf4'</i>	2187/653	<i>P. chrysogenum</i> , Pc06g01020	92/96	PRK08310, amidase	4e-30
<i>orf5'</i>	2352/783	<i>P. chrysogenum</i> , Pc06g01000	94/96	pfam04082, fungal specific transcription factor domain	2e-16
<i>orf6'</i>	807/157	<i>P. chrysogenum</i> , Pc06g00990	90/94	No conserved domain detected	
<i>orf7'</i>	1346/411	<i>P. chrysogenum</i> , Pc06g00980	94/96	pfam07942, N2227-like protein	1e-85
<i>orf8'</i>	2268/702	<i>P. chrysogenum</i> , Pc06g00970	98/98	COG0443, DnaK, Molecular chaperone	2e-17

Genus abbreviation: *A.* - *Aspergillus*, *B.* - *Botryotinia*, *C.* - *Chaetomium*, *G.* - *Giberella*, *N.* - *Nectria*, *P.* - *Penicillium*.

[‡] based on P06g01030 conceptual translation as predicted by FGENSEH, different from the version in GenBank. Sequence of the *gcf* gene cluster was submitted to GenBank under the accession GU574478.