Published in final edited form as:

Environ Microbiol. 2018 September ; 20(9): 3325–3342. doi:10.1111/1462-2920.14333.

Enfumafungin synthase represents a novel lineage of fungal triterpene cyclases

Eric Kuhnert^{1,2}, Yan Li^{1,3}, Nan Lan¹, Qun Yue^{1,3}, Li Chen¹, Russell J. Cox², Zhiqiang An¹, Kenichi Yokoyama⁴, and Gerald F. Bills^{1,*}

¹Texas Therapeutics Institute, the Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA.

²Leibniz University Hannover, Institute for Organic Chemistry and BMWZ, Hannover, Germany.

³Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China.

⁴Duke University School of Medicine, Department of Biochemistry, Durham, NC, USA.

Summary

Enfumafungin is a glycosylated fernene-type triterpenoid produced by the fungus *Hormonema carpetanum*. Its potent antifungal activity, mediated by its interaction with β -1,3-glucan synthase and the fungal cell wall, has led to its development into the semi-synthetic clinical candidate, ibrexafungerp (=SCY-078). We report on the preliminary identification of the enfumafungin biosynthetic gene cluster (BGC) based on genome sequencing, phylogenetic reconstruction, gene disruption, and cDNA sequencing studies. Enfumafungin synthase (*efuA*) consists of a terpene cyclase domain (TC) fused to a glycosyltransferase (GT) domain and thus represents a novel multifunctional enzyme. Moreover, the TC domain bears a phylogenetic relationship to bacterial squalene-hopene cyclases (SHC) and includes a typical DXDD motif within the active centre suggesting that *efuA* evolved from SHCs. Phylogenetic reconstruction of the GT domain indicated that this portion of the fusion gene originated from fungal sterol GTs. Eleven genes flanking *efuA* are putatively involved in the biosynthesis, regulation, transport and self-resistance of enfumafungin and include an acetyltransferase, three P450 monooxygenases, a dehydrogenase, a desaturase and a reductase. A hypothetical scheme for enfumafungin assembly is proposed in which the E-ring is oxidatively cleaved to yield the four-ring system of enfumafungin. *EfuA* represents the first member of a widespread lineage of fungal SHCs.

Introduction

Based on the success and the limitations of the echinocandin class of antifungals, fungal natural products in the fernene class, represented by enfumafungin produced by *Hormonema carpetanum*, are highly promising novel anti-fungals (Fig. 1) (Peláez *et al.*, 2000; Schwartz *et al.*, 2000; Bills *et al.*, 2004). Echinocandins and enfumafungin target β -1,3-glucan synthase, a multiprotein complex required for normal formation of the fungal cell wall (Onishi *et al.*, 2000), but enfumafungin apparently binds at a location distinct from the

*For correspondence. billsge@vt.edu or gerald.f.bills@uth.tmc.edu; Tel. +1 713 486-2344.

binding site for echinocandins. Mutations in glucan synthase in *Candida albicans* and *C. glabrata* that result in resistance to the echinocandins are located in FKS1 domains that differ from those associated with decreased susceptibility to ibrexafungerp, supporting the lack of cross-resistance between the two different agents *in vitro*. As a result, enfumafungin does not exhibit cross resistance with echinocandins, and remains active against pathogens that have acquired resistance against echinocandins (Jiménez-Ortigosa *et al.*, 2017; Pfaller *et al.*, 2017). Chemical derivatives of enfumafungin also show *in vitro* and *in vivo* activity against multi-drug resistant pathogens, including azole- and echinocandin-resistant fungi. Furthermore, the experimental drug is effective against the emerging multidrug-resistant pathogen *C. auris*, classified by the CDC as an emerging global health threat (Berkow *et al.*, 2017; Larkin *et al.*, 2017).

The oral availability and the activity against drug resistant pathogens have made enfumafungin an excellent lead for next generation antifungal therapeutics. Merck and Scynexis generated series of enfumafungin derivatives, and identified ibrexafungerp that has advanced to phase 2 clinical trials as oral formulations with indications for invasive and vulvovaginal candidiasis, and as a step-down therapy from intravenous echinocandins in patients with invasive candidiasis (Wring *et al.*, 2017a, b). Such an experimental drug would offer a major advantage over the echinocandin class of antifungals because of its oral availability and its effectiveness against emerging echinocandin-resistant strains.

Triterpenoid natural products derived from squalene are cornerstones for diverse biological activities, including anticancer, antibacterial, antiviral and antifungal activities, as well as forming structural steroids, such as cholesterol and ergosterol. Similar fungal-produced triterpenes, some with antifungal effects, have been described elsewhere, including retigeric acids (Sun *et al.*, 2009; Chang *et al.*, 2012), lobarialides (Wang *et al.*, 2009), polytolypin (Gamble *et al.*, 1995), peniciside (Yuan *et al.*, 2012), fuscoatroside (=FR207944) (Joshi *et al.*, 2002; Kobayashi *et al.*, 2005), hyalodendrosides (Bills *et al.*, 2000) and the kolokosides (Deyrup *et al.*, 2007) (Fig. 2).

Despite the importance of enfumafungin as the starting material for the semi-synthesis of ibrexafungerp and its unique combination of structural features, no information regarding the mechanism of its biosynthesis has been reported. Enfumafungin is structurally characterized by the presence of four rings with the six-membered D-ring, and glucose attached to 3'-OH (Fig. 1). Based on its ring structures, enfumafungin was classified as a fernene-type triterpene (Schwartz *et al.*, 2000). Our current bio-synthetic understanding of fungal triterpenes is limited to lanostanes (Schmidt-Dannert, 2015), and little is known about fungal fernene-type triterpenes. However, a small group of fungal hopene-squalene-type terpene cyclases were noted to be phylogenetically allied to those of bacteria (Frickey and Kannenberg, 2009), but at the time, the products of these anomalous hopene-squalene-type terpene synthases were unknown. We show here that the enfumafungin synthase, efuA, is a member of this cryptic lineage of eukaryotic squalene-hopene cyclases (SHCs), and recognition of the linkage will guide recognition of other biosynthetic products of this widespread class of fungal terpenes. Therefore, elucidation of the steps involved in enfumafungin biosynthesis will expand our understanding of fungal triterpene biosynthesis.

Herein, we report the identification of the enfumafungin biosynthetic gene cluster (BGC) in *H. carpetanum* based on genome sequencing, phylogenetic reconstruction, gene disruption and cDNA sequencing studies.

Results

Hormonema carpetanum genome characteristics

Genome sequencing of the wild-type strain of *H. carpetanum* ATCC 74360 yielded a genome length of 32,775,731 bp. The genome sequence was assembled into 129 contigs with an average size of 254,075 bp (1028–1,757,996 bp) and a N50 of 537,720 bp. The average GC percentage was 49.96%. A total of 9976 putative genes were predicted leading to a gene density of 304 genes Mbp⁻¹ and a total gene size of 17,735,073 (54.1% of the genome size). Among those, 151 tRNA and 65 rRNA genes were identified. Additional information on the classification of predicted protein functions is summarized in Supporting Information Fig. S1.

Bioinformatic prediction of the enfumafungin gene cluster

To identify genes responsible for enfumafungin biosynthesis, the genome was analyzed for the presence of secondary metabolite BGCs using the antiSMASH platform (Blin *et al.*, 2017). In total, 31 tentative BGCs were predicted, including 10 BGCs containing putative polyketide synthase genes (PKS), 6 with putative nonribosomal peptide synthetase (NRPS) genes, 1 putative NRPS-PKS hybrid, eight non-classified biosynthetic core enzymes and 6 putative terpene biosynthetic enzymes.

Enfumafungin is a triterpene glycoside with distinctive structural features that include an acetate group at the C-2 position of the triterpene backbone, a glucose unit *O*-linked to the C-3 position, a hemiacetal functionality and a carboxylic acid group at C-18 (Fig. 1A). In eukaryotes, including fungi, all triterpenes with a known biosynthetic pathway so far are biosynthesized from 2,3-oxidosqualene through the action of a lanosterol synthase or a close orthologue. The only previously characterized enzymes that use squalene as a substrate are squalene-hopene-cyclases exclusively found in bacteria. Therefore, in the fungal kingdom, lanosterol synthases, are the only enzymes, thus far, identified to form cyclic triterpenoids, and lanostanes are the most common backbone for such triterpenes (Ríos and Andújar, 2016). Thus, we initially looked for BGCs with a lanosterol synthase homologue accompanied by genes that might encode backbone-modifying enzymes, such as a GT, acetyltransferase and oxygenases. The genome indeed contained a putative terpene BGC with a core gene with 52% identity and 98% query coverage to the lanosterol synthase of *Pneumocystis carinii* (Q96WJ0) (Milla *et al.*, 2002). This putative BGC also coded for an oxidoreductase and a cysteine synthase. However, it was not accompanied by a GT or oxygenase genes that would be required to modify lanosterol to the enfumafungin backbone.

Among the six putative terpene BGCs in the genome, only one other BGC contained a putative triterpene cyclase. Surprisingly, this putative cyclase gene (*efuA*) showed sequence similarity to genes encoding bacterial SHCs, including some of the same sequences noted in a prior study on the phylogeny of triterpenes cyclases (Frickey and Kannenberg, 2009).

Moreover, similarity searches and domain prediction tools revealed that the C-terminus of the cyclase is fused to a complete GT domain (26% identity to the sterol 3- β -glucosyltransferase of *Aspergillus fumigatus*, Q4WID6) (Nierman *et al.*, 2005), suggesting the presence of a TC-GT fusion protein. This predicted TC-GT gene, *efuA*, is flanked by genes for a putative acetyltransferase, cytochrome P450s and other oxygenases that would be consistent with the biosynthetic logic needed to assemble enfumafungin, as well as transporters and transcriptional regulators (Fig. 3). Therefore, we reasoned these genes most likely constitute the enfumafungin BGC.

Phylogenetic analysis of enfumafungin synthase (*efuA*)

Domain structure analysis predicted the core gene (*efuA*) with a length of 3997 bp which includes a terpene synthase domain fused to a GT domain in one open reading frame. The gene contains a single short intron (52 bp) and encodes a 1314 aa (amino acid) long protein. Because TC–GT fusion proteins have not been reported before, we searched databases with the protein sequence to determine whether this type of enzyme might be present, but as-yet unrecognized, in genome sequences of other organisms. Triterpene cyclase sequences of fungi, bacteria, plants and animals were aligned and used to calculate a maximum likelihood backbone tree (Fig. 4). *EfuA* clustered with high bootstrap support (100%) with a set of proteins encoded by the genomes of fungi from the Leotiomycetes, Eurotiomycetes, Sordariomycetes, Dothideomycetes of the Pezizomycotina (Ascomycota), including *Rhynchosporium agropyri*, *R. commune*, *R. secalis*, *Endocarpon pusillum*, *Diaporthe ampelina*, *Valsa mali*, *V. mali* var. *pyri* and *Phialocephala scopiformis*. The function of these proteins is unknown. After comparing and re-annotating the respective amino acid sequences, all of these predicted proteins were found to have a similar length and included a GT domain (Fig. 5). The identity between the predicted protein sequences ranged from 39.4% to 98.0% (Supporting Information Table S1) identity with respect to amino acid content. Therefore, prediction of a TC–GT fusion proteins was unlikely to be a sequencing artefact because orthologous fusion proteins were also found in many other fungi. In fact, such enzymes eventually may be found to be common because genome sampling of the fungi remains limited. However, to date, none of the corresponding strains or species has been studied for their respective secondary metabolites, and thus the products of their TC–GT proteins remain unknown. Within the phylogenetic tree (Fig. 4), two conspicuous clades were exclusively formed by bacterial orthologues of SHCs.

A larger sister clade to the TC–GT clade included triterpene cyclases with significant identity to the *efuA* terpene cyclase domain, but all lacked a GT domain (Fig. 4). This larger clade encompassed an even broader range of fungi, including some species of the Basidiomycota. Their similarity with fungal terpene cyclases in the phylogenetic reconstruction suggested they might share a common ancestor rather than being evolved independently (Frickey and Kannenberg, 2009). This broad array of sequences were more closely rooted with bacterial SHC orthologues as noted previously (Frickey and Kannenberg, 2009), suggesting that they encode for triterpenes with hopene-type ring systems. Indeed some of the fungi in these clades, for example, *Aschersonia* species, have been reported to produce hopane-type terpenoids (Isaka *et al.*, 2009, 2010).

With the exception of the lanosterol synthases, including ERG7 (Milla *et al.*, 2002) which have been proved to be involved in the formation of different natural products (e.g., ergosterol, ganoderic acid) (Zhang *et al.*, 2017), only two additional fungal triterpene cyclases have been identified thus far, which are responsible for the biosynthesis of helvolic acid (Mitsuguchi *et al.*, 2009) and clavatic acid (Godio *et al.*, 2007; Godio and Martín, 2009). Helvolic acid and clavatic acid both have lanostane backbones, and the amino acid sequences of their synthases forms a clade with high bootstrap support (100%) that includes fungal lanosterol synthases, including the predicted ERG7-like gene from the *H. carpetanum* genome, suggesting that the functions of these triterpene synthases are likely very similar to lanosterol synthases (Fig. 4).

To infer the relationship of the GT domains from the predicted TC–GTs, a phylogenetic tree with related amino acid sequences was calculated (Fig. 6). The *efuA* GT domain forms a highly supported (100%) delimited clade with all other GT domains that are putatively fused to a terpene cyclase. The clade groups with a characterized sterol 3- β -glucosyltransferase from *Arabidopsis thaliana* (Q9XIG1) (DeBolt *et al.*, 2009), but the clade lacks bootstrap support. Furthermore, the enfumafungin-type GTs are nested between a cluster containing a functionally characterized sterol GT of *Saccharomyces cerevisiae* (Q06321) (Warnecke *et al.*, 1999) and various clades with GTs that show homology to the latter. These data imply that the enfumafungin GT is derived from a sterol GT ancestor. The phylogenetic analysis of *efuA* indicated that TC–GT fusion proteins are a previously unrecognized and distinctive lineage of fungal triterpene cyclase that is highly conserved across at least four ascomycete classes.

Composition of the enfumafungin BGC

Eleven genes proximal to *efuA* are likely to be involved in enfumafungin biosynthesis, regulation, transport and possibly self-resistance (Fig. 3, Table 1). Three genes encode P450 monooxygenases (*efuB*, *efuG* and *efuH*). A similarity search against the Swissprot database in Gen-bank reveals a similarity of *efuH* with the *tenB* monooxygenase (A0JJT9, 88% query coverage, 24% identity) responsible for *N*-hydroxylation of the 2-pyridone nitrogen during tenellin biosynthesis (Halo *et al.*, 2008). Based on the structure of enfumafungin, minimally three P450 enzymes are required for the functionalization of C-2, and C-25 and the cleavage of the E-ring (see ‘Discussion’ for biosynthetic hypothesis). In particular, the hydroxylation of the C-2 position is a prerequisite for the site of acetate esterification, and critical for potent anti-fungal activity. In addition, the C-3 position may also need to be hydroxylated prior to the coupling of the sugar moiety in the case that squalene serves as the substrate for *efuA*.

The acetyltransferase gene (*efuC*; 1656 bp; no introns) likely encodes an enzyme for the acetylation at C-2. Acetyltransferase genes have been identified as common tailoring enzymes of terpene biosynthesis gene clusters in various fungi including trichothecenes, botrydial, helvolic acid and PR toxin (Kimura *et al.*, 1998; Pinedo *et al.*, 2008; Mitsuguchi *et al.*, 2009; Hidalgo *et al.*, 2014).

Domain searches for the predicted protein of *efuE* revealed a SerA (phosphoglycerate dehydrogenase or related dehydrogenase) or FDH-GDH-like (formate/glycylate

dehydrogenase) superfamily domain respectively. This kind of enzyme can catalyze various reactions of which the removal of hydrogen atoms from its substrate (oxidation) is the most common. Even though dehydrogenases have been frequently found in terpene gene clusters, their role in terpene biosynthesis has only been partially elucidated as is the case for helvolic acid, terretinin and PR toxin (Mitsuguchi *et al.*, 2009; Guo *et al.*, 2012; Hidalgo *et al.*, 2014). A possible function in enfumafungin biosynthesis might be the oxidation of a putative hydroxy group at C-25, which could trigger a spontaneous hemiacetal formation if C-23 is hydroxylated as well (e.g., by a monooxygenase).

Two other redox enzyme genes were identified for possible participation in the gene cluster. The predicted protein efuI shows similarity with the GA4 desaturase of *Fusarium fujikuroi* (Q96UG5, 87% query coverage, 30% identity), which is responsible for gibberellin biosynthesis and catalyzes α -ketoglutarate dependent oxidation of the C-1/C-2 bond into an olefin (Tudzynski *et al.*, 2003). The function of this enzyme is difficult to predict, but it might play a role in the establishment of the farnene-type backbone by isomerization of the double bond. The efuL peptide sequence resembles those of oxidoreductases of the short-chain dehydrogenase reductase (SDR) superfamily and exhibits similarities with isoflavone reductase of the plant *Pisum sativum* (P52576, 76% query coverage, 28% identity) involved in phytoalexin biosynthesis (Paiva *et al.*, 1994). SDRs are commonly found in fungal terpene biosynthetic pathways and are responsible for backbone modifications, for example, formation of a ketone functionality in helvolic acid biosynthesis (Mitsuguchi *et al.*, 2009). Due to complexity and temporal permutations of possible reactions, how efuL participates in enfumafungin biosynthesis cannot yet be predicted.

The gene cluster contained four non-enzymatic genes that may be involved in the transportation, self-resistance and regulation of enfumafungin biosynthesis. The gene *efuF* encodes a 493-amino-acid protein with similarities to the UhpC superfamily (sugar phosphate permease). This putative transporter enzyme might facilitate the transport of the glucose units to the subcellular site of enfumafungin biosynthesis. The putative transcription regulator gene efuD encodes a transcription initiation factor IIE domain and is possibly responsible for the transcription regulation of one or more genes within the gene cluster. However, this type of putative transcription factor has not been associated with the specific regulation of biosynthetic genes in any organism, and therefore, its role will require experimental confirmation. The putative transporter gene efuK codes for a pimeloyl-ACP methyl ester carboxylesterase (abhydrolase superfamily) domain and oxysterol-binding protein domain. The latter portion shares 40% identity with the kes1 protein of *Schizosaccharomyces pombe* (O74178), an enzyme that is thought to be a non-vesicular sterol transporter (Schulz *et al.*, 2009). Consequently, efuK might be involved in transport of enfumafungin to and across organelle membranes. The last gene proposed to be part of the enfumafungin gene cluster is efuJ. Domain structure analysis of its predicted peptide reveals no known domains, but 21% of the aa sequence show high identity to the SED1 cell wall protein of *Saccharomyces cerevisiae*, which is required for compensation of cell wall instability (Q01589, 51% identity) (Shimoi *et al.*, 1998; Hagen *et al.*, 2004). Judging from these results, efuJ probably does not encode any steps in enfumafungin biosynthesis, but because it is related to a fungal cell wall structural protein, we speculate that its product might play a role in self-resistance toward enfumafungin.

EfuA is essential for enfumafungin biosynthesis

To verify that *efuA* encodes the terpene synthase responsible for assembling the enfumafungin scaffold, hygromycin-resistant disruption mutants were created via protoplast transformation (Fig. 7A). Transformants successfully grew in hygromycin-amended media (Fig. 7C), and the disruption of *efuA* was verified by PCR (Fig. 7B). Production of enfumafungin in transformants and wild-type strain was compared by growing both strain types in BRFT medium. Culture extracts were analyzed by HPLCMS and screened for antifungal activity against *C. albicans*. The HPLC/MS analyses demonstrated the absence of enfumafungin in the transformants, whereas the wild-type strain retained its normal production levels (Fig. 8A). The growth-inhibition assay confirmed the loss of the antifungal activity in transformants, while the wild-type extracts caused strong inhibition zones (Fig. 8B). Therefore, we concluded that *efuA* is the core terpene cyclase gene of the enfumafungin BCG.

EfuA is a bifunctional protein

Bifunctional enzymes comprising of fused prenyl transfer-ase and terpene cyclase domains have been discovered from fungi, such as the fusicoccadiene synthase from *Phomopsis amygdali* (Toyomasu *et al.*, 2007) and ophiobolin sesterterpene synthases from various fungi (Chiba *et al.*, 2013; Brill *et al.*, 2016). However, to the best of our knowledge GTs have not yet been reported as part of multifunctional proteins. To verify that a single gene encodes the TC and GT domains, RNA was obtained from *H. carpetanum* under enfumafungin-producing conditions and reverse-transcribed into cDNA. Sequencing of the resulting *efuA* cDNA confirmed that the TC and GT domains are linked in a single open reading frame. The predicted single intron was verified as well.

Discussion

Fungal terpenoids are a diverse class of natural products with a manifold of biological activities both beneficial and detrimental for humans. Due to the advances of genome sequencing technologies, terpene synthase genes in fungi have been identified and characterized for all of major lineages (Quin *et al.*, 2014; Schmidt-Dannert, 2015). Although triterpene synthases are abundant in fungal genomes, thus far, only three have been associated with secondary metabolites. In addition, the lanosterol synthase (LS) is essential for the growth of the organism as it presents the key enzyme of ergosterol bio-synthesis (Parks and Casey, 1995). A single copy of the latter gene is present in the genome of *H. carpetanum* which we assume to be responsible for ergosterol biosynthesis. Recently, the LS in *Ganoderma lingzhi*, besides its role in ergosterol biosynthesis, has been shown to provide the intermediate for the biosynthesis of ganoderic acids (Zhang *et al.*, 2017). Another related gene tagged as oxidosqualene–clavarinone cyclase (*occ*) was revealed to encode the core cyclase for the clavatic acid biosynthetic pathway in the basidiomycete *Hypholoma sublateritium* (Godio *et al.*, 2007; Godio and Martín, 2009). Respective disruption mutants did not require sterols for growth, hence, it was concluded that *occ* was not involved in ergosterol biosynthesis. Interestingly, *occ* showed a high identity in the amino acid sequence with the LS of *G. lingzhi* (Fig. 4) suggesting that both catalyze the formation of the lanosterol backbone (Schmidt-Dannert, 2015). A protostadienol synthase

(AfuOSC3) was identified as the core cyclase in the helvolic acid biosynthetic pathway in *A. fumigatus*. For the first time, it was shown that fungal triterpene cyclases are able to form slightly different carbon backbones (Mitsuguchi *et al.*, 2009). However, the close phylogenetic relationship of the latter enzyme with other LSs point to an origin from a common lineage of oxidosqualene cyclases. In contrast, efuA is the first example of a fungal triterpene synthase that is unrelated to LSs. The amino acid sequence of efuA only shares a 15%–20% identity with other fungal LSs, but shares a 35%–37% identity with bacterial SHCs. Therefore, it is tempting to speculate that efuA ancestors might share a common origin with bacterial hopene cyclases, as it has been previously hypothesized (Frickey and Kannenberg, 2009).

Enfumafungin represents a class of glycosylated triterpenoids that are thought to be related to a fernene-type carbon skeleton (Schwartz *et al.*, 2000) (Fig. 2). Similar compounds have been described including peniciside from *Penicillium* sp. (Yuan *et al.*, 2012), fuscoatroside from *Humicola fuscoatra* (Joshi *et al.*, 2002) and the identical FR207944 from *Chaetomium* sp. (Kobayashi *et al.*, 2005), WF11605 from an unknown soil fungus (Shigematsu *et al.*, 1992), hyaldendrosides A and B from *Hyalodendron* sp. (Bills *et al.*, 2000) and kolokosides A-D from *Xylaria* sp. (Deyrup *et al.*, 2007). All these metabolites, with the exception of the kolokosides, exhibited some degree of antifungal activity. Hence, it is unclear if acetylation at the C-2 position is required for the antifungal activity, because structurally similar unacetylated and unglycosylated fernene-type terpenoids, for example, retigeric acid B, the lobarialides and polytolypin also are antifungal (Gamble *et al.*, 1995; Sun *et al.*, 2009; Wang *et al.*, 2009). However, with the exception of retigeric B (Chang *et al.*, 2012), the modes of action of these metabolites have not been determined, and other inhibitory mechanisms may be involved.

Comparisons of the structural features of glycosylated fernene-type terpenoids (Fig. 2) reveal a striking difference in the number of rings established in the core scaffold, which range from four (enfumafungin, fuscoatroside, WF11605, kolokoside A) to five (peniciside, hyaldendroside B, kolokoside C-D), and up to six rings (hyaldendroside A). These differences raise the question of whether these compounds are derived from different core terpene structures or if they share the same biosynthetic core enzymes. As none of the producer organisms has been genome-sequenced, homology searches with the enfumafungin gene cluster is not yet possible. However, a partial hypothesis can be developed by analyzing the structural variations among the different kolokosides (Deyrup *et al.*, 2007), which are likely products of the same biosynthetic pathway. In order to form kolokoside C or D, respectively, the five-membered ring of kolokoside B is proposed to undergo ring expansion catalyzed by a Baeyer–Villiger monooxygenase in a similar fashion as reported for the biosynthesis of the meroterpenoid austinol (Matsuda *et al.*, 2013). The 6-membered ring of kolokoside D could then be enzymatically cleaved to create the carboxylic acid moiety of kolokoside A. In contrast, the sixth ring in hyaldendroside A is probably formed by hydroxylation at C-28 and C-29 of hyaldendroside B, followed by oxidation of the hydroxy group at C-28 and a subsequent spontaneous hemiacetal formation (Fig. 9). Therefore, it is plausible that glycosylated fernene-type triterpenoids are derived from the same precursor (fernane) and share similar BCG architecture.

Finally, through retrospective prospecting of public genomes, we suggest that polytolypin (Fig. 2) can be tentatively connected to its BGC from *Polytolypa hystrix* (NCBI PRJNA234736) (Muñoz *et al.*, 2018). The BGC (Supporting Information Fig. S2) contains a terpene cyclase gene with strong identity to the *efuA*-TC (Fig. 4), but lacks the GT domain, which is consistent with the unglycosylated triterpene product. Furthermore, the polytolypin BGC harbours four P450 oxygenases which would be consistent with high degree of oxygenation of the carbon backbone.

Based on the predicted genes included in the enfumafungin BGC and the structural similarities between enfumafungin and related compounds, a biosynthetic scheme is proposed (Fig. 9). EfuA has the canonical DXDD motif (Fig. 5) conserved among bacterial SHCs and would be responsible for the activation of the C-2/C-3 bond to initiate the cation-mediated cyclization of squalene. Therefore, squalene could be a substrate of this enzyme. However, we also consider a possibility, in which *efuA* acts on 2,3-oxidosqualene. Bacterial SHCs are capable of cyclizing 2,3-oxidosqualene to give C-3-hydroxylated hopene (Gao *et al.*, 2012), although to our knowledge, no SHC is known to use 2,3-oxidosqualene as a physiologically relevant substrate. In fungi, 2,3-oxidosqualene is the primary precursor of triterpenes, including ergosterol, and therefore, is likely to be abundant. Thus, it is conceivable that 2,3-oxidosqualene is the physiological substrate for enfumafungin synthase. The reliance on 2,3-oxidosqualene as a substrate is also attractive in that the product of the cyclization is already hydroxylated at the C-3 position, thus facilitating the GT domain to act directly on the product of the TC domain. Therefore, this mechanism offers a biochemical rationale to fuse TC and GT domains to facilitate the metabolism flux. Moreover, unglycosylated fernene-type compounds, such as retigeric acid B, the lobarialides and polytolypin contain a C-3 oxygen, indicating a 2,3-oxidosqualene precursor is plausible.

The cyclization of squalene or 2,3-oxidosqualene accompanies a series of H and Me migrations to yield the precursor with the methylation pattern and stereo-chemistry consistent with enfumafungin. Bacterial SHCs usually conclude the cyclization cascade without H or Me migration by a hydration of the C-22 cation or a deprotonation at C-29. On the other hand, lanosterol synthase catalyzes H and Me migrations, similar to what we propose for enfumafungin synthase. Considering that SHCs have been considered as evolutionary ancestors of lanosterol synthases, it appears that *efuA*, which catalyzes hopene-type ring formation while catalyzing H and Me migrations, resides in the middle of such an evolutionary transition. Biochemical characterization of *efuA* is underway in our group to understand its function, mechanism and evolutionary relationships with SHC and lanosterol synthase.

If 2,3-oxidosqualene is considered as the substrate of the GT domain, then *efuA* most likely catalyzes glycosylation immediately after cyclization. The fernene glycoside then could be processed by *efuI* which catalyzes isomerization of a double bond established by *efuA* to form the core structure (Fig. 9). The latter would then undergo a series of hydroxylations in unknown order at C-2, C-19, C-23 and C-25, which would be catalyzed by two of the three monooxygenases (*efuB*, *efuG* or *efuH*). The hydroxy-group at C-25 becomes oxidized by *efuE* to enable a spontaneous, non-enzymatic hemiacetal formation with C-23. After hydroxylation at C-2, acetylation by *efuC* takes place. The final steps in enfumafungin

Author Manuscript

biosyn-thesis require expansion of the 5-membered ring by lactonization via a Baeyer–Villiger reaction mediated by one of the BGC’s monooxygenases (efuB, efuG or efuH) followed by ring cleavage (Fig. 9). CYP85A2 from *Arabidopsis thaliana* and CYP85A3 from tomato are cytochrome P450s that mediate Baeyer–Villiger oxidations resulting in a steroidal lactone ring during the conversion of castasterone to brassinolide (Kim *et al.*, 2005; Nomura *et al.*, 2005), so the involvement of one of BCG’s P450 monooxygenases is plausible. This type of reaction would establish a double bond between C-20 and C-21 which could be reduced by the reductase enzyme efuL to form the final product. Functional analysis studies of the respective genes of the enfumafungin BGC will be necessary to establish functions of each of genes and the sequence of the catalytic reactions.

Author Manuscript

Hormonema carpetanum is the only organism known to produce enfumafungin to date. Therefore, whether other fungi with related TC–GTs are able to biosynthesize enfumafungin-like metabolites will be of great interest, particularly if such metabolites affect fungal β -1,3-glucan synthase and are orally available. An analysis of the respective gene clusters the strains in the TC–GT terpene cyclase clade (Fig. 4) revealed that all of them contained two P450 monooxygenases adjacent to the core enzyme arranged in the same order as in efuB, efuG and efuH. However, no acetyltransferase, dehydrogenase, desaturase or reductase could be identified (Supporting Information Fig. S3). Hence, these strains are predicted to produce glycosylated farnene-type triterpenoids other than enfumafungin.

Author Manuscript

The discovery of the enfumafungin BGC reveals a new lineage of fungal triterpene synthases and the first hopene-type cyclase in the Fungi linked to its natural product. The gene cluster also appears to encode novel glycosylation and oxidation reactions, including a potentially unprecedented ring opening mechanism. Furthermore, understanding the sequence and timing of reactions in the pathway will enable the engineering of the pathway to generate new derivatives as starting points for semi-synthesis directed at improved antifungal potency and a broader spectrum of activity. Moreover, simplified starting material for semi-synthesis of the clinical candidate ibrexafungerp could be made. For example, the synthesis of the drug candidate starts with acid-mediated deglycosylation and deacetylation (Heasley *et al.*, 2012), which could be circumvented by simultaneously mutating the efuA GT domain and disrupting the BGC’s acetyltransferase gene. Overexpression of the BCG’s transcription factor might also enhance the fermentation yields of enfumafungin, thus improving industrial scale production of ibrexafungerp.

Experimental procedures

Fungal and bacterial strains, vectors and other reagents

Author Manuscript

The first strain reported to produce enfumafungin, *H. carpetanum* ATCC 74360 (Peláez *et al.* 2000), was used for genetic manipulations and production of authentic enfumafungin. Antifungal activity assays were carried out with *C. albicans* (ATCC 10231). Stellar™ Competent Cells (Clontech Laboratories, Inc., Mountain View) were used for vector cloning. The plasmid pAg1-H3 served as template for vector construction (Zhang *et al.* 2003). The PrimeSTAR® GXL DNA polymerase (Clontech Laboratories) was used for PCR. Restriction endonucleases were purchased from New England Biolabs (Beverly, Massachusetts). Plasmids were purified with the QIAprep Spin Miniprep Kit (Qiagen, Venlo, The

Netherlands). PCR products and digested vectors were purified with NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany).

Production, isolation and confirmation of enfumafungin

Mycelium of ATCC 74360 was cultured in solid rice medium (BRFT) (Peláez *et al.*, 2000) for 14 days at 23°C to prepare a standard of enfumafungin and to verify its production in wild-type and mutant strains. After growth, cultures were extracted with acetone for 2 h in an ultrasonic bath, and the organic phase was vacuum-filtered through filter paper. The extract was concentrated and dried by vacuum evaporation in a rotovap. The dried extract was partitioned in a hexane/methanol/water (+0.1% formic acid) mixture (8:6:1), and the lower layer was evaporated to dryness. The crude extract was pre-fractionized with Sephadex LH-20 as described previously (Schwartz *et al.*, 2000), and the active fraction inhibiting growth of *C. albicans* ATCC 10231 was separated by preparative high performance liquid chromatography (HPLC) with an Agilent 1260 HPLC equipped with a diode array detector (DAD) using the following conditions: mobile phase of acetonitrile supplemented with 0.1% formic acid (solvent A) and deionized water (Merck Millipore, Billerica, MA) supplemented with 0.1% formic acid (solvent B); an Agilent Zorbax SB-C18 column (5 µm, 9.4 × 250 mm) was used as the stationary phase with a linear gradient from 40% to 60% solvent A in 25 min with a flow rate of 2 ml min⁻¹, and with detection at 200±10 nm.

The structure of the isolated enfumafungin was confirmed by collecting NMR data using a Bruker AVANCE-500 MHz spectrometer with a 5-mm triple-resonance cryoprobe at 298 K (Supporting Information Figs. S4–S8) and comparing the spectra with those from literature (Schwartz *et al.*, 2000). For analytical-scale monitoring of enfumafungin, the same solvents and HPLC coupled to an Agilent 6120 single quadrupole mass detector were employed using an Agilent Zorbax Eclipse Plus C18 column (5 µm, 4.6 × 150 mm) and a linear gradient from 5% to 100% solvent A in 20 min, followed by 8 min isocratic conditions at a flow rate of 0.6 ml min⁻¹. Enfumafungin was detected at 200 ± 10 nm at a retention time of 16.2 min. Mass signals were recorded in API-ES (atmospheric pressure ionization–electrospray) negative ionization mode (3500 V capillary voltage), where the target compound exhibited a weak signal at *m/z* 707.

DNA isolation and sequencing

Genomic DNA of *H. carpetanum* ATCC 74360 was extracted using a previously described protocol (Zhang *et al.*, 2003). Library construction and genome sequencing was carried out by MR. DNA LABS (Shallowater, TX). The library was prepared using Nextera DNA Sample preparation kit (Illumina, San Diego, CA) with an average fragment size of 771 bp and clustered with the Illumina cBot. Sequencing of the clusters was conducted with an Illumina HiSeq 2500 sequencing system with paired end reading for 500 cycles. *De novo* assembly of the sequences was performed with the SeqMan NGen software (ver. 12, DNASTAR, Madison, WI).

Genome annotation

Genes from ATCC 74360 were predicted with the annotation tool Augustus (Stanke and Morgenstern, 2005) using *Aureobasidium namibiae* CBS 147.97 (GenBank [PRJNA292600](#)) as a reference genome. Predicted functions and EC numbers of proteins were assigned using UniProt BLAST search (<http://www.uniprot.org/>). Protein families were evaluated against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using KEGG Automatic Annotation Server (KAAS: <http://www.genome.jp/kegg/kaas/>). Carbohydrate-active enzymes (CAZymes) were predicted with the automated carbohydrate-active enzyme annotation database (dbCAN, <http://csbl.bmb.uga.edu/dbCAN/>) (Yin *et al.*, 2012) using the family classification from CAZy database (<http://www.cazy.org/>) (Lombard *et al.*, 2014). Transfer RNAs (tRNA) were identified with ARAGORN Ver. 1.2.36 (Laslett and Canback, 2004) and ribosomal RNA (rRNA) with Barnap (<http://www.vicbioinformatics.com/software.barnap.shtml>). Secondary metabolite gene clusters were predicted with antiSMASH (<http://fungismash.secondarymetabolites.org#!/start>) (Blin *et al.*, 2017). GenBank accession numbers for the enfumafungin BGC are listed in Table 1.

Phylogenetic analyses of enfumafungin terpene cyclase and glycosyltransferase domains

The predicted protein sequence of *efuA* was divided into its terpene cyclase (TC) and glycosyltransferase (GT) domains by alignment to corresponding amino acid (aa) sequences of functionally verified fungal TCs and GTs from GenBank. Each portion was searched against the GenBank database, and highly similar sequences were used for phylogenetic reconstruction. Protein sequences of the eight closest hits for the TC were re-annotated from the original genome sequence to predict complete genes. Additional identified GT domains from putative fusion proteins were included in the GT analysis (in case of *Diaporthe ampelina* and *Phialocephala scopiformis*). Sequences were aligned with the MUSCLE alignment tool of Geneious using default settings and 100 iterations. Amino acid sequences were truncated to unambiguously aligned regions leaving a maximum of 736 characters (minimum 469) for TC analysis and 561 characters (minimum 392) for GT analysis. Phylogenetic reconstruction was carried out with the RAXML plugin of Geneious under the GAMMA model of rate heterogeneity, BLOSUM62 substitution matrix, and 100 rapid bootstrap inferences. For the TC phylogeny, a lanosterol synthase-type gene identified from the *H. carpetanum* genome and five closely related GenBank sequences were used as the outgroup. The GT phylogeny was rooted with a *Bacillus subtilis* glycosyltransferase (AIY96331).

Disruption of the enfumafungin synthase gene, *efuA*

To verify the function of *efuA*, a hygromycin resistance-based construct was generated and inserted into the target gene via protoplast transformation. Two distantly positioned fragments of the targeted gene with a length of 2191 bp (EFTC1) and 2162 bp (EFTC2, partially contains non-coding downstream region), respectively, were amplified with primer sets containing a 15 bp overlap region necessary for directional cloning (see Supporting Information Table S3). PCR conditions and purification of the products followed the manufacturer's recommendations. Amplification products were visualized by gel electrophoresis (1% agarose in TAE buffer). Fragment sizes were estimated with the iVDye

1 kb DNA ladder (GenDEPOT, Katy, TX). Fragments were cloned into the pAg1-H3 vector containing the hygromycin resistance gene (Zhang *et al.*, 2003) with the In-Fusion® HD Cloning Kit (Clontech) using the PvuII and ApaI restriction sites for EFTC1 and AscI and SbfI restriction sites for EFTC2. The cloning and transformation of the competent cells was achieved following the supplier's protocols. Positive colonies were verified by PCR using a third primer set possessing binding sites within EFTC1 and EFTC2 (EFTC1-F2, EFTC2-R2, Supporting Information Table S3). Corresponding plasmids were isolated and used as a template to amplify the disruption cassette with the EFTC1-F/EFTC2-R primer set. The constructs were purified and stored at -80°C.

To transform ATCC 74360, a 50-ml seed culture in SMY medium (40 g maltose, 10 g yeast extract, 10 g Bacto neopeptone in 1 l deionized H₂O) was incubated on an orbital shaker for 3–4 days at 25°C and 220 rpm. Under these conditions, the strain forms mainly yeast-like conidia. The culture was then filtered through a layer of Miracloth (EMD Millipore) to remove hyphae. Cells were centrifuged for 5 min at 3000g and washed twice with sterile water. Washed cells were centrifuged again and resuspended in 10 ml of SMY medium. The protoplasting solution was prepared according to the protocol described by (Oakley *et al.*, 2012) with the modification of adding 15 mg of yeast lytic enzyme (Alfa Aesar, Haverhill, MA) after filtration. The cell suspension and 10 ml of protoplasting solution were combined and incubated on a shaker at 30°C and 120 rpm. Protoplasts formed after 1–2 h with up to 50% of the cells forming protoplasts. The protoplasts were harvested by centrifugation at 3000g at room temperature for 3 min. The supernatant was removed and cells were washed twice in 15 ml of STC buffer [1.2 M sorbitol, 10 mM Tris-HCl (pH 7.5), 50 mM CaCl₂]. Afterward, the solution was centrifuged again, and cells were resuspended in 1 ml of STC buffer. About 100 µl of the suspension were transferred to a 1.5-ml tube and mixed by pipetting with approximately 5 µg of construct DNA. The mixture was incubated on ice for 10 min. Next, 50 µl of PEG solution (25% PEG, 100 mM CaCl₂, 0.6 M KCl, 10 mM Tris-HCl, pH 7.5) was added, mixed and incubated on ice for 20 min. In the last step, 1 ml of PEG solution was combined with the previous solution and incubated at room temperature for 30 min. After incubation, the mixture was transferred to 50 ml of liquid minimal medium (MM) (Oakley *et al.*, 2012) supplemented with 0.8% agar and 1 M sucrose, and a thin layer was poured on top of MM plates containing 1.5% agar, 0.2 M sucrose and 100 µg ml⁻¹ hygromycin. Transformants appeared after 10–14 days of incubation at 27°C.

Verification of mutants

Transformants were transferred to SMY plates supplemented with hygromycin and incubated at room temperature for 7 days. DNA was isolated from mycelia with the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's recommendation. To verify whether the construct was inserted into the genome at the designated position, a diagnostic PCR reaction was employed. The corresponding primer pair contained a binding position upstream of the *efuA* (*efuA*-US-F) and within the TrpC promoter of the hygromycin resistance gene (HygR-R, Supporting Information Table S3). PCR conditions and visualization of the corresponding products were as described above.

Seed cultures were prepared from fresh mycelia of transformants by inoculating 50 ml of liquid SMY media and were agitated at 220 rpm and 25°C for 5 days. About 1 ml of seed culture was used to inoculate BRFT production media. Fermentation, extraction and HPLC analyses were as described above. Chromatograms of transformants were compared with those of the wild-type strain grown under the same conditions to verify presence or absence of enfumafungin. The same extracts were tested for inhibition of *C. albicans* ATCC 10231 in a zone of inhibition assay.

RNA isolation and reverse transcription

Hormonema carpetanum was grown in 250-ml flasks containing 50 ml of production medium (50 g glucose, 1 g tryptophan, 10 g yeast extract, 33 g N-Z-amine-AS, 5 g (NH₄)₂SO₄, 9 g KH₂PO₄, in 1 l of deionized H₂O adjusted to pH 6.2 before autoclaving) for 8 days at 25°C and 220 rpm (Liesch *et al.*, 1998). RNA was isolated from mycelia using the ZR Fungal/Bacterial RNA MiniPrep (Zymo Research, Irvine, CA) following the manufacturer's protocol. RNA was reverse transcribed into cDNA applying the High Capacity RNA-to-cDNA Kit (Applied Biosystems, Foster City, CA). The *efuA* cDNA was sequenced with four sets of primers (efuA_1F–efuA_4R, see Supporting Information Table S3).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

EK's work at the University of Texas Health Science Centre at Houston was supported by a German Academic Exchange Service (DAAD) Postdoctoral Fellowship. In addition, work was supported by a Robert Welch Foundation Grant AU-20030616 (to ZA), the University of Texas Health Science Centre at Houston new faculty start-up funds, the Kay and Ben Fortson Endowment, and the National Institute of General Medical Sciences R01GM121458 (to GB). KY was supported by the Duke University Medical Center and the National Institute of General Medical Sciences R01GM115729. RJC thanks the Deutsche Forschungsgemeinschaft for LCMS equipment (INST 187/621).

References

- Berkow EL, Angulo D, and Lockhart SR (2017) *In vitro* activity of a novel glucan synthase inhibitor, SCY-078, against clinical isolates of *Candida auris*. Antimicrob Agents Chemother 61: e00435–17. [PubMed: 28483955]
- Bills G, Dombrowski A, Morris SA, Hensens O, Liesch JM, Zink DL, et al. (2000) Hyalodendrosides A and B, antifungal triterpenoid glycosides from a lignicolous hyphomycete, *Hyalodendron species*. J Nat Prod 63: 90–94. [PubMed: 10650085]
- Bills GF, Collado J, Ruibal C, Peláez F, and Platas G (2004) *Hormonema carpetanum* sp. nov., a new lineage of dothideaceous black yeasts from Spain. Stud Mycol 50: 149–157.
- Blin K, Wolf T, Chevrette MG, Lu XW, Schwalen CJ, Kautsar SA, et al. (2017) AntiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res 45: W36–W41. [PubMed: 28460038]
- Brill ZG, Grover HK, and Maimone TJ (2016) Enantio-selective synthesis of an ophiobolin sesterterpene via a programmed radical cascade. Science 352: 1078–1082. [PubMed: 27230373]
- Chang W, Li Y, Zhang L, Cheng A, and Lou H (2012) Retigeric acid B attenuates the virulence of *Candida albicans* via inhibiting adenylyl cyclase activity targeted by enhanced farnesol production. PLoS One 7: e41624. [PubMed: 22848547]

- Chiba R, Minami A, Gomi K, and Oikawa H (2013) Identification of ophiobolin F synthase by a genome mining approach: a sesterterpene synthase from *Aspergillus clavatus*. *Org Lett* 15: 594–597. [PubMed: 23324037]
- DeBolt S, Scheible W-R, Schrick K, Auer M, Beisson F, Bischoff V, et al. (2009) Mutations in UDP-glucose:sterol glucosyltransferase in *Arabidopsis* cause transparent testa phenotype and suberization defect in seeds. *Plant Phys* 151: 78–87.
- Deyrup ST, Gloer JB, O'Donnell K, and Wicklow DT (2007) Kolokosides A-D: triterpenoid glycosides from a Hawaiian isolate of *Xylaria* sp. *J Nat Prod* 70: 378–382. [PubMed: 17284074]
- Frickey T, and Kannenberg E (2009) Phylogenetic analysis of the triterpene cyclase protein family in prokaryotes and eukaryotes suggests bidirectional lateral gene transfer. *Environ Microbiol* 11: 1224–1241. [PubMed: 19207562]
- Gamble WR, Gloer JB, Scott JA, and Malloch D (1995) Polytolypin, a new antifungal triterpenoid from the coprophilous fungus *Polytolypa hystricis*. *J Nat Prod* 58: 1983–1986. [PubMed: 8691217]
- Gao Y, Honzatko RB, and Peters RJ (2012) Terpenoid synthase structures: a so far incomplete view of complex catalysis. *Nat Prod Rep* 29: 1153–1175. [PubMed: 22907771]
- Godio RP, and Martín JF (2009) Modified oxidosqualene cyclases in the formation of bioactive secondary metabolites: biosynthesis of the antitumor clavarinic acid. *Fungal Genet Biol* 46: 232–242. [PubMed: 19130892]
- Godio RP, Fouces R, and Martín JF (2007) A squalene epoxidase is involved in biosynthesis of both the antitumor compound clavarinic acid and sterols in the basidiomycete *H.sublateritium*. *Chem Biol* 14: 1334–1346. [PubMed: 18096502]
- Guo C-J, Knox BP, Chiang Y-M, Lo H-C, Sanchez JF, Lee K-H, et al. (2012) Molecular genetic characterization of a cluster in *A. terreus* for biosynthesis of the meroterpenoid terretonin. *Org Lett* 14: 5684–5687. [PubMed: 23116177]
- Hagen I, Ecker M, Lagorce A, Francois JM, Sestak S, Rachel R, et al. (2004) Sed1p and Srl1p are required to compensate for cell wall instability in *Saccharomyces cerevisiae* mutants defective in multiple GPI-anchored mannoproteins. *Mol Microbiol* 52: 1413–1425. [PubMed: 15165243]
- 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. [PubMed: 19067514]
- Heasley BH, Pacofsky GJ, Mamai A, Liu H, Nelson K, Coti G, et al. (2012) Synthesis and biological evaluation of antifungal derivatives of enfumafungin as orally bioavailable inhibitors of β -1,3-glucan synthase. *Bioorg Med Chem Lett* 22: 6811–6816. [PubMed: 22672801]
- Hidalgo PI, Ullán RV, Albillos SM, Montero O, Fernández-Bodega MA, [notdef] García-Estrada C, et al. (2014) Molecular characterization of the PR-toxin gene cluster in *Penicillium roqueforti* and *Penicillium chrysogenum*: cross talk of secondary metabolite pathways. *Fungal Genet Biol* 62: 11–24. [PubMed: 24239699]
- Isaka M, Hywel-Jones NL, Sappan M, Mongkolsamrit S, and Saidaengkham S (2009) Hopane triterpenes as chemotaxonomic markers for the scale insect pathogens *Hypocrella* s.Lat. and *Aschersonia*. *Mycol Res* 113: 491–497. [PubMed: 19422075]
- Isaka M, Yangchum A, Rachtawee P, Komwijit S, and Lutthisungneon A (2010) Hopane-type triterpenes and binaphthopyrones from the scale insect pathogenic fungus *Aschersonia paraphysata* BCC 11964. *J Nat Prod* 73: 688–692. [PubMed: 20364867]
- Jiménez-Ortigosa C, Perez WB, Angulo D, Borroto-Esoda K, and Perlin DS (2017) De novo acquisition of resistance to SCY-078 in *Candida glabrata* involves FKS mutations that both overlap and are distinct from those conferring echinocandin resistance. *Antimicrob Agents Chemother* 69: e00833–00817.
- Joshi BK, Gloer JB, and Wicklow DT (2002) Bioactive natural products from a sclerotium-colonizing isolate of *Humicola fuscoatra*. *J Nat Prod* 65: 1734–1737. [PubMed: 12444718]
- Kim T-W, Hwang J-Y, Kim Y-S, Joo S-H, Chang SC, Lee JS, et al. (2005) *Arabidopsis* CYP85A2, a cytochrome P450, mediates the Baeyer-Villiger oxidation of castasterone to brassinolide in brassinosteroid biosynthesis. *Plant Cell* 17: 2397–2412. [PubMed: 16024588]
- Kimura M, Kaneko I, Komiyama M, Takatsuki A, Koshino H, Yoneyama K, and Yamaguchi I (1998) Trichothecene 3-O-acetyltransferase protects both the producing organism and transformed yeast

from related mycotoxins: cloning and characterization of Tri101. *J Biol Chem* 273: 1654–1661. [PubMed: 9430709]

- Kobayashi M, Yoshimura S, Kinoshita T, Hashimoto M, Hashimoto S, Takase S, et al. (2005) FR207944, an antifungal antibiotic from *Chaetomium* sp. no.217 II. Isolation and structure elucidation. *Biosci Biotechnol Biochem* 69: 1029–1032. [PubMed: 15914927]
- Larkin E, Hager C, Chandra J, Mukherjee PK, Retuerto M, Salem I, et al. (2017) The emerging pathogen *Candida auris*: growth phenotype, virulence factors, activity of antifungals, and effect of SCY-078, a novel glucan synthesis inhibitor, on growth morphology and biofilm formation. *Antimicrob Agents Chemother* 61: e02396–02316. [PubMed: 28223375]
- Laslett D, and Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32: 11–16. [PubMed: 14704338]
- Liesch JM, Meinz MS, Onishi J, Morris SA, Schwartz RE, Bills GF, Giacobbe RA, Horn WS, Zink DL, Cabello A, Díez MT, Martín I, Peláez F, and Vicente F (1998) Antifungal agent obtained from *Hormonema*. US Patent 5,756,472.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, and Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42: D490–D495. [PubMed: 24270786]
- Matsuda Y, Awakawa T, Wakimoto T, and Abe I (2013) Spiro-ring formation is catalyzed by a multifunctional dioxygenase in austinol biosynthesis. *J Am Chem Soc* 135: 10962–10965. [PubMed: 23865690]
- Milla P, Viola F, Bosso SO, Rocco F, Cattel L, Joubert BM, et al. (2002) Subcellular localization of oxidosqualene cyclases from *Arabidopsis thaliana*, *Trypanosoma cruzi*, and *Pneumocystis carinii* expressed in yeast. *Lipids* 37: 1171–1176. [PubMed: 12617471]
- Mitsuguchi H, Seshime Y, Fujii I, Shibuya M, Ebizuka Y, and Kushiro T (2009) Biosynthesis of steroidal antibiotic fusidanes: functional analysis of oxidosqualene cyclase and subsequent tailoring enzymes from *Aspergillus fumigatus*. *J Am Chem Soc* 131: 6402–6411. [PubMed: 19415934]
- Muñoz JF, McEwen JG, Clay OK, and Cuomo CA (2018) Genome analysis reveals evolutionary mechanisms of adaptation in systemic dimorphic fungi. *Sci Rep* 8: 4473. [PubMed: 29540755]
- Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, et al. (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438: 1151–1156. [PubMed: 16372009]
- Nomura T, Kushiro T, Yokota T, Kamiya Y, Bishop GJ, and Yamaguchi S (2005) The last reaction producing brassinolide is catalyzed by cytochrome P-450s, CYP85A3 in tomato and CYP85A2 in *Arabidopsis*. *J Biol Chem* 280: 17873–17879. [PubMed: 15710611]
- Oakley CE, Edgerton-Morgan H, and Oakley BR (2012) Tools for manipulation of secondary metabolism pathways: rapid promoter replacements and gene deletions in *Aspergillus nidulans*. *Methods Mol Biol* 944: 143–161. [PubMed: 23065614]
- Onishi J, Meinz M, Thompson J, Curotto J, Dreikorn S, Rosenbach M, et al. (2000) Discovery of novel antifungal (1,3)-beta-D-glucan synthase inhibitors. *Antimicrob Agents Chemother* 44: 368–377. [PubMed: 10639364]
- Paiva NL, Sun YJ, Dixon RA, Vanetten HD, and Hrazdina G (1994) Molecular cloning of isoflavone reductase from pea (*Pisum sativum* L): evidence for a 3R-isoflavanone intermediate in (+)-pisatin biosynthesis. *Arch Biochem Biophys* 312: 501–510. [PubMed: 8037464]
- Parks LW, and Casey WM (1995) Physiological implications of sterol biosynthesis in yeast. *Annu Rev Microbiol* 49: 95–116. [PubMed: 8561481]
- Peláez F, Cabello A, Platas G, Díez MT, Gonzalez A, Basilio A, et al. (2000) The discovery of enfumafungin, a novel antifungal compound produced by endophytic *Hormonema* species, biological activity, and taxonomy of the producing organisms. *Syst Appl Microbiol* 23: 333–343. [PubMed: 11108011]
- Pfaller MA, Messer SA, Rhomberg PR, and Castanheira M (2017) Differential activity of the oral glucan synthase inhibitor SCY-078 against wild-type and echinocandin-resistant strains of *Candida* spp. *Antimicrob Agents Chemother* 61: e00161–17. [PubMed: 28533234]

- Pinedo C, Wang C-M, Pradier J-M, Dalmais B, Choquer M, Le Pêcheur P, et al. (2008) Sesquiterpene synthase from the botrydial biosynthetic gene cluster of the phytopathogen *Botrytis cinerea*. *ACS Chem Biol* 3: 791–801. [PubMed: 19035644]
- Quin MB, Flynn CM, and Schmidt-Dannert C (2014) Traversing the fungal terpenome. *Nat Prod Rep* 31: 1449–1473. [PubMed: 25171145]
- Ríos J-L, and Andújar I (2016) Lanostanoids from fungi as potential medicinal agents In *Fungal Metabolites*, Mérillon J-M, and Ramawat KG (eds). New York, NY: Springer International Publishing, pp. 1–34.
- Schmidt-Dannert C (2015) Biosynthesis of terpenoid natural products in fungi. *Adv Biochem Eng Biotechnol* 148: 19–61. [PubMed: 25414054]
- Schulz TA, Choi MG, Raychaudhuri S, Mears JA, Ghirlando R, Hinshaw JE, and Prinz WA (2009) Lipid-regulated sterol transfer between closely apposed membranes by oxysterol-binding protein homologues. *J Cell Biol* 187: 889–903. [PubMed: 20008566]
- Schwartz RE, Smith SK, Onishi JC, Meinz M, Kurtz MB, Giacobbe RA, et al. (2000) The isolation and structure determination of enfumafungin, a triterpene glycoside antifungal agent from the fermentation of a *Hormonema* sp. *J Am Chem Soc* 122: 4882–4886.
- Shigematsu N, Tsujii E, Kayakiri N, Takase S, Tanaka H, and Tada T (1992) WF11605, an antagonist of leukotriene B₄ produced by a fungus II Structure determination. *J Antibiot* 45: 704–708. [PubMed: 1320602]
- Shimoi H, Kitagaki H, Ohmori H, Iimura Y, and Ito K (1998) Sed1p is a major cell wall protein of *Saccharomyces cerevisiae* in the stationary phase and is involved in lytic enzyme resistance. *J Bacteriol* 180: 3381–3387. [PubMed: 9642191]
- Stanke M, and Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res* 33: W465–W467. [PubMed: 15980513]
- Sun L, Sun S, Cheng A, Wu X, Zhang Y, and Lou H (2009) *In vitro* activities of retigeric acid B alone and in combination with azole antifungal agents against *Candida albicans*. *Antimicrob Agents Chemother* 53: 1586–1591. [PubMed: 19171796]
- Toyomasu T, Tsukahara M, Kaneko A, Niida R, Mitsuhashi W, Dairi T, et al. (2007) Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. *Proc Natl Acad Sci U S A* 104: 3084–3088. [PubMed: 17360612]
- Tudzynski B, Mihlan M, Rojas MC, Linnemannstöns P, Gaskin P, and Hedden P (2003) Characterization of the final two genes of the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi*: des and P450-3 encode GA4 desaturase and the 13-hydroxylase, respectively. *J Biol Chem* 278: 28635–28643. [PubMed: 12750377]
- Wang X-N, Zhang H-J, Ren D-M, Ji M, Yu W-T, and Lou H-X (2009) Lobarialides A–C, antifungal triterpenoids from the lichen *Lobaria kurokawae*. *Chem Biodivers* 6: 746–753. [PubMed: 19479840]
- Warnecke D, Erdmann R, Fahl A, Hube B, Müller F, Zank T, et al. (1999) Cloning and functional expression of UGT genes encoding sterol glucosyltransferases from *Saccharomyces cerevisiae*, *Candida albicans*, *Pichia pastoris*, and *Dictyostelium discoideum*. *J Biol Chem* 274: 13048–13059. [PubMed: 10224056]
- Wring SA, Randolph R, Park S, Abruzzo G, Chen Q, Flattery A, et al. (2017a) SCY-078 a first in class orally active antifungal glucan synthesis inhibitor: pre-clinical pharmacokinetics and pharmacodynamic target in murine models of disseminated candidiasis. *Antimicrob Agents Chemother* 61: 11.
- Wring SA, Randolph R, Park S, Abruzzo G, Chen Q, Flattery A, et al. (2017b) Preclinical pharmacokinetics and pharmacodynamic target of SCY-078, a first-in-class orally active antifungal glucan synthesis inhibitor, in murine models of disseminated candidiasis. *Antimicrob Agents Chemother* 61: e02068–02016. [PubMed: 28137806]
- Yin Y, Mao X, Yang J, Chen X, Mao F, and Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40: W445–W451. [PubMed: 22645317]
- Yuan X-H, Xu G-B, Wu W-L, Yang T, and Li G-Y (2012) Peniciside, a new triterpenoid glycoside, from the fungus *Penicillium* sp. 169. *Arch Pharm Res* 35: 311–314. [PubMed: 22370784]

- Zhang A, Lu P, Dahl-Roshak AM, Paress PS, Kennedy S, Tkacz JS, and An Z (2003) Efficient disruption of a polyketide synthase gene (pks1) required for melanin synthesis through Agrobacterium-mediated transformation of *Glaea lozoyensis*. Mol Genet Genomics 268: 645–655. [PubMed: 12589439]
- Zhang D-H, Jiang L-X, Li N, Yu X, Zhao P, Li T, and Xu J-W (2017) Overexpression of the squalene epoxidase gene alone and in combination with the 3-hydroxy-3-methylglutaryl coenzyme a gene increases ganoderic acid production in *Ganoderma lingzhi*. J Agric Food Chem 65: 4683–4690. [PubMed: 28530827]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Structure of enfumafungin and producing *Hormonema carpetanum* strain. A. Numbering system and structural features of enfumafungin—I: glucose unit, II: acetate group, III: hemiacetal function, IV: carboxylic acid function. B. *H. carpetanum* ATCC 74360 on SMY media, 16 days old.

Fig. 2.
Some fungal triterpenes with fenene-type core scaffold (red) related to enfumafungin.

Fig. 3.
Schematic diagram of the enfumafungin biosynthetic gene cluster. Numbers below indicate the contig's relative nucleotide position. Refer to Table 1 for proposed gene functions and GenBank accession numbers.

Fig. 4. Phylogenetic relationship among enfumafungin terpene cyclase (efuA), other squalene-hopene cyclases (SHC) and lanosterol synthases (LS) as inferred from the respective protein sequences. Maximum likelihood (ML) bootstrap support values above 50%, from 100 RAXML replicates are assigned to the nodes of the tree topology of the most likely tree found by RAXML. GenBank accession numbers are followed by species names. EfuA is highlighted in red, the putative polytolypin synthase in blue and the *H. carpetanum* LS in green. TC-GT, terpene cyclase/glycosyltransferase fusion proteins.

Fig. 5.

Phylogenetic relationship of enfumafungin glycosyltransferase and its closest BLAST hits as inferred from the respective protein sequences. Likelihood (ML) bootstrap support values above 50%, from 100 RAXML replicates are assigned to the nodes of the tree topology of the most likely tree found by RAXML. GenBank accession numbers are followed by species names.

Fig. 6.

Alignment of related terpene cyclase-glycosyl transferase fusion protein sequences. Homologies are colour-coded based on polarity (red—acidic aa, blue—basic aa, green—polar, yellow—nonpolar). The terpene cyclase domain is highlighted in red and the glycosyl transferase domain in blue. The black box indicates the conserved DXDD active site motif of the terpene cyclase.

Fig. 7.

Gene disruption construct of *efuA* and mutant verification. A. Scheme of the gene knock-out construct and PCR amplified regions for mutant verification (Ct1 and Ct2: control PCR1 and 2). B. PCR detection of positive *efuA* disruption mutants. L, 1 kb ladder; 1, wild type Ct1; 2, wild-type Ct2; 3, mutant Ct1; 4, mutant Ct2. C. Growth of wild-type (WT) and mutant (*efuA*) on SMY media without (–Hyg) or supplemented with hygromycin (+Hyg). (Note: the wild-type proliferates mainly in the hyaline yeast-like stage on unamended medium, whereas the mutant grows mainly as filamentous melanized hyphae in the hygromycin-amended medium).

Fig. 8.

Verification of enfumafungin production in wild-type and mutant strains of *H. carpetanum*. A. Comparison of HPLC-UV chromatograms derived from wild-type (red line) and *efuA* mutant (green line) extracts. The enfumafungin standard is depicted in blue (Note: under regular HPLC conditions enfumafungin appears in two conformations yielding two overlapping peaks). B. Agar-well-diffusion assays of wild-type and mutant extracts against *C. albicans* (1–3: wild-type, 4–6: *efuA* mutant, 7: 70 µM enfumafungin, 8: 54 µM amphotericin B, 9: methanol).

Fig. 9.

Proposed schematic biosynthetic pathway for enfumafungin. Enzymes that are proposed to be involved in the biosynthesis are indicated above the arrows.

Table 1.

Proteins located in and near the proposed enfumafungin biosynthetic gene cluster, their predicted characteristics, predicted functions and closest BLAST hits for their respective predicted amino acid sequences

Gene	GenBank accession	G + C (%)	Predicted protein (aa/kDa) ^a /pI ^a	Conserved domain ^b	Proposed function	Closest BLAST matches		
						Ident/ QC (%)	GenBank accession	Organism
<i>efuB</i>	MF611883	51.8	505/57.410/7.18	p450 super family	P450 monooxygenase	63/97	XP_013348741.1	<i>Aureobasidium subglaciale</i>
<i>efuC</i>	MF611884	64.0	551/60.251/6.65	αβ-Hydrolase super family	Acetyltransferase	78/100	KEQ84840.1	<i>Aureobasidium pullulans</i>
<i>efuD</i>	MF611885	58.3	454/50.061/4.27	TFIIIE super family	Transcription initiation factor	53/100	KEQ84839.1	<i>Aureobasidium pullulans</i>
<i>efuE</i>	MF611886	50.0	400/43.876/5.32	Mand_dh_like	Dehydrogenase	72/82	XP_007584832.1	<i>Neofusicoccum parvum</i>
<i>efuF</i>	MF611887	53.0	493/53.856/8.15	UhpC super family	Sugar phosphate permease	69/92	XP_018189485.1	<i>Xylona heveae</i>
<i>efuA</i>	MF611888	51.8	1,314/146.756/6.05	ISOPREN_C2_like super family, GT1_Grf_like	Squalene-hopene cyclase, glycosyl transferase	48/97	CZT12320.1	<i>Rhynchosporium agropyri</i>
<i>efuG</i>	MF611889	51.1	563/63.784/8.04	p450 super family	P450 monooxygenase	60/97	KU 161982.1	<i>Valsa mali</i> var. <i>pyri</i>
<i>efuH</i>	MF611890	49.4	565/63.221/8.72	p450 super family	P450 monooxygenase	54/92	KU 161983.1	<i>Valsa mali</i> var. <i>pyri</i>
<i>efuI</i>	MF611891	51.9	295/33.240/4.89	N/A	Desaturase	46/96	XP_016246718.1	<i>Cladophialophora immunda</i>
<i>efuJ</i>	MF611892	54.9	614/62.375/8.10	N/A	SED1-like protein, Structural cell wall protein, putative resistance gene	38/61	XP_016761069.1	<i>Sphaerulina musiva</i>
<i>efuK</i>	MF611893	49.3	696/78.433/5.83	Oxysterol binding protein, MhpC	Sterol transport	65/53	KEQ81548.1	<i>Aureobasidium pullulans</i>
<i>efuL</i>	MF611894	51.4	312/34.665/5.52	PCBER_SDR_a	Flavin-dependent reductase	61/98	KKY19647.1	<i>Diplodia seriata</i>

^aMolecular weight and isoelectric point are estimated by Geneious software package (ver. 9.1.7).

^bIdentified by the NCBI Conserved Domain Search tool (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).