Aspergillus fumigatus Secondary Metabolite Clusters: Conservation Across Fungi and Examination of the Interrelatedness of Two Clusters

By

Kurt O. Throckmorton

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 8/10/2015

The dissertation is approved by the following members of the Final Oral Committee:

Nancy P. Keller, Professor, Bacteriology and Medical Microbiology and Immunology
Xin Sun, Professor, Medical Genetics
Chris T. Hittinger, Assistant Professor, Genetics
Jon P. Woods, Professor, Medical Microbiology and Immunology
James E. Gern, Professor, Medicine

ABSTRACT

Aspergillus fumigatus is an important opportunistic human pathogen and a prolific producer of secondary metabolites (SMs). SMs, or natural products, are small molecules that afford the producing species protection against biotic and abiotic stressors and aid in niche securement. They are made by several types of synthases and 'decorated' by other enzymes encoded together in an SM gene cluster. Here, we focus on the SMs produced by *A. fumigatus* and the conservation of the gene clusters that produce them in other fungi. In particular, we focus on SMs derived from non-reducing polyketide synthases (NR-PKSs) belonging to a group that lack the domain typically required for product release. We identify upwards of 150 uncharacterized PKSs belonging to this group based on analysis of publicly available sequence data, and leverage the existing knowledge of gene clusters from this group to predict products for many of them. We identify among this group the (*tpc*) gene cluster responsible for production of the cytotoxic spore metabolite, trypacidin. Further, we characterize an unusual redundancy in the synthesis of another spore metabolite, endocrocin, by the *tpc* cluster and the previously described *enc* cluster in a strain of *A. fumigatus*.

ACKNOWLEDGEMENTS

I would first like to thank my mentor, Dr. Nancy Keller, for accepting me into her lab and helping me reach this goal. Thanks to her and the opportunities she has provided me, I have learned a lot over the past few years. I would also like to thank my committee members, past and present, for their mentorship and guidance, including Drs. Xin Sun, Chris Hittinger, James Gern, Jon Woods, Grace Boekhoff-Falk, and Francisco Pelegri. Thank you to the administrative staff in the Departments of Genetics and Medical Microbiology and Immunology for facilitating all of the paperwork and reservations for travel, talks, and transitions between grants and labs.

I would like to thank the members of the Keller lab for their assistance and camaraderie, especially Fang Yun Lim and Dr. Philipp Wiemann with whom I worked closely on multiple projects; Tsokyi Choera, Sachin Jain, Dr. Pin Mei Wang, and Joe Spraker with whom I shared an office and many lunches during my time in the Keller lab; and Drs. Jin Woo Bok and Jon Palmer for providing guidance and help troubleshooting. I would like to thank my classmates, especially Dr. Kelsey McIntyre and Richard Wang, for their friendship and willingness to endure long conversations.

Lastly, I would like to thank my family for their support and for raising me to work hard; my friends back in Ohio for always making themselves available on the rare occasions that I visit; and Krista Morin for always being supportive and uplifting throughout this journey.

TABLE OF CONTENTS

Section **Link to page**

ABSTRACT I

ACKNOWLEDGEMENTS II

TABLE OF CONTENTS III

CHAPTER 1 1

Introduction 1

- **Aspergilli and their impacts** 1
- **Secondary metabolism in filamentous fungi** 2
- **Phylogeny of NR-PKSs** 3
- **Partial redundancy of two *A. fumigatus* group V NR-PKSs** 4
- **References** 4

CHAPTER 2 8

Evolution of chemical diversity in a group of non-reduced polyketide gene clusters: Using phylogenetics to inform the search for novel fungal natural products 8

- **Abstract** 9
- **Introduction** 9
- **Group V NR-PKS phylogeny** 13
- **Phylogenetics directs product prediction** 14
  - **Group V1** 14
    - Endocrocin-like clusters 16
    - Monodictyphenone-like clusters 17
    - Trypacidin, geodin, pestheic acid-like clusters 19
  - **Group V2** 21
    - Asperthecin-like clusters 22
    - TAN-1612-like clusters 22
    - Neosartoricin-like clusters 23
    - Viridicatumtoxin-like clusters 23
  - **Group V3: Griseofulvin and various alternariol-like pathways.** 24
- **Uncharacterized clusters in symbionts and pathogens** 26
  - **Plant pathogens** 26
  - **Mycorrhizal fungi** 27
  - **Fungal pathogens** 28
  - **Animal pathogens** 28

Entomopathogenic fungi. 28
Dermatophytes. 29
Experimental procedures **29**
Conclusions **31**
Acknowledgments **34**
Tables **35**
Table 1: Group V PKSs identified and analyzed in this study. *35*
Figures **39**
Figure 1: Diagram of the domains of different types of PKSs. *39*
Figure 2: Phylogenetic tree of all NR-PKSs analyzed in this study. *40*
Figure 3: Phylogenetic tree of all group V PKSs analyzed in this study. *41*
Figure 4: The relationships between group V1 clusters and the aflatoxin gene cluster. *42*
Figure 5: tpc-like clusters of interest. *44*
Figure 6: nsc-like clusters in Trichophyton and vrt-like clusters in Metarhizium. *46*
Figure 7: gsf-, pkg-, PksI-, and SNOG_15820-like clusters. *48*
References **50**
Appendix **58**
Supplemental tables **58**
Table S1: Composite multiFASTA file used for group V1 MultiGeneBLAST. *58*
Table S2: Composite multiFASTA file used for group V2 MultiGeneBLAST. *64*
Table S3: Composite multiFASTA file used for group V3 MultiGeneBLAST. *68*
Supplemental figures **73**
Figure S1. .... *73*
Figure S2. .... *74*
References **78**
CHAPTER 3 .... **80**
Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus* **80**
Abstract **81**
Introduction **81**
Results **84**
Identification of the trypacidin-producing gene cluster **84**
Endocrocin is produced by both the tpc and enc clusters through distinct routes **86**
The contribution of the tpc and enc clusters to virulence in Toll-deficient Drosophila melanogaster **90**
Discussion **90**
Experimental procedures **95**
Fungal strains and growth conditions **95**

Construction of mutants **96**

Comparative genomic analysis **97**

SM extraction and Chromatography **98**

Drosophila assays **99**

Acknowledgements 100

Figures **101**

Figure 1: The relationships between characterized group V1 gene clusters. *101*
Figure 2: HPLC analysis of tpc cluster mutants and regulators. *103*
Figure 3: Proposed biosynthetic pathway of trypacidin. *105*
Figure 4: HPLC analysis of enc and tpc PKS deletants and model for an endocrocin shunt. *106*
Figure 5: HPLC analysis of mutants of mdpH homologs in the tpc and enc pathways. *108*

References **110**

Appendix **116**

Supplemental tables **116**

Table S1: Product template domains used in the construction of the group V1 phylogenetic tree. *116*
Table S2: Comparison of the tpc and ged gene clusters. *122*
Table S3: SNPs identified in the CEA10 tpc cluster. *123*
Table S4: Strains used in this study. *124*
Table S5: Primers used in this study. *126*
Table S6: Plasmids used in this study. *131*

Supplemental figures **132**

Figure S1: Confirmation of mutants by Southern blot. *132*
Figure S2: TLC analysis of standards used in this study. *134*
Figure S3: TLC analysis of tpc cluster mutants. *135*
Figure S4: TLC analysis of tpc cluster regulators. *136*
Figure S5: Semi-qRT-PCR analysis of LaeA and BrlA regulation of the enc and tpc gene clusters. *137*
Figure S6: TLC analysis of enc and tpc PKS deletants. *138*
Figure S7: TLC analysis of mdpH homolog mutants in the enc and tpc gene clusters. *139*
Figure S8: Assessment of virulence of enc and tpc cluster mutants in Toll-deficient Drosophila. *140*

CHAPTER 4 .... **141**

Concluding remarks and future directions **141**

Group V NR-PKSs **141**

Product prediction directed by phylogenetics **142**

Toward rational design of SMs **144**

Endocrocin, trypacidin, and group V1 .... **145**

References 147

CHAPTER 1

Introduction

Aspergilli and their impacts

Aspergillus is a genus of filamentous fungi named for the resemblance of their asexual spore-bearing structure, the conidiophore, to a holy water sprinkler, or aspergillum (Ainsworth, 1976). The genus contains a number of species of scientific, economic, and medial importance, including the genetic model organism *A. nidulans* (Goldman and Kafer, 2004), the producer of the antihypercholesterolemic drug lovastatin, *A. terreus* (Hendrickson et al., 1999), the plant pathogens and producers of the highly carcinogenic metabolite aflatoxin, *A. flavus* and *A. parasiticus* (Yu, 2012; Roze et al., 2013), and the opportunistic human pathogen, *A. fumigatus* (Latgé, 1999). All produce numerous secondary metabolites, SMs; small molecules produced by many plants, bacteria, and fungi that afford them protection from biotic and abiotic stressors, such as fungivorous insects or ultraviolet radiation, and facilitate niche securement (Rohlfs and Churchill, 2011; Allam and Abd El-Zaher, 2012; Scherlach et al., 2013). Here, we focus on the SMs produced by *A. fumigatus*, related SMs in fungi in general, and a pair of *A. fumigatus* SMs with an unusual partial redundancy in their biosynthetic pathways.

A. fumigatus, like many other aspergilli, is a ubiquitous species of saprotrophic fungus (Bennett, 2009). The asexual spores, or conidia, it produces are abundant in atmospheric air and, as such, are regularly inhaled by humans (Goodley et al., 1994; Hospenthal et al., 1998). In immunocompetent individuals, these inhaled spores encounter macrophages and neutrophils in the lung and are eradicated. However, in immunocompromised individuals, such as those undergoing chemotherapy, recipients of organ transplants, and AIDS patients, these spores can

germinate, penetrate the airway epithelium, and become disseminated via the bloodstream. When this occurs, mortality varies from 30-85% for different patient subpopulations (Brakhage, 2005; Maschmeyer et al., 2007).

Though several species of aspergilli are capable of causing disease (Ozcan et al., 2003; Hedayati et al., 2007), *A. fumigatus* is the most common cause of pulmonary disease (Latgé, 1999). What makes *A. fumigatus* especially able to infect immunocompromised humans is a subject of ongoing study. Pathogenicity is multifactorial, i.e. no single factor accounts for this, but one hypothesis is that the unique suite of secondary metabolites produced by this species might facilitate its ability to colonize the human host. It is estimated that *A. fumigatus* contains about 35 SM gene clusters (Inglis et al., 2013). Many of these produce compounds that are toxic or immunomodulatory (Amitani et al., 1995; Bok et al., 2005; Bok et al., 2006; Spikes et al., 2008; Lodeiro et al., 2009; Gauthier et al., 2012; Berthier et al., 2013; Chooi et al., 2013; Yin et al., 2013), and several of these are species-specific.

Secondary metabolism in filamentous fungi

Many members of the class Ascomycota, particularly the genera *Penicillium* and *Aspergillus*, are prolific producers of SMs. The backbones or scaffolds of these metabolites are produced by one of several known types of synthase, including non-ribosomal peptide synthases (NRPSs), dimethylallyltryptophan synthases (DMATs), terpene cyclases (TCs), and polyketide synthases (PKSs). Here, we focus on PKSs, which iteratively condense malonyl-CoA to form a carbon backbone. PKSs come in several flavors that vary in the catalytic domains they harbor. So-called highly reducing (HR) PKSs have domains that reduce the nascent carbon chain,

whereas non-reducing (NR) PKSs lack these domains. NR-PKSs additionally have a product template (PT) domain that determines which carbons will be involved in subsequent cyclization of the carbon backbone. The genes encoding these large synthases are typically clustered with other genes encoding enzymes that 'decorate' the carbon backbone to arrive at the final metabolite. These clusters also often include $Zn(II)_2Cys_6$ transcription factors that regulate the expression of the gene cluster.

Phylogeny of NR-PKSs

NR-PKSs have been divided into eight groups (Liu et al., 2015). One of these groups, group V, consists of PKSs that lack the thioesterase/Claisen cyclase domain usually required for product release. Instead, this activity is encoded in a physically discrete metallo-β-lactamase type thioesterase (Awakawa et al., 2009). This group produces a wide variety of compounds including some of the longest polyketide chains (Chooi et al., 2010; Li et al., 2011), several known toxins (Gauthier et al., 2012; Demuner et al., 2013), and a historically widely used antifungal drug (Finkelstein et al., 1996). A recent study demonstrated that the catalytic pocket of group V PKSs have, on average, the highest volume of any of the eight groups of PKSs and a set of specific cavity-lining residues (Liu et al., 2015). The diversity of compounds produced by PKSs from this group can likely be attribute, in part, to these characteristics.

Previous studies have characterized a total of twelve SM gene clusters associated with PKSs from group V (Szewczyk et al., 2008; Chiang et al., 2010; Chooi et al., 2010; Li et al., 2011; Ahuja et al., 2012; Lim et al., 2012; Saha et al., 2012; Chooi et al., 2013; Nielsen et al., 2013; Xu et al., 2014; Chooi et al., 2015) and identified a two subgroups, called V1 and V2. Not

only the PKSs, but also the MβLs have co-diverged along these lines (Li et al., 2011). This led us to hypothesize that, guided by the sequence of the PKSs alone, one might be able to identify SM gene clusters that are highly related to characterized gene clusters. If so, these might be predictable by comparison, or interesting for the dissimilarities. We explore this hypothesis as a means to extract information from the abundance of fungal genome sequence data available, and as a tool to direct future research. We also discuss our results as they apply to the overarching goal of predicting the products of uncharacterized SM clusters both to identify novel compounds that might be of pharmacological interest and to avoid repeated characterization of previously studied or disinteresting gene clusters or compounds.

Partial redundancy of two *A. fumigatus* group V NR-PKSs

A. fumigatus has three PKSs belonging to group V, two of which had previously been characterized, the neosartoricin (*nsc*) and endocrocin (*enc*) clusters (Lim et al., 2012; Chooi et al., 2013). Here, we show that the third is the PKS responsible for the production of the cytotoxic spore metabolite, trypacidin, known to be a product of *A. fumigatus* for nearly 50 years (Balan et al., 1963). We also show that, unusually, this cluster contributes to the production of another spore metabolite, endocrocin, along with the previously mentioned *enc* cluster. These clusters are both regulated by the global regulator of SM in ascomycetes, LaeA, and the conidiation-specific transcription factor, BrlA.

References

Ahuja, M., Chiang, Y.M., Chang, S.L., Praseuth, M.B., Entwistle, R., Sanchez, J.F. et al. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J Am Chem Soc* **134**: 8212-8221.

Ainsworth, G.C. (1976) *Introduction to the history of mycology*. Cambridge: Cambridge Univ. Press.

Allam, N.G., and Abd El-Zaher, E.H.F. (2012) Protective role of *Aspergillus fumigatus* melanin against ultraviolet (UV) irradiation and *Bjerkandera adusta* melanin as a candidate vaccine against systemic candidiasis. *African J Biotech* **11**: 6566-6577.

Amitani, R., Taylor, G., Elezis, E.N., Llewellyn-Jones, C., Mitchell, J., Kuze, F. et al. (1995) Purification and characterization of factors produced by *Aspergillus fumigatus* which affect human ciliated respiratory epithelium. *Infect Immun* **63**: 3266-3271.

Awakawa, T., Yokota, K., Funa, N., Doi, F., Mori, N., Watanabe, H., and Horinouchi, S. (2009) Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem Biol* **16**: 613-623.

Balan, J., Ebringer, L., Nemec, P., Kovac, S., and Dobias, J. (1963) Antiprotozoal Antibiotics. II. Isolation and characterization of trypacidin, a new antibiotic, active against *Trypanosoma cruzi* and *Toxoplasma gondii*. *J Antibiot (Tokyo)* **16**: 157-160.

Bennett, J.W. (2009) *Aspergillus*: a primer for the novice. *Med Mycol* **47 Suppl 1**: S5-12.

Berthier, E., Lim, F.Y., Deng, Q., Guo, C.J., Kontoyiannis, D.P., Wang, C.C. et al. (2013) Low-volume toolbox for the discovery of immunosuppressive fungal secondary metabolites. *PLoS Pathog* **9**: e1003289.

Bok, J.W., Balajee, S.A., Marr, K.A., Andes, D., Nielsen, K.F., Frisvad, J.C., and Keller, N.P. (2005) LaeA, a regulator of morphogenetic fungal virulence factors. *Eukaryot Cell* **4**: 1574-1582.

Bok, J.W., Chung, D., Balajee, S.A., Marr, K.A., Andes, D., Nielsen, K.F. et al. (2006) GliZ, a transcriptional regulator of gliotoxin biosynthesis, contributes to *Aspergillus fumigatus* virulence. *Infect Immun* **74**: 6761-6768.

Brakhage, A.A. (2005) Systemic fungal infections caused by *Aspergillus* species: epidemiology, infection process and virulence determinants. *Curr Drug Targets* **6**: 875-886.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microbiol* **76**: 2067-2074.

Chooi, Y.H., Cacho, R., and Tang, Y. (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* **17**: 483-494.

Chooi, Y.H., Muria-Gonzalez, J.M., Mead, O.L., and Solomon, P.S. (2015) SnPKS19 encodes the polyketide synthase for alternariol mycotoxin biosynthesis in the wheat pathogen *Parastagonospora nodorum*. *Appl Environ Microbiol*.

Chooi, Y.H., Fang, J., Liu, H., Filler, S.G., Wang, P., and Tang, Y. (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **15**: 780-783.

Demuner, A.J., Almeida Barbosa, L.C., Mendes Miranda, A.C., Geraldo, G.C., da Silva, C.M., Giberti, S. et al. (2013) The Fungal Phytotoxin Alternariol 9-Methyl Ether and Some of Its Synthetic Analogues Inhibit the Photosynthetic Electron Transport Chain. *Journal of Natural Products* **76**: 2234-2245.

Finkelstein, E., Amichai, B., and Grunwald, M.H. (1996) Griseofulvin and its uses. *Int J Antimicrob Agents* **6**: 189-194.

Gauthier, T., Wang, X., Sifuentes Dos Santos, J., Fysikopoulos, A., Tadrist, S., Canlet, C. et al. (2012) Trypacidin, a spore-borne toxin from *Aspergillus fumigatus*, is cytotoxic to lung cells. *PLoS One* **7**: e29906.

Goldman, G.H., and Kafer, E. (2004) *Aspergillus nidulans* as a model system to characterize the DNA damage response in eukaryotes. *Fungal Genet Biol* **41**: 428-442.

Goodley, J.M., Clayton, Y.M., and Hay, R.J. (1994) Environmental sampling for aspergilli during building construction on a hospital site. *J Hosp Infect* **26**: 27-35.

Hedayati, M.T., Pasqualotto, A.C., Warn, P.A., Bowyer, P., and Denning, D.W. (2007) *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. *Microbiology* **153**: 1677-1692.

Hendrickson, L., Davis, C.R., Roach, C., Nguyen, D.K., Aldrich, T., McAda, P.C., and Reeves, C.D. (1999) Lovastatin biosynthesis in *Aspergillus terreus*: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem Biol* **6**: 429-439.

Hospenthal, D.R., Kwon-Chung, K.J., and Bennett, J.E. (1998) Concentrations of airborne *Aspergillus* compared to the incidence of invasive aspergillosis: lack of correlation. *Med Mycol* **36**: 165-168.

Inglis, D.O., Binkley, J., Skrzypek, M.S., Arnaud, M.B., Cerqueira, G.C., Shah, P. et al. (2013) Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. *BMC Microbiol* **13**: 91.

Latgé, J.P. (1999) *Aspergillus fumigatus* and aspergillosis. *Clin Microbiol Rev* **12**: 310-350.

Li, Y., Chooi, Y.H., Sheng, Y., Valentine, J.S., and Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**: 15773-15785.

Lim, F.Y., Hou, Y., Chen, Y., Oh, J.H., Lee, I., Bugni, T.S., and Keller, N.P. (2012) Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* **78**: 4117-4125.

Liu, L., Zhang, Z., Shao, C.L., Wang, J.L., Bai, H., and Wang, C.Y. (2015) Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. *Sci Rep* **5**: 10463.

Lodeiro, S., Xiong, Q., Wilson, W.K., Ivanova, Y., Smith, M.L., May, G.S., and Matsuda, S.P. (2009) Protostadienol biosynthesis and metabolism in the pathogenic fungus *Aspergillus fumigatus*. *Org Lett* **11**: 1241-1244.

Maschmeyer, G., Haas, A., and Cornely, O.A. (2007) Invasive aspergillosis: epidemiology, diagnosis and management in immunocompromised patients. *Drugs* **67**: 1567-1601.

Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., Larsen, T.O., and Mortensen, U.H. (2013) Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. *PLoS One* **8**: e72871.

Ozcan, M., Ozcan, K.M., Karaarslan, A., and Karaarslan, F. (2003) Concomitant otomycosis and dermatomycoses: a clinical and microbiological study. *Eur Arch Otorhinolaryngol* **260**: 24-27.

Rohlfs, M., and Churchill, A.C. (2011) Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genet Biol* **48**: 23-34.

Roze, L.V., Hong, S.Y., and Linz, J.E. (2013) Aflatoxin biosynthesis: current frontiers. *Annu Rev Food Sci Technol* **4**: 293-311.

Saha, D., Fetzner, R., Burkhardt, B., Podlech, J., Metzler, M., Dang, H. et al. (2012) Identification of a polyketide synthase required for alternariol (AOH) and alternariol-9-methyl ether (AME) formation in *Alternaria alternata*. *PLoS One* **7**: e40564.

Scherlach, K., Graupner, K., and Hertweck, C. (2013) Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu Rev Microbiol* **67**: 375-397.

Spikes, S., Xu, R., Nguyen, C.K., Chamilos, G., Kontoyiannis, D.P., Jacobson, R.H. et al. (2008) Gliotoxin production in *Aspergillus fumigatus* contributes to host-specific differences in virulence. *J Infect Dis* **197**: 479-486.

Szewczyk, E., Chiang, Y.M., Oakley, C.E., Davidson, A.D., Wang, C.C., and Oakley, B.R. (2008) Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microbiol* **74**: 7607-7612.

Xu, X., Liu, L., Zhang, F., Wang, W., Li, J., Guo, L. et al. (2014) Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte *Pestalotiopsis fici*. *Chembiochem* **15**: 284-292.

Yin, W.B., Baccile, J.A., Bok, J.W., Chen, Y., Keller, N.P., and Schroeder, F.C. (2013) A nonribosomal peptide synthetase-derived iron(III) complex from the pathogenic fungus *Aspergillus fumigatus*. *J Am Chem Soc* **135**: 2064-2067.

Yu, J. (2012) Current understanding on aflatoxin biosynthesis and future perspective in reducing aflatoxin contamination. *Toxins (Basel)* **4**: 1024-1057.

CHAPTER 2

Evolution of chemical diversity in a group of non-reduced polyketide gene clusters: Using phylogenetics to inform the search for novel fungal natural products

This work has been published as:
Throckmorton, K., Wiemann, P., and Keller, N.P. (2015) **Evolution of chemical diversity in a group of non-reduced polyketide gene clusters: Using phylogenetics to inform the search for novel fungal natural products.** *Toxins*.
It has been slightly modified from the published version.

Author Contributions

The text was written by KT, PW, and NPK. Sequences were obtained and processed by PW and KT. Phylogenetic trees were constructed by KT. MultiGeneBLAST analyses were carried out by KT and PW. Figures were constructed by KT and PW. General oversight and guidance was provided by NPK.

Abstract

Fungal polyketides are a diverse class of natural products, or secondary metabolites (SMs), with a wide range of bioactivities often associated with toxicity. Here, we focus on a group of non-reducing polyketide synthases (NR-PKSs) in the fungal phylum Ascomycota that lack a thioesterase domain for product release, group V. Although widespread in ascomycete taxa, this group of NR-PKSs is notably absent in the mycotoxigenic genus *Fusarium* and, surprisingly, found in genera not known for their secondary metabolite production (e.g. the mycorrhizal genus *Oidiodendron*, the powdery mildew genus *Blumeria*, and the causative agent of white-nose syndrome in bats, *Pseudogymnoascus destructans*). This group of NR-PKSs, in association with the other enzymes encoded by their gene clusters, produces a variety of different chemical classes including naphthacenediones, anthraquinones, benzophenones, grisandienes, and diphenyl ethers. We discuss the modification of and transitions between these chemical classes, the requisite enzymes, and the evolution of the SM gene clusters that encode them. Integrating this information, we predict the likely products of related but uncharacterized SM clusters, and we speculate upon the utility of these classes of SMs as virulence factors or chemical defenses to various plant, animal, and insect pathogens, as well as mutualistic fungi.

Introduction

The Kingdom Fungi is well known for its ability to synthesize bioactive secondary metabolites (SMs), also known as natural products. The dominant taxa producing SMs belong to several filamentous ascomycete genera, many of which are renowned plant, insect, and/or human pathogens. Virulence is often associated with the production of toxic SMs in these fungi (Perez-Nadales et al., 2014). In nature, where studied, SMs afford various fitness advantages to the

producing species ranging from protection from fungivory and physical insults (e.g. UV light) to competition with other microbes for niche securement (Rohlfs and Churchill, 2011; Allam and Abd El-Zaher, 2012; Scherlach et al., 2013).

SMs can be classified according to chemical types. For example, polyketides are derived from acetyl/malonyl-CoAs, non-ribosomal peptides from amino acids, and terpenes from isoprene. The reader is referred to several recent reviews for in-depth coverage of each chemical class (Bushley and Turgeon, 2010; Chiang et al., 2010a; Wawrzyn et al., 2012). Each class is associated with a defining synthase, polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS) and terpene cyclase/synthase, respectively. Ribosomal peptides (Umemura et al., 2014) and hybrid synthases, commonly hybrid PKS-NRPS enzymes, have also been described. The synthases contain conserved catalytic domains easily detectable by statistical analysis, i.e. hidden Markov models, thus making them highly suitable for phylogenetic analyses and evolutionary inferences (Fedorova et al., 2012; Gallo et al., 2013; Finn et al., 2015). Typically, the genes encoding these synthases are physically clustered with additional enzymatic genes required to form the end metabolite; these are termed SM gene clusters.

Polyketides in particular have drawn considerable interest due to their impact on human and plant health both positively (e.g. lovastatin) and detrimentally (e.g. aflatoxin). The first fungal PKS to be identified and characterized was 6-methylsalicylic acid synthetase found in several *Penicillium* and *Aspergillus* spp. (Beck et al., 1990; Wang et al., 1991; Fujii et al., 1996). Shortly thereafter, PKSs required for spore pigmentation as well as the mycotoxins aflatoxin and sterigmatocystin were characterized in several *Aspergillus* spp. (Mayorga and Timberlake, 1992;

Trail et al., 1995; Yu et al., 1995; Yu and Leonard, 1995). Additional PKSs involved in toxin synthesis include Pks1 and Pks2 (T toxin in *Cochliobolus heterostrophus*, (Yang et al., 1996; Baker et al., 2006)), FusA (fusarin production in *Fusarium fujikuroi*, (Díaz-Sánchez et al., 2012) and *F. moniliforme, F. venenatum* (Song et al., 2004)), Zea1 and Zea2 (zearalenone production in *Gibberella zeae*, (Kim et al., 2005; Gaffoor and Trail, 2006)), Fum1 (fumonisin production in *F. verticillioides*, *F. fujikuroi*, (Proctor et al., 1999; Yu et al., 2005)), Pks-CT (citrinin production in *Monascus purpureus*, (Shimizu et al., 2005)), and NhPKS1 (bostrycoidin and fusarubin in *Nectria haematococca*, (Awakawa et al., 2012)). PKSs and hybrid PKS-NRPSs have also been associated with the development of fungal spores and overwintering structures such as sclerotia (Cary et al., 2014; Schindler and Nowrousian, 2014; Cary et al., 2015).

Most fungal PKSs are multi-functional enzymes known as iterative type I PKS where each catalytic domain is encoded in one gene, though a few fungal PKS are of type III (Hashimoto et al., 2014). There are two main classes of type I PKS known as non-reducing (NR) and highly reducing (HR). All PKS harbor three essential domains – the β-ketoacyl synthase (KS), malonyl-CoA:acyl carrier protein transacylase (MAT), and acyl carrier protein (ACP) domains – however, NR- and HR-PKS vary in their catalytic domains that impact the reduction or dehydration steps of the growing carbon chain (Fig. 1). The minimal architecture of HR-PKSs is typically composed of keto reductase (KR), dehydratase (DH), and enoyl reductase (ER) domains, thereby resembling fatty acid synthases. Another key difference is the presence of a C-methyltransferase (CMeT) domain found in most HR-PKS and only in one subset of NR-PKS (e.g. *A. nidulans* AfoE, (Chiang et al., 2009)). However, despite the presence of CMeT domains

in HR-PKS, analysis in *Fusarium spp.* suggests that the domain can be non-functional (Brown et al., 2012).

NR-PKSs are characterized by the absence of KR, DH, and ER domains and the presence of starter-unit:ACP transacylase (SAT) and product template (PT) domains. While N-terminal SAT domains of most NR-PKS accept acetyl-CoA as starter unit (Crawford et al., 2006) there are certain examples where the starter unit is either provided by dedicated FAS-like proteins (Watanabe and Townsend, 2002) or HR-PKS (Zhou et al., 2010) which are usually encoded by genes located within the co-regulated cluster region. While the KS domain largely controls chain length (Liu et al., 2014), the PT domain, which is always located between the MAT and ACP domains, determines the cyclization pattern of the polyketide product (Crawford et al., 2008). The generated cyclized products of NR-PKSs are released by a variety of mechanism, that in some cases contribute to cyclization patterns as well (Gerber et al., 2009). The thioesterase (TE)-mediated product release by a canonical TE domain is the most common mechanism (e.g. orsellinic acid/F9775 (Bok et al., 2009; Schroeckh et al., 2009; Sanchez et al., 2010)). This domain often extends to a domain capable of C-C Claisen cyclization (a TE/CYC domain) (e.g. aflatoxin and sterigmatocystin PKS and many pigment PKSs, (Fujii et al., 2001; Korman et al., 2010)). In some cases, NR-PKSs contain a reductase-releasing (R) domain (e.g. asperfuranone (Chiang et al., 2009)); in others, e.g., the PKS involved in asperthecin biosynthesis in *A. nidulans*, there is no releasing domain. These latter NR-PKSs are usually coupled with metallo-β-lactamase-type TE proteins (MβL-TE) that allow for release of the nascent polyketide chain. These MβL-TEs are encoded by distinct genes located within their respective gene clusters, as illustrated for asperthecin (AptB), endocrocin (EncB), viridicatumtoxin (VrtG), anthrotainin

(AdaB), geodin (GedB), monodictyphenone (MdpF), pestheic acid (PtaB), trypacidin (TpcB), and neosartoricin (NscB) (Szewczyk et al., 2008; Awakawa et al., 2009; Chiang et al., 2010b; Chooi et al., 2010; Lim et al., 2012; Chooi et al., 2013; Xu et al., 2014; Throckmorton et al., 2015). Here we conducted a phylogenetic analysis of PKS genes from currently available fungal genomes, with a focus on TE-less NR-PKSs, group V. Our data strengthen predictions of subdivisions within this group, and, moreover, we expand upon the predictive power of this analysis to suggest directions for future research.

Group V NR-PKS phylogeny

Previous phylogenetic analyses have used the whole PKS, the KS domain, or the PT domain for comparison, but these have been noted to reflect one another, indicating their coevolution (Ahuja et al., 2012). These studies classified the NR-PKSs first into three subclades (Kroken et al., 2003), then further into five (Li et al., 2010), seven (Ahuja et al., 2012), and most recently eight groups (Liu et al., 2015). These NR-PKSs are present in many ascomycetes and some basidiomycetes, though only members of group VIII have been identified in basidiomycetes (Liu et al., 2015) with only one characterized example to date (Lackner et al., 2012). These groups generally represent unique combinations of product length and cyclization register. Of the eight current groups we focus on NR-PKSs belonging to the TE-less group V in this study (Fig. 2). Group V has twelve characterized gene clusters with eleven described products including endocrocin, monodictyphenone, trypacidin, geodin, pestheic acid, asperthecin, TAN-1612, neosartoricin, viridicatumtoxin, griseofulvin, and alternariol (Table 1, Fig. 3). All group V PKSs lack a domain for product release but they vary in the length (hepta- to decaketide) of their products; most make C6-C11 connections, but a couple of exceptional cases

make C8-C13 connections. Generally, the thioesterase activity is encoded in a separate but adjacent MβL-TE. Based on our analysis of the KS domains of 908 fungal PKSs identified through the NCBI's BLAST utility and manually using AspGD (Fig. S1), we have identified 188 PKSs belonging to group V (Fig. S2). These PKSs are derived from 88 species of 39 genera distributed across five classes of ascomycetes (Table 1). Notably, no group V PKSs were identified in the well characterized *Fusarium* spp. (Hansen et al., 2015). This phylogenetic tree allowed us to visualize relationships between these unknown PKSs and the twelve examples, thus defining the best contexts in which to deduce the functions of the putative clusters to which these PKSs belong.

Phylogenetics directs product prediction

Directed by our phylogenetic analysis, we sought to predict the products of the uncharacterized members of group V by comparison to the dozen characterized gene clusters. Coevolution of the genes in SM gene clusters allows that phylogenetic analysis of a constituent gene, or the protein it encodes, can be used to inform the search for clusters with interesting similarities or differences to known clusters. We applied the existing knowledge of the unique attributes of group V gene clusters to the clusters identified in our phylogenetic tree, and discuss examples from groups of gene clusters that are potentially of interest or that might be insignificantly different from known clusters. These results should aid in identifying interesting targets for future study and in avoiding duplication of efforts in the SM research community.

<u>Group V1</u>

Within the group V PKSs, a subset known as group V1 consists of octaketide synthases with C6-C11 cyclization. Five clusters from this subgroup have been characterized and their products determined. These include endocrocin, monodictyphenone, trypacidin, geodin, and pestheic acid, produced by *A. fumigatus*, *A. nidulans*, *A. fumigatus*, *A. terreus*, and *Pestalotiopsis fici*, respectively (Chiang et al., 2010b; Lim et al., 2012; Nielsen et al., 2013; Xu et al., 2014; Throckmorton et al., 2015). Many of the initial studies characterizing these clusters relied on earlier biochemical characterization of the geodin and aflatoxin biosynthetic pathways (Huang et al., 1995; Couch and Gaucher, 2004; Ehrlich et al., 2005; Henry and Townsend, 2005b, a; Cary et al., 2006; Awakawa et al., 2009; Ehrlich et al., 2010). Group V1 is notable for the large number of aflatoxin homologs its clusters contain. Some clusters belonging to groups V2 and V3 have homologs of *aflL* (*vrtK*) and *alfO* (*gsfD*), but the characterized clusters in group V1 collectively contain homologs of as many as seven *afl* or *stc* (sterigmatocystin) cluster genes. These include homologs of *aflR*, *aflS*, *aflX*, *aflY*, *aflM*, *hypC*, and *stcT* (Fig. 4A). The trypacidin and geodin clusters have previously been noted to contain partial aflatoxin clusters (Carbone et al., 2007), but this is true of group V1 clusters in general. Manual analysis of our MultiGeneBLAST (MGB) (Medema et al., 2013) results additionally revealed a gene with significant similarity to versicolorin B-synthase (Vbs, AflK) associated with several uncharacterized clusters (Fig. S2). We speculate that the high number of *afl*/*stc* gene homologs reflects that group V1 clusters ultimately derived from portions of the *afl*/*stc* clusters.

All characterized group V1 clusters produce metabolites with an anthraquinone skeleton, such as endocrocin, emodin, and versicolorin A, by action of anthrone oxidases, e.g. HypC, StcM, EncC, MdpH2, and TpcL. Except for endocrocin which represents an end-product, a

subset may then be processed into an open-ringed benzophenone structure, like monodictyphenone, by a Baeyer-Villiger oxidase (BVO), an NADH-dependent oxidoreductase (NOR), and potentially a glutathione *S*-transferase (GST). These reactions are similar to early steps in aflatoxin biosynthesis and this is reflected in the conservation of homologs for these key enzymes in group V1 clusters, though the sets of enzymes required between the aflatoxin/sterigmatocystin and group V1 cluster pathways are only partially overlapping (Fig. 4B). Benzophenones may further be converted into closed ring structures by spontaneous dehydration to xanthones, such as sterigmatocystin, or enzymatically to grisandienes, like trypacidin and geodin, by a multicopper oxidase (MCO), e.g. TpcJ. We leverage the knowledge of these conserved enzymes that catalyze the conversions between these chemical classes to make predictions about the products of related but uncharacterized clusters in this group.

Endocrocin-like clusters

The smallest characterized cluster from group V1, the four-gene endocrocin cluster was characterized as a virulence factor in *A. fumigatus* (Lim et al., 2012; Berthier et al., 2013) and is also found in *Neosartorya fischeri*. A similar PKS is present in *A. terreus* (EAU37396), but appears to not be surrounded by any decorating genes, suggesting that the cluster might be a remnant of an endocrocin-like cluster. The endocrocin cluster encodes only two other enzymes in addition to the PKS (EncA) and MβL-TE (EncB) characteristic of group V clusters, an anthrone oxidase (EncC) and a 2-oxoglutarate-Fe(II)-type oxidoreductase (EncD). This minimal complement of decorating enzymes enables the production of the simple anthraquinone endocrocin. Closely related PKSs in a series of other aspergilli are associated with homologs of *encB*, but no homologs of *encC* or *encD* were found in clusters that were available for analysis

by MGB or AspGD. Similar to what was observed in *A. terreus*, it appears that the clusters in this clade may be remnants of endocrocin-like clusters (Fig. S2). It is intriguing to speculate that cluster duplications from more complex clusters present in group V1 (see below) and subsequent deterioration led to this sub-group. Recent evidence for chemical redundancy between endocrocin and intermediates from the trypacidin cluster in *A. fumigatus* could explain the potential decay of the endocrocin and endocrocin-like clusters (Throckmorton et al., 2015).

Monodictyphenone-like clusters

The monodictyphenone-producing cluster in *A. nidulans* consists of 12 genes, notably including homologs of the three genes required for endocrocin production in the *enc* cluster. These encode a PKS (MdpG), an MβL-TE (MdpF), and an anthrone oxidase, MdpH2. Though the latter is annotated as part of a larger gene, *mdpH*, studies of the trypacidin- and geodin-producing clusters (Nielsen et al., 2013; Throckmorton et al., 2015) suggest that *mdpH* is actually two separate genes, herein referred to as *mdpH1* and *mdpH2*. The *mdp* cluster is capable of producing endocrocin, but only in the absence of *mdpH* (Chiang et al., 2010b), suggesting that the other half of this gene, *mdpH1*, encodes a decarboxylase. This cluster additionally produces prenylated xanthones with the activity of prenyltransferases encoded outside of the cluster itself (Sanchez et al., 2011). Among the proteins encoded by characterized group V cluster genes, MdpB, C, D, and I are unique to the *mdp* cluster. The presence of the corresponding genes can be used to differentiate *mdp*-like clusters from clusters more similar to the other members of group V1. Partial *mdp* clusters are found in two fungi closely related to *A. nidulans*, *A. sydowii* and *A. versicolor*.

Although not products of a group V PKS, the polyketide mycotoxins aflatoxin and sterigmatocystin share decorating enzymes to all group V1 pathways. Biosynthesis of aflatoxin and sterigmatocystin involves the conversion of the anthraquinone precursor versicolorin A to the xanthone demethylsterigmatocystin through a benzophenone-like intermediate and involves the actions of a cytochrome P450 monooxygenase (P450), AflN/StcS (Keller et al., 1994; Henry and Townsend, 2005b, a), a ketoreductase (KR), AflM/StcU (Skory et al., 1992), an NOR, AflX/StcQ (Skory et al., 1992; Cary et al., 2006), and a BVO, AflY/StcR (Ehrlich et al., 2005). In the biosynthesis of monodictyphenone and prenyl xanthones in *A. nidulans*, a similar ring-opening reaction involving the conversion of the anthraquinone chrysophanol to the benzophenone aldehyde arugosins was recently proposed to be carried out by a glutathione *S*-transferase (GST), MdpJ, an NOR, MdpK, and a BVO, MdpL (Chiang et al., 2010b; Simpson, 2012). The genes encoding these enzymes bear significant similarity to StcT, AflX, and AflY, respectively (Fig. 4). No known role is proposed for StcT in sterigmatocystin biosynthesis and no homolog of *stcT* is present in the aflatoxin gene cluster. Interestingly, action of only the BVO MdpL followed by hydrolysis is sufficient to produce the benzophenone carboxylic acid monodictyphenone as a shunt product. The *mdp* cluster also contains an *aflM* homolog in *mdpC*, but MdpC has been speculated to catalyze the conversion of emodin to chrysophanol in combination with MdpB and not to be involved in the ring-opening step (Simpson, 2012). The biosyntheses of the related compounds trypacidin and geodin in *A. fumigatus* and *A. terreus*, respectively, involve ring-opening conversion of the anthraquinone questin to the benzophenone desmethylsulochrin speculated to be catalyzed by a BVO (TpcI/GedK) and potentially an NOR (TpcG/GedF) and GST (TpcF/GedE) (Henry and Townsend, 2005b, a; Simpson, 2012; Nielsen et al., 2013; Throckmorton et al., 2015). In pestheic acid biosynthesis in *Pestalotiopsis fici*, the

ring-opening of the anthraquinone physcion to the benzophenone desmethylisosulochrin is similarly proposed to be mediated by a BVO, PtaJ, and an NOR, PtaF, but this cluster encodes no GST, suggesting that this enzymatic activity might not be required for this transition (Xu et al., 2014). In summary, these anthraquinone ring-opening reactions to form benzophenones all involve BVOs, NORs, and potentially GSTs. This is a similar but distinct set of enzymes required for the analogous chemical reactions in the biosynthesis of aflatoxin and sterigmatocystin. Taken together, it is remarkable that, despite the obvious differences between the aflatoxin/sterigmatocystin PKSs (belonging to group IV) and the PKSs present in group V, this subset of enzymatic genes catalyzing ring-opening reactions is shared. Similar to this set of enzymatic genes, homologs of the two regulatory proteins AflR/S are also conserved in group V1 (Fig. 4). It is noteworthy that, unlike the ring-opening enzymes, genes encoding AflR/S homologs can be found in other group IV clusters, i.e. the fusarubin cluster (Studt et al., 2012), and even in group III clusters, i.e. the bikaverin cluster (Wiemann et al., 2009).

One cluster from *A. ustus*, speculated to produce monodictyphenone in a recent study (Pi et al., 2015) due to its close phylogenetic relationship with that cluster, also has an MCO similar to that of the trypacidin, geodin, and pestheic acid clusters (see below). This suggests that this cluster might produce a chemical structure more similar to these latter clusters than to monodictyphenone and thereby exemplifies the need for a close evaluation of the whole cluster architecture.

Trypacidin, geodin, pestheic acid-like clusters

As noted above, the monodictyphenone cluster shares many similarities with the trypacidin and geodin biosynthetic pathways in *A. fumigatus* and *A. terreus*, respectively, and pestheic acid biosynthesis in *P. fici*. All of these clusters catalyze anthraquinone to benzophenone ring-opening reactions using a BVO and an NOR. In the biosyntheses of trypacidin and geodin, the transition from the open benzophenone to closed grisandiene is catalyzed by a MCO. Specifically, in trypacidin biosynthesis TpcJ converts monomethylsulochrin to trypacidin, and, in geodin biosynthesis, GedJ converts dihydrogeodin to geodin (Chooi et al., 2010; Cacho et al., 2013).

The phylogenetic alignment of the trypacidin, pestheic and geodin clusters to other fungi allow us to speculate on products from undefined fungal clusters. A PKS (AEO66245) encoded by *Thielavia terrestris* is closely related to the geodin and trypacidin PKSs, and the cluster to which it belongs has homologs to eight of the thirteen genes in the trypacidin cluster (Fig. 5). Importantly, these include genes encoding the enzymes required for the anthraquinone to benzophenone transition, a BVO and an NOR, as well as the benzophenone to grisandiene transition, which is catalyzed by a MCO in this subgroup of clusters. The presence of genes encoding these key enzymes suggests that this cluster might ultimately produce a grisandiene (trypacidin or geodin-like molecule). This fungus is a little-known species that can cause human infections (Liu, 2011). Interestingly, trypacidin is a toxic spore metabolite produced by the opportunistic pathogen *A. fumigatus* (Gauthier et al., 2012), which may suggest that the metabolite produced by the *T. terrestris* cluster could play a role in pathogenicity.

Inspection of another of the PKSs closely related to the trypacidin, geodin, and pestheic acid PKSs, GAM37897.1 from *Talaromyces cellulolyticus* (now recognized as synonymous with *T. pinophilus*) (Fig. 5), showed that the gene encoding this protein is part of a thirteen-gene cluster with homologs of nine of the thirteen genes in the trypacidin-producing gene cluster. This cluster also has genes encoding the key BVO and NOR enzymes and so likely produces a grisandiene, whether as an intermediate or an end-product. This species is known to produce many secondary metabolites including austin, mitorubrins, penicillides/purpactins/vermixocins, rubropunctatin, vermicellin, vermiculin, vermistatin and (3-*O*-methyl-, 3-*O*-methyl-5,6-epoxy-) funicones, MC-141, pestalacin A, stromemycin, dinapinone A1 and A2, and monoapinone A-E (Yilmaz et al., 2014). The structures of vermixocins and purpactins suggest they are products of this cluster. These compounds have grisandiene- and depsidone-like scaffolds, which are known or speculated to derive from the biosynthetic pathways of group V clusters such as geodin and pestheic acid (Adeboya et al., 1996; Nielsen et al., 2013). Notably, these compounds appear to be prenylated despite the lack of a prenyltransferase identified in the cluster; however, the modification of SMs by prenyltransferases encoded outside of the gene cluster has been observed in monodictyphenone derivatives (Sanchez et al., 2011).

Group V2

Another subgroup of group V is group V2, which includes nona- and decaketide synthases with C6-C11 cyclization. Characterized examples include asperthecin (Szewczyk et al., 2008), viridicatumtoxin (Chooi et al., 2010), anthrotainin (Li et al., 2011), and neosartoricin (Chooi et al., 2013). Enzymatic activities unique to this group include a fourth-ring cyclization facilitated by a flavin-monooxygenase and MβL-TE combination. In contrast to the tricyclic

(anthracene) backbones produced by most group V PKSs, two characterized examples in group V2 have the ability to generate tetracyclic (naphthacenedione) backbones, anthrotainin and viridicatumtoxin. This ability depends on several factors including the ability to synthesize a long, i.e. nona- or decaketide, backbone and the presence of both an MβL-TE with Claisen-cyclase activity and a unique flavin-monooxygenase (Li et al., 2011). Though VrtA is only a nonaketide synthase, it accepts the very unusual malonamoyl-CoA starter unit produced by VrtB and VrtJ, and thus has a long enough chain for a fourth cyclization (Chooi et al., 2010). Phylogenetic analysis of prenyltransferases associated with these clusters was previously used to identify a group of clusters with this triad of a unique PKS, MβL-TE, and FMO in dermatophytic fungi (Chooi et al., 2012) (Fig. 6).

Asperthecin-like clusters

Asperthecin, associated with sexual spore color in *A. nidulans* (Palmer and Keller unpublished data), is produced from a three-gene cluster encoding the NR-PKS, the MβL-TE and a FAD-dependent oxidoreductase (Szewczyk et al., 2008). As seen in Figure S2, the two close relatives *A. versicolor* and *A. sydowii* contain the same cluster, which we hypothesize is also likely to be associated with ascospore color in these two fungi.

TAN-1612-like clusters

TAN-1612, identified in *A. niger* (Li et al., 2011) and also present in *A. kawachii*, is a five-gene cluster, three of which are homologous to the *apt* cluster genes. In addition to the three described proteins above, this cluster contains a methyltransferase and a GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein. It is tempting to

speculate that TAN-1612 might be associated with ascospore pigmentation however the sexual stage of neither *A. niger* nor *A. kawachii* has been described yet for assessment of such a hypothesis.

Neosartoricin-like clusters

Neosartoricin is produced by a six-gene cluster in *A. fumigatus* and *N. fischeri* (Chooi et al., 2013). This cluster was also identified as producing the related fumicyclines (König et al., 2013). The cluster's PKS, NscA/FccA, produces a decaketide chain, the longest known of all PKSs so far described along with the TAN-1612-producing PKS, AdaA (Chooi and Tang, 2012). Five of the six genes in the *nsc*/*fcc* cluster are conserved in the dermatophytic genera, *Trichophyton*, *Arthroderma*, and *Microsporum* (Fig. 6) some of which have previously been noted (Chooi et al., 2012). Notably, the gene for which there is no conserved homolog in these species, *nscE*/*fccE*, has no proposed role in neosartoricin or fumicycline biosynthesis in *A. fumigatus* (Chooi et al., 2013; König et al., 2013). Compared to the *nsc* cluster, these conserved clusters in the Arthrodermataceae appear to have two two-gene inversions, but are otherwise syntenically conserved. Fumicyclines are induced in the presence of *Streptomyces rapamycinicus* and neosartoricin has demonstrated immunosuppressive activity (Chooi et al., 2013; König et al., 2013). This potentially suggests an important role for this cluster in virulence of dermatophytes.

Viridicatumtoxin-like clusters

Viridicatumtoxin is a tetracyclic mycotoxin produced by *Penicillium* species. The 14-gene *vrt* cluster contains homologs of the 5 conserved *nsc* genes mentioned above. We have identified a group of *Metarhizium* species that have gene clusters with homologs of ten of the

vrt-cluster genes (Fig. 6). *Metarhizium* species are entomopathogenic fungi in the Clavicipitaceae family. These *Metarhizium* clusters contain two regions of conserved synteny with four and five gene regions of the *vrt* cluster and are predicted to yield a tetracyclic polyketide. Interestingly, in *A. ustus* a *vrt*-like cluster contains homologs of *abr2* and *pksP* of the conidial pigment biosynthetic gene cluster, suggesting that this may be one large cluster or two interwoven clusters perhaps similar to that of the intermingled fumagillin/pseurotin supercluster in *A. fumigatus* (Wiemann et al., 2013). Several of these species have been noted to have *vrt*-like clusters in recent studies (Gibson et al., 2014; Pi et al., 2015).

<u>Group V3: Griseofulvin and various alternariol-like pathways.</u>

This group V subgroup includes heptaketide synthases catalyzing an unusual C8-C13 cyclization. Characterized examples include alternariol and griseofulvin. Alternariol is an important mycotoxin produced by members of *Alternaria*, *Aspergillus*, and *Phaeosphaeria* (Ahuja et al., 2012; Saha et al., 2012; Chooi et al., 2015). The metabolite is a fairly common crop contaminant with carcinogenic, phytotoxic, and antifungal activity. Despite its importance, genetic studies characterizing the biosynthesis of alternariol were only recently undertaken. To date, three gene clusters have been implicated in the synthesis of alternariol, one in *Alternaria alternata*, one in *Aspergillus nidulans*, and one in *Phaeosphaeria nodorum* (syn. *Parastagonospora nodorum*) (Fig. 7). Initially, PksJ (AFN68301) was identified in *A. alternata* using siRNA and gene deletion approaches as the PKS primarily responsible for alternariol production. However, two other PKSs, PksH (AFN68299) and PksI (AFN68300), were shown to be affected by the knockdown of *pksJ* expression. Notably, no MβL-TE was identified adjacent to PksJ in this study (Saha et al., 2012). In *A. nidulans*, promoter replacement experiments were

used to show that PkgA (CBF79143) and PkgB produce alternariol and coumarins (Ahuja et al., 2012). Most recently, in *P. nodorum*, SNOG_15829 (EAT76667) was also found to produce alternariol. The cluster associated with this NR-PKS includes a gene encoding an MβL-TE (SNOG_15826), but it bears little similarity to the other MBL-TEs of group V clusters (Chooi et al., 2015), which could be due to poor sequence quality. Interestingly, the NR-PKS EAT76667 is most similar to PksI from *A. alternata*, suggesting that PksI, and not PksJ or PksH, is the alternariol-producing PKS in *A. alternata*. These clusters were not characterized further than the identification of a PKS and an MβL-TE; the additional genes analyzed by MGB as part of the PkgA (CBF79143) and SNOG_15829 (EAT76667) clusters were included based only on their reported co-regulation with the PKS- and MβL-TE-encoding genes (Andersen et al., 2013; Chooi et al., 2015).

Griseofulvin, produced by *Penicillium* species, is an antifungal drug widely used against dermatophytic infections (Finkelstein et al., 1996; Aggarwal and Goindi, 2013; Margarido, 2014). Despite its resemblance to other grisandienes like trypacidin and geodin, the griseofulvin biosynthetic pathway (Chooi et al., 2010) is quite unique and constitutes an interesting example of convergent evolution at the biochemical level. The NR-PKS responsible for griseofulvin production, GsfA (ADI24953), generates a benzophenone directly and by a different cyclization register than that of the group V1 PKSs MdpG, TpcC, GedC, and PtaA, i.e. C1-C6 as opposed to C6-C11 for group V1 PKSs or even C4-C9 for the aflatoxin-producing group IV PKS, AflC (Cacho et al., 2013). Thus, whereas the biosynthetic pathways of group V1 clusters proceed through an anthraquinone intermediate to a benzophenone intermediate by action of a BVO, GsfA synthesizes a benzophenone as its initially released product. This is made possible by the

unusual C8-C13 connection it catalyzes. This connection is also observed in the biosynthesis of alternariol, but the two differ in their initial cyclization, i.e. C1-C6 for griseofulvin and C2-C7 for alternariol (Chooi et al., 2015). Further, the benzophenone to grisandiene transition in griseofulvin biosynthesis is catalyzed by a cytochrome P450 as opposed to a MCO in the cases of trypacidin, geodin, and pestheic acid (Cox, 2014). Curiously, the release mechanism of GsfA has yet to be elucidated, as it lacks the MβL-TE that is characteristic of group V. We speculate that the unique cyclization catalyzed by these PKSs might obviate the MβL-TE activity and explain the lack of an MβL-TE in the *gsf* cluster and the dissimilarity of the MβL-TE of the *P. nodorum* alternariol-producing cluster. It has been noted that the close relationship of the griseofulvin- and alternariol-producing PKSs is likely not coincidental and that unusual modes of cyclization may be unique to this clade (Liu et al., 2015).

Uncharacterized clusters in symbionts and pathogens

As expected, considering the number of sequenced *Aspergillus* species, many of the NR-PKSs identified from this study are from *Aspergillus* species. However, of the non-*Aspergillus* genera and species, many are symbiotic, including pathogenic fungi associated with specific hosts. These findings support reviews of pathogenic fungi which highlight the potential role of secondary metabolites in virulence (Perez-Nadales et al., 2014). Below we touch on this emerging theme from our analysis.

Plant pathogens

Several plant pathogenic fungi are present in two main groupings in Fig S2. Notably, the subclade to which the griseofulvin and alternariol PKSs belong contains many plant pathogenic

species (Fig. 7), including the pine pathogen *Grosmannia clavigera*, the sorghum pathogen *Colletotrichum sublineola*, the wheat pathogen *Gaeumannomyces graminis*, the maize pathogen *Bipolaris maydis*, the grape pathogen *Botrytis cinerea* and two grass powdery mildew species from the genus *Blumeria*. This latter finding is especially intriguing as the genomes of obligate biotrophs such as powdery mildews contain few secondary metabolite genes (Duplessis et al., 2011; Wicker et al., 2013). As alternariol produced by related clusters from *Alternaria*, *Aspergillus*, and *Phaeosphaeria* spp. is known to be an important phytotoxin, it is possible these NR-PKS produce a metabolite important in the fungal/plant host interaction.

Several plant pathogenic fungi also are present in the subclades producing trypacidin, geodin, pestheic acid and monodictyphenone. The genus *Bipolaris* and allied genera *Setophaeria* and *Pyrenophora* – all grass pathogens – are particularly well represented in this clade. Another grass pathogen, *Claviceps purpurea*, is also present in this subclade, although it is taxonomically distant from the *Bipolaris* species and is best known for its suite of alkaloid-producing clusters (Schardl et al., 2013). It would be interesting to see if these clusters produce a metabolite specialized to interactions with grass hosts.

Mycorrhizal fungi

Oidiodendron maius belongs to the rare group of fungi establishing mycorrhizal relationships with the plant family Ericaceae (heather) (Khouja et al., 2014; Elmore et al., 2015). Most unusually, this species contains four NR-PKSs associated with nearly every group V subgroup. Poor annotation of the genome prevented full assessment of the associated clusters; hopefully future sequencing efforts will provide data for accurate cluster predictions.

Fungal pathogens

The mycoparasitic genus *Trichoderma* is represented twice in our analysis, once where they contain a gene cluster quite similar to the *pkg* cluster (PKS=CBF79143) and one similar to the pestheic acid cluster of group V1. Considering the mycoparasitic lifestyle of these fungi, the presence of these clusters raises the question whether their respective products could play a role in mycoparasitism.

Animal pathogens

Several animal pathogens, from insect to human, contain NR-PKS also found throughout group V.

Entomopathogenic fungi.

The genus *Metarhizium,* along with *Beauveria bassiana* not represented in group V, is well known for its potential in biological control of various insects (Liu et al., 2013). A putative group V1 cluster was identified in *B. bassiana* including EJP67854, but it did not include an annotated PKS and so was excluded from further analysis. It is unclear whether this is related to any of the several NR-PKSs that have previously been noted in *B. bassiana* (Punya et al., 2015). As mentioned earlier, several species in *Metarhizium* contain clusters with significant similarity to the *Penicillium* viridicatumtoxin cluster (Fig. 6). The *Metarhizium* genus is also well represented in group V1, suggesting the ability to produce a metabolite similar to these compounds. Toxicity is associated with all of these metabolites and may afford virulence properties to these insect pathogens.

Dermatophytes.

Several dermatophytic genera, *Trichophyton*, *Arthoderma* and *Microsporum* contain the neosartoricin gene cluster. This compound exhibits antiproliferative activity which may be suggestive of an immunosuppressive role in human infection by dermatophytic fungi (Chooi et al., 2013). The causal agent of a devastating bat disease in North America known as white nose syndrome, *Pseudogymnoascus destructans*, is a dermatophyte of bats (Chibucos et al., 2013). This species and other members of the genus are found in two sections of group V, both grouping to the trypacidin, geodin, pestheic acid, and monodictyphenone clades. Possibly metabolites produced by these clusters could play a role in virulence of *P. destructans*.

Experimental procedures

To retrieve the amino acid sequences of NR-PKSs in our initial search, the monodictyphenone PKS, MdpG (Chiang et al., 2010b), was analyzed with the NCBI's BLASTP (Altschul et al., 1990) against fungi (taxid: 4751) with max target sequences increased to 1000 and other parameters set to defaults. A similar search was executed on AspGD (aspgd.org) and 25 non-duplicate PKSs were identified from the top 50 hits and added to the list retrieved from the NCBI. The human fatty acid synthase, FASN, was added as a marker for the outgroup, which also included 70 HR-PKSs and hybrid PKS/NRPSs. The KS domains for these 908 PKSs were retrieved using the NCBI's Conserved Domain Database (CDD) utility (Yadav et al., 2009; Finn et al., 2014; Marchler-Bauer et al., 2015) and aligned using MAFFT (Katoh and Standley, 2013) with default parameters. Alignment columns with greater than 40% gaps were removed using TrimAl (Capella-Gutiérrez et al., 2009), and the sequences were realigned. A maximum likelihood phylogenetic tree was constructed using FastTree (Price et al., 2010) in Geneious

8.1.5 (Kearse et al., 2012) with 1000 bootstrap replicates and otherwise default parameters. No bootstrap value threshold was applied. This tree was modified for presentation using FigTree (Rambaut, 2007) (Figs. 2, S1). To extract the sequences for the group V KS domains, a neighbor-joining tree was constructed in Geneious with default parameters, and the sequences for the smallest monophyletic group containing the characterized group V PKSs were selected and exported as a sub-alignment. These 188 sequences were realigned with MAFFT and used to construct a maximum likelihood phylogenetic tree using FastTree as above. This tree was modified for presentation using FigTree (Rambaut, 2007) (Figs. 3, S2). Excerpts of this tree were used to create Figures 4-7.

MultiGeneBLAST (Medema et al., 2013) (MGB) was used to facilitate the analysis of uncharacterized clusters associated with the PKSs identified by the methods described above. MGB architecture searches were executed with the synteny conservation weight set to 0, the percent identity threshold set to 25%, the maximum intergenic distance threshold set to 25 kb, and otherwise default parameters. MGB searches were carried out with multifasta files containing one representative of each group of orthologous genes from a given subclade of group V, e.g. the group V1 search used a file containing one PKS, one MβL-TE, etc. (Tables S1-3). If multiple genes in a single cluster encoded similar types of proteins, e.g. C6TFs, TFs, P450s, or MTs, they were considered as separate and included in the multifasta file for that subgroup (i.e. V1, V2, or V3) despite their potential redundancy. This was done to identify potential patterns in the distribution of orthologs across other clusters, i.e. if a given ortholog was more similar to one or the other of the potentially redundant genes. Alternatively, if the potentially redundant genes were not differentiable with our MGB parameters they and all their orthologs were considered to

encode the same type of protein and represented as such in the cluster diagrams and corresponding color keys. The database available online (http://multigeneblast.sourceforge.net/index.html) containing all GenBank entries, updated 1/2015, was used to obtain the majority of the cluster diagrams (Figs. S2, 4-7). Others were obtained from a custom fungal database generated by downloading 578 annotated fungal genomes from NCBI using a custom python script (https://github.com/nextgenusfs/NR-PKS_ms/get_ncbi_genomes.py). These genomes were then incorporated into a MGB database using the `makedb` program from the command line distribution of MGB. Still others, such as the *T. cellulolyticus* (GAM37897) and *A. terreus* (EAU31624) clusters (Figs. 5, S2), were manually created or modified (to reflect re-annotation of ATEG_08457 (Nielsen et al., 2013)), respectively. The protein descriptors in the keys or adjacent to the cluster diagrams (Figs. 4-7, S2) were derived from the function or conserved domain of the protein according to published studies describing group V SM gene clusters (Szewczyk et al., 2008; Chiang et al., 2010b; Chooi et al., 2010; Li et al., 2011; Ahuja et al., 2012; Lim et al., 2012; Saha et al., 2012; Simpson, 2012; Chooi et al., 2013; König et al., 2013; Nielsen et al., 2013; Xu et al., 2014; Chooi et al., 2015; Throckmorton et al., 2015) or individual BLAST searches.

Conclusions

The potential to synthesize polyketides is widespread in the fungal taxa ascomycetes and basidiomycetes but examination of biochemical PKS classes indicate taxonomic specificities (Liu et al., 2015). Illustrating this point, the β-lactamase releasing NR-PKSs described in this work and others (Ahuja et al., 2012; Liu et al., 2015), group V, are notably absent in certain mycotoxigenic genera (e.g. *Fusarium*) and basidiomycetes but, surprisingly, found in genera not

noted for secondary metabolism (e.g. *Oidiodendron* and *Blumeria*) (Figs. 5-7, S2). Many of the described products derived from these NR-PKSs exhibit toxic activities. For example, questin, trypacidin and endocrocin have been assessed for their impact on virulence as they are produced by the human pathogen *A. fumigatus*. Trypacidin and endocrocin in particular have been shown to exhibit toxic and neutrophil inhibitory properties respectively in pathogenicity studies (Gauthier et al., 2012; Berthier et al., 2013). The other *A. fumigatus* group V metabolite, neosartoricin, exhibits T-cell antiproliferative activity, which may be suggestive of an immunosuppressive role in human infection (Chooi et al., 2013). Furthermore, griseofulvin is a potent antifungal (Finkelstein et al., 1996), viridicatumtoxin is a mycotoxin (Hutchison et al., 1973; Bendele et al., 1984), and alternariol is known for its phytotoxic properties (Pero et al., 1973; Demuner et al., 2013). Thus it is not improbable that the products of the NR-PKS clusters identified in the pathogenic and symbiotic fungi in this study (Fig. S2) could impact fungal/host interactions, as virulence factors or signaling molecules. However, it is also possible that such roles could be coincidental or in addition to other functions in fungal biology.

As predicted by their structure, it is likely all are UV absorbing pigmented molecules observed in the visible yellow-orange spectrum. Possibly one significant role of β-lactamase PKSs may lie in protection or development of spores. Asexual spores are common air dispersed spores essential for propagation and dispersal of the kingdom Fungi and must be equipped with defenses against abiotic stresses, such as UV radiation, oxidative stress, and desiccation. Several studies have shown that loss of pigmentation of spores leads to reduced viability and/or virulence in pathogenic fungi (Wang and Casadevall, 1994; Tsai et al., 1998; Romero-Martinez et al., 2000; Esbelin et al., 2013). Sexual spores are the product of meiosis and essential for genetic

recombination and species diversity and, in several species, are also airborne spores that would be exposed to UV radiation similarly as asexual spores (Guenther and Trail, 2005; Fourie et al., 2013). Both questin/trypacidin and endocrocin are produced in the asexual spore (Gauthier et al., 2012; Berthier et al., 2013) and our studies (Palmer and Keller, unpublished data) – supported by an earlier investigation (Brown and Salvo, 1994) – suggest that asperthecin is the red pigment characterizing the color of *A*. *nidulans* sexual ascospores. Assessment of spore viability of NR-PKS fungal mutants under UV conditions might shed light on a conserved role of these molecules to protect from specific abiotic stresses.

In this study we have examined all of the group V NR-PKSs available from NCBI and all of the corresponding gene clusters that were readily available using MultiGeneBLAST. The relatively high number of studies characterizing SM gene clusters from group V (Szewczyk et al., 2008; Chiang et al., 2010b; Chooi et al., 2010; Li et al., 2011; Ahuja et al., 2012; Lim et al., 2012; Saha et al., 2012; Simpson, 2012; Chooi et al., 2013; König et al., 2013; Nielsen et al., 2013; Xu et al., 2014; Chooi et al., 2015; Throckmorton et al., 2015), has enabled us to predict the products of uncharacterized clusters in this group (Table 1). For group V1 this is largely based on the presence or absence of genes encoding key enzymes known to catalyze particular reactions, e.g. anthraquinone ring opening by BVOs and NORs or benzophenone ring-closure by MCOs. For groups V2 and V3 these predictions are based only on the presence of homologs of a majority of the genes in the most closely related characterized group V cluster, e.g. neosartoricin is predicted to be produced from clusters containing homologs of five of the six *nsc* cluster genes. This study demonstrates that, by comparison to characterized examples in a given group of NR-PKS, predictions can be made from phylogenetic analysis and used to help choose SM

gene clusters to characterize. For example, future research efforts might be better spent studying SM gene clusters of interest highlighted by this study and not on those that are highly similar to characterized clusters and therefore likely to produce a known compound. Application of these methods to other groups of NR-PKSs might provide a similarly beneficial perspective.

Acknowledgments

The authors would like to thank Dr. Jonathan Palmer for assistance with the computational analysis associated with the construction of phylogenetic trees and MultiGeneBLAST database construction. We would also like to thank our funding sources, grants from the National Science Foundation (EFRI-1136903), the United States Fish and Wildlife Service (F14AP00618), and the National Institutes of Health training grant to The University of Wisconsin-Madison Genetics Department (5T32GM07133).

Tables

Table 1: Group V PKSs identified and analyzed in this study.

Class	Genus	Species	PKS	Sub-group	Characterized product	Predicted product	Reference
Dothidiomycete	*Alternaria*	*alternata*	AFN68300	V3	Alternariol	Characterized	(Saha et al., 2012)
	Aureobasidium	*melanogenum*	KEQ66138	V1	None	Emodin	This study
	Bipolaris	*maydis*	EMD89515	V3	None	No data	This study
			ENH99769	V3	None	No data	This study
			ENI07798	V1	None	Other anthraquinone	This study
			AAR90273	V1	None	No data	This study
			AAR90274	V3	None	No data	This study
		oryzae	EUC41199	V1	None	Benzophenone	This study
			EUC45162	V1	None	Other anthraquinone	This study
		sorokiniana	EMD62925	V1	None	Other anthraquinone	This study
			EMD66882	V1	None	Benzophenone	This study
		victoriae	EUN25734	V1	None	Other anthraquinone	This study
		zeicola	EUC29913	V1	None	Other anthraquinone	This study
	Macrophomina	*phaseolina*	EKG11397	V1	None	No product	This study
			EKG18431	V1	None	No data	This study
	Phaeosphaeria	*nodorum*	EAT76667	V3	Alternariol	Characterized	(Chooi et al., 2015)
	Pseudocercospora	*fijiensis*	EME79056	V1	None	Emodin	This study
	Pseudogymnoascus	*destructans*	ELR08155	V1	None	Endocrocin	This study
		pannorum	KFY01830	V1	None	Endocrocin	This study
			KFY04191	V1	None	Other anthraquinone	This study
			KFY04767	V1	None	Endocrocin	This study
			KFY14212	V1	None	Emodin	This study
			KFY24933	V1	None	Endocrocin	This study
			KFY28347	V1	None	Endocrocin	This study
			KFY28376	V1	None	Endocrocin	This study
			KFY41668	V1	None	Emodin	This study
			KFY50141	V1	None	Grisandiene	This study
			KFY61750	V1	None	Endocrocin	This study
			KFY73941	V1	None	Endocrocin	This study
			KFY81274	V1	None	Other anthraquinone	This study
			KFY90954	V1	None	Other anthraquinone	This study
			KFY97098	V1	None	Endocrocin	This study
			KFZ02364	V1	None	Endocrocin	This study
			KFZ03027	V1	None	Other anthraquinone	This study
			KFZ03785	V1	None	Endocrocin	This study
			KFZ09857	V1	None	Endocrocin	This study
	Pyrenophora	*teres*	EFQ95560	V1	None	Endocrocin	This study
		tritici	EDU45231	V1	None	Endocrocin	This study
	Setosphaeria	*turcica*	EOA88807	V1	None	Other anthraquinone	This study
	Sphaerulina	*musiva*	EMF17386	V1	None	Other anthraquinone	This study
	Verruconis	*gallopava*	KIW05310	V1	None	No product	This study
Eurotiomycete	*Arthroderma*	*benhamiae*	EFE32713	V2	None	Neosartoricin	This study
		otae	EEQ30779	V2	None	Neosartoricin	This study
			EEQ31623	V3	None	Alternariol	This study

	Aspergillus	*acidus*	Aspfo1_0068040	V2	None	No data	This study
			Aspfo1_0069798	V1	None	No data	This study
		aculeatus	Aacu16872_063282	V1	None	No data	This study
			Aacu16872_063333	V1	None	No data	This study
		brasiliensis	Aspbr1_0070836	V1	None	No data	This study
			Aspbr1_0071307	V2	None	No data	This study
		clavatus	EAW13612	V1	None	No data	This study
		flavus	EED53479	V1	None	No data	This study
			KJJ30826	V3	None	No data	This study
		fumigatus	EAL84397	V1	Endocrocin	Characterized	(Lim et al., 2012)
			EAL84875	V2	Neosartoricin, fumicycline	Characterized	(Chooi et al., 2013; König et al., 2013)
			EAL89339	V1	Trypacidin	Characterized	(Throckmorton et al., 2015)
			EDP47078	V1	Endocrocin	Characterized	(Lim et al., 2012)
			EDP47964	V2	None	Neosartoricin, fumicycline	This study
			EDP50840	V1	None	Trypacidin	This study
			KEY78897	V1	None	Endocrocin	This study
			KEY82310	V1	None	Trypacidin	This study
			KEY82351	V2	None	Neosartoricin, fumicycline	This study
		glaucus	Aspgl1_0045725	V1	None	No data	This study
		kawachii	GAA85937	V2	None	TAN-1612	This study
			GAA88581	V1	None	No data	This study
		nidulans	CBF70387	V2	Asperthecin	Characterized	(Szewczyk et al., 2008)
			CBF79143	V3	Alternariol	Characterized	(Ahuja et al., 2012)
			CBF90097	V1	Monodictyphenone	Characterized	(Chiang et al., 2010b)
		niger	EHA20150	V1	None	No data	This study
			XP_001402309	V1	None	No product	This study
			AEN83889	V2	TAN-1612	Characterized	(Li et al., 2011)
			CAK40778	V2	None	TAN-1612	This study
			CAK47960	V1	None	No data	This study
		ochraceoroseus	KKK15179	V1	None	Emodin	This study
			KKK17199	V1	None	No data	This study
		oryzae	BAE58990	V1	None	No data	This study
			BAE62229	V3	None	Alternariol	This study
			KDE80734	V3	None	Alternariol	This study
			KDE81226	V1	None	No data	This study
			XP_001820992	V1	None	No data	This study
		parasiticus	KJK64046	V3	None	No data	This study
		rambellii	KKK15908	V1	None	No data	This study
			KKK27047	V1	None	Other anthraquinone	This study
		ruber	EYE98259	V1	None	Endocrocin	This study
		sydowii	Aspsy1_0090693	V1	None	No data	This study
			Aspsy1_0144848	V1	None	No data	This study
			Aspsy1_0151845	V2	None	No data	This study
			Aspsy1_0157033	V2	None	No data	This study
			Aspsy1_1049255	V1	None	No data	This study
		terreus	EAU31624	V1	Geodin	Characterized	(Couch and Gaucher, 2004; Nielsen et al., 2013)
			EAU37396	V1	None	No data	This study

			BAB88752	V1	None	No data	This study
		tubingensis	Asptu1_0059858	V2	None	No data	This study
			Asptu1_0123892	V2	None	No data	This study
		ustus	KIA75323	V1	None	Benzophenone	This study
			KIA75530	V2	None	Asperthecin	This study
			KIA75835	V2	None	Viridicatumtoxin	This study
		versicolor	Aspve1_0089706	V2	None	No data	This study
			Aspve1_0122449	V1	None	No data	This study
			Aspve1_0886277	V2	None	No data	This study
		wentii	Aspwe1_0034272	V1	None	No data	This study
		zonatus	Aspzo1_2112764	V1	None	No data	This study
	Capronia	*epimyces*	EXJ89638	V1	None	Other anthraquinone	This study
	Cladophialophora	*carrionii*	ETI19899	V1	None	Benzophenone	This study
		yegresii	EXJ61970	V1	None	Benzophenone	This study
	Endocarpon	*pusillum*	ERF75912	V1	None	No product	This study
	Microsporum	*gypseum*	EFR03594	V2	None	Neosartoricin	This study
	Neosartorya	*fischeri*	EAW20700	V2	Neosartoricin, fumicycline	Characterized	(Chooi et al., 2013; König et al., 2013)
			EAW24682	V1	None	Benzophenone	This study
			EAW24697	V1	None	Grisandiene	This study
			EAW25724	V1	None	No data	This study
	Penicillium	*aethiopicum*	ADI24926	V2	Viridicatumtoxin	Characterized	(Chooi et al., 2010)
			ADI24953	V3	Griseofulvin	Characterized	(Chooi et al., 2010)
		expansum	KGO43750	V1	None	Benzophenone	This study
			KGO65172	V1	None	Emodin	This study
		oxalicum	EPS34273	V1	None	Benzophenone	This study
	Talaromyces	*cellulolyticus*	GAM33809	V1	None	No data	This study
			GAM37897	V1	None	Grisandiene	This study
			GAM40075	V3	None	No data	This study
			GAM42425	V1	None	No data	This study
			GAM43179	V1	None	No data	This study
		islandicus	CRG83532	V1	None	No data	This study
			CRG86674	V3	None	No data	This study
			CRG92129	V1	None	No data	This study
		marneffei	KFX46552	V1	None	Grisandiene	This study
			KFX52365	V1	None	Other anthraquinone	This study
			ADH01670	V1	None	No data	This study
			ADH01674	V1	None	No data	This study
		stipitatus	EED18910	V1	None	Emodin	This study
			EED18976	V1	None	Benzophenone	This study
	Trichophyton	*equinum*	EGE06343	V2	None	Neosartoricin	This study
		interdigitale	KDB28089	V2	None	Neosartoricin	This study
		rubrum	EZG11077	V2	None	Neosartoricin	This study
			KDB38561	V2	None	Neosartoricin	This study
		soudanense	EZF78756	V2	None	Neosartoricin	This study
		tonsurans	EGD99348	V2	None	Neosartoricin	This study
		verrucosum	EFE44835	V2	None	Neosartoricin	This study
Lecanoromycete	*Usnea*	*longissima*	AGI60157	V1	None	No data	This study
Leotiomycete	*Blumeria*	*graminis*	CCU75801	V3	None	No data	This study
			EPQ66189	V3	None	No data	This study
	Botrytis	*cinerea*	AAR90250	V3	None	No data	This study
			EMR83380	V3	None	No data	This study
			XP_001553397	V3	None	No data	This study
	Oidiodendron	*maius*	KIM92894	V1	None	Grisandiene	This study
			KIM93459	V2	None	No data	This study
			KIM96903	V2	None	No data	This study
			KIM99919	V3	None	No data	This study
	Sclerotinia	*borealis*	ESZ98980	V1	None	Benzophenone	This study
Sordariomycete	*Claviceps*	*purpurea*	CCE31584	V1	None	Benzophenone	This study
	Colletotrichum	*graminicola*	EFQ33703	V3	None	Alternariol	This study

		sublineola	KDN60962	V3	None	No data	This study
			KDN62802	V3	None	No product	This study
	Diaporthe	*ampelina*	KKY32371	V1	None	Benzophenone	This study
			KKY34489	V1	None	Endocrocin	This study
			KKY37364	V1	None	No product	This study
	Eutypa	*lata*	EMR67234	V1	None	Benzophenone	This study
	Gaeumannomyces	*graminis*	EJT69423	V3	None	Alternariol	This study
	Grosmannia	*clavigera*	EFX02748	V1	None	Benzophenone	This study
			EFX04268	V3	None	No product	This study
	Metarhizium	*acridum*	EFY89907	V2	None	Viridicatumtoxin	This study
		album	KHN97625	V2	None	Viridicatumtoxin	This study
		anisopliae	KID64953	V2	None	Viridicatumtoxin	This study
			KID65337	V1	None	Grisandiene	This study
			KJK89454	V2	None	Viridicatumtoxin	This study
			KJK95825	V1	None	Grisandiene	This study
		brunneum	KID65403	V1	None	Grisandiene	This study
			KID75074	V2	None	Viridicatumtoxin	This study
		guizhouense	KID84413	V1	None	Grisandiene	This study
			KID87428	V2	None	Viridicatumtoxin	This study
		majus	KID93731	V1	None	Grisandiene	This study
			KID97686	V2	None	Viridicatumtoxin	This study
		robertsii	EXU97310	V2	None	Viridicatumtoxin	This study
			EXV00352	V1	None	Grisandiene	This study
	Myceliophthora	*thermophila*	AEO58356	V1	None	Benzophenone	This study
	Pestalotiopsis	*fici*	AGO59040	V1	Pestheic acid	Characterized	(Xu et al., 2014)
			ETS76950	V1	Pestheic acid	Characterized	This study
	Phaeomoniella	*chlamydospora*	KKY14863	V3	None	No data	This study
	Stachybotrys	*chartarum*	KEY72288	V1	None	Benzophenone	This study
			KFA45466	V3	None	No data	This study
			KFA54249	V1	None	Benzophenone	This study
			KFA70942	V3	None	No data	This study
			KFA71118	V1	None	Benzophenone	This study
	Thielavia	*terrestris*	AEO66245	V1	None	Grisandiene	This study
	Trichoderma	*atroviride*	EHK46042	V1	None	No data	This study
		harzianum	KKO99957	V1	None	Other anthraquinone	This study
			KKP03797	V3	None	No data	This study
		reesei	ETR97748	V3	None	Alternariol	This study
		virens	EHK20655	V3	None	Alternariol	This study

Table 1: List of the classes, genera, and species with PKSs belonging to group V, their accession numbers, their subgroup, characterized products, and predicted products based. These predictions are based on the presence or absence of homologs of the requisite genes to produce these compounds as determined by previous studies of group V SM gene clusters.

Figures

Figure 1: Diagram of the domains of different types of PKSs.

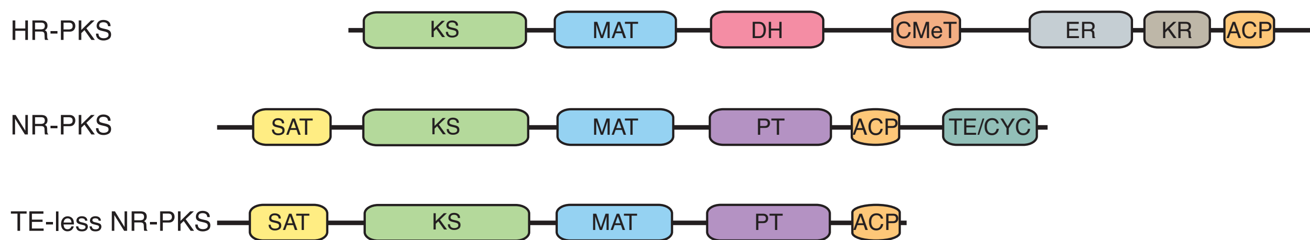


Figure 1: A diagram showing the domains of typical highly reducing (HR), non-reducing (NR), and group V NR type I fungal polyketide synthases. SAT=starter-unit:ACP transacylase, KS=β-ketoacyl synthase, MAT=malonyl-CoA:ACP transacylase, DH=dehydratase, CMeT=C-methyltransferase, ER=enoyl reductase, KR=keto reductase, PT=product template, ACP=acyl carrier protein, TE=thioesterase, CYC=Claisen cyclase.

Figure 2: Phylogenetic tree of all NR-PKSs analyzed in this study.

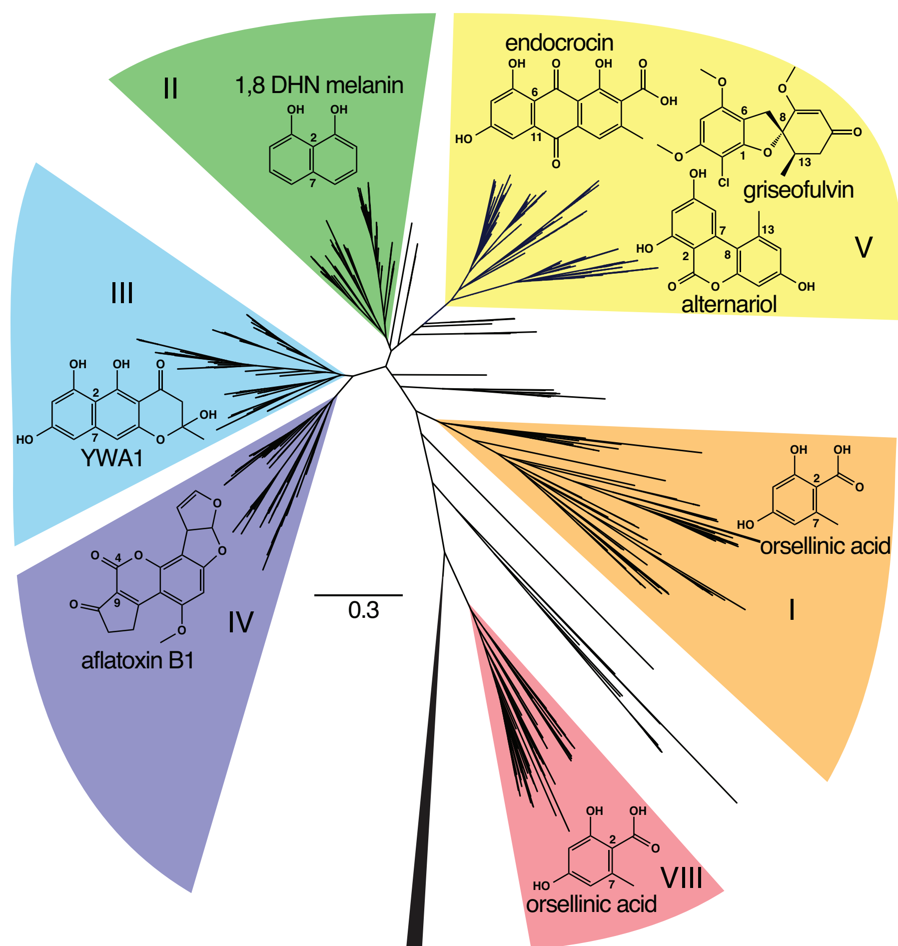


Figure 2: A maximum likelihood phylogenetic tree constructed with FastTree (Price et al., 2010) using the KS domains of 908 NR-PKSs. Clades corresponding to characterized groups I-VIII are highlighted and labeled. No members of groups VI and VII were identified by our methods. Examples of structures produced by each group are shown in their respective highlighted regions and the numbering of the carbon-carbon bonds indicates the mode of cyclization of the PKSs that produce these compounds. The outgroup, consisting of the KS domains of 70 HR- and hybrid PKS/NRPSs, as well as the human fatty acid synthase, FASN, is collapsed.

Figure 3: Phylogenetic tree of all group V PKSs analyzed in this study.

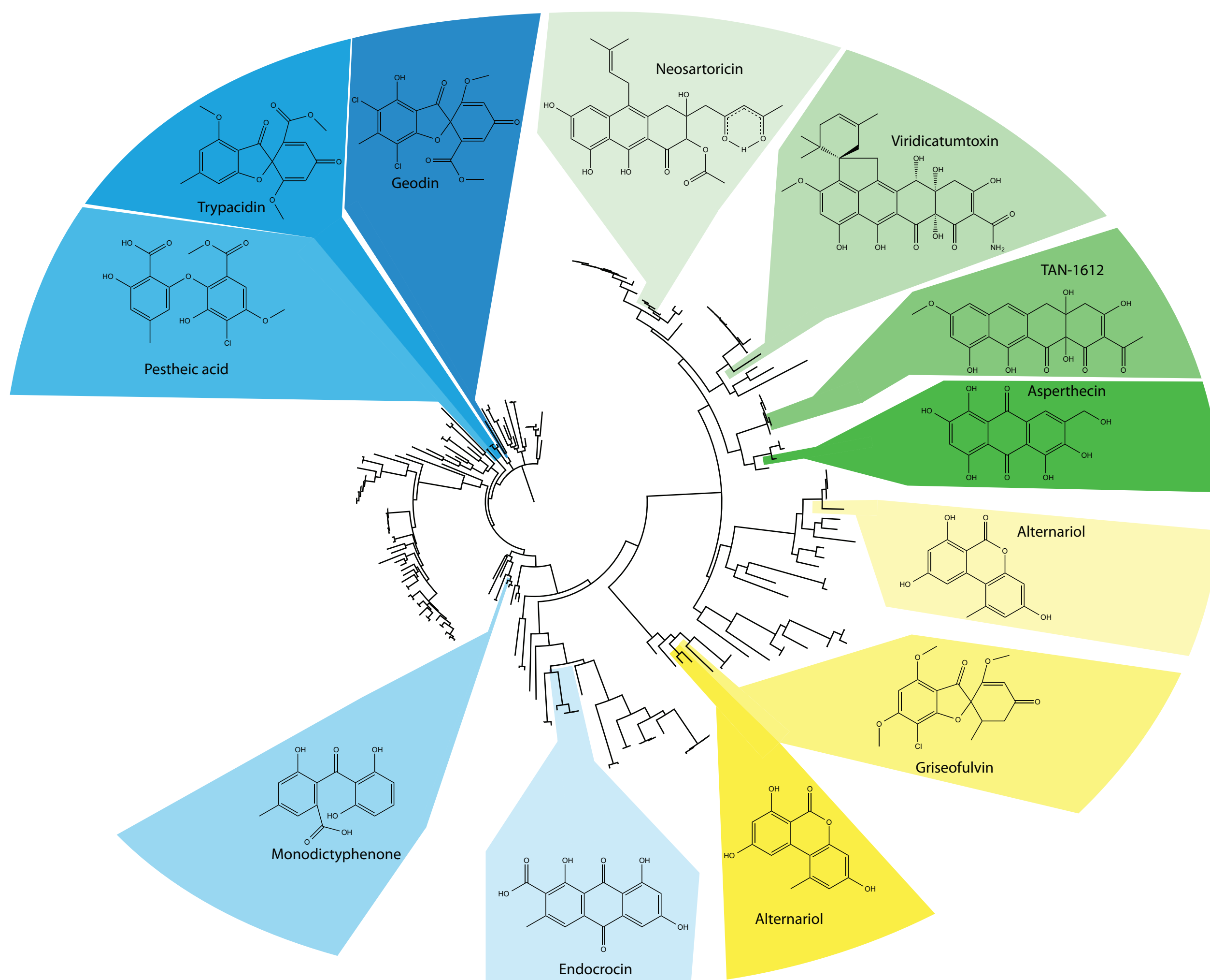


Figure 3: A maximum likelihood phylogenetic tree constructed with FastTree (Price et al., 2010) using the KS domains of a subset of 188 group V NR-PKSs extracted from the broader set used above. The leaves corresponding to characterized PKSs are highlighted in shades of blue for group V1, green for group V2, and yellow for group V3. The structures produced by these PKSs and their associated decorating enzymes are shown adjacently.

Figure 4: The relationships between group V1 clusters and the aflatoxin gene cluster.

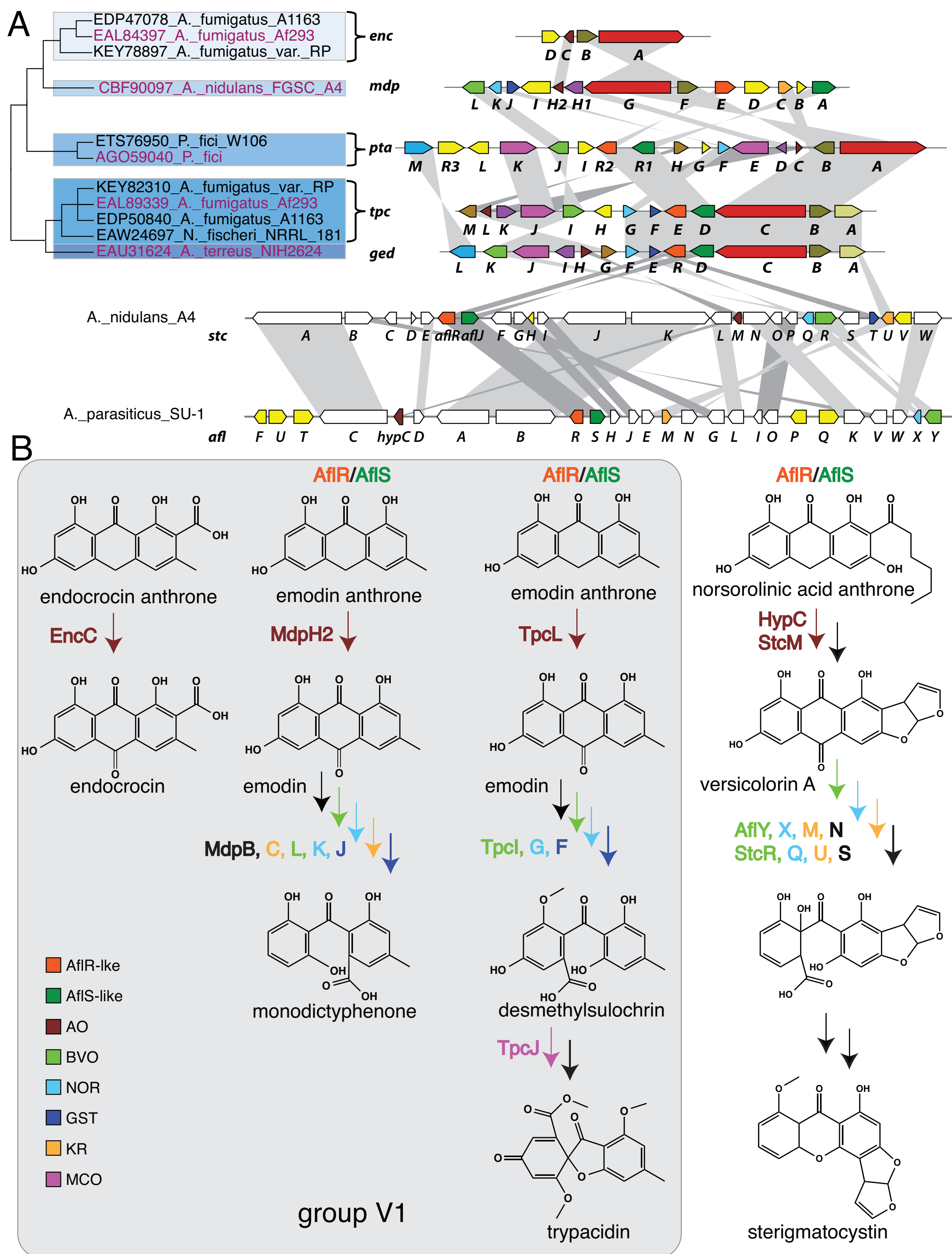


Figure 4: A) A phylogenetic tree created from the group V maximum likelihood tree (Fig. S2) showing just the relationships between the characterized group V1 PKSs. The gene cluster diagrams next to brackets depict the cluster corresponding to the PKS with its accession number highlighted in red, but all of the bracketed PKSs belong to clusters which are identical in terms of the presence and synteny of their group V-cluster homologs. Genes are represented as arrows with a color corresponding to their ortholog group and these are connected by shaded regions. Genes colored in yellow are unique among clusters shown here. Genes with no color in the *afl* and *stc* clusters do not have a homolog in the group V1 clusters shown. B) A comparison of the analogous reactions catalyzed by the enzymes encoded by homologs of *afl* cluster genes in the endocrocin, monodictyphenone, trypacidin, and aflatoxin pathways. The reactions of the trypacidin pathway are representative of the geodin and pestheic acid biosynthetic pathways. Pathways of group V1 clusters are enclosed in a grey box. The enzymes catalyzing each reaction are shown to the left of the arrows and the color of the text and arrows matches Fig. 4A. Arrows in black represent reactions not shown or reactions for which the enzymes, also labeled in black, are not homologous, except for AflN and StcS, which are homologous.

Figure 5: tpc-like clusters of interest.

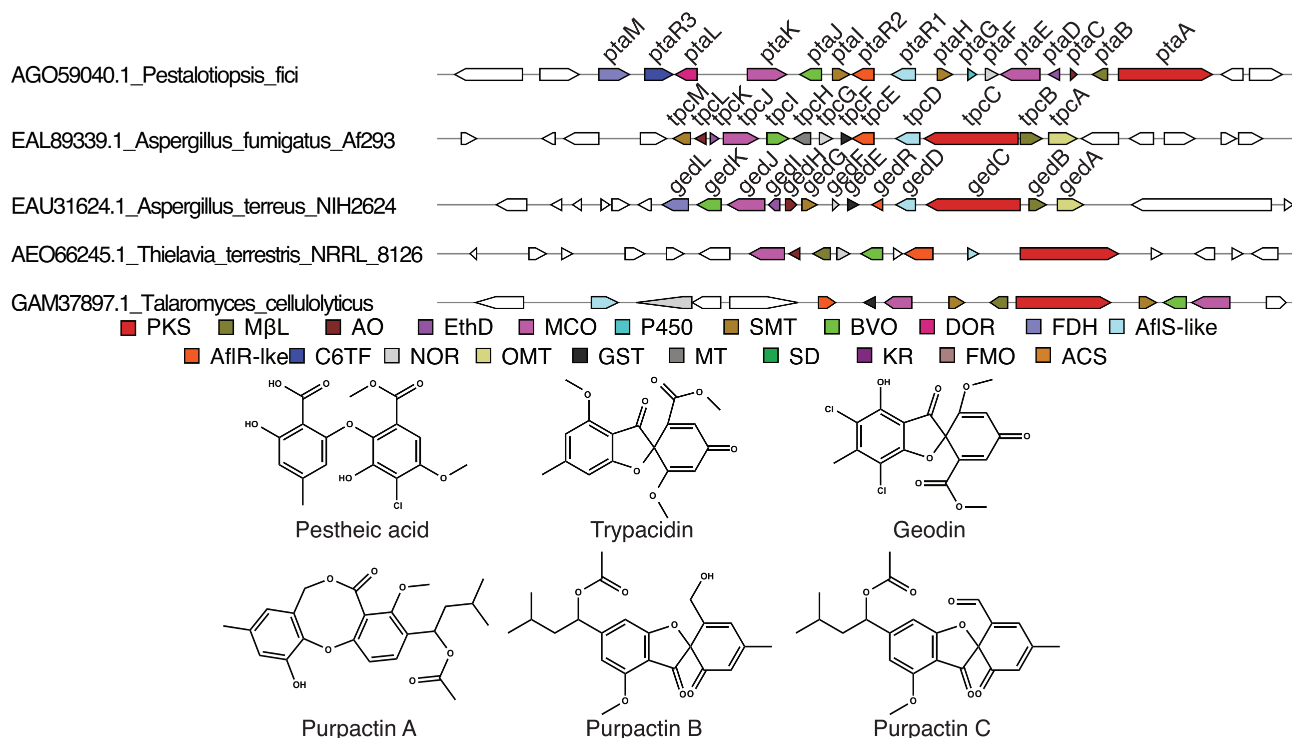


Figure 5: Examples of *tpc*-like clusters of interest. At top, the gene cluster diagrams are shown for the characterized *pta*, *tpc*, and *ged* clusters as well as two uncharacterized gene clusters from the group V1 chosen from the larger group V phylogenetic tree, Fig. S2. Genes are represented as arrows with a color corresponding to the proteins they encode which are detailed in the color key below the cluster diagrams. Genes with no color were not identified as homologous to any group V1 cluster gene. The structures of the characterized products from this clade, pestheic acid, trypacidin, and geodin, are shown below the color key. Below these are three metabolites reported to be produced by *T. cellulolyticus* (now considered synonymous with *T. pinophilus*), purpactins A-C (Yilmaz et al., 2014). PKS=Polyketide synthase, MβL=Metallo-β-lactamase-type thioesterase, AO=Anthrone oxidase, EthD=EthD domain-containing protein, a putative decarboxylase (Chiang et al., 2010b; Throckmorton et al., 2015), MCO=multicopper oxidase, P450=cytochrome P450, SMT=S-adenosylmethionine-dependent methyltransferase,

BVO=Baeyer-Villiger oxidase, DOR=Pyridine nucleotide-disulfide oxidoreductase, FDH=Flavin-dependent halogenase, AflS=Transcriptional co-regulator of the aflatoxin biosynthetic gene cluster (Ehrlich et al., 2012), AflR=Transcriptional regulator of the aflatoxin biosynthetic gene cluster, C6TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, NOR=NADH-dependent oxidoreductase, OMT=*O*-methyltransferase, GST=Glutathione *S*-transferase, MT=Methyltransferase, SD=Scytalone dehydratase, KR=Ver-1-like ketoreductase (Skory et al., 1992; Simpson, 2012), FMO=Flavin-dependent monooxygenase, ACS=Acyl-CoA synthase.

Figure 6: nsc-like clusters in Trichophyton and vrt-like clusters in Metarhizium.

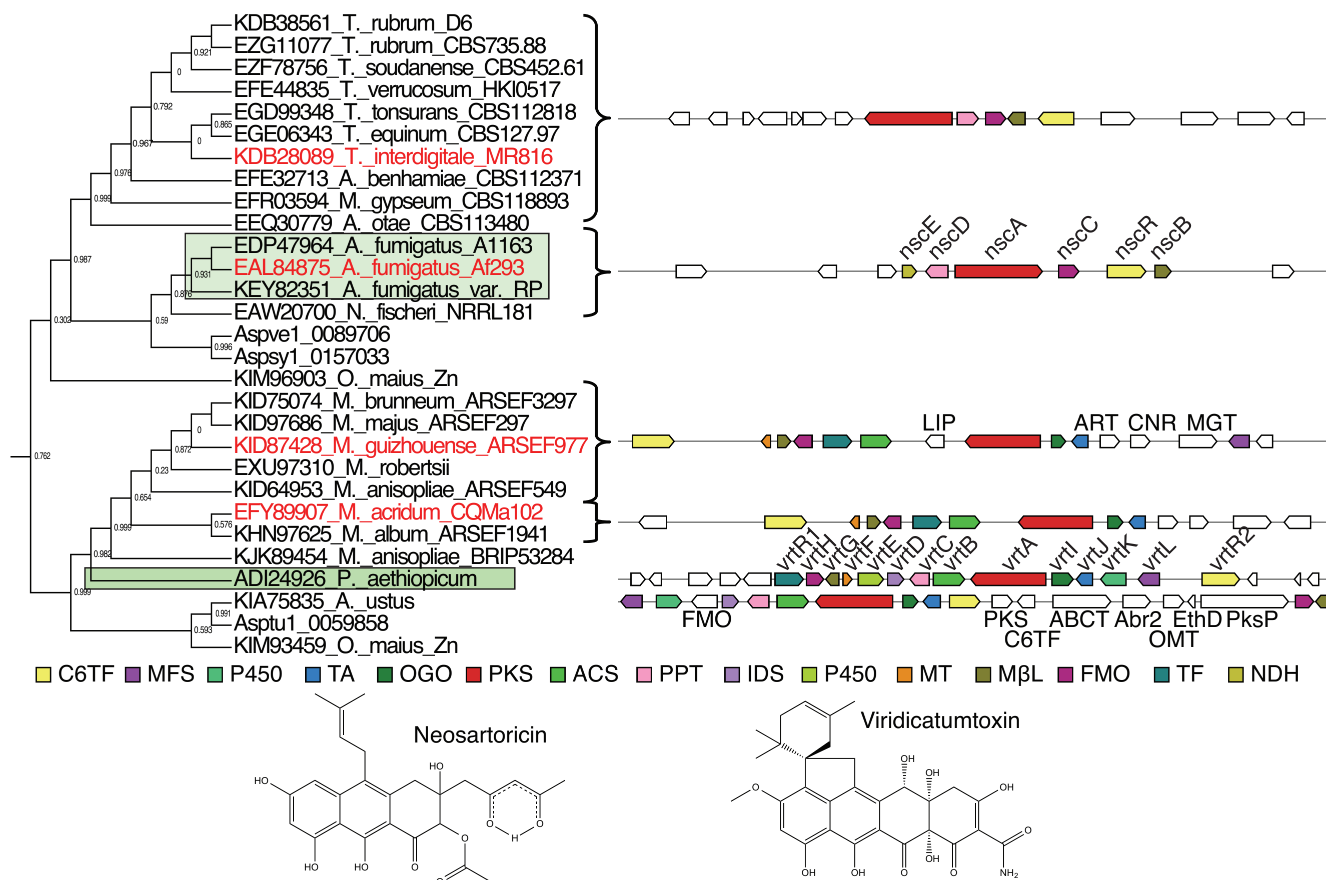


Figure 6: A clade of *nsc*-like clusters in *Trichophyton*, a clade of *vrt*-like clusters in *Metarhizium*, and other closely related clusters. An excerpt of the group V phylogenetic tree made with FastTree (Price et al., 2010), Fig. S2, containing the PKSs from the neosartoricin-producing cluster, NscA (EAL84875) (Chooi et al., 2013), and the viridicatumtoxin-producing cluster, VrtA (ADI24926) (Chooi et al., 2010), and groups of related uncharacterized PKSs, primarily in *Trichophyton* and *Metarhizium*, respectively, is shown at top left. The bootstrap values are presented next to their corresponding nodes. The green boxes indicate PKSs from the same species in which the characterized clusters were originally described. Next to the tree are the gene clusters corresponding to the PKSs that were identifiable through MultiGeneBLAST analysis. Gene cluster diagrams next to brackets depict the cluster corresponding to the PKS with

its accession number highlighted in red, but all of the bracketed PKSs belong to clusters which are identical in terms of the presence and synteny of their *nsc*- or *vrt*-cluster homologs. Genes are represented as arrows with a color corresponding to the proteins they encode which are detailed in the color key below the tree and cluster diagrams. Genes with no color were not identified as homologous to any group V2 cluster gene. The products of the characterized examples from this clade, neosartoricin and viridicatumtoxin, are shown at bottom. C6TF=GAL4-like Zn(II)$_2$Cys$_6$-domain and fungal-specific transcription factor domain-containing protein, VrtR1-like, MFS=Major Facilitator Superfamily transporter, P450=Cytochrome P450, TA=Threonine aldolase, OGO=2-oxoglutarate-Fe(II)-type oxidoreductase, PKS=Polyketide synthase, ACS=Acetoacetyl-CoA synthase, PPT=Polycyclic prenyltransferase, IDS=Isoprenyl diphosphate synthase, MT=Methyltransferase, MβL=Metallo-β-lactamase-type thioesterase, FMO=Flavin-dependent monooxygenase, TF=GAL4-like Zn(II)$_2$Cys$_6$-domain and fungal-specific transcription factor domain-containing protein, VrtR2-like, NDH=NAD-dependent dehydratase, LIP=Secretory lipase, ART=Arrestin, CNR=Copper-containing nitrite reductase, MGT=Magnesium transporter, ABCT=ABC transporter, Abr2=Conidial pigment laccase, OMT=*O*-methyltransferase, EthD=EthD domain-containing protein, putative decarboxylase, PksP=Conidial pigment polyketide synthase.

Figure 7: gsf-, pkg-, PksI-, and SNOG_15820-like clusters.

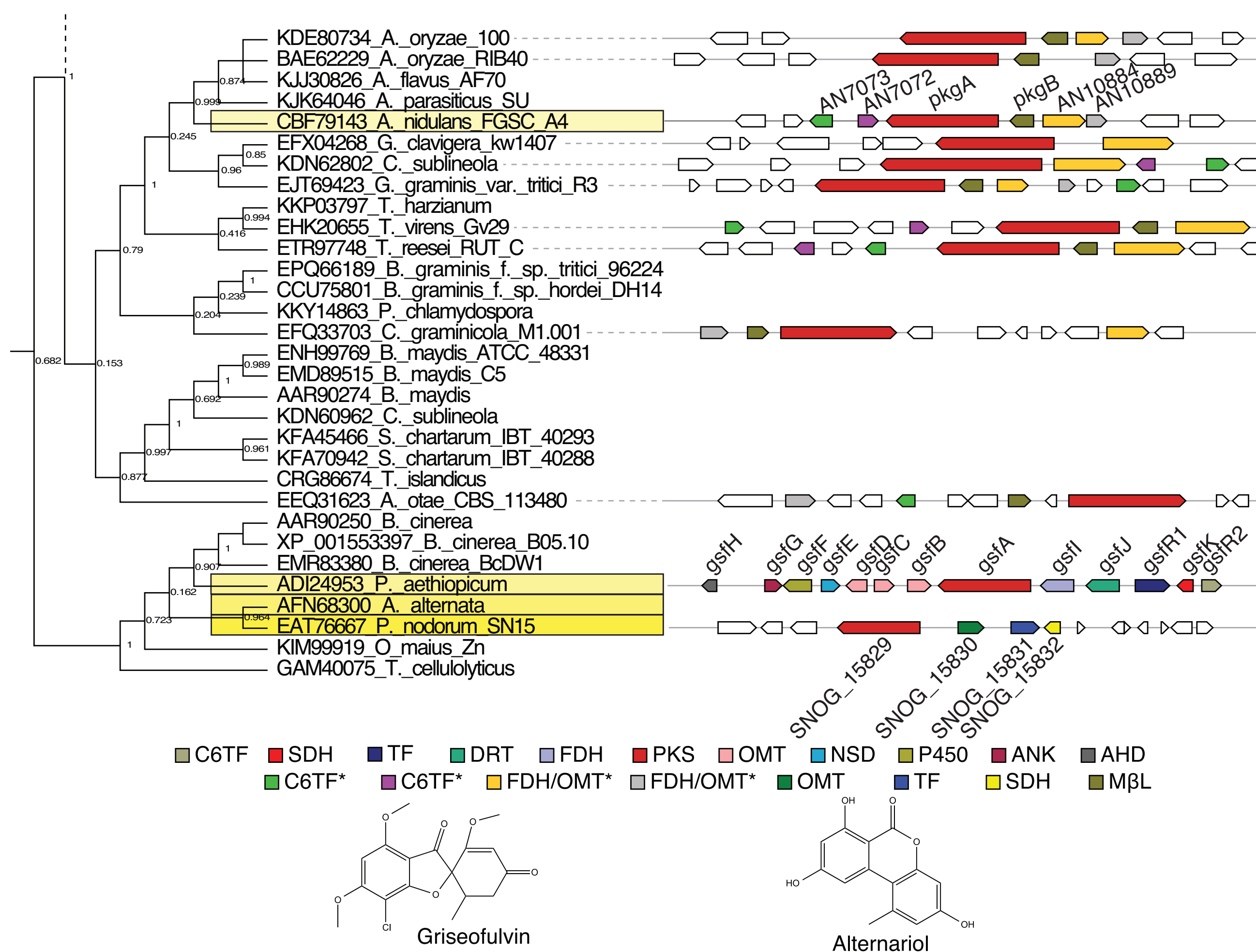


Figure 7: A clade of *gsf*-, *pkg*-, PksI-, and SNOG_15820-like clusters. An excerpt of the group V phylogenetic tree made with FastTree (Price et al., 2010), Fig. S2, containing the PKSs from the griseofulvin-producing cluster, GsfA (ADI24953) (Chooi et al., 2010), three alternariol-producing clusters (Ahuja et al., 2012; Saha et al., 2012; Chooi et al., 2015), and a group of related uncharacterized PKSs, is shown at top left. The bootstrap values are presented next to their corresponding nodes. The yellow boxes indicate PKSs from the same species in which the characterized clusters were originally described. Next to the tree are the gene clusters corresponding to the PKSs that were identifiable through MultiGeneBLAST analysis. Genes are represented as arrows with a color corresponding to the proteins they encode which are detailed

in the color key below the tree and cluster diagrams. Asterisks signify potential gene truncation due to misannotation. Genes with no color were not identified as homologous to any group V3 cluster gene. The products of the characterized examples from this clade, griseofulvin and alternariol, are shown at bottom. C6TF=GAL4-like Zn(II)$_2$Cys$_6$-domain and fungal-specific transcription factor domain-containing protein, GsfR2-like, SDH=Short chain dehydrogenase, TF=GAL4-like Zn(II)$_2$Cys$_6$-domain and fungal-specific transcription factor domain-containing protein, GsfR1-like, DRT=Drug resistance transporter, EmrB subfamily, FDH=Flavin-dependent halogenase, PKS=Polyketide synthase, OMT=*O*-methyltransferase, NSD=Nucleoside-diphosphate-sugar dehydratase, P450=Cytochrome P450, ANK=Ankyrin repeat-containing protein, AHD=YcaC-related amidohydrolase, FDH/OMT=Flavin-dependent halogenase and *O*-methyltransferase bifunctional protein, MβL=Metallo-β-lactamase-type thioesterase.

References

Adeboya, M.O., Edwards, R.L., Lassoe, T., Maitland, D.J., Shields, L., and Whalley, A.J.S. (1996) Metabolites of the higher fungi. 29. Maldoxin, maldoxone, dihydromaldoxin, isodihydromaldoxin and dechlorodihydromaldoxin. A spirocyclohexadienone, a depsidone and three diphenyl ethers: Keys in the depsidone biosynthetic pathway from a member of the fungus genus *Xylaria*. *Journal of the Chemical Society-Perkin Transactions 1*: 1419-1425.

Aggarwal, N., and Goindi, S. (2013) Preparation and *in vivo* evaluation of solid lipid nanoparticles of griseofulvin for dermal use. *J Biomed Nanotechnol* **9**: 564-576.

Ahuja, M., Chiang, Y.M., Chang, S.L., Praseuth, M.B., Entwistle, R., Sanchez, J.F. et al. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J Am Chem Soc* **134**: 8212-8221.

Allam, N.G., and Abd El-Zaher, E.H.F. (2012) Protective role of *Aspergillus fumigatus* melanin against ultraviolet (UV) irradiation and *Bjerkandera adusta* melanin as a candidate vaccine against systemic candidiasis. *African J Biotech* **11**: 6566-6577.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Andersen, M.R., Nielsen, J.B., Klitgaard, A., Petersen, L.M., Zachariasen, M., Hansen, T.J. et al. (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. *Proc Natl Acad Sci U S A* **110**: E99-107.

Awakawa, T., Kaji, T., Wakimoto, T., and Abe, I. (2012) A heptaketide naphthaldehyde produced by a polyketide synthase from *Nectria haematococca*. *Bioorg Med Chem Lett* **22**: 4338-4340.

Awakawa, T., Yokota, K., Funa, N., Doi, F., Mori, N., Watanabe, H., and Horinouchi, S. (2009) Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem Biol* **16**: 613-623.

Baker, S.E., Kroken, S., Inderbitzin, P., Asvarak, T., Li, B.Y., Shi, L. et al. (2006) Two polyketide synthase-encoding genes are required for biosynthesis of the polyketide virulence factor, T-toxin, by *Cochliobolus heterostrophus*. *Mol Plant Microbe Interact* **19**: 139-149.

Beck, J., Ripka, S., Siegner, A., Schiltz, E., and Schweizer, E. (1990) The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. *Eur J Biochem* **192**: 487-498.

Bendele, A.M., Carlton, W.W., Nelson, G.E., Peterson, R.E., and Grove, M.D. (1984) Viridicatumtoxin mycotoxicosis in mice and rats. *Toxicol Lett* **22**: 287-291.

Berthier, E., Lim, F.Y., Deng, Q., Guo, C.J., Kontoyiannis, D.P., Wang, C.C. et al. (2013) Low-volume toolbox for the discovery of immunosuppressive fungal secondary metabolites. *PLoS Pathog* **9**: e1003289.

Bok, J.W., Chiang, Y.M., Szewczyk, E., Reyes-Dominguez, Y., Davidson, A.D., Sanchez, J.F. et al. (2009) Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5**: 462-464.

Brown, D.W., and Salvo, J.J. (1994) Isolation and characterization of sexual spore pigments from *Aspergillus nidulans*. *Appl Environ Microbiol* **60**: 979-983.

Brown, D.W., Butchko, R.A., Baker, S.E., and Proctor, R.H. (2012) Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. *Fungal Biol* **116**: 318-331.

Bushley, K.E., and Turgeon, B.G. (2010) Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. *BMC Evol Biol* **10**: 26.

Cacho, R.A., Chooi, Y.H., Zhou, H., and Tang, Y. (2013) Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. *ACS Chem Biol* **8**: 2322-2330.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972-1973.

Carbone, I., Ramirez-Prado, J.H., Jakobek, J.L., and Horn, B.W. (2007) Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. *BMC Evol Biol* **7**: 111.

Cary, J.W., Ehrlich, K.C., Bland, J.M., and Montalbano, B.G. (2006) The aflatoxin biosynthesis cluster gene, *aflX*, encodes an oxidoreductase involved in conversion of versicolorin A to demethylsterigmatocystin. *Appl Environ Microbiol* **72**: 1096-1101.

Cary, J.W., Harris-Coward, P.Y., Ehrlich, K.C., Di Mavungu, J.D., Malysheva, S.V., De Saeger, S. et al. (2014) Functional characterization of a *veA*-dependent polyketide synthase gene in *Aspergillus flavus* necessary for the synthesis of asparasone, a sclerotium-specific pigment. *Fungal Genet Biol* **64**: 25-35.

Cary, J.W., Uka, V., Han, Z., Buyst, D., Harris-Coward, P.Y., Ehrlich, K.C. et al. (2015) An *Aspergillus flavus* secondary metabolic gene cluster containing a hybrid PKS-NRPS is necessary for synthesis of the 2-pyridones, leporins. *Fungal Genet Biol* **81**: 88-97.

Chiang, Y.M., Oakley, B.R., Keller, N.P., and Wang, C.C. (2010a) Unraveling polyketide synthesis in members of the genus *Aspergillus*. *Appl Microbiol Biotechnol* **86**: 1719-1736.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Keller, N., Oakley, B.R., and Wang, C.C. (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J Am Chem Soc* **131**: 2965-2970.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010b) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microbiol* **76**: 2067-2074.

Chibucos, M.C., Crabtree, J., Nagaraj, S., Chaturvedi, S., and Chaturvedi, V. (2013) Draft Genome Sequences of Human Pathogenic Fungus *Geomyces pannorum Sensu Lato* and Bat White Nose Syndrome Pathogen *Geomyces* (*Pseudogymnoascus*) *destructans*. *Genome Announc* **1**.

Chooi, Y.H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: from genes to molecules. *J Org Chem* **77**: 9933-9953.

Chooi, Y.H., Cacho, R., and Tang, Y. (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* **17**: 483-494.

Chooi, Y.H., Muria-Gonzalez, J.M., Mead, O.L., and Solomon, P.S. (2015) SnPKS19 encodes the polyketide synthase for alternariol mycotoxin biosynthesis in the wheat pathogen *Parastagonospora nodorum*. *Appl Environ Microbiol*.

Chooi, Y.H., Wang, P., Fang, J., Li, Y., Wu, K., and Tang, Y. (2012) Discovery and characterization of a group of fungal polycyclic polyketide prenyltransferases. *J Am Chem Soc* **134**: 9428-9437.

Chooi, Y.H., Fang, J., Liu, H., Filler, S.G., Wang, P., and Tang, Y. (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **15**: 780-783.

Couch, R.D., and Gaucher, G.M. (2004) Rational elimination of *Aspergillus terreus* sulochrin production. *J Biotechnol* **108**: 171-178.

Cox, R. (2014) Oxidative rearrangements during fungal biosynthesis. *Nat Prod Rep* **31**: 1405-1424.

Crawford, J.M., Dancy, B.C., Hill, E.A., Udwary, D.W., and Townsend, C.A. (2006) Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc Natl Acad Sci U S A* **103**: 16728-16733.

Crawford, J.M., Thomas, P.M., Scheerer, J.R., Vagstad, A.L., Kelleher, N.L., and Townsend, C.A. (2008) Deconstruction of iterative multidomain polyketide synthase function. *Science* **320**: 243-246.

Demuner, A.J., Almeida Barbosa, L.C., Mendes Miranda, A.C., Geraldo, G.C., da Silva, C.M., Giberti, S. et al. (2013) The Fungal Phytotoxin Alternariol 9-Methyl Ether and Some of Its Synthetic Analogues Inhibit the Photosynthetic Electron Transport Chain. *Journal of Natural Products* **76**: 2234-2245.

Duplessis, S., Cuomo, C.A., Lin, Y.C., Aerts, A., Tisserant, E., Veneault-Fourrey, C. et al. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A* **108**: 9166-9171.

Díaz-Sánchez, V., Avalos, J., and Limón, M.C. (2012) Identification and regulation of *fusA*, the polyketide synthase gene responsible for fusarin production in *Fusarium fujikuroi*. *Appl Environ Microbiol* **78**: 7258-7266.

Ehrlich, K.C., Montalbano, B., Boué, S.M., and Bhatnagar, D. (2005) An aflatoxin biosynthesis cluster gene encodes a novel oxidase required for conversion of versicolorin a to sterigmatocystin. *Appl Environ Microbiol* **71**: 8963-8965.

Ehrlich, K.C., Li, P., Scharfenstein, L., and Chang, P.K. (2010) HypC, the anthrone oxidase involved in aflatoxin biosynthesis. *Appl Environ Microbiol* **76**: 3374-3377.

Ehrlich, K.C., Mack, B.M., Wei, Q., Li, P., Roze, L.V., Dazzo, F. et al. (2012) Association with AflR in endosomes reveals new functions for AflJ in aflatoxin biosynthesis. *Toxins (Basel)* **4**: 1582-1600.

Elmore, M.H., McGary, K.L., Wisecaver, J.H., Slot, J.C., Geiser, D.M., Sink, S. et al. (2015) Clustering of two genes putatively involved in cyanate detoxification evolved recently and independently in multiple fungal lineages. *Genome Biol Evol* **7**: 789-800.

Esbelin, J., Mallea, S., Ram, A.F., and Carlin, F. (2013) Role of pigmentation in protecting *Aspergillus niger* conidiospores against pulsed light radiation. *Photochem Photobiol* **89**: 758-761.

Fedorova, N.D., Moktali, V., and Medema, M.H. (2012) Bioinformatics approaches and software for detection of secondary metabolic gene clusters. *Methods Mol Biol* **944**: 23-45.

Finkelstein, E., Amichai, B., and Grunwald, M.H. (1996) Griseofulvin and its uses. *Int J Antimicrob Agents* **6**: 189-194.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F. et al. (2015) HMMER web server: 2015 update. *Nucleic Acids Res*.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R. et al. (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**: D222-230.

Fourie, P., Schutte, T., Serfontein, S., and Swart, F. (2013) Modeling the effect of temperature and wetness on *Guignardia pseudothecium* maturation and ascospore release in citrus orchards. *Phytopathology* **103**: 281-292.

Fujii, I., Watanabe, A., Sankawa, U., and Ebizuka, Y. (2001) Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*. *Chem Biol* **8**: 189-197.

Fujii, I., Ono, Y., Tada, H., Gomi, K., Ebizuka, Y., and Sankawa, U. (1996) Cloning of the polyketide synthase gene *atX* from *Aspergillus terreus* and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. *Mol Gen Genet* **253**: 1-10.

Gaffoor, I., and Trail, F. (2006) Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. *Appl Environ Microbiol* **72**: 1793-1799.

Gallo, A., Ferrara, M., and Perrone, G. (2013) Phylogenetic study of polyketide synthases and nonribosomal peptide synthetases involved in the biosynthesis of mycotoxins. *Toxins (Basel)* **5**: 717-742.

Gauthier, T., Wang, X., Sifuentes Dos Santos, J., Fysikopoulos, A., Tadrist, S., Canlet, C. et al. (2012) Trypacidin, a spore-borne toxin from *Aspergillus fumigatus*, is cytotoxic to lung cells. *PLoS One* **7**: e29906.

Gerber, R., Lou, L., and Du, L. (2009) A PLP-dependent polyketide chain releasing mechanism in the biosynthesis of mycotoxin fumonisins in *Fusarium verticillioides*. *J Am Chem Soc* **131**: 3148-3149.

Gibson, D.M., Donzelli, B.G., Krasnoff, S.B., and Keyhani, N.O. (2014) Discovering the secondary metabolite potential encoded within entomopathogenic fungi. *Nat Prod Rep* **31**: 1287-1305.

Guenther, J.C., and Trail, F. (2005) The development and differentiation of *Gibberella zeae* (anamorph: *Fusarium graminearum*) during colonization of wheat. *Mycologia* **97**: 229-237.

Hansen, F.T., Gardiner, D.M., Lysøe, E., Fuertes, P.R., Tudzynski, B., Wiemann, P. et al. (2015) An update to polyketide synthase and non-ribosomal synthetase genes and nomenclature in *Fusarium*. *Fungal Genet Biol* **75**: 20-29.

Hashimoto, M., Nonaka, T., and Fujii, I. (2014) Fungal type III polyketide synthases. *Nat Prod Rep* **31**: 1306-1317.

Henry, K.M., and Townsend, C.A. (2005a) Synthesis and fate of o-carboxybenzophenones in the biosynthesis of aflatoxin. *J Am Chem Soc* **127**: 3300-3309.

Henry, K.M., and Townsend, C.A. (2005b) Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. *J Am Chem Soc* **127**: 3724-3733.

Huang, K.X., Fujii, I., Ebizuka, Y., Gomi, K., and Sankawa, U. (1995) Molecular cloning and heterologous expression of the gene encoding dihydrogeodin oxidase, a multicopper blue enzyme from *Aspergillus terreus*. *J Biol Chem* **270**: 21495-21502.

Hutchison, R.D., Steyn, P.S., and Van Rensburg, S.J. (1973) Viridicatumtoxin, a new mycotoxin from *Penicillium viridicatum* Westling. *Toxicol Appl Pharmacol* **24**: 507-509.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S. et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647-1649.

Keller, N.P., Kantz, N.J., and Adams, T.H. (1994) *Aspergillus nidulans verA* is required for production of the mycotoxin sterigmatocystin. *Appl Environ Microbiol* **60**: 1444-1450.

Khouja, H.R., Daghino, S., Abbà, S., Boutaraa, F., Chalot, M., Blaudez, D. et al. (2014) OmGOGAT-disruption in the ericoid mycorrhizal fungus *Oidiodendron maius* induces reorganization of the N pathway and reduces tolerance to heavy-metals. *Fungal Genet Biol* **71**: 1-8.

Kim, Y.T., Lee, Y.R., Jin, J., Han, K.H., Kim, H., Kim, J.C. et al. (2005) Two different polyketide synthase genes are required for synthesis of zearalenone in *Gibberella zeae*. *Mol Microbiol* **58**: 1102-1113.

Korman, T.P., Crawford, J.M., Labonte, J.W., Newman, A.G., Wong, J., Townsend, C.A., and Tsai, S.C. (2010) Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc Natl Acad Sci U S A* **107**: 6246-6251.

Kroken, S., Glass, N.L., Taylor, J.W., Yoder, O.C., and Turgeon, B.G. (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci U S A* **100**: 15670-15675.

König, C.C., Scherlach, K., Schroeckh, V., Horn, F., Nietzsche, S., Brakhage, A.A., and Hertweck, C. (2013) Bacterium induces cryptic meroterpenoid pathway in the pathogenic fungus *Aspergillus fumigatus*. *Chembiochem* **14**: 938-942.

Lackner, G., Misiek, M., Braesel, J., and Hoffmeister, D. (2012) Genome mining reveals the evolutionary origin and biosynthetic potential of basidiomycete polyketide synthases. *Fungal Genet Biol* **49**: 996-1003.

Li, Y., Image, I.I., Xu, W., Image, I., and Tang, Y. (2010) Classification, prediction, and verification of the regioselectivity of fungal polyketide synthase product template domains. *J Biol Chem* **285**: 22764-22773.

Li, Y., Chooi, Y.H., Sheng, Y., Valentine, J.S., and Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**: 15773-15785.

Lim, F.Y., Hou, Y., Chen, Y., Oh, J.H., Lee, I., Bugni, T.S., and Keller, N.P. (2012) Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* **78**: 4117-4125.

Liu, D. (2011) *Myceliophthora* and *Thielavia*. In *Molecular Detection of Human Fungal Pathogens* Liu, D. (ed). 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742: CRC Press, Taylor & Francis Group, pp. 445-448.

Liu, L., Zhang, Z., Shao, C.L., Wang, J.L., Bai, H., and Wang, C.Y. (2015) Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. *Sci Rep* **5**: 10463.

Liu, Q., Ying, S.H., Li, J.G., Tian, C.G., and Feng, M.G. (2013) Insight into the transcriptional regulation of Msn2 required for conidiation, multi-stress responses and virulence of two entomopathogenic fungi. *Fungal Genet Biol* **54**: 42-51.

Liu, T., Sanchez, J.F., Chiang, Y.M., Oakley, B.R., and Wang, C.C. (2014) Rational domain swaps reveal insights about chain length control by ketosynthase domains in fungal nonreducing polyketide synthases. *Org Lett* **16**: 1676-1679.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y. et al. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Research* **43**: D222-D226.

Margarido, L.a.C. (2014) Oral treatments for fungal infections of the skin of the foot. *Sao Paulo Med J* **132**: 127.

Mayorga, M.E., and Timberlake, W.E. (1992) The developmentally regulated *Aspergillus nidulans wA* gene encodes a polypeptide homologous to polyketide and fatty acid synthases. *Mol Gen Genet* **235**: 205-212.

Medema, M.H., Takano, E., and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* **30**: 1218-1223.

Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., Larsen, T.O., and Mortensen, U.H. (2013) Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. *PLoS One* **8**: e72871.

Perez-Nadales, E., Nogueira, M.F., Baldin, C., Castanheira, S., El Ghalid, M., Grund, E. et al. (2014) Fungal model systems and the elucidation of pathogenicity determinants. *Fungal Genet Biol* **70**: 42-67.

Pero, R.W., Posner, H., Blois, M., Harvan, D., and Spalding, J.W. (1973) Toxicity of metabolites produced by the "Alternaria". *Environ Health Perspect* **4**: 87-94.

Pi, B., Yu, D., Dai, F., Song, X., Zhu, C., Li, H., and Yu, Y. (2015) A genomics based discovery of secondary metabolite biosynthetic gene clusters in *Aspergillus ustus*. *PLoS One* **10**: e0116089.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.

Proctor, R.H., Desjardins, A.E., Plattner, R.D., and Hohn, T.M. (1999) A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. *Fungal Genet Biol* **27**: 100-112.

Punya, J., Swangmaneecharern, P., Pinsupa, S., Nitistaporn, P., Phonghanpot, S., Kunathigan, V. et al. (2015) Phylogeny of type I polyketide synthases (PKSs) in fungal entomopathogens and expression analysis of PKS genes in *Beauveria bassiana* BCC 2660. *Fungal Biol* **119**: 538-550.

Rambaut, A. (2007). FigTree. URL http://tree.bio.ed.ac.uk/software/figtree/

Rohlfs, M., and Churchill, A.C. (2011) Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genet Biol* **48**: 23-34.

Romero-Martinez, R., Wheeler, M., Guerrero-Plata, A., Rico, G., and Torres-Guerrero, H. (2000) Biosynthesis and functions of melanin in *Sporothrix schenckii*. *Infect Immun* **68**: 3696-3703.

Saha, D., Fetzner, R., Burkhardt, B., Podlech, J., Metzler, M., Dang, H. et al. (2012) Identification of a polyketide synthase required for alternariol (AOH) and alternariol-9-methyl ether (AME) formation in Alternaria alternata. *PLoS One* **7**: e40564.

Sanchez, J.F., Entwistle, R., Hung, J.H., Yaegashi, J., Jain, S., Chiang, Y.M. et al. (2011) Genome-based deletion analysis reveals the prenyl xanthone biosynthesis pathway in Aspergillus nidulans. *J Am Chem Soc* **133**: 4010-4017.

Sanchez, J.F., Chiang, Y.M., Szewczyk, E., Davidson, A.D., Ahuja, M., Elizabeth Oakley, C. et al. (2010) Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Mol Biosyst* **6**: 587-593.

Schardl, C.L., Young, C.A., Hesse, U., Amyotte, S.G., Andreeva, K., Calie, P.J. et al. (2013) Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. *PLoS Genet* **9**: e1003323.

Scherlach, K., Graupner, K., and Hertweck, C. (2013) Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu Rev Microbiol* **67**: 375-397.

Schindler, D., and Nowrousian, M. (2014) The polyketide synthase gene *pks4* is essential for sexual development and regulates fruiting body morphology in *Sordaria macrospora*. *Fungal Genet Biol* **68**: 48-59.

Schroeckh, V., Scherlach, K., Nützmann, H.W., Shelest, E., Schmidt-Heck, W., Schuemann, J. et al. (2009) Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *Proc Natl Acad Sci U S A* **106**: 14558-14563.

Shimizu, T., Kinoshita, H., Ishihara, S., Sakai, K., Nagai, S., and Nihira, T. (2005) Polyketide synthase gene responsible for citrinin biosynthesis in *Monascus purpureus*. *Appl Environ Microbiol* **71**: 3453-3457.

Simpson, T.J. (2012) Genetic and biosynthetic studies of the fungal prenylated xanthone shamixanthone and related metabolites in *Aspergillus* spp. revisited. *Chembiochem* **13**: 1680-1688.

Skory, C.D., Chang, P.K., Cary, J., and Linz, J.E. (1992) Isolation and characterization of a gene from *Aspergillus parasiticus* associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. *Appl Environ Microbiol* **58**: 3527-3537.

Song, Z., Cox, R.J., Lazarus, C.M., and Simpson TJ, T.J. (2004) Fusarin C biosynthesis in *Fusarium moniliforme* and *Fusarium venenatum*. *Chembiochem* **5**: 1196-1203.

Studt, L., Wiemann, P., Kleigrewe, K., Humpf, H.U., and Tudzynski, B. (2012) Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. *Appl Environ Microbiol* **78**: 4468-4480.

Szewczyk, E., Chiang, Y.M., Oakley, C.E., Davidson, A.D., Wang, C.C., and Oakley, B.R. (2008) Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microbiol* **74**: 7607-7612.

Throckmorton, K., Lim, F.Y., Kontoyiannis, D., Zheng, W., and Keller, N.P. (2015) Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*. In *Environ Microbiol*. Madison, U.o.W.-. (ed).

Trail, F., Mahanti, N., Rarick, M., Mehigh, R., Liang, S.H., Zhou, R., and Linz, J.E. (1995) Physical and transcriptional map of an aflatoxin gene cluster in *Aspergillus parasiticus* and functional disruption of a gene involved early in the aflatoxin pathway. *Appl Environ Microbiol* **61**: 2665-2673.

Tsai, H.F., Chang, Y.C., Washburn, R.G., Wheeler, M.H., and Kwon-Chung, K.J. (1998) The developmentally regulated *alb1* gene of *Aspergillus fumigatus*: its role in modulation of conidial morphology and virulence. *J Bacteriol* **180**: 3031-3038.

Umemura, M., Nagano, N., Koike, H., Kawano, J., Ishii, T., Miyamura, Y. et al. (2014) Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. *Fungal Genet Biol* **68**: 23-30.

Wang, I.K., Reeves, C., and Gaucher, G.M. (1991) Isolation and sequencing of a genomic DNA clone containing the 3' terminus of the 6-methylsalicylic acid polyketide synthetase gene of *Penicillium urticae*. *Can J Microbiol* **37**: 86-95.

Wang, Y., and Casadevall, A. (1994) Decreased susceptibility of melanized *Cryptococcus neoformans* to UV light. *Appl Environ Microbiol* **60**: 3864-3866.

Watanabe, C.M., and Townsend, C.A. (2002) Initial characterization of a type I fatty acid synthase and polyketide synthase multienzyme complex NorS in the biosynthesis of aflatoxin B(1). *Chem Biol* **9**: 981-988.

Wawrzyn, G.T., Bloch, S.E., and Schmidt-Dannert, C. (2012) Discovery and characterization of terpenoid biosynthetic pathways of fungi. *Methods Enzymol* **515**: 83-105.

Wicker, T., Oberhaensli, S., Parlange, F., Buchmann, J.P., Shatalina, M., Roffler, S. et al. (2013) The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nat Genet* **45**: 1092-1096.

Wiemann, P., Guo, C.J., Palmer, J.M., Sekonyela, R., Wang, C.C., and Keller, N.P. (2013) Prototype of an intertwined secondary-metabolite supercluster. *Proc Natl Acad Sci U S A* **110**: 17065-17070.

Wiemann, P., Willmann, A., Straeten, M., Kleigrewe, K., Beyer, M., Humpf, H.U., and Tudzynski, B. (2009) Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. *Mol Microbiol* **72**: 931-946.

Xu, X., Liu, L., Zhang, F., Wang, W., Li, J., Guo, L. et al. (2014) Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte Pestalotiopsis fici. *Chembiochem* **15**: 284-292.

Yadav, G., Gokhale, R.S., and Mohanty, D. (2009) Towards prediction of metabolic products of polyketide synthases: an *in silico* analysis. *PLoS Comput Biol* **5**: e1000351.

Yang, G., Rose, M.S., Turgeon, B.G., and Yoder, O.C. (1996) A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin. *Plant Cell* **8**: 2139-2150.

Yilmaz, N., Visagie, C.M., Houbraken, J., Frisvad, J.C., and Samson, R.A. (2014) Polyphasic taxonomy of the genus Talaromyces. *Stud Mycol* **78**: 175-341.

Yu, F., Zhu, X., and Du, L. (2005) Developing a genetic system for functional manipulations of *FUM1*, a polyketide synthase gene for the biosynthesis of fumonisins in *Fusarium verticillioides*. *FEMS Microbiol Lett* **248**: 257-264.

Yu, J., Chang, P.K., Cary, J.W., Wright, M., Bhatnagar, D., Cleveland, T.E. et al. (1995) Comparative mapping of aflatoxin pathway gene clusters in *Aspergillus parasiticus* and *Aspergillus flavus*. *Appl Environ Microbiol* **61**: 2365-2371.

Yu, J.H., and Leonard, T.J. (1995) Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. *J Bacteriol* **177**: 4792-4800.

Zhou, H., Qiao, K., Gao, Z., Meehan, M.J., Li, J.W., Zhao, X. et al. (2010) Enzymatic synthesis of resorcylic acid lactones by cooperation of fungal iterative polyketide synthases involved in hypothemycin biosynthesis. *J Am Chem Soc* **132**: 4530-4531.

Appendix

Supplemental tables

Table S1: Composite multiFASTA file used for group V1 MultiGeneBLAST.

>gi|514825998|gb|AGO59040.1| PtaA (Pestalotiopsis fici)
MSDNSGSGTSPWGSLNTPVGPPKVTLAYFSNEFPPDDLNFIVRKLFDRTSKGPFCSIDGV
LLCAIQFANLIGHYETTDHLFPFGSSIASVAGLGIGLVAAAAVSVTPSLADLPVAGAEAV
RIAFRLGVLVDGVSQNLQPRDRSTTGTPDSWAYVIPDVSPEVVQKELDEIHSREKTPIPSK
IFVSALSRTSVTISGPPARLRSLFRLSDFFRDRKFVALPVYGGLCHAGHIYEQRHVQEVVE
KSVLDETHVRYSPSVRLFSTSTGKPFLSTSVTNLFEQVVGEILTQKIQWDKVVKGVLERI
QELSATEVEVLVFRDSLPVHELVKALKSADSGLQTTTEDLLQWLHQSRERLQGPRGSLQ
SKIAIVGMSCRMPSGATDTEKFWELLEKGLDVHRKIPADRFDVETHHDPTGKRVNTSITP
YGCFIDEPGLFDAGFFNMSPREAQQTDPMQRLALVTAYEALERAGYVANRTSATNLHRI
GTFYGQASDDYREVNTAQEISTYFIPGGCRAFGPGRINYFFKFSGPSYSIDTACSSSLATIQ
AACTSLWNGDTDTVVAGGMNVLTNSDAFAGLGNGHFLSKTPNACKTWDCEADGYCR
ADGIGSIVMKRLEDAEADNDNILGVILGAGTNHSADAISITHPHAPSQAFLYRQILRDAA
LDPFDVSFVEMHGTGTQAGDSEEMQSVTEVFAPIANKRRTSKQPLHIGAVKSNVGHGE
AVAGVTALIKVLLMFQKEAIPPHAGIKNSINPGFPKDLDKRNINIPYQKTAWPRSTDRKRI
AVVNNFSAAGGNTTIAIEEGPLRQTIGHDPRTTHLIPISAKSKVSLKGNIQRLIDYLEVSPD
VSLADLSYSLTARRYHHSHRVAITTSDVAHLKKQLRSQLDSADSHKPIVAAAGPPPVAF
AFTGQGASYGTMDLELYHESKYFRDQILQLDSFAQGQGFPSFVPAIDGSFPKEHTHRPV
VTQLALLCTEIALAKYWASLGVKPDVVIGHSLGEYAALHVAGVLSASDAIFLVGQRAL
MLEKKCQAGSHKMLAVRASLAQVQEAAGELPYEVACINGQKDTVLSAAKDDIDKLAS
VLESAGYKCFSLDVAFAFHSAQTDPILDDFESVSRTGVLFQAPNLPVISPLLGKVVFNDK
TINANYVRRATRESVDFLSALEAAQKISIIDESTTWIEIGPHPVCMGFIRSAVPSIKVASPSI
RRGENNWQTLVQTLGALHLAGIPVDWNEYHRPFEQALRLLDLPTYSWNDKTYWIQYN
GDWALTKGNTFYDAEKAAKAPRVGGDLPPSPISTSTVHRVIGETFDGTAGTVDIQSDLM
QQDFHDAAYGHKMNNCGVVTSSIHADIVYTIGRYLHTKLKPGVKDIHMNISNLEVVKG
LVAQKNRDVPQLIQVSISTEDISSGTAQVTWFNVLPDGGLDEPFATATLFYGKANDWLQ
SWIPTTHLVLGRVHELERLAEQGVANRFSRNMAYGLFARNLVDYADKYRGMQSVVLH
GLEAFADVELTKEKGGTWTVPPFFIDSVAHLAGFIMNVSDAVDTANNFCVTPGWESMR
FARPLLAGARYRSYVKMIPTEEDAGVFLGDVYIFQDNKIIGQVRGIKFRRYPRLLLDRFF
SAPDAAKHGGKHAPAVKAAIPPALEKKSAVVVAQVPVVDKPPPTKENAVAAPAAKSPE
PVAAAAVNEDSITVKAMALVAAEAALDVSELEDDVQFANIGVDSLMSLVIAEKFRETLG
VTISGSLFLEYPAVGDLRAWLLEYYG

>gi|514825999|gb|AGO59041.1| PtaB (Pestalotiopsis fici)
MVGKGGYRQINKALNICAFEDYLDGQQKSLPPLNDVEQISPNVLRVLGQNPGKFTLQGT
NTYIIGTGEKRLLIDTGQGIPEWADLISSTLANSSIRLSAVLLSHWHGDHTGGVPDLLRLY
PHLSDSIYKHSPSKGQQPIEDGQVFEVEGATVRAVHAPGHSHDHMCFVIEEENAMFTGD
NVLGHGTAAVELLSTWMATLRLMQSHNCGRGYPAHGEVIPNLNAKISGELASKERRER
QVLQHLNRIRKEEQGGKGSATVQRLVVEMYGDTDQQMREQALEPFIDEVLRKLAEDEK
VAFQLRAGEKTWFAIALE

>gi|514826001|gb|AGO59043.1| PtaC (Pestalotiopsis fici)
MMGLPLMAVPMLLDTGADPVYLARQWARMYYYGVRTMPPLAITTFILYVWTIIRRRSQ
HQAWYILAVAAVVTMGMIPFTWYVLAPTNNALFRLAEGPEAASGTTAGSLEEVTELLV
RWNKLHIARSLFPLTGVVIALSDAM

>gi|514826002|gb|AGO59044.1| PtaD (Pestalotiopsis fici)
MAEMSKNPASTLKPSRYLCLTICGYRKPGMSEEDYRYHMTQVSAPMTKDLMAKYGVK
RWTMNYMIRRWPSCPTLTASAKLCLRTLEDYKRMKQDPWYKEHLIGDHEKFADTKRS
KMTIGWIEEWISDGKPVDGLEFKS

>gi|514826000|gb|AGO59042.1| PtaE (Pestalotiopsis fici)
MVIRTPFSIRLYESRPGMMFAMFQSILFLAFYGRPVFGSAAARDYACVNTAESRDCWKD
GFNIETDYYGKEEAPEGKLVEESTVKNNFTENYNGTAVHWHGIRQKETNWLDGVPGVT
QCPITPGDSQVYEFRVTQYGTSWYHSHYSLQLMELMLMQSQGPIVIHGPSSANWDVDL
GPWLLSDWYHDDAFALDHVGITTNRAAIPKSSLINGKGYYECDPTNDAKCTGTRDYYE
VVLKQGTKYKFGIINTSTILTYTFWIDGHNFTIIAIDFVPIEPLTVDTLNVGIGQRYEIIIETN
PDFDDDSSFWMHAQYCFINQTDIVDDKVGIVRYESAGSSDPPYINKSDYHLNFGCADPK
PESLVPILKQQVGAQANPLAAEDYFRVGLGNFTWPDATNSTGSVFLWFLQKLPLYVNW
SEPSVKKLTIDETADFPPNSRPIELDYETGQWVYFVIESDWDPAGAVDQYGQEIRVEPSV
HPFHLHGHDFLILAQGLGKFTSDIQPNLDNPPRRDTVDVEPLGYVWIAFQIDNPGAWLF
HCHIAFHSSDGIAIQFLEQPSKLKPIMEEAGVLGDFADRCNKWDDWYQAVNIPHNATQA
DSGV

>gi|514825993|gb|AGO59035.1| PtaF (Pestalotiopsis fici)
MSRYAILGSTGNCGTALIENVLDSSMTEVHAFCRNQEKLERLVPRVISDARVKVFVGGI
GDTETLAACLHGCNAVFLCITTNDNVPGCRVAQDTALGVVKVLERSRADGFLPMPKLV
LLSSATIDDVLSRNTPWVLRSILLKSASHVYEDLRKTEILLRAEQDWLTTIFIKPGALSVDI
QRGHALSLTDEDSPVSYLDLAAAMIEAVNDPQGRYDMRNVGVVNTHGRANFPSGTPLC
IAVGLLSHFAPFLHPYLPSGTGPR

>gi|514826003|gb|AGO59045.1| PtaG (Pestalotiopsis fici)
MASTGSAQKETLNKFISGWKNANAEEMLAVASDDYTQQTLPFSLGHDVRPKQVAEVM
LPKLYSILENYELKIHQVVHDVENQKAAVYAISKADTPFGFPWLNEFSAFITFNNAGDKV
VNVQEMVDTEFFQKFFPAYQSFLSQNK

>gi|514825994|gb|AGO59036.1| PtaH (Pestalotiopsis fici)
MSTNDEVFAKDNEFWKTYLRGRAQPPESFFERIFRYHEDHGGHFGTVHDCGAGNGPYS
QKLRSRFKHVIVSDVAPGNVELAKERLGNDGFSFRVARVEDFDDIPTGSVDLVFATNVM
HWVEPSRGAKAIVSQLKSGGTFIAAGFGPARFEDQKVQDIWTRISQSGGRRLIMKADDP
TKILKVAVRSSRYYDVAPTDTSLFVPGTQRIHLNMNNGGLTDIVYPEDYVAAAEPSYTG
PQDDEIFESEDGWSFETDLEGVKDHFATFPFSKEDPEVFAELWAELEKYVADGRPIRGC
WPAKIILATRV

>gi|514826008|gb|AGO59050.1| PtaR1 (Pestalotiopsis fici)
MSSVPIDSVQAAQQVVSTKIKDKSWIMPTDPAEFLQQIAYQSQLLACLHWLGEFQILAC
VPLSGSVPIKDVADLAGVPVSQLAHVIRFMATAGFMKEPRRGEVAHTPQSAAFVTDPSF

LDAGIFLAQVSARSARKMAQNSAISTLMGGDGANNGSDFDNGELLKSTSESPRVQREVT
AYLHYVVNEVSDTANLLAQLDWRKLGSSSVVEIRADRIYPASLVLTELHSTPRFTVQTF
QEDSVEQGTAVTTTATTSSSFISKSSEEPSPKRIKPGITYQKRSLGSPQVVTDATMYVMRL
ERPHSTSVQRSILDEGQIVSELRAHLGVLKHNSNATLILIGPLLPEAGAIDAKAEMVVRFR
DLSLNQLTSEREMEVGELVDIIGDVQDESGCLVVVNKLYSRTSSTVALEVRYELYNYRK
G

>gi|514825995|gb|AGO59037.1| PtaR2 (Pestalotiopsis fici)
MEAADPNNNLTITSPSTLLSNPTQPPAQPLKLRDSCHACASSKVKCHKEKPTCSRCRKRG
ITCEYFAHRRPGRKQENRAKDTTNHVERQENTTAVEMLDLNWPAPDFSTQTSIANDNL
DVFHDIFVPPDQLNNGLTDFTIDFDDFDIQSDPAEIASLPDTSSLESMFVTSPTAPTDTITP
NVITPNVGLSVLEGLPDTTHHTQAINLASYIQTPPTEKTPDSRIKHLEENKDPCMTRALSF
LTQLSESTSRICKTSETGCSGTNKKSLPESLDGIIAENRRLLEAMSNILQCRCSEDDDLLCI
QAIVASKILNLYASAIEIKPSPARVGSGVSTHTTAGQYEPQVEQQLSTRTHPQLASGRDPI
RMAAQSVLGELHRVQRLLSQMLQKSKDNETMRRKGSENGLRAVADKVPLTSGVSFGSI
EADLRYKLGKLSIEIITLLRGA

>gi|514825996|gb|AGO59038.1| PtaI (Pestalotiopsis fici)
MPSPTNPAKVETPFTTAALGNGVDFWKAYVENRPHPSDSFFELISEYHHSHGDSAAQSAI
AHDVGTGPGNIAEKLLRHFDHVVGSDVNEQALAAAPALLPADSIKRMTFVKSSAEDLA
SANIPESVGKGQTDLILVSECIPLLDISKAFAAFRALLRPGGTLAIYFYSRPIFTGDNEAEL
NQLYDRIATRVCQFLLPFKGTPGFPIHYRAAEAMSSGLDSIPFDPEAWQDVVRYKWNAD
VPLTFNSKEGYDFEVEPVDRRDHSTEITKEITDRDFWAEEWDIGRVASFLDSVFPNYRNK
AGDKFEEVQSLFTELETALGGPKATRKVSFPVVLLLATRK

>gi|514825997|gb|AGO59039.1| PtaJ (Pestalotiopsis fici)
MAASTAAQVQLSEEALGLARIFENPKGSLEAASKLLQKNHDEFHVFWRDVGGHNHIPH
SVLSILALGGGPAELQRAWDDGVAIQRPTPPLDEDVVKKLENPAEFRARIGSIPNYTNFL
HFFRNQMDKKGWQAVVSEYAFSRTPLAETIFAQLFEGAYHPFIHIGFGIEFNLPSIIAEGL
AQAATHDSAGIEGFFLEAERQAAQSKGPGKSLVQLLDEVRTTEKIKTAARLPDGPVRVR
DGVIGRAGAEIAALASQFRVPADQLSRGAAESINISAYTAGAAQRAGKARKIDFFHMHN
TTSSLFLTVFLNQPWISTEDKVRIVEWKGRLDLVWYAACSAPDLNVDHVIGYKPAQSAG
WGWKELYEAINVAHDDGHLAKIVRALKNGEEVSRPFESGEGAEAFPIKGDSWLKLAQM
SYDTTLDLPDDDKWIWGAGFLPLWNKVPSL

>gi|514826010|gb|AGO59052.1| PtaK (Pestalotiopsis fici)
MKSLFLTGLLSALTWASEFDYLYTTEPVTPLPQGYPWGSKTANDSHPEGKLVTVHNDLE
EGTALHWHAFLQKETPWQDGVPGITQCPIAPGACFTYTFVADSYGTSWYHSHYSAQYA
DGILGPIIVHGHPTVPYDIVLEHLFDEDFAVVVSKSFADSRNATNIRQTKPANNNLINGRN
SWNCTLKDLGDDTPCQSNAPLSEFRLTPGKKHRLRILNVGGSAIQKFSLDGHKLQVIAH
DFVPVLPYEVEFLTLGVGQRADVIVEALANGTGTYTMRATIPPAPCANSVDHDATALVH
YGNTTSTFSNSSSEAWPSFIEALGVCDGLPTEEITPWYAIPAPEAPATTQIINVTLAQNETG
QYLFYMDNSSFRVNYNHPVLLLSNLGNNSYPDDPEWNVYNFGSNNSIRIVMYNNAIRT
HPIHLHGHNFFVEAVGLGEWDGHVDHPENPVRRDTAMLPQGGYMVISFNADNPGAWP
LHCHVAWHVSSGFYVTVLERPDEIAEYKIPSVVGQTCRDWWGYTNHTIVNQIDSGL

>gi|514826005|gb|AGO59047.1| PtaL (Pestalotiopsis fici)
MKHIVIIGGGFAGVSTAHRFLKNVGKSTTAPYKVTLVSRDSHFFWNIAAPRGIIPGQIPEE
KLFQPIAEGFSQYGPDKFEFVLGTATDLDVGGKTLVVDVDGKATRISYDYLIIGSGSRTKI
PGPFKSDGSTDGVKQTIHDFQERVKAAKTIVVVGAGPTGVETAGELAFEYGTSKKIILIS
GGPTVLENRPASVTKTALKQLETLNVDVRVNTKAKDPVTLPDGKKELTLSGGEKLVVD
LYIPTFGVLPNSSFVPSQYLDSNGFVQVDQYFQVKGAEGVFAIGDVSDSEAPQFWFVEK
QSVHIAKNLILSLSGKAPTPYKASATGMMGLQIGKNSGTGHFGNFKLPGFLVKTIRKTLF
VENLPKTVDGSML

>gi|514826006|gb|AGO59048.1| PtaR3 (Pestalotiopsis fici)
MATPKRMWPTDLKHIVRYTSRRTSSRYRVPPGQKANPKLAPKVHAAASTPSTNTSHSTE
TTPPSDNGFYDTAESPLSPNGAQHLTIATIPAEDSILLMHFLDKVFPLQYPMYRPDILEGG
RGWLLALLLQTKSLYHAALALSSYHRRMLVFERISEQCRATAAVQQEKHLETCLNEVR
QAMVILDQRTRQRKSYDGMGTVTSIVQLVFFELFAGQDHAWRTHLNAAIDVYDQNCR
DKLEHLDLSEASKTILRNDQRLAVDGALVTQEVTTFRFMGGSIIWLDILSSLAAGSVPRL
LSYHQGVLDAASQVKLENIMGCKNWLMCQIGRIAALQGHRRQDGWTSQRHGIKLHSIA
ADIKSEIESGMAREALESLNIQTSDSCGINNSSTNSVTLTTRMFAFMAIIHLHLVTHDFER
LDNLRETIADAIRLLQSQVPCSMIPVIVAPLFIIGCVAAQGDEQSLFRASLASDTSQHRLY
RHRKDVLSALEEIWSKRQTSTDYTWNDVLDMGQHKYLLFL

>gi|514826004|gb|AGO59046.1| PtaM (Pestalotiopsis fici)
MSVPAQTSVLIVGGGPAGSYAATVLAREGVDVVLLEAEKFPRYHIGESMLASIRFFLRFV
ELEEEFDRHGFEKKYGATFKITEKNPAYTDFAASLGEGGYSWNVVRSESDEIIFRYAGKC
GAKTFDGTKVESLTFEPYPHEGFDESVHLANPGRPVSANWSRKDGSSGVIKFDYIIDGSG
RNGLISTKYLKNRSFNQGLKNIANWTYWKGAKRFNVGEKNENSPLFEALKDGSGWVW
AIPLHNDTISVGVVARQDAFFEKKKESGLSGEAFYKEYLKLAPQIKNELLRDATIVSDIK
QATDWSYSASAYAGPNFRLIGDAGCFVDPYFSSGCHLAMTSALSASVSIQAVRRGQCDE
LTGAKWHTTKVAEGYTRFLLLVMTVQRQLRMKDKNIISTDEEEGFDMAFKKIQPVIQG
VADTRTEDEQTQRRAAEAVDFSLESFEITPEKQAAVISKIERSQAEPELLEKLTPEEVHIL
GNIVNRTFEREKDELNLTHFTGDMIDGYSAKLEHGNIGLYKREKALLNGTASRAAAVLK
SIHQVA

>Afu4g14580 COORDS:Chr4_A_fumigatus_Af293:3847775-3849411W, translated using codon table 1 (482 amino acids)
MERQPKSLCDATQLLETANIISDTVQTIIAEWSAEAKAPQGSGKQNAPMLPSRELFDAQR
TILAAVGKLTELVSDPSARILEVATQFQESRSLYIAAERRIPDLLAAGDEGGVHIDQISQK
AKIEPRKLARILRYLCSIGIFKQTGPDTFANNRISAALVSNEPLRAYVQLVNSEGFTASDR
LPHTLLHPDTGPSYDVAKTAWQNAVCTKKTRWEWLEERVAPEQLLESGGHYPGIPSLV
MGLPPREDDGLVARPELEIMGLSMVGGGRVFGTAHVYDFPWASLGDALVVDVGGGVG
GFPLQLSKVYPQLRFIVQDRGPVVKQGLEKVWPRENPEALHQGRVQFVEHSFFDTNPTE
GADIYFLRYVLHDWSDDYCVRILAAIRSSMAAHSRLLICDQVMNTTIGDPDLDSAPSPLP
ANYGYHTRFSHSRDITMMSCINGIERTPAEFKGLLQAAGLKLKKIWDCRSQVSLIEAVLP
EMNGFR*

>Afu4g14530|gstC COORDS:Chr4_A_fumigatus_Af293:3835440-3836117W, translated using codon table 1 (225 amino acids)

MPDIQPITVYGKGGPNPPRVAIILAELDLPHKVIEVPLSKVKEPDYVAINPNGRIPAIYDPN
TDLTLWESGAIVEYLVSHYDPDHRISFPAGSNLAALATQWLFFQASGQGPYYGQASWF
KKFHHEKVPSAIERYVKEINRVTGVLEGHLSRQKVAADGDGPWLVGGKCSFADLAWIP
WQVIVTAIIQPEDGYTVEDYPHVKNWLDRMMARPGVQKGMADIFPST*

>Afu4g14510 COORDS:Chr4_A_fumigatus_Af293:3833644-3832621C, translated using codon table 1 (282 amino acids)
MLEKVFHEKSFADQYTYGAKISELYAETLITESGIAKSHQRPLIILDNACGTGSISSTLQRT
LDERNKRSLKLTCGDLSEGMVDYTKQRMQAEGWNNAEAKIVNAQDTGLPSDHYTHVY
TAFAFNMFPDYKAALRECLRILQPGGTLATSTWKTANWCTIMKPVIATMPGQLSYPTM
DEINTMLNKGWDRESDVRAEFEQAGFDHVNITTVEKQCLLPVQEFGEACKILLPYILSKF
WTQEQRDQYEADVPSYLMRYLEREYGKDGLAPMKGVAIIASGRKP*

>ATEG_08460 (conserved hypothetical protein)
MLASMRFFLRFIDLEEQFDAYGFQKKYGATFKINSKREAYTDFSASLGPGGYAWNVIRS
EADDLIFRYAGEQGAHIFDGTKVDDIEFLSYDGADGANFTPAAFLVNPGRPVAATWSRK
DGTRGRIKFDYLIDASGRAGIISTKYLKNRTVNEGLRNIANWSYWKGAKVYGEGSDQQ
GSPFFEALTDGSGWCWAIPLHNGTLSVGVVMRQDLFFGKKKAAGSPGSLEMYKLCLQS
VPGISGLLEDAEIVSDVKMASDWSYSASAYAGPHFRVAGDAGCFIDPYFSSGVHLALVG
GLSAATTIQAVRRGETSEFSAAKWHSSKVTEGYTRFLLVVMAVLRQLRKQNAAVITDD
KEEGFDTAFGLIQPVIQGQADTGESEQQRMVAGVQFSLERFGQATPEAQRALLDKVQFA
GQNAEELEKLTADELAVLHNIIGRQLKMTKVEKNLDNFTRDVIDGWAPRVERGKLGLQ
RADTSIMTAEMKDLFQLNRSLDSTKAGIQLPA*

>AN10049|mdpB COORDS:ChrVIII_A_nidulans_FGSC_A4:4457666-4458357W, translated using codon table 1 (214 amino acids)
MTLQPTFEGRTPEQCLNVHTDSHPDITGCQAALFEWAESYDSKDWDRLKQCIAPFLRID
YRAFLDKLWEKMPAEEFVAMVSHPHFLGNPLLKTQHFVGTMKWEKVDDSKIVGYHQ
MRVAHQKHLDSQMKEVVAKGHGHGSATVTYRKINGEWKFAGIEPNIRWTEFGGEGIFG
PPEKEENGVAADDQVMNSNGSSEVEERNGHVVNKAVEVRSV*

>AN0146|mdpC COORDS:ChrVIII_A_nidulans_FGSC_A4:4456446-4457370W, translated using codon table 1 (289 amino acids)
MSPAIQRLSLVSSHLNSNVSALPKMTATTHAPYRLEGKVALVTGSGRGIGAAMALELGR
LGAKVVVNYANSREPAEKLVQEIKELGTDAIALQANIRNVSEIVRVMDDAVAHFGGLDI
VCSNAGVVSFGHLGEVTEEEFDRVFSLNTRAQFFVAREAYRHLNTHGRIILMSSNTAKE
FSVPRHSVYSGSKGAIESFVRVMAKDCGDKQITVNAVAPGGTVTDMFYDVAQHYIPNG
EKHSAEELQKMAATVSPLKRNGFPVDIAKVVGFLASREAEWVNGKIITVDGGAA*

>AN0147|mdpD COORDS:ChrVIII_A_nidulans_FGSC_A4:4454244-4455887W, translated using codon table 1 (521 amino acids)
MTHFPVNIASDKQEFDPERWAKTPTTESSVNGENGTAPTSGLPSRHPSTGISVLIVGAGM
GGLMTALECWRKGHDVAGILERSEGPVYSGDIIVMQPSAVSIIRHWPDMLHDMKAEQV
HAVVSYETHDGRHIYGPTVPSFNDPEHLETRKGPFVAPAQVRRKFYRMLLRQVARCGL
RVEYGKTVKSYFEDEKDGKGGVIIATTGEAEVRVADIVVAADGLKSPSEILIAGQHVPPR
SSGLSIYRTAFPKDLAMQNELVRKRWSDSPPIWEYWLGPGMYLGVFVGDDIISFGFTPR

DDIVEGTATESWEPDTDPETVAQAMLSGAGDWDPAVLALIRSAPKGAIVHWPLLWRDL
RREWTSPAGRVVQVGDSAHSFIPTSGNGGSQALEDAITLATCLQLAGSSQRAYLGTKIY
NLLRYERVSCAQKMSFVNSQLKTGTDWDAIWKDPAKIRTRFPKWIFQHDPEAYAYEKF
GEAFAHLLDGREFVNTNYPPGHEFRAWTVEEVWRNIADGKRVEDLLDGDWS*

>AN10035|mdpI COORDS:ChrVIII_A_nidulans_FGSC_A4:4441618-4439762C, translated using codon table 1 (592 amino acids)
MSVSRSCFRPFLPAEIDGGHLPVDPSVFTHIERGLHQNPQGFAIQSTHQQPCHFSALVQTG
SGTENGGAPNYDAVEREPGTCLAWTYTQLHHAALRIAAGLLARNAQPSTRMLLLIPNG
AEFCLLLWTAVVLRVTIVCLDEELLNVEQHDELRRMLKTINPRVIVVQDVKGADVIDVA
LRNLPLDPDILKITLSELAGSQPDSAWRSLLSLSLTPALSASETESLLSSARWDSSNAART
YSILYTSGTSGVPKGCPLHISGMSYVLQSQSWLVNAENCTRALQQAHPCRGIAIAQTLQT
WREGGTVVMTGNGFNAGDLVHAVKRHAVSFVVLTPAMVHPVADELKGRNGAADSVR
TVQIGGDAVTRGALEICTRLFPKARVVVNHGMTEGGGAFVWPFNRPRDIPFYGEMSPV
GSVARGAAVRIRGANATVARGELGELHVSCPSIIPGYLGGVSAQSFHDEDGRRWFKTGD
VGLMDKQGVVFILGRMKDMINGKVMPAPIESCLEKYTSVQTCVVNAGGPFAVLARYT
GKKEAQIRRHVVRALGKSNALNGVIYLHQLGLERFPVNGTHKIARGDVEGAMLAYLQT
EPTSR*

Table S1: The composite multifasta file used for the group V1 MultiGeneBLAST search. This list contains one member of each group of orthologous genes in all group V1 clusters.

Table S2: Composite multiFASTA file used for group V2 MultiGeneBLAST.

>gi|297242539|gb|ADI24939.1| VrtR2 (Penicillium aethiopicum)
MPSLSSKTSTMQRSCRPQMSACPNQQQKDRPVPQLSCVLCRDRKLKCDKLDPCSNCTSS
GVACTPIYRPRLPRGRHARTVQTKASTPPDTRRRGSSNESTTAPAPDDGGLGTHIDQLDN
LVQDREVSKLGLSGEGNGLQELISLVSDETMPATAWSTHCFGTISRILSSRIRRLESLVQE
TARIQTPKRARKPMTPVVVQWYSAPLAGCNWNRMVVQTPQGLEVQQFPAPPTSYSAR
RSPELSGNDIWADLMDHDMHDPPQYNALELPPDLTNEGGVDNMGSSGRDDPINNGFNA
LRLLGINNSLSPSFISLPRDRLSASKLCQVYLQNVDPIIKILHRPSLSRWMVDGAPTYLGSS
EDDYAVKALESAVCYTAANTMTEHQCQAAFQKTKSSIMAVRRKMCEDALENAGLLTT
RDMTVLQAFILYLVTPTDLSKIGRRSEDKDTAVWALVALAIRLIKAMGLNQEPSEGARK
GESFFQQQLRLRLWLTACLIDLQASFAQATDPLITHRDAACAVPYVANINDSFDVDTA
HPVASHEELTDTTFALVTYRVQVAGRLFNFGPGCSTAAERHKLAQEVQQQVFTLLHYC
DPESSSYAWFTWHSTQSIIFAVRLSELLPFRCGQPGGHVPPPSPRAEGDTTLLWRALQNL
EKAQLIRADPRGDGFRWYITTPWLALSTAISECNSCTDVALVCRAWPVIEISYRQHEELQ
ISDECQLPQGPLVHLMNKTREKLAPLLQEGGARLSDSQTVDRASADSLQPPVPVGSIPID
PLLTNGSLGADTAMSEASSIGSLPPFEQQCWKQMTMPTDGAPVRDGVVFTSELYNPLQS
DFLNSHG
>gi|297242538|gb|ADI24938.1| VrtL (Penicillium aethiopicum)
MSKLSDNHSSASEGEKEAGDLESGPTAISSEPSFDDADRDPNLITWDGPKDPENPKNWP
KGLRWKNTWTVSLFVFISPVSSSMIAPAMSDLAKSLGMHAEIEIYLSLSIFILAYSIGPIFF
GPASELYGRVRLLQISNVWYLAWNLGCGFATTKGQLFAFRFLAGIGGSAPLAIGGGAIS
DMWTAEERGKAMGVYTLGPLLGPVVGPIAGGFIAEYSTWRWVFWSTSAAALAVQVVG
FFWLQECHPGTLLRKRRDRLAKETGNENLHTAEKVETLGYKLLHAFERPVKMFTTQPIV
FCMAIYMAYLFGISYLMFATFPEIWTVVYHESPGIGGLNYLSIAIGSFIGLFFNLKLVDRIY
RSLKARNNGVGKPEYRMPSLAVGSVISTIGLFWYGWSIGNTHWIMPNIGALIFAMGTISC
LQGMQTYIVDSYQTYAASAMAACAVLRSLCGFGFPLFAPYMYNSLGYGWGTSLLAFIT
MVVGWGAPFAFWHFGPRLRAMSKYASG
>gi|297242537|gb|ADI24937.1| VrtK (Penicillium aethiopicum)
MAFSTYLGSLESSLVLKGLAGVWLVWYIGRVFYNIFLHPLANVPGPLLCKFSKIPWDYW
QWTGRLPQNTAKVHAKYGEIVRIGPNELSFTNNAAWNDIFAKVPGRAQWPRHPKRVPQ
GKNGPQSIMNTAGTYHARFRRLLNHAFSEKGLQEQQDLITKYIDIFVSKVDGFARTGQS
LDVTKWFVMVGFDVISDLGWSEPFNCVENGEVHEWMKTFAETAFDTQLKFLFRERGL
MFLAPYLVPMKLQLARLNNFKYARARVEERIKTGGTRGDFWDKISVKSAGDNASGEGL
TKEEMVVAAVTLVGTGSHTISTLLTGLAYFLGTNPHTMKKLVDEIRTSFNSPEEIDLVSV
HKLKYLTACLNETMRLYPPVINMLWRTPPQGGGHASGIFIPEGTGCNMSFFGIAQNPDY
FTRPLDFCPERFLPDPPAEFRDDNHEAYHPFSLGAYNCLGQNLANAESRLIMTKLLWYF
DFELDGTVDKDWLDQKSYGVFIKKELPVKFHPGPNAVRHVANGNGVATNGHANGHAN
GHARINTK
>gi|297242536|gb|ADI24936.1| VrtJ (Penicillium aethiopicum)
MTNSPIADLVHHPERVQSPSLVNSKMNGDAKAVTEWTEPGPAAFDFRSDTVTRPTEQM
LAAIAATTLQDDDFRQDPTTLGLEAWMAELTGKAAGLFVVSGTMGNQLGVRAHLQSPP
HSVLCDARSHLVTHEAGGVASLSGAMVSCVTPVNGRYMTQADLEAHVNRGTLITDCPT
RLVVLEIPLGGVILPLDKCRRISEWARAQGIALHLDGARLWEAVAAGAGSLRDYCACFD
SVSLCFSKGLGAPIGSVLVGSETLRERARWIRKSIGGGMRQAGVVCAAARVAVEATFLG
GLLKRSHARARDIATFWEIHGGRLTYPTETNMVWLDLEAVGWTPERLIRRGAELGLRF

MGARLVVHYQIGDEAIGRLQDLMLEILVSGLVDHPRDS
>gi|297242535|gb|ADI24935.1| VrtI (Penicillium aethiopicum)
MGARSKEARHRWFSRLPPDASSSFLPCKRPKSVAFVLWIRPSCFLRESLPAELSGLPGLPL
PPQAPTNTMNCGLPEAALSVIDWGRLKSGDVNEGARLLSACDDQGFFYLDLSSEPSFLH
DHKSVLHFMDQYFHQGLADKMKDDRQSDTHGYEPVATSTGALNTLPDYYESLKASRD
ELHGDGRNLAPAVCDRQDLFLRFSDMMHQMVIAILQELDCQLGFGGDRASFQDFHRKD
AESLTTLSMFRYPKQETLDLGVGHNKHTDIGTLTFLLCDQWGLQVLSKDPAGWRFVAP
REGHAVINVGDTLRFLSGNRFRSAVHRVIPTQRLQHEDRYSIAYFLRAANDTQFTDSAG
RQVSAKQWHDEKFDVFRETHEEQEKMPILTGGMERLENLPGIWCGVRLTDGICHRLGA
EDC
>gi|297242526|gb|ADI24926.1| VrtA (Penicillium aethiopicum)
MLHPTEVEKFPPTLLYFGNEFPNDDVNELFRRLQQHSKDRRFRLLNAYLEESILVLQDEV
AKLPHHIKSRVPYFDNIVTLSEHGYLRDLGLGAAMESAFLLILQLGLFIGNHEAVDRELN
LPKNVTTVAGLSVGLFSAAAIALSASLAEVVRNGAECLRVSFRLGVYAGDFSSSLEAPQP
EGMLASWAHVVTGMTEESVQSELTRVNEDLGNPETSKVFISAADKSSVSVSGPPSRIKA
AFLQSSDLRYSKSLPLPVYDGLCHATHIYSQDDVNTVLEISESLIPATRPFQLSVISSRTGV
PFTATTASDLLSEIATELVMGTIYLDNIIEGIVRHIGAFPGAQSCRIDSFRTSIIFKGILEAIAT
DHPDLTIEKNDLVDWVHQDFGTRRRNDPANSKLAIVGMACRMPGGANNVEEFWQLLE
QGRDACTTVPPDRFDLETHYDPTGKTENAAQTPYGNFIDRPGYFDAAFFAMSPKEAEQT
DPMQRLAIVTAYEAMEMAGLVIGRTQSTRRDRIGSYYGQASDDWRELNASQNIGTYAV
PGGVRGFTVGRINYFFKLSGPCLCIDTACSSSMAAVHAACTALWAGDVDVALAGGVNII
TDPDNYAGLGNAHFLSPTGQCKVWDKGADGYCRAEGIGSVVIKRLEDAEADNDNILAV
VLSAATNHCADAISITHPHAGHQKDNCRRVLRKAGVSPMQVSYVEMHGTGTQAGDAIE
SESVLDVFAPLKPLRRPDQRLHLGAVKSNIGHGEAAAGISSLIKMLLMFQKNAIPPHIGIR
TEMNPQLPKDLGRRNAGLVFETTPWLRPEGKKRISVVNSFGAHGGNTTLLLEDAPERHR
QRISPESTDGRSVYAISVSAKSKKSLQGNLSSLLGYLEQHPDTDLADLSYTTCARRTHHN
LRVATVVSSVSALQKFLRSAIDSNIATTVQSVPSNIPSVVFTFTGQGASDRGVRQELFDDF
PAFRTQVLQLDQLVQRLGFPSVVPALRGSTDEEVLSPVVSQLSIVVLEIALSRFWRLLGIR
PSAVIGHSLGEYAALEVAGVLSAADVLYLVGRRAQITEQRCTPYGHSMLSVLATPDEID
RVVRRVPETSNVEYEVSCQNTHEDTVLGGAKADIEAIRKVLETTSYKCVPLAIPFAYHTS
QMDVVVDELEEIAKNIPFKAPSIPVLSTMLGTVVFDGKTINPTYLRRQTRGTVKFVAAVE
TARDLGLIDEKTVWVDLGPHPVCVGFIRKLSPESRIAASCRRNEENLSTITKSLVTLHLAG
ATPLWNEFFRPNEQVYRLLNLPKYSWNETNYWIPYLGTWALDKALLKYGITPVGAKAP
ATLPAAGLRTSTIHQTTLETIDSMTATLHVLSDMQAPEFRAAVYGHTMNNCGVATSSIW
TDMALAVGEYLYRKLVPQAKEVHMNVCDLEVLHAQVISKVKGCSQPLALEAHLDLDM
QYMSLKWYNTNAATGERAPEWFASAAVRFENPDAWTAEWNRTGHLVLGRIETLRRLA
ADGVANRISKRLAYTLFKNVVDYSDWYRGIDDVIMNDYEAVANVTLIPDRHGTWHTPP
HWIDSVCHLAGLIMNGSDASNTQDFFYVTPGSDSFRLLKPLAPGAKYISYVRMFPLSAE
AGNMYAGDVYILKDDVIVGVLCQIRFRRVPRLLMDRFFSPPTADNAGVHGTPGSQRSQ
APPAATHAATKSPQKTLQVPSGHVPNKASIVDTNHVQASHPSAVPVRHDQSNGVSGVS
DSDSSTSITSSNSTADTSTTPTESEDADSGLVGQCIKIISRETNLDMSELTPDATFAQLGVD
SLMSLVLSEKFRNELGIDVKSSLFLECPTIGEVKEWIDQNC
>gi|297242527|gb|ADI24927.1| VrtB (Penicillium aethiopicum)
MAPYLQHFIAWFGKDSILSKLIQIFNPPSHSDSAKIIISDDDGRSPEPPGGEEILWRHSAPKS
TQMWAFLQKVNKKYGYSLRTYQDLHHWSITHRGQFWGEAWDYCGIRHSRQYDEVVD
EAAAMWPRPAWFRGARLNFAENLLFPKVPQAVSDEAIAVISVDEGGKRDFITWQDLRE

RVRRCQSGMRALKIQPSDRVAGYVANHSNALVAMLATASLGAIWTAVSPDTGVIGAV
DRLVQIEPRLLFTDNAVIYNGRIHPVLTKTQEIITALPSLEAVVVMRTVAGVQEDLPSSPR
SPAKRLTYNSFLAQGPETQKLEFVHMPAEHPLYILYSSGTTGPPKCIVHGAIGTLMQHKK
EHMLQSDIQPGDRLCFVTTCMWMMWHWLVSGLASGATVVLYNGSPFYYAPSSHGTSA
KDDLAMPKLVDELGITQFGASATYFSMLERRKHLPRSLTQGRLSLETLKAVYSTGSPLA
PSTFRYIYQAFGSQVNLGSISGGTDIIADFGVPSPLQAVVAGEIQVIALGMAVQAWGPTG
MDLSVTGEPGELVQFWGPSGASKYESSYFAKYPGVWAHGDHIQINPRTGGLIMLGRSD
GTLNPKGVRFGSAEIYHVLQYHFAAQVEDALCVGRRRRQDTDEMVVLFVKMRQDQL
WSLALAEAIRRTVREELSPRHVPELIIECPEIPVTANGKKVEVLVKRIVSGVEVPAAGGSG
TVNADCLQWFQEWATQN
>gi|297242528|gb|ADI24928.1| VrtC (Penicillium aethiopicum)
MGDATATEVFPVMDSIAVAMGDVESQLRVFHNVSRFLPTDDANQVFWWRTTGRHFAI
MMHEARYSEARQVELLLFYRFVIAPRLGPRPTSATPWFHSRVAPGIGDGSPIGYSWRWG
TGPDTKPLIRHYIEAIGPLTGTTADPLNEFAAKEMLYQLGQLVPGVELPLAWKFAAHIRP
SLTDEPTRAVAGSSILIGLQCAPESAGIEVMAGLMTRSPAQVPELLHSIFPRAMRDAYGP
DASLDGLNMVRDFVCHDPQGQYLTILGTTAIDCCAAASSRFKVYVTTTNTSFAHLAAV
MTLGGRKPEAPESLTQLQELWYALKGLDPEFPVTAEPLSSVCGAANGTASGNPNANVS
GVTFYFDIHPKYPFPHIKLQVDISKHTISDLGAINAVTEFLARRGQAADAQAYLNVVRA
MVPDEELRTRRGLQAFFAFAFKNGAVDITSYFLPQIYRRYAEVQAELEPRKDCQGRSEL
SSKLQRRSRFDSY
>gi|297242529|gb|ADI24929.1| VrtD (Penicillium aethiopicum)
MATSTTTSLKEFLSVFPQLVADLRALCLEEYQLPACVWDRFESTLNHNTLGGKCNRGLS
VIDSVRLLRDGLELSPAEYFDAAVLGWLVELLQATMLVLDDIMDGSPTRRGKPSWYRV
PGVGMAAVNDATMLESAIYMLLKKYFAGRAIYLPVVDLFHETALQIELGQAFDMLIAN
EGTPDLTTFVPATYSQIVTYKTAFYSFYLPVALALHAVDAATPTNLAAARAILVPMGEY
FQVQDDYLDCFADPTVLGKVGTDIIEGKCSWLVVQALQRASTDQAQLLAENYGSASGE
SSVKALYSELDLESVYRTFEEQRVAELRTLITGLDESQGLRKSVFEELLGKIYQRRK
>gi|297242530|gb|ADI24930.1| VrtE (Penicillium aethiopicum)
MFIAAVTHNWMERLAALSLLHYVLGAIFLLLLFHMLSNFFHPGLVDVPGPFAAKFTDL
WRLFKVWQRRFKEDLPGLHASHRSTLIRIGPRMVSCSDPRAVELIYGFHTEFSKSDMVK
AMAPIYKGKKQPTMFAAADNKTHARIRKPVAGAYAMTSIMQRMDELFIRPKRACDIHN
WVQYFAFDMVLEMTMSRNLGFMKAGGDVDGVLKQLQKDLDYRGIALAMPIIDRIWRL
NPVSKFFKPKQSGHFAMRCKRILEDRMAYEKSLDSRTQQQQDQKPHDFAHRFLEAQRK
DPSISDGQLIGYMQANLIAGSDTTAVVMRTAIYYTLKQPWILQRLVTELDQYHGPLPVPF
RIARFEMPFCGAIVREALRRHFAFIGMMERQTPPCGVVMPDGRRLPGGVVIGMHGDLIG
RDRAIFGEDADEFNPLRWLARPGEPEAKYQERLRAMNAHDLAFGHGPRGCIGKHVAEM
EIYKFIPTFFALIQPRFMRPEQSWTVRQLFVFKQSGMDMMLDWRQGKGLQSMA
>gi|297242531|gb|ADI24931.1| VrtF (Penicillium aethiopicum)
MTTTTTTDDTQKLDPSASDEVIYKSWDLLIYEIWVLGIVSTWAWGCSTTEYLLPQFRAN
VGTNHLDVGSGTGYYLRKGGIPASTRLTLLDLERPALDLGLQRCGRSDARGLQADILQP
LPVIDKFDSVSMYYLLHCIPASVEDKCAIFKHIKHNMTPDGVIHGANVLGKGVRNDGHF
AAYVRRGVLKAGIFHNLDDNAYDFEHALRMNFEEVETRVVGSVFIFRASRPKLDEGDL
LET
>gi|297242532|gb|ADI24932.1| VrtG (Penicillium aethiopicum)
MATRIPFDESYWEEYLSGQEASLPALPAVTQLSPRVTRLLAGNPGIMQLQGTNTYLVGT
GPARILIDTGEGRPVWHATLAEHLRTHHLTLEYILLTHWHGDHTGGIPDLIAHDPTLQSRI

YKHHPDRGQRPIRDGQRFTVTGATVRAVFTPGHAIDHMCFLIEEEKALLTGDNVLGHGF
AIVQDLAEYMASLARMAALGCERGYPAHGAVIENLPAKMQLYIHHNEVRVQQVITALA
SVVKLPGKRVGMTVPEIGRAIYGEVPREIVENAIVPFLSQVLWKLAEDRKVGFEPGEAN
KRRWFGLVTQQ
>gi|297242534|gb|ADI24934.1| VrtH (Penicillium aethiopicum)
MQRANHTRPVLIIGAGLSGLAIGRLLTNNGIANIVFEASPPERSQGFAISLHDWGYSLLLE
ALGGLSLRAMTKAVAPDRFIGGTGWVDLIMRDNTTGKVLVEPDVDARPAVIRANRNSL
RAWMADCGDDELDVRYGHRLKSISGSVGNVQAVFENGAEYRGSIVIAADGVHSAVRS
QVLPHIVPEVLPVVVYHGEFQVSHDEYDRCVRPVIGTANILAGVGDGFNTPITVCNITKT
QVHLDWSYSRPARGENDPLFSTKTPEDQTRDLPQALLEELASRQLAEPWAKYINPETIQ
QHSVFRWISRCVYMPTADALHAAQAGVVFIGDAWHAMPIFGGEGGNHALVDSVELAA
AMVKEANVERAVAVYYEGAARRCQEAVRRSRSRFYVLHRPMAEWRDIAEKRRAKAA
LEQKH
>gi|297242533|gb|ADI24933.1| VrtR1 (Penicillium aethiopicum)
MEDTTETTETTDTTAVSRLVPLAPAPARAPSMDAVNGSFDSELTTARRFNCQSCVRKKI
KCNRAVPTCASCSKAKLHCVYQSRPPRKRKRSRGEEDVYERLAQYERILHDHNLLQAA
AASTPSGRDTETSAISTRAPTPVLPDAQTHTKAGKVLLSTDGRSRYIDNVLLLDAGEGDL
CELPESEQEDYNHDETSPDESTPTGLLGALAAHTISGAIIGNTQSLTNLHPTYEQAAKLW
QAYVKNVEPLCKILHVPTVAKMFDTVSKQPAAVSKNDECLMFVIYYFAVFSMSDDECL
HEFNYPRAQLLSRYQTTVIQALVNASWLKTTAMPVLQAYTLFLIALRTQIDSHTFWILTG
IAVRLAQRMGLHRDGESLGLPPFEVQMRRRLFWQLLPLDSYAGQTSGTGISISPSSWDT
KQPLNINDDQIFPGMTQPPCEQRGASEMIFCLSRMELSNFYIRTGVKLKEHGDTIQFRDA
EDIERLIDEVEDLIETKFLRYCDILNPLHFLTTGIVRSAIDAVRLRARMPLLKQQTITDAQR
RRLCALAEKVLDTNSTIFSNPSTQNFRWQMQAFFLWDALLCILRNIAEVGFYSPSELAAA
WSKVANVYANHDELVKARRTLYVTIAKVTLKAWLANPPRDSSPQPAFITALLTQHEPK
GINQQQNSVLSDDKAADGASLFDEFFDNMNGTDLDINNAFNLDSSSDWLFWDQICRGT
SLS
>Afu7g00180|nscE COORDS:Chr7_A_fumigatus_Af293:44976-44029C, translated using codon table 1 (315 amino acids)
MHIFITGATGFIGRVVSELAIQQGHTVHGLSRSPQGDEILTSIGAIPIRGDLATHNILREQSA
KADAVFHLAFDHDFGKSYDQIIKLDTEAVDALAAPLVGTSKPLIAASGILTVRPDQGDC
VVNESAPYTKNTRVRRHVCEENALSWAERGVRVNVVRLPPYVYGRANETGFAARMVR
MAVDNGVSGYIASVKDRCVTSVYVDDAAALFLLLASDKTVKAGEIFHGTADWDTTYG
MLAKAIGRAVGVPVRQFEREEAEERWGAFLSSFFGLIIRASNGKAVEKLGWKPSGPSLV
EELETGSYRQVAERFKQMKN*

Table S2: The composite multifasta file used for the group V2 MultiGeneBLAST search. This list contains one member of each group of orthologous genes in all group V2 clusters.

Table S3: Composite multiFASTA file used for group V3 MultiGeneBLAST.

>gi|297242553|gb|ADI24952.1| GsfR2 (Penicillium aethiopicum)
MPPLYRRSCITCVQSKRKCDQGLPKCQRCLAKNIHCEYNPRYPNRRRQTTERNVDENVS
LVEPIAEEPSRGCQLQRSPARPTSPTHSPHANDIFFNFANDPFNLESIPQDNFLNSTIFEDV
VTQQAPNDTERITSDTTAQARVEFAAKKLSVIPKIFSQQGQTMFIHRQLFQDRAPPALQD
ALSACALYCLKSTENQTLVFRNLEHKRKQLISSIDPLLASKLDLLEALQALVLYQIISLFD
GDIRLRAQAEADEPVLLMWAAQLTLRTPQFQPPLGLSNPQSLAGSASMDWGRWLIEESS
RRTLITASMLKGVYSFVKLGYDTVPDMRMSFTAQAVLWNSQSEISWRRAYKEKERLEI
QVTHWDETIAKAKANDLEELGVLIMVMLKGTGATGEWLGHSQNIRYGLEEAYYGSV
>gi|297242552|gb|ADI24951.1| GsfK (Penicillium aethiopicum)
MPATVLITGGNRGLGKGLVATYLSTPDTTVIATVRDPSKCESLSALPKALGSNLLLVKLE
VTSKDSITTAIGTLDTHNIGSIDVVIANAGISGPTSSLAEAPVSELQRYIDVNAYGPFELFK
AVLPLLRSSNSGNKAKFVCISSAGGSLAAMYNFMPISAYGASKALANFLVKWLALDNK
DIIIWAQNPGSVDTDMARDGLDLAKSLGFDLSSLSFTSPEESACAIKKLIDGATTEMSGKF
LDHDGSELAW
>gi|297242551|gb|ADI24950.1| GsfR1 (Penicillium aethiopicum)
MSDGPETAEGDTDDAVQDAAVNSRVASESSARSQPRATVVGCLPPTPFAIHTTPAPSEH
SKEKNVSRRLPTEKTPSRLATPQFPPTPVSSRGSIAEPSAYPSITQALRSCLPPQKDIEILLS
NLSSMSIFCYKSSFKLCSSWPSEMTEEQIPIANLLYSETHPVLLARHMLLFAVGLQHLSPT
KAIPGLTRHHRAIMEQLADSAIKLVNTDDVLLGTLEGLENLILESFYHIDGGNIRRAWIT
MRRAVMTAQLLGLHRPGHYRFKTVNKQNDLDPAVMWACIVSTEQFLCLLLGLPTSTSG
ASFTIPRATSACVESGNLPVLIPDVVRKIIERNQTHVPQEALDMTQKIDHELLGVVKQWP
PAFWRPLQLSGLEVDSADAFWETRRAWDHIFYYSLVNQLHLPYMLNPSHVSQKVYSRI
ACASASREILIRQIAIRTFNPVTAGCRMGDFVALIAGMTLMLAHILSHCSKGTENLLVHQ
RVGDRATVERALECMESMSEQHEDILTAKCAALLKNLLDIEAGPAEARSDDGQKDDQN
VLVVKVPHVGAIKIARDGISITPFDTEQEQVSHDGVTIGGFGSIHVSTPHDSDRDGDHQA
DVVTPNDTASQATTQVVRPASARKQWRPVSQRSSGDVFTLPEETFPDASAGMEEWLFQ
GLDTAFFDVLISGAGEQPLNSTDTEGWNFVMSP
>gi|297242550|gb|ADI24949.1| GsfJ (Penicillium aethiopicum)
MDKQNTDSEKRATAEAPPQSLCTLVSSSEQGGMDITNHNAKAGAADEYPHGVRLAAV
VFSLMLGMFLVALDNTILGTAIPKITDEFHDLNKVSWYGSAYLMTFGCGFQSTWGKFY
KYFPIKVWFLVAVFIFEVGSLICAVAQNPTTLIVGRAIAGFGGSGVGVGIFTIIGFAAPPEN
RPQLLGFTGATYGIAAVLGPLIGGAFTDKCFYINLPIGGVAAGTIFLLFKPPTSASPAKATP
KEKFLQMDLVGATLMMGLIVSYILALQYGGQTHSWKSSEVIGLLVGFFLFVLAFVTWEI
YQKERAMIVPRLFMRRYISVGSIYMFFFSGAYFIILYYLPIYFQSVYNSSPIGSGVKMLALI
IPLTLAAIVQGWALSKIRIVPLFWIIGGALGTVGCGLFYTFDTETSVGKWVGYQIIVGFST
GWTFQIAMSNAQVHAPPEDMSQATAIVNFFMTVGGAFFISAAQCAFSNQLIKTITKNLPE
LDPTVAISTGATQIREAFTASQVPIVVDAYMVGLKAVFAITIAAFGVATVIGFFGSWKKL
LADELEKATGGVA
>gi|297242549|gb|ADI24948.1| GsfI (Penicillium aethiopicum)
MAIPQSCTVLVAGGGPGGSYTAAALAREGVDVVLLEADCHPRYHIGESLLPSMRYLLRF
IDLEDTFEQHGFQKKLGAFFKLNAKSAGYTDFIRANGPNGYSWNVVRSESDEILFRHAT
KSGAKTFENVSLKSVNFEPYENDKFTSQDKLTNPGRPVSAEWKTKDGCSGTISFDYLVD
ATGRVGILSTKYLKNRKFNESFRNIAMWGYFKGNIPPSPGTDRENQPISEGMRDGSGWV
WMLPLHNGTVSIGAVVRKDIFQAKKKALPEGTTEAQTLASLVALCPTISSYLEPAELASG

IRQAADYSYSANAYAGPNFRIVGDAGCFIDPFFSSGHHLALSSALAAATSINACIRGDCN
EFDASRWFAKKVDEGYTLFLVVVMAALKQIRMQEQPILSDLDEEGFDRAFTILRPVIQG
AADKETAPKAKGESITETIDLCLTALNDLHDTELQRKLTSIVEAKGTPEQEQLLGKLSPD
ETAALHRMRAMHSILPMGELEDFENSNIDGFKARLEKGSLGLRRERALCRDHAGDLQM
>gi|297242554|gb|ADI24953.1| GsfA (Penicillium aethiopicum)
MTSAKVLYFSGEIPQGDPEGDQRTLFRKLHLLSKERDHVVLASLLECVTLTLKDECSKLS
PQYRDLLPPFESVLDLTDHVVQLRKTPLGGAIERVLVLVFQLGSLVAYHEAHPLEYNFTP
ASTVIIGRGSGLLSAAAIGLSPSIVMVPSIAKEIARISFRFGLVVDKVCRSLEVSSDEINSDG
AWVYCVHGIGEKEARDAVNQFNEIKAYPSTNGASVFNVDDAGNSVSIGGPPKTLEALFS
ESNIFKKTKNVAMRKIQGMWHTDRVYGPEHVEQIVPKIESARELHVPLISPVSGDPFRET
EAGPLLEQIMEEILMERVRWDMIIETVSKQLKQLMPKSVQLVSIQPSHYNQNMLERWKS
ELPDAAVSGLTMMPAILELALEQSPPKDTRSSKIAVVGMSCRFPGSDTTEEFWERLMLG
EDMHRHIPPDRFDVETHVDPTGKRHNTSKTSYGCFVDNPGLFDAMFFGMSPREAEQTD
PMQRLALVTAYEALEKAGYVDGRGVIHRKRVGTFYGQASDDYREVNSGQEVGTYFIPG
GCRAFGPGRINYFLNFWGPSFSVDTACSSSLAAIQAACSSLWSGDIDMAITGGMNILSNS
DVYAGLSQGHFLSPTGGCKTWDEGADGYCRSDGVGSVVLKRLEDAEADNDNILAVVL
SAATSHSAEAVSITHPHDAAQALLYNQIVRRAGIDPLEVGYVEMHGTGTQAGDPTEMRS
VTSVFAPPHIQGSRPIPLHVGSVKANMGHGEAAAGIMAFVKTMLVFQNGIIPPHIGVKTG
LNPALPDLDKAGVVIPFRAANWRPTGTKKRLAMVNNFGAAGGNTAMIIEEAKARPRLC
EDIREAHAITISAKTAVSLSLNIKRLVEYIESAQDLSLADVAYTVSARRRHYEYRKSVVV
RSLAEAIKHLQPHIETAKSQTPTLVKRPPVAFAFAGQGTFYVGIAAQIYRDSPFFRAQIDQ
FDNLARRQNFPSFLPAINKTCAHEDLPASSIHLAIVCVEVALARMCMTFGIKPCAVIGHSL
GEYAALAVAEVLSDSDTVFLVGTRATILESNCSPYTHGMISVRASVDDISREADGLPFEV
SCINGPNETVIGGTVENLEAVADRLSKVGYRKTRLDVPHAYHTAQMDNVVNELIRQSQ
GIAYNTPKIPIMSPRDSSVIETGANIDSSYLPTSLKKAVDFAGALNAAWEAGVVSKSTVW
LELSHHPVCSGFINRTLPNTSLTCSTLHRDSDNWTSLLKTLSSLYEVGLNIDWNEYHRPF
EHALRLVSAPTYAWNNKDYWIQYRGDWNLTKGQVLPEAELPAVSGFRTSSIHRLYSEN
YDSSTAHLLGECNMTDLSLKGVIEGHAMNGYGVASSFLHAEMAFTLARRIQEKASLSTF
TGMGINVTNFEYHDPVVKDASSLDPYPIVVDAEANLEMGEVQIKWFNPAIEKWYCHAI
AYYEDPSTWLSNWSRTTRLVTSRIDALVAMSNKGMANKLTTSLAYTLFGKLVDYSSMY
HTMQWVILNEDEAVAEVVFPADTQGDWAVPPHFIDGVVSLSGFILNGGTHFDNVNNFFI
TPSWKSMRFAKPLAPGGRYLTYVRMIPEGVDDKGRLGSYVGDVYILQDGEIVGVVEAIL
FRQWPRIMLNRFFQPVGMAPPAPRVEKKRDAGRGTLPSSSSLQEKTTATAVTAKITARF
PGSVITPSRSAPISKSGSSPKIVPQLDYSLLTPRTSPNSDERIEKTDSDSGFEEADGANDVTS
RAVEILAEELAVDKGLLTDECEIADIGVDSLMSLVISQKLREDLGIEVRDAFYLEVTTIGD
LKKLLS
>gi|297242555|gb|ADI24954.1| GsfB (Penicillium aethiopicum)
MASNTSRSAHLAQIITENTANIETYRREQGLPPLSLGPDAPLDVKYPPNVEKCRRAVIDA
TLELGELATGPVELRLVPGWAIMTMFGVTQFICDFDIARQIPLAGDISYEDLSKAINVAVP
VLRQVLRAGMPYHMFYESRPGHVAHTATTKVMASESLISDWTSLYTDVLFPASAGLSK
ALREEPTASDPSKTGFMVTKGDGESGMYMYFEKHPEEARRFAGVMEAFQKDEAYAVR
HLTDSWPSDSQTGKLVDLGGSTGAVAFALAEKFPGLEIVVQDLPGALEAAHVREGKNV
SFMPHDFFNEQPVKDADVYMFRWILHNWPDGHVQRILRALVPSLKPGAKVIVFDEIMPP
AGTLPLSIERYQRNIDFGMLTLFNSKIRDIVEWKEIITQSDQRFNVTGVRYPENSRLSLIEI
VWQP
>gi|297242556|gb|ADI24955.1| GsfC (Penicillium aethiopicum)

MTLDQISRIQALVAEIADLSQTFEGPSERFMVDGLPFLRKSEELVTLVQSPAEHATSLIIKT
METAVIRTLLSLNVLQTIPTTGSITLQSLAVATETQESLLERLLRVVTKTGFIIRENGGYSH
SHTSLAYAGPLGALFAPCYDEGIRALVRLPEYLSVKDKEEAKNARYSLFTWNEGQEGK
ATFEILSTMPARTEGIHTLAMNVQHLRPYTGFFDYSKLVSEDRERPVFVDVGGGNGHVI
KEILQAFPQIRPEQCVLEDRAETLELARTTGLLPAGVQLLEHDYLTRQPVSNAKAYHLRA
VAYNLGDAELVQLLKQIVPVMGADSKVLIAENILFDDNSTVFSTVSDMIMLGIGGKERT
EQNFREVLYEAGLTIEGIHRAPGLEYGIMEASLATS
>gi|297242557|gb|ADI24956.1| GsfD (Penicillium aethiopicum)
MSTPEQWIQEFEALCSRASILFPSGIEDENIRIRALRIAEKAVHQLHTPMTFAEAQTWAPL
ELFGAGVACEMGIFDVLSKSSVPLSPVDIATELNTDPALVARIMRLLDAHYMVDQVTLG
QYAANAITRDYVQPYRKGNVMTQVALMPSYFALPAWLRDNDYKVRPDANHCAWQV
GANTTKTFWEMPRTTQEDDFVTFYPFEAVFSTSNVDDILFVDIGGGLGHQAMRVRSAFP
RSRGRIIVQDLPQVTNKITTASLPDVEIMDHDMADPQPVKGARVYYLRGVLHNHADHIS
IKYLSQFAAAMSPESRLLIHEALATDLNPTKNITRFDLSMLASCGGAQRSEAEQKALLEK
VGLEVSGVWSTPRDWSIMEARLKRE
>gi|297242558|gb|ADI24957.1| GsfE (Penicillium aethiopicum)
MPKTAFITGANGLSGSAIVEYLCNTTTSDDWGSIIVTSRSPFKSTVMDPRIKFIALDFVND
VSSLVETMKEVCGAVTHAYFCSYLHKDDFAESYTVNKALFENFIAAIDKAAPKLENVTL
QTGGKYYNLHVEPVPSPARENDPRRYGPFENFYFTQEDTLAEMQRGKTWSWNVIRPEA
IIGANSQPYGLNVALTIAMYFLICRELGSASPMPTNQRYWEGTDDVSYAPLIADLTIFVST
RKSCANEAFNVTNGDYFTWRYMWPRLAASLGAKADSQQCFEKPMPGEGELQLDWSL
AEWCKDKRKVWEDLCDRQGLPGAKATFDLAGWAVGDFLYQRTWSATLSVNKARRFG
WTGHMDSYQSFVDTFDKFRQLGLIPK
>gi|689554835|gb|ADI24958.2| GsfF (Penicillium aethiopicum)
MTVLFILSAGLVAVFGYLVSWFIYCRTLHPLSKVPGPFWPSVTRLWLTYAVSRGELDVV
QRDLHRRYGPLVRIAPDEIACADPEAIRKIYSTTSPLNKSDFYHIWDVGAFSKYPNAFAIV
DENMHFERRRIVSSVYSMSTVLTLEPYIDNCSRLFVKRMTERTVPHEAIDLGDWFLWYA
YDVIGELFFGHSLGFIENRGDEGGFLASLEVMLPVLTIAAASSPLVRGLIMGLFTLSSTAR
KGLKGMNHIIETARASVDKRASAVAEPGKGERKDLLHNLLNIVSSKGDKLDFGIEDVKN
EAFAALTAGADATMIELQAIFYYLVKDRSVYEELRKEVDQAVETGKLSEFPSYSEVVQL
PLLKATIKEALRLHPAVGFTMPRVVGQAGIELLGMYIPPGWKVGMNAAVVGRDEGVY
GTDANTFRPERWIERTDSDMDRCNNLVFGAGTRTCIGKQIALSEIYKMVPLLIRKFDFAL
VDPSKSWTTHDYFFNKQSGVQVKVTVRGL
>gi|297242560|gb|ADI24959.1| GsfG (Penicillium aethiopicum)
MLCSSCGVPTSTQDPSSPRAKAAAADRRREQNRRAQKRFRQKHREQKATPEQDQSQPQ
KRASDAMDTVGPTTSKESTPLTPPSCYEERQADSVVDYDCLLDPSDRCVMNKWTDDF
WTSTLEETAPFAATALVLPSERKDSQEILTMRDFDDSVDKNQNGRVPAGPTVLHRAVQT
GNSKVVGLLLEHNANCNTKDNTGLTPLLCAVIGGHEEVLELLLSHGASIGHVDDAHWS
ALHWAVFHKRHRILERLLRCCSGDSSLLNIRNKDGETPLSVAVSAGSEVAVKLLLEFGA
TVNIEQSS
>gi|297242561|gb|ADI24960.1| GsfH (Penicillium aethiopicum)
MFISIGIKAVLLLVASIPAQAWNRLDKDNAALLIIDHQVGLAQVVRDYNTNDFRNNILGH
AALGNVFNLPTVLTTSSDAGPNGLMLKEITDMHPNATFVRRQGEVNAWDNADFRAAV
EATGKKQLIIAGIVTEVCTSFLALSLVDAGYEVFANTDASGTFDARLAEDANRRMEKAG
VTLMGLFGIVCDLMRDWRKTPGLPEVLPFLDKYQFGYGLVARHHAGAIQNGTFFPVEG
ALI

>gi|111055548|gb|EAT76668.1| hypothetical protein SNOG_15830 (Phaeosphaeria nodorum SN15)
MAYTPMEQPLQFSVTRIGQDLKIFEILTNAGPQSLQELQILTHAHPITLALSLTGRLLRYM
SSVGLIHETGVNQFEASKKSQNLATPEAEKIVTHFFENCGPLFQEMPAFLRSKKYQDITD
GQNTVFQPAYNTDLDTNEWFSQNPLHMGARIKYMAMEQAVRGRWLNVYPIQSQLKG
WNPELPAFVDVGGSVGHYCTMFKNNFPDVPGRVILQDLPDTLSHALKTPGVELMGHDF
FQPQPIKGAKFYHLGWILNNWNDDESKQILRQIRSAMNAESVVLINDVILPEAGVPSFAA
SLDLVMLGACGSRERMRKEWDEILADVGLIVKDCIVYHHELCHGIIGAKLA
>gi|111055549|gb|EAT76669.1| hypothetical protein SNOG_15831 (Phaeosphaeria nodorum SN15)
MAETIKLEVKDSAHRRRSVPLLSSGERQPPAKTRNGDCIVLPSKTPTSLASTIVDDVLVD
CNGPSCTLMGKETRERLAIGQSDALLVMFYLDNLFPFLFPFYRPSLLEGGRTWILEMMIC
SPVVKQALLCQSSCFMSMAQGMANWEMVLEQTRDAFKVLRLSLQVISDAGVTEHIHGT
VRILASIVQVQRFEVAVLSFDNCQAHLNVALSLFSQLLESTDDTNTACPSSSFNAVLNQL
KPKTWSASTGVFQVPIAEQAAFRFSSALLILDDIIAATILQEQPRLYNYHRSLLCTPNSSG
AVVNLEETVGCQNWALLYIGEIATLDAWKQQYQHTGSLDLMELVRRATDIEALVEGHV
LQLESELLGASSEGASLLDMFDSDSRQTKTSASQITLVTRVWAHAALIYLSVVVSGWQP
ANLDVRYHVSQTFELLTQQIVPPALLRSMVWPFCVAGCMAAPEQEAQLRTMVQVLQPP
SMFGTVRKALEIMEDVWSKRGMGDAGKRDLATCFKNQGDLVLLV
>gi|160705606|gb|EAT76670.2| hypothetical protein SNOG_15832 (Phaeosphaeria nodorum SN15)
MLPISSYISRLSCFGTKPKPKPKPIKSRISKLSTQTCFTITKDTNMATVVLILGSGPRVGAS
VAQTFKESGYSVALASRKGSNSVTDDGYFSIKADLASPASILGVFSTIKSKFGAAPSVVIY
NAAALTPPSDPSSVLSVSTEAVTKDLNINVVSPYVAAQQAIAAWETLPQDAKKTFIYTG
NVLNELIVPIPMFLTLGIGKAASAYWVGLADTLYKARGYRFFHADERHPDGKNKGMAL
DGPAHGEYYHELAQHPDGVPWQATFVKGQGYRSFK
>AN7070|pkgB COORDS:ChrIV_A_nidulans_FGSC_A4:1131258-1129968C, translated using codon table 1 (340 amino acids)
MSGGFYSSPFWAGYLETQRSRLPVLPEIDDGLSHCVVRFLGYNPGSMQLQGTNTYLVGT
GSTRILIDTGEGAPQWAVSVTRYLEDHDISISHVLLTHWHKDHTGGVADLLAHDPSIIVY
KHAPDPGQQAIANGQTFKTQGATLRAVLTPGHAVDHMCFLLEEENALFTGDNVLGHGY
SVAEDLETYTASLRLMAGLKCSVGYPGHGDAILNLPQTIARYISQRVAREKKIYAILALH
ACSCSSRNGGSTSSIGSVSESGDSDEEDNNMKTSRPAMQGLSTAEIGGLVYGESVKNSPT
FDSAVGPLLNQVLYMLLEQGKCCDHVSILVIFQKPGFFSIPVI*
>AN7073 COORDS:ChrIV_A_nidulans_FGSC_A4:1120517-1119359C, translated using codon table 1 (321 amino acids)
MEAYKTQPPVRKLKDSCDVCSASKLRCDKQKPTCARCANLNRPCTYSPARRGGRPHRV
RRDRSKSQSQSQSSRQFFGMPDANTSSSPFAEPTRVPSQTDRGMSCSNDWFLRTQSRVH
DHIQDSQPSAQSAKVSPSPCKNPMNTRLAAETAETDMDCTRVALSIVEQLERSQEQRPR
STAPTYTHGGLTATEACQRLLTILMCPCSDQAEVALLVASGCISLMDVIHSSAGFASESL
GHDGSVSSCNSPPISSEQDPLIRSWSQPQSISRSCSLASDSRSQVGDLSKIAKVIVQFTERY
CQDAKVAAEPRAHWVTRYESLSR*
>AN7072 COORDS:ChrIV_A_nidulans_FGSC_A4:1121945-1122955W, translated using codon table 1 (336 amino acids)
MELPAESELQYAGECLSLPGTFLEPPIEDPPSSVLNLLNLSQVDFNSYDFSSLGSREFSSK
WQTNTPLCTDSLSDESAPGLLTEDMGISPIPMPAAEATCPQESEDRLCRNPQGRCISLAT

GILGSMHAGSNSCILQVATSDQGGASDRQPQQSRAADAILSMNQSALRTVRSILNCSCY
ESPQVLLLVTVMCSRITAWYWRIADIYSYSHGNPTAGSPRAALPTSVGSRAETRRRDFFI
GNHRLDREVETVVIRHVLLGMLQELQLVIRDFAGQAGQSPAGTVDTDDPTSTSDLMLSG
MRARVVAFLRKQLHSLTSALDHTDSGFGTMGPHVSHY*
>AN10884 COORDS:ChrIV_A_nidulans_FGSC_A4:1131717-1134001W, translated using codon table 1 (647 amino acids)
MMNVPEKCKVLVVGGGPAGSYAASALAREGIDVVLLEAEKFPRYHIGESMLPSMRHFL
KFIDAYDKWDAHGFNIKKGGAFRLNWSRPETYTDFIAAGGPGGYAWNVIRSEADELLF
KHAAECGVKTFDETKVASIEFSSPDLSSGGTHPFGRPVSATWTRKDGTSGTISMDYIVDA
SGRNGLISTKYLKNRSYNKGLKNVASWGYWRGGGVHGVGTHKEGAPYFEALKDASG
WVWFIPLHNGTHSVGVVQNQEMATEKKRKMAEPSSKGFYLESLEFVPGIKELLANAELI
SEVKSASDWSYSASSYAFPGVRIAGDAGSFIDPFFSSGVHLALSGGLSAATTIAAAIRGDC
DENVAASWHDKKTSESYTRFLLVVSSALKQIRSQDEPVISDFDEGSFERAFDLFRPIIQGQ
ADADAKGKLTQAEISKTVEFCFRAFAHVSFEQKEALVQKLKSLGHDGDAYDENNRKAL
EEIEKQLTPEEQTILKTLKGRRMVRPEDSLNIDNFTLDSIDGLAPRLEKGKLGLSAAKKAE
LKFTAHDPLSFLNGEAMAAQKTSPNGNLEINGHTQTNGNHLANGHGEVNGHSNAGASS
AKSCMADLIAAENDSSQPSFDEATRHRLISSLQQSAEELETPYDTVLRYVNAVSQIN*
>AN10889 COORDS:ChrIV_A_nidulans_FGSC_A4:1134037-1135110W, translated using codon table 1 (337 amino acids)
MLTALSKQGRQTALVCIGGDLGIFKSLAESKAPLSSKQLAEATMADPLLVSRIMRYLVA
SRLVGETAPDQYVATKKTYVFADPRFEEPIRFFHAVSNRAFQALPEFLKETGYQNETQRS
AFQKGLGTELQLYPWLKQHPDVLKNFQAAMRLTKDANGVGVMPLDSSVSSGHEGVMF
VDIGGNTGHQAAEVLSQHPELAGRVTVQDRGEVIKSAPEMKGIQWMEHDFFDVQPVKG
AKYYYLRAILHNWDDDHAVQILANIVPAMSADSLVAIDEVVVPDRDAHLWPAGLDLQ
MYTIFGTRERTAAQWDAILDRAGLRAVAVKRYAPVMQSSVIFAAAK*

Table S3: The composite multifasta file used for the group V3 MultiGeneBLAST search. This list contains one member of each group of orthologous genes in all group V3 clusters.

Supplemental figures

Figure S1.

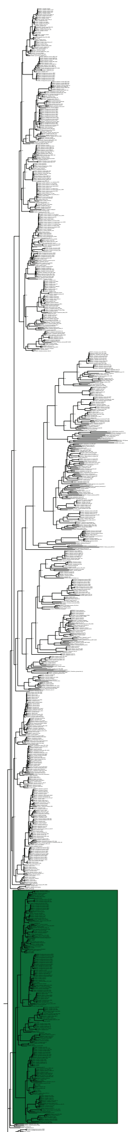


Figure S1: A maximum likelihood phylogenetic tree constructed with FastTree (Price et al., 2010) from the KS domains of the 908 PKSs used in this study. Bootstrap values are shown next to the internal nodes of the tree. The clade corresponding to the 188 group V PKSs is highlighted in green. The leaves of the tree are labeled with the accession number of the PKS and its corresponding species.

Figure S2.

Figure S2: A maximum likelihood phylogenetic tree constructed with FastTree (Price et al., 2010) from the KS domains of the 188 group V PKSs extracted from the larger set used above. The bootstrap values are presented next to their corresponding nodes and the leaves are labeled with the accession number of the PKS and the corresponding species. The characterized PKSs belonging to groups V1, V2, and V3 are shaded in blue, green, and yellow, respectively (Szewczyk et al., 2008; Chiang et al., 2010; Chooi et al., 2010; Li et al., 2011; Ahuja et al., 2012; Lim et al., 2012; Saha et al., 2012; Simpson, 2012; Chooi et al., 2013; König et al., 2013; Nielsen et al., 2013; Xu et al., 2014; Chooi et al., 2015; Throckmorton et al., 2015). These shaded boxes include closely related PKSs from the same species in which the characterized cluster was originally described. Next to the tree are the gene clusters corresponding to the PKSs that were identifiable through MultiGeneBLAST analysis. Gene cluster diagrams next to brackets depict the cluster corresponding to the PKS with its accession number highlighted in red, but all of the bracketed PKSs belong to clusters which are identical in terms of the presence and synteny of their group V-cluster homologs. Genes are represented as arrows with a color corresponding to the proteins they encode which are detailed in the color key to the right of the cluster diagrams. These color keys are subgroup-specific, and black lines extending from the tree to the color keys delimit the subgroups. Genes with no color were not identified as homologous to any group V1 cluster gene. Colorless genes that were noted as occurring frequently in related clusters are labeled adjacently. The products of the characterized clusters are shown at right. Group V1: PKS=Polyketide synthase, MβL=Metallo-β-lactamase-type thioesterase, AO=Anthrone oxidase, OGO=2-oxoglutarate-Fe(II)-type oxidoreductase, EthD=EthD domain-containing protein, a putative decarboxylase (Chiang et al., 2010; Throckmorton et al., 2015), MCO=multicopper oxidase, P450=cytochrome P450, SMT=S-adenosylmethionine-dependent

methyltransferase, BVO=Baeyer-Villiger oxidase, DOR=Pyridine nucleotide-disulfide oxidoreductase, FDH=Flavin-dependent halogenase, AflS=Transcriptional co-regulator of the aflatoxin biosynthetic gene cluster (Ehrlich et al., 2012), AflR=Transcriptional regulator of the aflatoxin biosynthetic gene cluster, C6TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, NOR=NADH-dependent oxidoreductase, OMT=*O*-methyltransferase, GST=Glutathione *S*-transferase, MT=Methyltransferase, SD=Scytalone dehydratase, KR=Ver-1-like ketoreductase (Skory et al., 1992; Simpson, 2012), FMO=Flavin-dependent monooxygenase, ACS=Acyl-CoA synthase, Kelch=Kelch domain-containing protein, SDR=Short-chain dehydrogenase, GA4D=GA4 desaturase family protein, MFS=Major Facilitator Superfamily transporter, GMCOR=Glucose-methanol-choline oxidoreductase, *vbs*-like, MFS=Major Facilitator Superfamily transporter, PPT=Polycyclic prenyltransferase, AKOR=Aldo/keto oxidoreductase, SH=Salicylate hydroxylase. Group V2: C6TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, VrtR1-like, MFS=Major Facilitator Superfamily transporter, P450=Cytochrome P450, TA=Threonine aldolase, OGO=2-oxoglutarate-Fe(II)-type oxidoreductase, PKS=Polyketide synthase, ACS=Acetoacetyl-CoA synthase, PPT=Polycyclic prenyltransferase, IDS=Isoprenyl diphosphate synthase, MT=Methyltransferase, MβL=Metallo-β-lactamase-type thioesterase, FMO=Flavin-dependent monooxygenase, TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, VrtR2-like, NDH=NAD-dependent dehydratase, LIP=Secretory lipase, ART=Arrestin, CNR=Copper-containing nitrite reductase, MGT=Magnesium transporter, ABCT=ABC transporter, Abr2=Conidial pigment laccase, OMT=*O*-methyltransferase, EthD=EthD domain-containing protein, putative decarboxylase (Chiang et al., 2010; Throckmorton et al., 2015), PksP=Conidial

pigment polyketide synthase. Group V3: C6TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, GsfR2-like, SDH=Short chain dehydrogenase, TF=GAL4-like $Zn(II)_2Cys_6$-domain and fungal-specific transcription factor domain-containing protein, GsfR1-like, DRT=Drug resistance transporter, EmrB subfamily, FDH=Flavin-dependent halogenase, PKS=Polyketide synthase, OMT=*O*-methyltransferase, NSD=Nucleoside-diphosphate-sugar dehydratase, P450=Cytochrome P450, ANK=Ankyrin repeat-containing protein, AHD=YcaC-related amidohydrolase, FDH/OMT=Flavin-dependent halogenase and *O*-methyltransferase bifunctional protein, MβL=Metallo-β-lactamase-type thioesterase.

References

Ahuja, M., Chiang, Y.M., Chang, S.L., Praseuth, M.B., Entwistle, R., Sanchez, J.F. et al. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J Am Chem Soc* **134**: 8212-8221.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microbiol* **76**: 2067-2074.

Chooi, Y.H., Cacho, R., and Tang, Y. (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* **17**: 483-494.

Chooi, Y.H., Muria-Gonzalez, J.M., Mead, O.L., and Solomon, P.S. (2015) SnPKS19 encodes the polyketide synthase for alternariol mycotoxin biosynthesis in the wheat pathogen *Parastagonospora nodorum*. *Appl Environ Microbiol*.

Chooi, Y.H., Fang, J., Liu, H., Filler, S.G., Wang, P., and Tang, Y. (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **15**: 780-783.

Ehrlich, K.C., Mack, B.M., Wei, Q., Li, P., Roze, L.V., Dazzo, F. et al. (2012) Association with AflR in endosomes reveals new functions for AflJ in aflatoxin biosynthesis. *Toxins (Basel)* **4**: 1582-1600.

König, C.C., Scherlach, K., Schroeckh, V., Horn, F., Nietzsche, S., Brakhage, A.A., and Hertweck, C. (2013) Bacterium induces cryptic meroterpenoid pathway in the pathogenic fungus *Aspergillus fumigatus*. *Chembiochem* **14**: 938-942.

Li, Y., Chooi, Y.H., Sheng, Y., Valentine, J.S., and Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**: 15773-15785.

Lim, F.Y., Hou, Y., Chen, Y., Oh, J.H., Lee, I., Bugni, T.S., and Keller, N.P. (2012) Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* **78**: 4117-4125.

Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., Larsen, T.O., and Mortensen, U.H. (2013) Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. *PLoS One* **8**: e72871.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.

Saha, D., Fetzner, R., Burkhardt, B., Podlech, J., Metzler, M., Dang, H. et al. (2012) Identification of a polyketide synthase required for alternariol (AOH) and alternariol-9-methyl ether (AME) formation in Alternaria alternata. *PLoS One* **7**: e40564.

Simpson, T.J. (2012) Genetic and biosynthetic studies of the fungal prenylated xanthone shamixanthone and related metabolites in *Aspergillus* spp. revisited. *Chembiochem* **13**: 1680-1688.

Skory, C.D., Chang, P.K., Cary, J., and Linz, J.E. (1992) Isolation and characterization of a gene from *Aspergillus parasiticus* associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. *Appl Environ Microbiol* **58**: 3527-3537.

Szewczyk, E., Chiang, Y.M., Oakley, C.E., Davidson, A.D., Wang, C.C., and Oakley, B.R. (2008) Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microbiol* **74**: 7607-7612.

Throckmorton, K., Lim, F.Y., Kontoyiannis, D., Zheng, W., and Keller, N.P. (2015) Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*. In *Environ Microbiol*. Madison, U.o.W.-. (ed).

Xu, X., Liu, L., Zhang, F., Wang, W., Li, J., Guo, L. et al. (2014) Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte Pestalotiopsis fici. *Chembiochem* **15**: 284-292.

CHAPTER 3

Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*

This work has been published as:
Throckmorton, K., Lim, F.Y., Kontoyiannis, D., Zheng, W., and Keller, N.P. (2015) **Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*.** *Environ Microbiol*.
It has been slightly modified from the published version.

Author contributions
KT, FYL, and WZ made the *A*. *fumigatus* mutants used in this study. KT carried out the chemical analysis of all mutants and the comparative genomic analysis. The lab of DK did all *Drosophila* assays. KT and FYL made the figures. KT, FYL, and NPK wrote the text.

Abstract

Filamentous fungi are renowned for the production of bioactive secondary metabolites. Typically, one distinct metabolite is generated from a specific secondary metabolite cluster. Here, we characterize the newly described trypacidin (*tpc*) cluster in the opportunistic human pathogen *Aspergillus fumigatus*. We find that this cluster as well as the previously characterized endocrocin (*enc*) cluster, both contribute to the production of the spore metabolite endocrocin. Whereas trypacidin is eliminated when only *tpc* cluster genes are deleted, endocrocin production is only eliminated when both the *tpc* and *enc* non-reducing polyketide synthase-encoding genes, *tpcC* and *encA*, respectively, are deleted. EncC, an anthrone oxidase, converts the product released from EncA to endocrocin as a final product. In contrast, endocrocin synthesis by the *tpc* cluster likely results from incomplete catalysis by TpcK (a putative decarboxylase), as its deletion results in a nearly 10-fold increase in endocrocin production. We suggest endocrocin is likely a shunt product in all related non-reducing polyketide synthase clusters containing homologs of TpcK and TpcL (a putative anthrone oxidase), e.g. geodin and monodictyphenone. This finding represents an unusual example of two physically discrete secondary metabolite clusters generating the same natural product in one fungal species by distinct routes.

Introduction

Fungi are well known for the ability to synthesize bioactive secondary metabolites (SMs) of diverse chemical structure and complexity. The dominant SM-producing taxa belong to several genera of filamentous ascomycetes. SMs are generally thought to contribute to the fitness of filamentous fungi, primarily as protection from abiotic stress (e.g. melanins) and/or for niche securement or defense in conflicts with other microbes or insects (e.g. toxins), reviewed in

(Rohlfs and Churchill, 2011; Scherlach et al., 2013). The same or similar metabolites can also serve as virulence factors in pathogenic fungi, including plant (Friesen et al., 2008) and animal (Dagenais and Keller, 2009) pathogens.

Aspergillus fumigatus is the primary causative agent of invasive aspergillosis (IA) among immunocompromised individuals (Brakhage, 2005; Maschmeyer et al., 2007). This species is notorious for producing a plethora of SMs (e.g. gliotoxin, helvolic acid, hexadehydroastechrome, trypacidin, endocrocin, neosartoricin, and fumagillin) shown to exhibit immunomodulatory and/or cytotoxic properties that are thought to facilitate the pathogenicity of this fungus (Amitani et al., 1995; Bok et al., 2005; Bok et al., 2006; Spikes et al., 2008; Lodeiro et al., 2009; Gauthier et al., 2012; Berthier et al., 2013; Chooi et al., 2013; Yin et al., 2013), although their ecological role may relate to their antimicrobial activity (Carberry et al., 2012; Kang et al., 2013) and UV protective capacity (Allam and Abd El-Zaher, 2012).

Secondary metabolism is spatially and temporally regulated in fungi with metabolites produced in different compartments of the fungal thallus (Lim and Keller, 2014; Kistler and Broz, 2015). Mycelial SMs such as gliotoxin and helvolic acid have been extensively studied over the last decades (Amitani et al., 1995; Bok et al., 2006; Spikes et al., 2008; Kwon-Chung and Sugui, 2009). However, with the exception of dihydroxynaphthalene (DHN) melanin (Jahn et al., 1997; Sugui et al., 2007; Bayry et al., 2014), the spore-borne SMs are relatively understudied, despite spores being crucial to the fungus for dispersal and colonization of new substrates.

Recent studies have identified two spore metabolites that may contribute to the virulence of *A. fumigatus.* Endocrocin, a spore-borne product of a non-reducing polyketide synthase (NR-PKS) in *A. fumigatus*, was demonstrated to inhibit neutrophil migration both *in vitro* and *in vivo*, and deletion of the endocrocin polyketide synthase-encoding *encA* yielded a less pathogenic strain using the *Toll*-deficient *Drosophila* model (Berthier et al., 2013). Another spore metabolite, trypacidin, was shown to exhibit cytotoxic properties against both the A549 pulmonary adenocarcinoma cell line and human bronchial epithelial cells (Gauthier et al., 2012). Interestingly, trypacidin, first isolated from *A. fumigatus*, was identified as an anti-protozoal agent (Balan et al., 1963; Turner, 1965) and the conidia of *A. fumigatus* have multiple reported defenses against predation by soil amoebae (Van Waeyenberghe et al., 2013; Hillmann et al., 2015).

Despite being isolated more than 50 years ago, the genes responsible for trypacidin production in *A. fumigatus* have not been identified. Trypacidin is similar in structure to polyketides that belong to the clade of NR-PKSs involved in the synthesis of anthraquinone-derivatives, which include geodin in *A. terreus* (Nielsen et al., 2013), monodictyphenone in *A. nidulans* (Chiang et al., 2010; Sanchez et al., 2011; Simpson, 2012), and endocrocin in *A. fumigatus* (Lim et al., 2012). These NR-PKSs lack the canonical thioesterase (TE) domain responsible for releasing the nascent products from the enzyme. Instead, a separate gene encoding a metallo-β-lactamase type thioesterase (MβL-TE) is located adjacent to these NR-PKSs (Awakawa et al., 2009). These two enzymes catalyze the first steps in the biosynthesis of polyketides of this TE-less NR-PKS clade, discussed in recent reviews (Ahuja et al., 2012; Chooi and Tang, 2012). Genome mining in *A. fumigatus* identified three NR-PKSs, two of which have

been characterized and found to synthesize endocrocin (Lim et al., 2012; Berthier et al., 2013) and neosartoricin (Chooi et al., 2013).

Here we found that the third TE-less NR-PKS is responsible for the biosynthesis of the polyketide trypacidin. The thirteen-gene trypacidin (*tpc*) cluster is nearly identical to the thirteen-gene *ged* NR-PKS cluster and several *tpc* genes also share homology with both the twelve-gene *mdp* cluster and the four-gene *enc* cluster. Unexpectedly, we found that both the *tpc* and *enc* clusters contribute to the production of endocrocin, with the former cluster ultimately producing trypacidin. Through *tpc* and *enc* gene deletions coupled with metabolite profiling of these strains, we characterized two distinct routes to endocrocin production differentiated by early enzymatic steps in their respective pathways.

Results

Identification of the trypacidin-producing gene cluster

The cluster containing the NR-PKS-encoding gene AFUA_4G14560 (hereafter referred to as *tpcC*) was first identified as one of the SM clusters regulated by the global regulator of secondary metabolism, LaeA (Perrin et al., 2007). Analysis of the product template (PT) domain of TpcC shows that it belongs to the same clade of NR-PKS as the monodictyphenone PKS of *A. nidulans* (Chiang et al., 2010), the endocrocin PKS of *A. fumigatus* (Lim et al., 2012), and the recently described geodin PKS of *A. terreus* (Nielsen et al., 2013) (Fig. 1A and Table S1). The PT domains of PKSs determine the regioselectivity of cyclization during the synthesis of the polyketide backbone (Crawford et al., 2009) and have been used for phylogenetic analyses. The phylogenetic group of PKSs to which TpcC belongs has previously been designated group V1

(Li et al., 2011), a subdivision of group V (Li et al., 2010) which, with a few exceptions, make a C6-C11 bond during the initial cyclization (Ahuja et al., 2012).

The set of genes surrounding *tpcC* also has significant identity to the monodictyphenone, endocrocin, and geodin biosynthetic clusters (Fig. 1B). Genetic characterization of the geodin-producing cluster from *A. terreus* (Nielsen et al., 2013) highlighted a high degree of similarity between these two clusters and their final products (Fig. 1B and Table S2). Given this similarity, we hypothesized that *tpcC* was the most likely candidate to initiate trypacidin synthesis in *A. fumigatus*. While initial *in silico* predictions of the trypacidin biosynthetic gene cluster spanned a 66 kb region with 31 genes (AFUA_4G14420 - AFUA_4G14730) (Khaldi et al., 2010; Medema et al., 2011; Inglis et al., 2013), homology with the geodin cluster suggests the *tpc* cluster consists of a 25 kb region with thirteen genes: AFUA_4G14580 through AFUA_4G14460 which we term *tpcA-M* (Fig. 1B and Table S2).

In order to confirm the assignment of these genes to trypacidin biosynthesis, we created deletion mutants of *tpcC*, encoding the NR-PKS, and *tpcB*, encoding the MβL-TE (Table S2). Deletion of either gene resulted in loss of trypacidin and pathway intermediates including questin as determined by comparison to purified standards (Fig. 2A-C, Fig. S2 and 3). Additionally, we identified two Zn_2Cys_6 transcription factor-encoding genes in the region of the *tpc* cluster, namely, *tpcE* and AFUA_4G14590 (Table S2). Deletion of *tpcE* but not AFUA_4G14590 resulted in loss of trypacidin and questin (Fig. 2D-F, Fig. S4). TpcE is a homolog of AflR, the sterigmatocystin/aflatoxin Zn_2Cys_6 transcription factor (Woloshuk et al., 1994). The coactivator AflS (formerly called AflJ) is also required for sterigmatocystin/aflatoxin synthesis (Woloshuk et

al., 1994; Ehrlich et al., 2012) and enhances AflR activity. An *aflS* homolog, *tpcD*, is located next to *tpcE*, and its deletion also resulted in loss of trypacidin (Fig. 2D-F, Fig. S4). We also confirmed regulation of the trypacidin and endocrocin clusters by LaeA and the conidiation-specific transcription factor BrlA (Mah and Yu, 2006) (Fig. 2G-I, Fig. S5) as reported or suggested by earlier studies (Perrin et al., 2007; Gauthier et al., 2012). Both production of metabolites and gene expression showed BrlA to more tightly regulate the clusters than LaeA.

We present a likely biosynthetic pathway for trypacidin synthesis (Fig. 3) based on these data and proposed and characterized polyketide biosynthetic pathways, primarily geodin, aflatoxin, monodictyphenone, and prenyl xanthones (Henry and Townsend, 2005b, a; Chiang et al., 2010; Sanchez et al., 2011; Simpson, 2012; Nielsen et al., 2013).

<u>Endocrocin is produced by both the tpc and enc clusters through distinct routes</u>

Because trypacidin and endocrocin are both regulated by LaeA and BrlA (Fig. 2G-I) and the PKSs involved in their syntheses belong to the same phylogenetic clade, group V (Li et al., 2010), we thought it possible that their biosyntheses might have a unique interrelationship. Therefore, we assessed the AF293 strain, which produces both endocrocin and trypacidin, to determine if any such relationship existed. In contrast to the Δ*encA* mutation in the CEA10 background that displayed a complete loss of endocrocin (Lim et al., 2012), we found that the AF293 Δ*encA* mutant surprisingly still produced endocrocin, although at lower levels than wild-type (Fig. 4A-C). Extracts of the Δ*tpcC* strain also appeared to show a reduction, but not elimination, of endocrocin relative to the wild-type control (Fig. 4A-C). Previous work had characterized the *enc* cluster in strains derived from the *A. fumigatus* isolate CEA10 (Lim et al.,

2012). However, this strain does not produce trypacidin or its intermediates, despite the presence of the *tpc* cluster (Fig. S6). Examination of the nucleotide sequence of the *tpc* cluster in CEA10 uncovered a single-nucleotide insertion in the fourth exon of *tpcC*. This is predicted to result in a frameshift and premature termination codon, which would explain the inability of this isolate to produce these metabolites (Table S3).

Given this, as well as the similarity of the genes encoding the initial enzyme activities of the two clusters (Lim et al., 2012), we hypothesized that the source of endocrocin in the AF293 Δ*encA* mutant might be the *tpc* cluster (Fig. 4G). We therefore created several Δ*encA*/Δ*tpcC* double deletion mutants and found that all of them exhibited a complete loss of endocrocin production (Fig. 4A-C). This observation confirms that the trypacidin biosynthetic cluster contributes to the production of endocrocin and indicates that, in wild-type AF293, endocrocin is derived from both the *tpc* and *enc* clusters. Conversely, it does not appear that the *enc* cluster contributes to the trypacidin pathway as demonstrated by complete loss of trypacidin and its precursors in *tpc* cluster deletants despite the presence of an intact *enc* cluster (Fig. 2A-F). This might suggest that the pathways are independent or perhaps compartmentalized within the cell.

Both trypacidin and endocrocin have been reported as spore metabolites. To more thoroughly assess tissue localization of endocrocin produced by either pathway, spore (Fig. 4D), conidiophore (containing fallen spores and some surface mycelium, Fig. 4E), and mycelia/agar (Fig. 4F) fractions were analyzed for endocrocin in both the Δ*encA* mutant and Δ*tpcC* mutants. Supporting earlier studies, endocrocin was primarily localized to the spore fraction (Fig. 4D). Lesser amounts were found in conidiophore and mycelial fractions (Fig. 4E-F). Interestingly, a

small amount of endocrocin was detectable in the mycelial fraction of the Δ*encA* mutant but not of the Δ*tpcC* mutant, suggesting more specific localization of the endocrocin produced by the *enc* cluster to the spore (Fig. 4F).

We asked whether the redundancy of endocrocin production could be attributed to utilization of an early intermediate common to both pathways by downstream enzymes in the endocrocin pathway. Specifically, we asked whether the monooxygenase EncC (the first decorating enzyme in the endocrocin biosynthetic pathway after release of the backbone from EncA by EncB, Fig. 4G) might use early pathway intermediates derived from either the *enc* or *tpc* clusters as a substrate to generate endocrocin. EncC possesses similarity to HypC, an anthrone oxidase involved in a similar conversion in aflatoxin biosynthesis (Ehrlich et al., 2010), and deletion of EncC results in loss of endocrocin in the CEA10 background (Lim et al., 2012). If EncC could utilize precursors from early steps in trypacidin synthesis, we hypothesized that deletion of *encC* would result in complete loss of endocrocin, much like the Δ*encA*/Δ*tpcC* mutants. However, endocrocin production in the AF293 Δ*encC* mutant was similar to that of the AF293 Δ*encA* mutant showing a reduction but not elimination of endocrocin (Fig. 4A-C). While these results cannot rule out the possibility of intermediate sharing between the two pathways, they are consistent with the hypothesis that endocrocin is produced as a shunt product of the trypacidin pathway.

The hypothetical endocrocin shunt would likely occur from the early steps in the trypacidin pathway. Previous work had shown that deletion of *mdpH* in *A. nidulans* resulted in accumulation of endocrocin (Chiang et al., 2010), and *tpcK* and *tpcL* are homologous to the 5’

and 3' halves of *mdpH*, respectively (Fig. 1B and Fig. 5A). TpcL also shows homology to the anthrone oxidases HypC and EncC, whereas TpcK is a putative decarboxylase, based on the prediction of the same activity for MdpH (Chiang et al., 2010). We therefore made single and double deletions of *tpcK* and *tpcL* to address if either of the encoding enzyme activities contributed to endocrocin production.

Deletion of *tpcK* increased endocrocin production nearly 10-fold and exhibited concomitant loss of downstream trypacidin pathway metabolites, emodin and questin (Fig. 5B-D, Fig. S7). This observation is consistent with the proposed role of TpcK as a decarboxylase (Fig. 2G and 4G) and similar to the result of *mdpH* deletion in *A. nidulans*. Deletion of *tpcL* resulted in modest decrease in the production of questin, but an approximately 2-fold increase in endocrocin and emodin (Fig. 5B-D, Fig. S7). This suggests that the proposed enzymatic step for TpcL (Fig. 2G and 4G) might be catalyzed by other oxidases in the genome or may occur spontaneously. Double deletion of *tpcK* and *tpcL* resulted in a similar phenotype to the Δ*tpcK* mutant, suggesting that *tpcK* is upstream of *tpcL* in the trypacidin pathway. A third putative anthrone oxidase-encoding gene, AFUA_4g09250, was found in the *A. fumigatus* genome, however, its deletion had no effect on the levels of metabolites in the endocrocin or trypacidin pathways (Fig. 5B-D, Fig. S7).

We also examined the metabolic output of deletions of *tpcK*, *tpcL*, *tpcK*/*L*, or Afu4g09250 in a Δ*encC* background. These mutants yielded approximately half the amount of endocrocin relative to the amount produced by their respective single deletants in a wild-type *encC* background, but no other significant changes (Fig. 5E-G). Thus, it appears that endocrocin

production from the trypacidin pathway results from incomplete conversion of atrochrysone carboxylic acid to downstream products, specifically by TpcK. Taken together, these results strongly support independent mechanisms leading to endocrocin biosynthesis by the *tpc* and *enc* clusters where one cluster yields endocrocin as a shunt metabolite (*tpc*) and one as an end metabolite (*enc*) (Fig. 4G).

The contribution of the tpc and enc clusters to virulence in Toll-deficient Drosophila melanogaster

An earlier study had shown that *encA* deletion resulted in a less virulent strain using the *Toll*-deficient *Drosophila* model of IA (Lionakis et al., 2005; Berthier et al., 2013). We again used this model to compare the pathogenicity of both single Δ*encA* and Δ*tpcC* deletion mutants and the double deletant to wild-type (Fig. S8). The results show that the Δ*encA* mutant has significantly decreased virulence in two of three trials, while the Δ*tpcC* mutant has no consistent effect. The Δ*encA*/Δ*tpcC* mutant is also significantly impaired in virulence in two of the three trials.

Discussion

The *tpc* cluster constitutes the fifth described example of a LaeA- and BrlA-regulated spore SM cluster alongside DHN melanin, fumigaclavines, endocrocin, and the fumiquinazolines (Coyle et al., 2007; Perrin et al., 2007; Twumasi-Boateng et al., 2009; Lim et al., 2012; Upadhyay et al., 2013; Lim et al., 2014). Although trypacidin was initially reported to be isolated from the mycelium (Turner, 1965), it was later found to localize to the conidia (Parker and Jenner, 1968; Gauthier et al., 2012), a finding supported by our work. In an unexpected twist, we

find that both the *tpc* and *enc* clusters contribute to the synthesis of endocrocin. To our knowledge, only two prior studies have described two physically discrete clusters contributing to the synthesis of the same metabolite in one fungal species (Forseth et al., 2013; Guo et al., 2015). This finding is in contrast to the recently discovered phenomenon of superclusters, in which a single co-regulated cluster can produce two distinct metabolites (e.g. fumagillin and pseurotin) (Wiemann et al., 2013), and thus expands our concept of SM plasticity in filamentous fungi.

The *tpc* and *enc* clusters, along with the *ged* and *mdp* clusters, contain a TE-less NR-PKS coupled with a MβL-TE, characteristic of this phylogenetic clade of PKSs (Fig. 1A). In three of these four clusters, these genes are present as a divergently oriented pair (Fig. 1B). In addition to maintaining this gene pair, the early biosynthetic steps proposed for all four metabolites are very similar, consisting of the PKS-MβL step followed by anthrone oxidase activity (Fig. 2G and 4G) (Chiang et al., 2010; Lim et al., 2012; Nielsen et al., 2013). Interestingly, *in vitro* expression of the geodin PKS-encoding gene, *gedC* (formerly called ACAS), and MβL-TE-encoding gene, *gedB* (formerly called ACTE), yielded endocrocin, emodin, and their precursors (Awakawa et al., 2009). Also, as mentioned, deletion of *mdpH* of the monodictyphenone cluster resulted in accumulation of endocrocin (Chiang et al., 2010). Coupling these data with our finding that the *tpc* cluster can produce endocrocin, likely as a result of incomplete TpcK activity, we suggest that endocrocin production may be generated as a shunt product from NR-PKS/MβL-TE clusters containing TpcK and TpcL homologs or produced directly by those clusters containing a TpcL (but not TpcK) homolog. The anthrone oxidase and decarboxylase activities are key in endocrocin metabolism, either directing to (TpcL/EncC) or diverting from (TpcK/MdpH) its synthesis.

Previous studies (Awakawa et al., 2009) suggest that the NR-PKS/MβL-TE pairing alone could produce some amount of endocrocin which may be oxidized non-enzymatically from an anthrone precursor (Fig. 4G). Deletion of the anthrone oxidase HypC in *A. parasiticus* revealed a similar situation where conversion of norsolorinic acid anthrone to norsolorinic acid also occurred non-enzymatically (Ehrlich et al., 2010). Deletion of *tpcL* in our study still allowed for emodin accumulation, presumably through processing of emodin anthrone either non-enzymatically or by another oxidase encoded in the genome. Interestingly, endocrocin production is reported not only from numerous fungi (Kurobane et al., 1979; Räisänen et al., 2000; Yang et al., 2013), but also insects (Kikuchi et al., 2011) and plants (Jan et al., 2015) which may reflect yet other biosynthetic pathways leading to endocrocin.

From where might the additional genes in the more complex NR-PKS/MβL-TE clusters derive? The *tpc* cluster was referenced in a 2007 study of the evolution of the aflatoxin gene cluster (Carbone et al., 2007). Aflatoxin is a polyketide generated from a large NR-PKS cluster containing several genes with significant similarity to *tpc* cluster genes, in addition to the already noted *hypC*, which was not yet discovered in 2007. In this paper, the authors identify modules, consisting primarily of highly correlated gene pairs that they propose are duplicated and evolved in secondary metabolite gene clusters. At the time, five genes in what we now know as the *tpc* cluster were speculated as having arisen from a hypothetical cluster ancestral to the aflatoxin cluster. These were *tpcD/E*, which are homologous to *aflR*/*S*; *tpcG*/*I,* which are homologous to *aflX/Y*; and *tpcC,* homologous to *aflC*. Expansion of an ancestral *enc*-like cluster can be envisioned by 'capture' of such modules. It is possible that the current *enc* and *tpc* clusters both

arose from a small ancestral cluster in *A. fumigatus*. Alternatively, the *enc* cluster, comprising fewer genes than the *tpc* cluster, may have arisen from loss of modules from an ancestral *tpc*-like cluster.

Though previous predictions of the trypacidin cluster range from AFUA_4G14420-AFUA_4G14730 (Khaldi et al., 2010; Medema et al., 2011; Inglis et al., 2013), the cluster is nearly identical to the geodin-producing cluster in *A. terreus* (Nielsen et al., 2013). Twelve of the thirteen genes in these clusters are homologous, with only the methyltransferase-encoding *tpcH* and halogenase-encoding *gedL* unique to each cluster. These two activities are reflected in the chemical structures of their respective end metabolites (Fig. 1B). Interestingly, these unique genes are on either side of a five-gene inversion, suggesting a possible mechanism for introduction and maintenance of variation in an ancestral SM-producing gene cluster. Variations of fungal SM clusters resulting from inversions and deletions are common and contribute to the production of biosynthetically related but different end metabolites, e.g. the aflatoxin and sterigmatocystin gene clusters of *A. flavus* and *A. nidulans* (Hodges et al., 2000; Lee et al., 2006; Carbone et al., 2007) and the sirodesmin and gliotoxin clusters of *Leptosphaeria maculans* and *A. fumigatus* (Gardiner et al., 2004; Fox and Howlett, 2008).

The *enc* and *tpc* clusters are both regulated by LaeA and BrlA, while the *tpc* cluster is further regulated by the TpcD/E pair. The acquisition or maintenance of *tpcD*/*E* might allow for more dynamic production of trypacidin and its intermediates, many of which are bioactive (Fujimoto et al., 1999; Ohashi et al., 1999; Choi et al., 2007; Shrimali et al., 2013). The presence of cluster-specific transcription factors potentiates a more nuanced activation of the cluster

perhaps to time the production of the metabolites. As suggested for aflatoxin biosynthesis (Meyers et al., 1998; Du et al., 2007; Roze et al., 2007), and consistent with observations of the geodin cluster (Nielsen et al., 2013), the AflR/S homologs may enable the uncoupling of early and late phases of the biosynthetic pathway. In total, the coordinate timing of this and the conidiation-specific level of regulation, involving BrlA, may facilitate the specific accumulation of the end metabolite, trypacidin, into the spore.

For this study we utilized *A. fumigatus* strain AF293, which produces trypacidin and endocrocin, however, the previous study characterizing the *enc* cluster used CEA10-derived strains, which do not produce trypacidin. There are other examples of differential production of a metabolite between strains of a given species, including *A. fumigatus* (Frisvad et al., 2009). In some cases, genetic defects responsible for this differential metabolite production have been described in the cluster. For example, in AF293, a point mutation in *ftmD* prevents the strain from producing fumitremorgins in contrast to CEA10 and *A. fumigatus* strain BM939 (Kato et al., 2013). Further examples have been suggested in *A. niger* (Andersen et al., 2011), and such defects are thought to have resulted in the loss of aflatoxin production in some *A. oryzae* strains (Lee et al., 2006; Kiyota et al., 2011).

In summary, the newly described trypacidin cluster and the endocrocin cluster both produce endocrocin through distinct routes in *A. fumigatus*. The trypacidin and endocrocin clusters are both regulated by LaeA and BrlA, and their end metabolites are localized to the asexual spores. Relatively little is known about the biological significance of endocrocin and trypacidin synthesis, but we speculate that they may have a role in protection of the spore from

abiotic stressors. Endocrocin and precursors of trypacidin (e.g. questin and emodin) are pigmented anthraquinones. Anthraquinones have been shown to provide protection from UV radiation to the producing organism (Nybakken et al., 2004) and may contribute to protection to *A. fumigatus* as well. Other than the trypanocidal activity of trypacidin (Balan et al., 1963), additional studies are required to confirm any biological function of these metabolites and the redundancy in endocrocin production for the fungus.

Experimental procedures

Fungal strains and growth conditions

All strains used in this study are listed in Table S4. Mutants were derived from *pyrG1* or *argB1*/*pyrG1* auxotrophic backgrounds of *Aspergillus fumigatus* strain AF293 (Osherov et al., 2001; Xue et al., 2004). Stocks of each strain were stored in 30% (v/v) glycerol in 0.01% (v/v) TWEEN-80 at -80 °C. Strains were activated and grown on solid glucose minimal medium (GMM) (Shimizu and Keller, 2001) with appropriate supplements at 37 °C for 3 days for spore collection. For *pyrG1* auxotrophs, 5.2 mM uridine and 5 mM uracil were added as supplements, and for *argB1* auxotrophs, 5.7 mM L-arginine was added as a supplement. Spores were collected in 0.01% (v/v) TWEEN-80 and enumerated using a hemacytometer. For isolation of genomic DNA for PCR and Southern blot, 10 mL of liquid minimal medium (LMM) (Shimizu and Keller, 2001) with yeast extract was inoculated with spores from solid medium and grown overnight at 37 °C. For assessment of metabolite production, 5 µL of a suspension of 2 x 10^6 spores/mL of each strain were point-inoculated on solid GMM and grown at 29 °C for 120 hours without light selection. For tissue-specific metabolite analysis, this time was increased to 192 hours. The conidial fraction of the culture was collected by tapping the plate with a spatula with the lid

down, the conidiophore/surface mycelial fraction (also containing residual spores) was collected by scraping into 5 mL 0.01% (v/v) TWEEN-80, and the mycelial fraction was collected by taking three 1.5 cm diameter cores after rinsing the plate with ddH_2O.

<u>Construction of mutants</u>

The mutants used in this study were created using the double-joint PCR method (Szewczyk et al., 2006). Genomic DNA was isolated as previously described (Shimizu and Keller, 2001). Generation of protoplasts and transformation were performed as previously described (Szewczyk et al., 2006). Primers (Table S5) were designed to amplify 800-1000 bp flanks with 20 bp of overlap with the selection cassette-containing plasmid (Table S6) with DNASTAR's Lasergene12 suite (DNASTAR; Madison, WI) and Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012). For deletion strains, the plasmid used was pJW24 (Calvo et al., 2004), encoding the *A. parasiticus pyrG* gene, or pJMP4 (Sekonyela et al., 2013), encoding the *A. fumigatus argB* gene. The *A. fumigatus akuA* gene was deleted in AF293.6 (Xue et al., 2004) using *A. parasiticus pyrG* amplified from pJW24 (Calvo et al., 2004) to produce TFYL43.2. To generate TFYL44.1, *A. parasiticus pyrG* was removed from the *akuA* locus by transforming strain TFYL43.2 with single-joint PCR product containing 0.5 - 1 kb gDNA fragments flanking the *pyrG* sequence. TFYL45.1 was generated from TFYL43.2 by replacing the *A. parasiticus pyrG* at the *akuA* locus with firefly luciferase constitutively expressed by the *A. nidulans gpdA* promoter amplified from pJMP147 (Palmer and Keller, unpublished). For both TFYL44.1 and TFYL45.1, transformants were plated onto sorbitol minimal media (SMM) supplemented with the above-mentioned concentrations of uridine, uracil, arginine, and sub inhibitory concentration (0.75 mg/ml) of 5-fluoroorotic acid (5-FOA). Both *encA* and *encC* were deleted in strain

TFYL45.1 using *A. parasiticus pyrG* amplified from pJW24 to generate strains TFYL68.1 and TFYL69.1 respectively. *tpcC* was deleted in strain TFYL45.1 using *A. fumigatus argB* amplified from pJMP4 to generate strain TFYL71.1. To generate the double *encA* and *tpcC* deletion mutant (TFYL72.1), *tpcC* was deleted from strain TFYL68.1 using *A. fumigatus argB* amplified from pJMP4 (Sekonyela et al., 2013). Both TFYL68.1 and TFYL69.1, arginine auxotrophs, were transformed with pJMP4 to generate the prototrophs TFYL73.1 and TFYL74.1, respectively, while TFYL71.1, a uridine and uracil auxotroph, was transformed with pKJA12.1 (Affeldt and Keller, unpublished) to generate the prototroph TFYL76.1. Transformants were screened for proper integration of the construct first by PCR, then by Southern blot (Fig. S1). For all auxotrophic mutants, maintenance of the mutant allele(s) was confirmed via PCR after complementation to prototrophy. The wildtype control used for all strains created in the Δ*akuA* background, TFYL81.1, was created from TFYL43.2 by complementation to prototrophy with *A. fumigatus argB*. This strain did not differ significantly from AF293 in production of any of the metabolites assessed in this study.

Comparative genomic analysis

Sequence information and genome annotation were derived from the *Aspergillus* Genome Database aspgd.org (Cerqueira et al., 2014). Comparison of orthologous genes in various fungal species was carried out using the Sybil browser hosted at AspGD. The NCBI's BLAST (Altschul et al., 1990) was used to manually supplement the annotations of AspGD and the predictions made by the Sybil browser. DNASTAR's Lasergene12 suite (including SeqBuilder, SeqMan Pro, and EditSeq) was used for sequence analysis (DNASTAR; Madison, WI). For phylogenetic tree construction Clustal Omega (Sievers et al., 2011) was used to align a large set of NR-PKSs

(Table S1) and UGene was used to extract the PT domain based on the canonical PT domain from the *A. parasiticus* PksA. These PT domains were then realigned and used as the basis for the phylogenetic tree with Clustal Omega and FigTree (Rambaut, 2007).

SM extraction and Chromatography

For tissue-specific metabolite extraction from the conidial fraction (see above), the spores were collected in 6 mL of 50:50 ethyl acetate:methanol. The conidiophore fraction was extracted with an equal volume (5 mL) of ethyl acetate. The mycelial fraction and all other samples were processed as follows: A 1.5 cm core was taken from each plate grown as described above, and homogenized in 2.5 mL of double-distilled water (ddH_2O) in a glass extraction vial with a Kinematica homogenizer. To each vial, 2.5 mL of ethyl acetate was added, vortexed, and allowed to steep overnight. The vials were then centrifuged at 1100 *g* for 3 minutes and 1.9 mL of the organic layer was transferred to a new vial. This extract was allowed to evaporate at room temperature. For thin-layer chromatography (TLC), the extracts were dissolved in 50 µL ethyl acetate and 10 µL was spotted to a 250 µm silica gel plate (Whatman, Cat#15-4410-222) which was then placed in a mobile phase of toluene, ethyl acetate, and formic acid in a 50:40:7 volume ratio. The plate was then imaged at 366 and 254 nm using a FOTO/Analyst® Investigator gel imaging system (Fotodyne Inc). For high performance liquid chromatography (HPLC) analysis, the extracts were resuspended in 50% (v/v) acetonitrile in ddH_2O and filtered using a 0.2 µm syringe filter. The samples were analyzed using a Perkin Elmer Flexar Instrument as previously described (Lim et al., 2014), but with 1% (v/v) formic acid in solvents A (ddH_2O) and B (acetonitrile), a 1.5 mL/min flow rate, and using the following modified program: equilibration in 20% B for 10 min before each sample injection, 20% B for 2 min, ramp to 50% B over 20

min, ramp to 100% B over 1 min, hold at 100% B for 5 min. Strains were analyzed in duplicate, the peak areas for each compound averaged and calculated as a proportion of their respective wild-type controls, either AF293 or TFYL81.1 (Table S4). TFYL81.1 did not differ significantly from AF293 in production of any of the metabolites assessed in this study. Statistical significance of pairwise comparisons between the metabolite profiles of different strains was assessed by multiple t-tests with a threshold of 0.05 and the Holm-Sidak correction for multiple comparisons in Prism (GraphPad Software, Inc.). Trypacidin standard was obtained from Dr. Olivier Puel (UMR1331 ToxAlim, French National Institute for Agricultural Research, INP, UPS, Toulouse, France), and endocrocin, emodin, and questin standards from Dr. Clay Wang (USC School of Pharmacy, 1985 Zonal Avenue, Los Angeles, California, U.S.A.) (Fig. S2).

Drosophila assays

Flies were generated by crossing a *Drosophila melanogaster* line carrying a thermosensitive allele of *Toll* (Tl r632) with a line carrying a null allele of *Toll* (Tl I-RXA) (Lionakis et al., 2005). Five to seven day old adult female *Toll*-deficient flies were used for all experiments. Twenty to thirty-two flies were infected with each *A. fumigatus* strain. *A. fumigatus* isolates were grown on yeast extract agar glucose at 25 °C. Conidia were collected from two day old cultures in sterile 0.9% (w/v) saline, enumerated using a hemacytometer, and adjusted to a concentration of 5×10^7 conidia/mL. The dorsal thorax of CO_2 anaesthetized flies was punctured using a sterile 10 µm needle dipped in this conidial suspension. As a negative control, flies were punctured with a needle that was not dipped in conidial suspension. Flies were monitored daily for survival over seven days. Death within three hours was considered to be a result of the injection procedure and these flies were not included in subsequent analysis. The flies were

housed in a 29 °C incubator to maximize expression of the Tl r632 phenotype (Lionakis et al., 2005). The *Toll*-deficient flies were transferred into fresh vials every three days. Each experiment was repeated three times on different days at a consistent time of day to eliminate variability due to circadian rhythm.

Acknowledgements

We would like to thank Dr. Clay Wang for endocrocin, emodin, and questin standards, and Dr. Olivier Puel for the trypacidin standard. We would like thank Nathaniel Albert for carrying out the *Drosophila* assays, and Dr. Philipp Wiemann for general consultation and proofreading. We would also like to thank our funding sources: grants from the American Asthma Foundation (11-0137), the USDA Hatch Formula Fund (WIS01710), and the NIH (R01 AI065728) to N.P.K. and the University of Wisconsin-Madison Genetics Department (5T32GM07133). The authors declare no conflict of interest.

Figures

Figure 1: The relationships between characterized group V1 gene clusters.

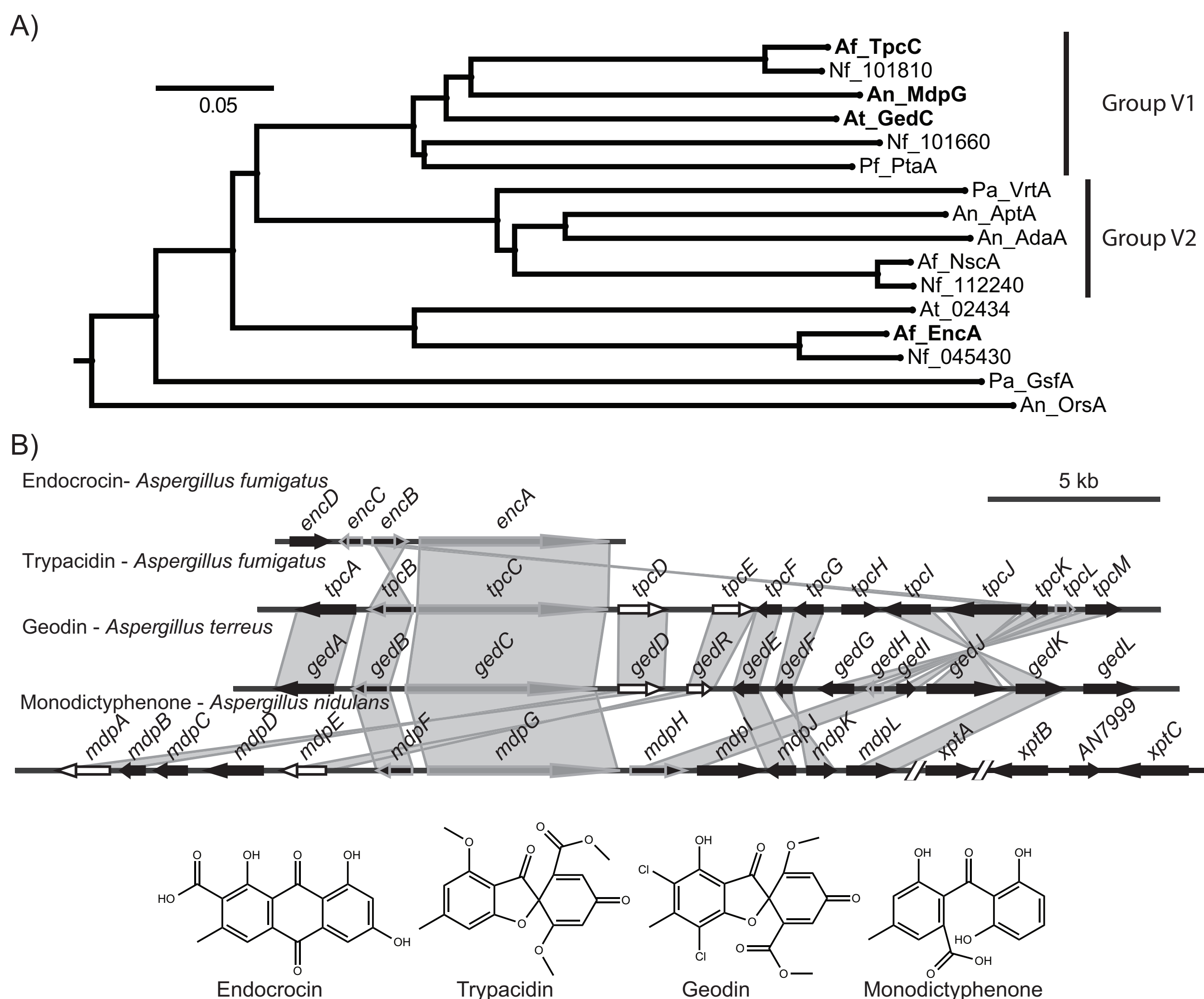


Figure 1: A) Phylogenetic tree showing the relationships between the group V PKSs in the aspergilli based on their product template (PT) domains. This is an excerpt of the tree described in the methods, chosen to include the group V PKSs, with the orsellinic acid PKS, OrsA, as an outgroup. The tree shows that these PKSs fall into two main subgroups, group V1 and V2 (Li et al., 2011), with EncA basal to both. The PKSs corresponding to clusters shown in Fig. 1B are in bold. B) The trypacidin gene cluster and its homologous characterized group V1 non-reducing polyketide synthase (NR-PKS) gene clusters in the aspergilli. The producing species are listed

above each cluster and their respective end products are displayed below. All but the endocrocin cluster also have homologs of the paired transcriptional regulators of the aflatoxin gene cluster, AflR/AflS, shown in white. PKS genes are shown in gray, genes conserved in all clusters shown with a gray outline, and all other genes in black. Shaded regions connect homologous genes, and noncontiguous segments are separated by slashes.

Figure 2: HPLC analysis of tpc cluster mutants and regulators.

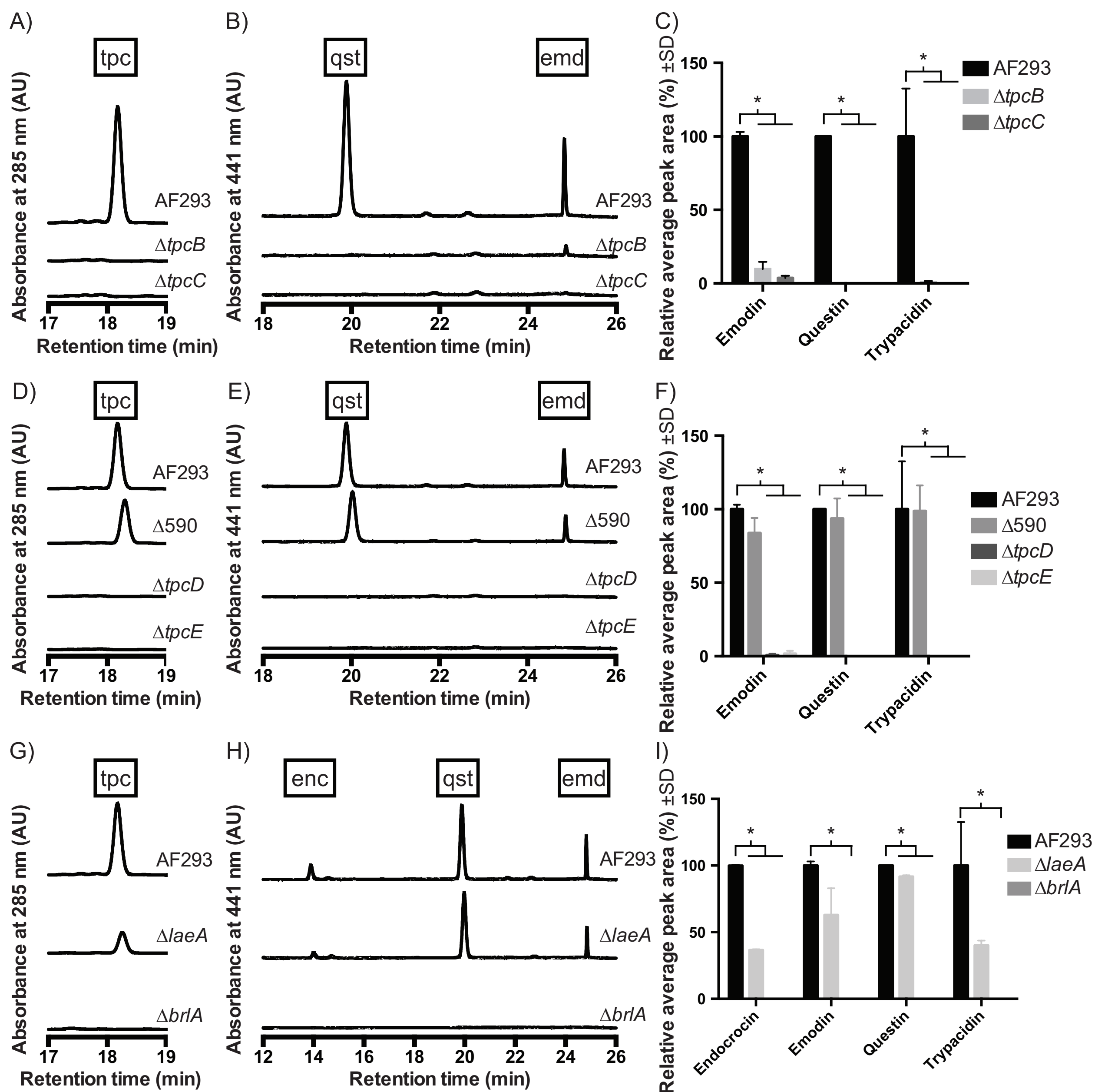


Figure 2: Representative chromatograms from High-Performance Liquid Chromatography (HPLC) analysis of wild-type (WT, AF293) and trypacidin cluster mutants (Δ*tpcB* and Δ*tpcC*, A-C; Δ*tpcE*, Δ*tpcD*, and ΔAFUA_4g14590 (Δ590), D-F; Δ*laeA* and Δ*brlA*, G-I) at 285 nm (A, D, and G) and 441 nm (B, E, and H). These data are trimmed to show only the relevant ranges of retention times for detection of trypacidin (tpc), its precursors questin and emodin (qst and emd,

respectively), and endocrocin (enc). Quantification of duplicate analyses is presented in the bar graphs (C, F, and I) and asterisks represent statistical significance of the indicated pairwise comparisons with $\alpha=0.05$.

Figure 3: Proposed biosynthetic pathway of trypacidin.

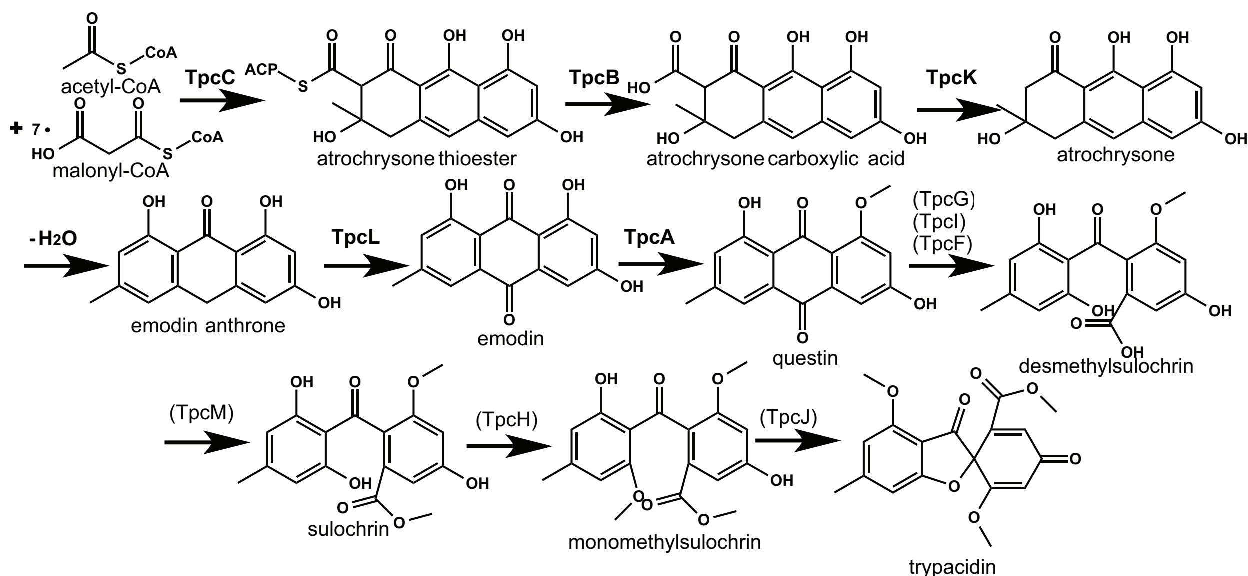


Figure 3: The proposed biosynthetic pathway of trypacidin. The protein putatively catalyzing each step is indicated over the arrow, and the dehydration potentially occurs spontaneously. Genes encoding the enzymes in bold were deleted in this study, whereas the assignment of roles for enzymes that are in regular font and parentheses were made based on comparison to similar characterized biosynthetic pathways. The assignment of the methyltransferases TpcA, TpcH, and TpcM, was determined by analysis of a Δ*tpcA* strain (Fig. S3) and comparison to the proposed cluster and biosynthetic pathway for geodin (Nielsen et al., 2013). Note the conversion of questin to desmethylsulochrin, and similar reactions in monodictyphenone and aflatoxin biosynthesis, have multiple proposed mechanisms (Henry and Townsend, 2005b, a; Simpson, 2012).

Figure 4: HPLC analysis of enc and tpc PKS deletants and model for an endocrocin shunt.

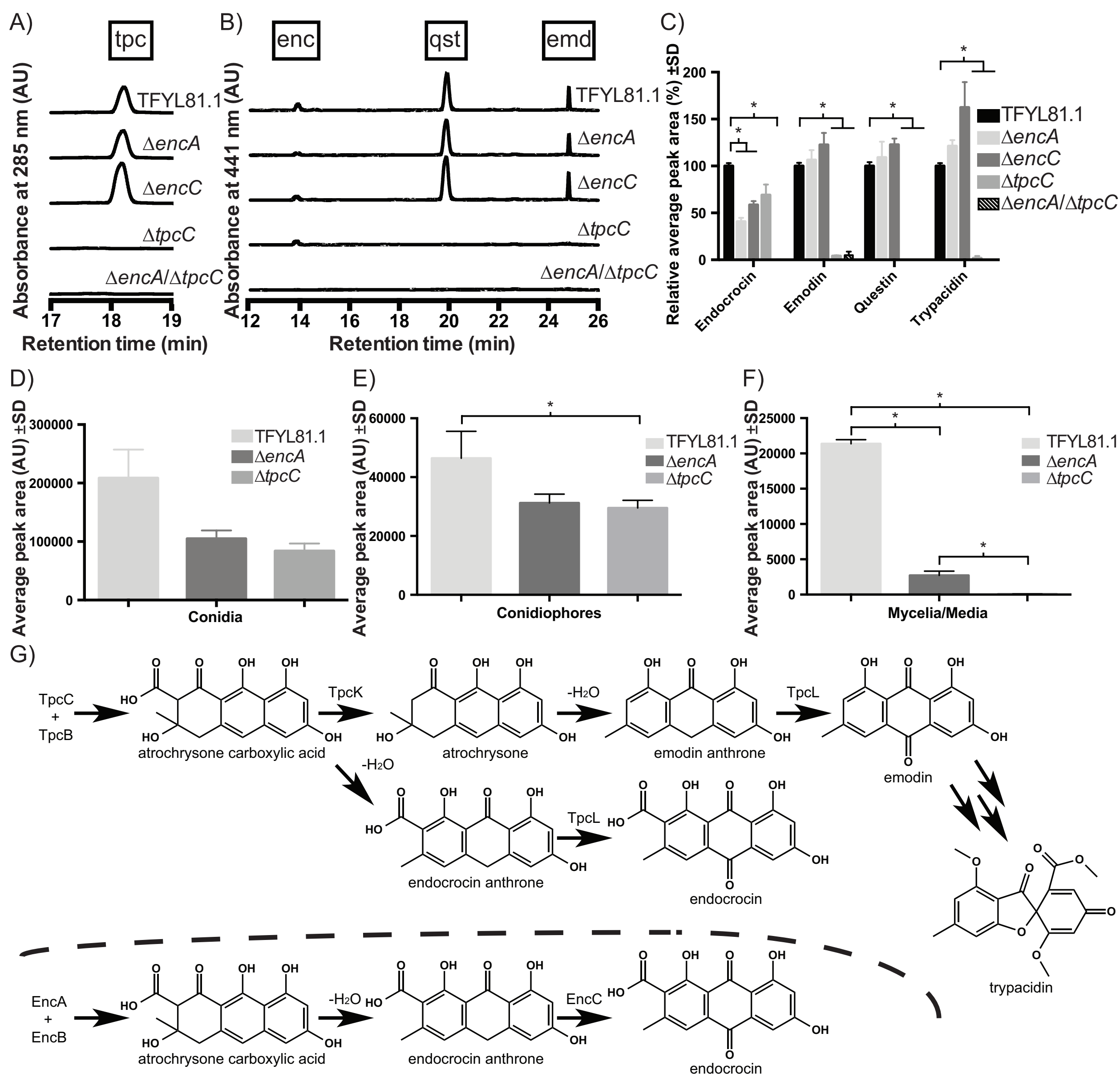


Figure 4: Representative chromatograms from HPLC analysis of WT (TFYL81.1), Δ*encA*, Δ*encC*, Δ*tpcC*, Δ*encA*/Δ*tpcC* mutants at 285 nm (A) and 441 nm (B). These data are trimmed to show only the relevant ranges of retention times for detection of trypacidin (tpc), its precursors questin and emodin (qst and emd, respectively), and endocrocin (enc). Quantification of duplicate analyses is presented in the bar graph (C). Quantification of triplicate analyses of the

tissue-specific production and localization of endocrocin (detected at 441 nm) to the conidia (D), conidiophores (E), and mycelia/media (F) assessed by HPLC analysis of WT (TFYL81.1), Δ*encA*, and Δ*tpcC* mutants. Asterisks (C-F) represent statistical significance of the indicated pairwise comparisons with $\alpha=0.05$. Comparison of the early steps of the predicted trypacidin and endocrocin biosynthetic pathways (D). The protein putatively catalyzing each step is indicated over the arrow, and the dehydration potentially occurs spontaneously.

Figure 5: HPLC analysis of mutants of mdpH homologs in the tpc and enc pathways.

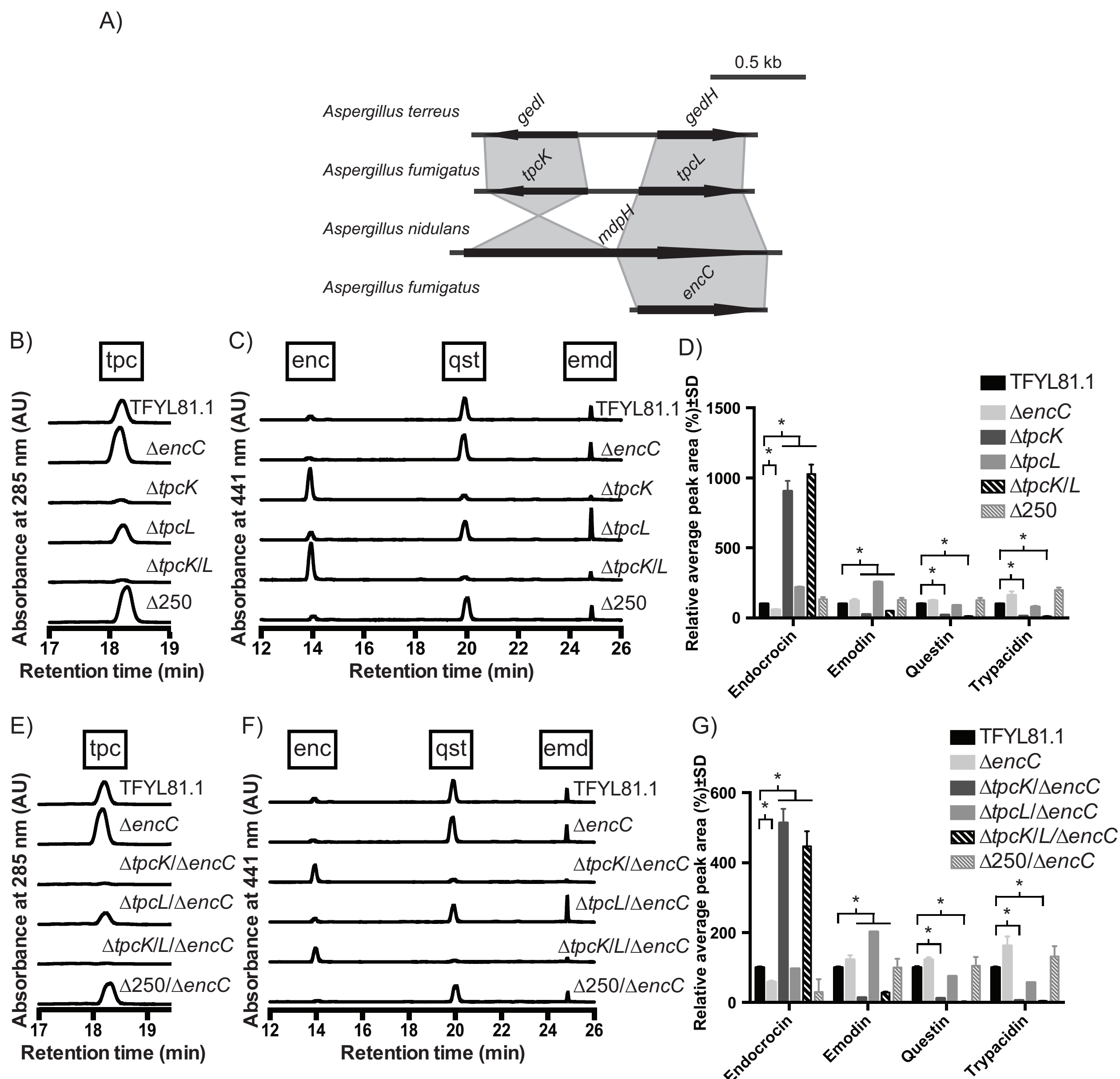


Figure 5: A diagram of the homologs of *tpcK* and *tpcL* in *A. fumigatus*, *A. terreus*, and *A. nidulans* (A). Representative chromatograms from HPLC analysis of WT (TFYL81.1) and endocrocin and trypacidin cluster mutants (Δ*encC*, Δ*tpcK*, Δ*tpcL*, Δ*tpcK*/ΔtpcL, and ΔAFUA_4g09250 (Δ250), B-D; Δ*encC*, Δ*tpcK*/Δ*encC*, Δ*tpcL*/Δ*encC*, Δ*tpcK*/ΔtpcL/Δ*encC*, and ΔAFUA_4g09250/Δ*encC* (Δ250/Δ*encC*), E-G) at 285 nm (B and E) and 441 nm (C and F).

These data are trimmed to show only the relevant ranges of retention times for detection of trypacidin (tpc), its precursors questin and emodin (qst and emd, respectively), and endocrocin (enc). Quantification of duplicate analyses is presented in the bar graphs (D and G) and asterisks represent statistical significance of the indicated pairwise comparisons with $\alpha=0.05$.

References

Ahuja, M., Chiang, Y.M., Chang, S.L., Praseuth, M.B., Entwistle, R., Sanchez, J.F. et al. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J Am Chem Soc* **134**: 8212-8221.

Allam, N.G., and Abd El-Zaher, E.H.F. (2012) Protective role of *Aspergillus fumigatus* melanin against ultraviolet (UV) irradiation and *Bjerkandera adusta* melanin as a candidate vaccine against systemic candidiasis. *African J Biotech* **11**: 6566-6577.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Amitani, R., Taylor, G., Elezis, E.N., Llewellyn-Jones, C., Mitchell, J., Kuze, F. et al. (1995) Purification and characterization of factors produced by *Aspergillus fumigatus* which affect human ciliated respiratory epithelium. *Infect Immun* **63**: 3266-3271.

Andersen, M.R., Salazar, M.P., Schaap, P.J., van de Vondervoort, P.J., Culley, D., Thykaer, J. et al. (2011) Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome Res* **21**: 885-897.

Awakawa, T., Yokota, K., Funa, N., Doi, F., Mori, N., Watanabe, H., and Horinouchi, S. (2009) Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem Biol* **16**: 613-623.

Balan, J., Ebringer, L., Nemec, P., Kovac, S., and Dobias, J. (1963) Antiprotozoal Antibiotics. II. Isolation and characterization of trypacidin, a new antibiotic, active against *Trypanosoma cruzi* and *Toxoplasma gondii*. *J Antibiot (Tokyo)* **16**: 157-160.

Bayry, J., Beaussart, A., Dufrêne, Y.F., Sharma, M., Bansal, K., Kniemeyer, O. et al. (2014) Surface structure characterization of *Aspergillus fumigatus* conidia mutated in the melanin synthesis pathway and their human cellular immune response. *Infect Immun* **82**: 3141-3153.

Berthier, E., Lim, F.Y., Deng, Q., Guo, C.J., Kontoyiannis, D.P., Wang, C.C. et al. (2013) Low-volume toolbox for the discovery of immunosuppressive fungal secondary metabolites. *PLoS Pathog* **9**: e1003289.

Bok, J.W., Balajee, S.A., Marr, K.A., Andes, D., Nielsen, K.F., Frisvad, J.C., and Keller, N.P. (2005) LaeA, a regulator of morphogenetic fungal virulence factors. *Eukaryot Cell* **4**: 1574-1582.

Bok, J.W., Chung, D., Balajee, S.A., Marr, K.A., Andes, D., Nielsen, K.F. et al. (2006) GliZ, a transcriptional regulator of gliotoxin biosynthesis, contributes to *Aspergillus fumigatus* virulence. *Infect Immun* **74**: 6761-6768.

Brakhage, A.A. (2005) Systemic fungal infections caused by *Aspergillus* species: epidemiology, infection process and virulence determinants. *Curr Drug Targets* **6**: 875-886.

Calvo, A.M., Bok, J., Brooks, W., and Keller, N.P. (2004) *veA* is required for toxin and sclerotial production in *Aspergillus parasiticus*. *Appl Environ Microbiol* **70**: 4733-4739.

Carberry, S., Molloy, E., Hammel, S., O'Keeffe, G., Jones, G.W., Kavanagh, K., and Doyle, S. (2012) Gliotoxin effects on fungal growth: mechanisms and exploitation. *Fungal Genet Biol* **49**: 302-312.

Carbone, I., Ramirez-Prado, J.H., Jakobek, J.L., and Horn, B.W. (2007) Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. *BMC Evol Biol* **7**: 111.

Cerqueira, G.C., Arnaud, M.B., Inglis, D.O., Skrzypek, M.S., Binkley, G., Simison, M. et al. (2014) The *Aspergillus* Genome Database: multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Res* **42**: D705-710.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microbiol* **76**: 2067-2074.

Choi, S.G., Kim, J., Sung, N.D., Son, K.H., Cheon, H.G., Kim, K.R., and Kwon, B.M. (2007) Anthraquinones, Cdc25B phosphatase inhibitors, isolated from the roots of *Polygonum multiflorum* Thunb. *Nat Prod Res* **21**: 487-493.

Chooi, Y.H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: from genes to molecules. *J Org Chem* **77**: 9933-9953.

Chooi, Y.H., Fang, J., Liu, H., Filler, S.G., Wang, P., and Tang, Y. (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **15**: 780-783.

Coyle, C.M., Kenaley, S.C., Rittenour, W.R., and Panaccione, D.G. (2007) Association of ergot alkaloids with conidiation in *Aspergillus fumigatus*. *Mycologia* **99**: 804-811.

Crawford, J.M., Korman, T.P., Labonte, J.W., Vagstad, A.L., Hill, E.A., Kamari-Bidkorpeh, O. et al. (2009) Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* **461**: 1139-1143.

Dagenais, T.R., and Keller, N.P. (2009) Pathogenesis of *Aspergillus fumigatus* in Invasive Aspergillosis. *Clin Microbiol Rev* **22**: 447-465.

Du, W., Obrian, G.R., and Payne, G.A. (2007) Function and regulation of *aflJ* in the accumulation of aflatoxin early pathway intermediate in *Aspergillus flavus*. *Food Addit Contam* **24**: 1043-1050.

Ehrlich, K.C., Li, P., Scharfenstein, L., and Chang, P.K. (2010) HypC, the anthrone oxidase involved in aflatoxin biosynthesis. *Appl Environ Microbiol* **76**: 3374-3377.

Ehrlich, K.C., Mack, B.M., Wei, Q., Li, P., Roze, L.V., Dazzo, F. et al. (2012) Association with AflR in endosomes reveals new functions for AflJ in aflatoxin biosynthesis. *Toxins (Basel)* **4**: 1582-1600.

Forseth, R.R., Amaike, S., Schwenk, D., Affeldt, K.J., Hoffmeister, D., Schroeder, F.C., and Keller, N.P. (2013) Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. *Angew Chem Int Ed Engl* **52**: 1590-1594.

Fox, E.M., and Howlett, B.J. (2008) Biosynthetic gene clusters for epipolythiodioxopiperazines in filamentous fungi. *Mycol Res* **112**: 162-169.

Friesen, T.L., Faris, J.D., Solomon, P.S., and Oliver, R.P. (2008) Host-specific toxins: effectors of necrotrophic pathogenicity. *Cell Microbiol* **10**: 1421-1428.

Frisvad, J.C., Rank, C., Nielsen, K.F., and Larsen, T.O. (2009) Metabolomics of *Aspergillus fumigatus*. *Med Mycol* **47** **Suppl 1**: S53-71.

Fujimoto, H., Fujimaki, T., Okuyama, E., and Yamazaki, M. (1999) Immunomodulatory constituents from an ascomycete, *Microascus tardifaciens*. *Chem Pharm Bull (Tokyo)* **47**: 1426-1432.

Gardiner, D.M., Cozijnsen, A.J., Wilson, L.M., Pedras, M.S., and Howlett, B.J. (2004) The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. *Mol Microbiol* **53**: 1307-1318.

Gauthier, T., Wang, X., Sifuentes Dos Santos, J., Fysikopoulos, A., Tadrist, S., Canlet, C. et al. (2012) Trypacidin, a spore-borne toxin from *Aspergillus fumigatus*, is cytotoxic to lung cells. *PLoS One* **7**: e29906.

Guo, C., Sun, W., Bruno, K.S., Oakley, B.R., Keller, N.P., and Wang, C. (2015) Spatial regulation of a common precursor from two distinct genes generates metabolite diversity In. *Chem. Sci.*

Henry, K.M., and Townsend, C.A. (2005a) Synthesis and fate of o-carboxybenzophenones in the biosynthesis of aflatoxin. *J Am Chem Soc* **127**: 3300-3309.

Henry, K.M., and Townsend, C.A. (2005b) Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. *J Am Chem Soc* **127**: 3724-3733.

Hillmann, F., Novohradská, S., Mattern, D.J., Forberger, T., Heinekamp, T., Westermann, M. et al. (2015) Virulence determinants of the human pathogenic fungus *Aspergillus fumigatus* protect against soil amoeba predation. *Environ Microbiol*.

Hodges, R.L., Kelkar, H.S., Xuei, X., Skatrud, P.L., Keller, N.P., Adams, T.H. et al. (2000) Characterization of an echinocandin B-producing strain blocked for sterigmatocystin biosynthesis reveals a translocation in the *stcW* gene of the aflatoxin biosynthetic pathway. *J Ind Microbiol Biotechnol* **25**: 333-341.

Inglis, D.O., Binkley, J., Skrzypek, M.S., Arnaud, M.B., Cerqueira, G.C., Shah, P. et al. (2013) Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. *BMC Microbiol* **13**: 91.

Jahn, B., Koch, A., Schmidt, A., Wanner, G., Gehringer, H., Bhakdi, S., and Brakhage, A.A. (1997) Isolation and characterization of a pigmentless-conidium mutant of *Aspergillus fumigatus* with altered conidial surface and reduced virulence. *Infect Immun* **65**: 5110-5117.

Jan, S., Kamili, A.N., Parray, J.A., and Bedi, Y.S. (2015) Differential response of terpenes and anthraquinones derivatives in *Rumex dentatus* and *Lavandula officinalis* to harsh winters across north-western Himalaya. *Nat Prod Res*: 1-5.

Kang, D., Son, G.H., Park, H.M., Kim, J., Choi, J.N., Kim, H.Y. et al. (2013) Culture condition-dependent metabolite profiling of *Aspergillus fumigatus* with antifungal activity. *Fungal Biol* **117**: 211-219.

Kato, N., Suzuki, H., Okumura, H., Takahashi, S., and Osada, H. (2013) A point mutation in *ftmD* blocks the fumitremorgin biosynthetic pathway in *Aspergillus fumigatus* strain Af293. *Biosci Biotechnol Biochem* **77**: 1061-1067.

Khaldi, N., Seifuddin, F.T., Turner, G., Haft, D., Nierman, W.C., Wolfe, K.H., and Fedorova, N.D. (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet Biol* **47**: 736-741.

Kikuchi, N., Teshiba, M., Tsutsumi, T., Fudou, R., Nagasawa, H., and Sakuda, S. (2011) Endocrocin and its derivatives from the Japanese mealybug *Planococcus kraunhiae*. *Biosci Biotechnol Biochem* **75**: 764-767.

Kistler, H.C., and Broz, K. (2015) Cellular compartmentalization of secondary metabolism. *Front Microbiol* **6**: 68.

Kiyota, T., Hamada, R., Sakamoto, K., Iwashita, K., Yamada, O., and Mikami, S. (2011) Aflatoxin non-productivity of *Aspergillus oryzae* caused by loss of function in the *aflJ* gene product. *J Biosci Bioeng* **111**: 512-517.

Koressaar, T., and Remm, M. (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**: 1289-1291.

Kurobane, I., Vining, L.C., and McInnes, A.G. (1979) Biosynthetic relationships among the secalonic acids. Isolation of emodin, endocrocin and secalonic acids from *Pyrenochaeta terrestris* and *Aspergillus aculeatus*. *J Antibiot (Tokyo)* **32**: 1256-1266.

Kwon-Chung, K.J., and Sugui, J.A. (2009) What do we know about the role of gliotoxin in the pathobiology of *Aspergillus fumigatus*? *Med Mycol* **47 Suppl 1**: S97-103.

Lee, Y.H., Tominaga, M., Hayashi, R., Sakamoto, K., Yamada, O., and Akita, O. (2006) *Aspergillus oryzae* strains with a large deletion of the aflatoxin biosynthetic homologous gene cluster differentiated by chromosomal breakage. *Appl Microbiol Biotechnol* **72**: 339-345.

Li, Y., Image, I.I., Xu, W., Image, I., and Tang, Y. (2010) Classification, prediction, and verification of the regioselectivity of fungal polyketide synthase product template domains. *J Biol Chem* **285**: 22764-22773.

Li, Y., Chooi, Y.H., Sheng, Y., Valentine, J.S., and Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**: 15773-15785.

Lim, F.Y., and Keller, N.P. (2014) Spatial and temporal control of fungal natural product synthesis. *Nat Prod Rep* **31**: 1277-1286.

Lim, F.Y., Ames, B., Walsh, C.T., and Keller, N.P. (2014) Co-ordination between BrlA regulation and secretion of the oxidoreductase FmqD directs selective accumulation of fumiquinazoline C to conidial tissues in *Aspergillus fumigatus*. *Cell Microbiol* **16**: 1267-1283.

Lim, F.Y., Hou, Y., Chen, Y., Oh, J.H., Lee, I., Bugni, T.S., and Keller, N.P. (2012) Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* **78**: 4117-4125.

Lionakis, M.S., Lewis, R.E., May, G.S., Wiederhold, N.P., Albert, N.D., Halder, G., and Kontoyiannis, D.P. (2005) *Toll*-deficient *Drosophila* flies as a fast, high-throughput model for the study of antifungal drug efficacy against invasive aspergillosis and *Aspergillus* virulence. *J Infect Dis* **191**: 1188-1195.

Lodeiro, S., Xiong, Q., Wilson, W.K., Ivanova, Y., Smith, M.L., May, G.S., and Matsuda, S.P. (2009) Protostadienol biosynthesis and metabolism in the pathogenic fungus *Aspergillus fumigatus*. *Org Lett* **11**: 1241-1244.

Mah, J.H., and Yu, J.H. (2006) Upstream and downstream regulation of asexual development in *Aspergillus fumigatus*. *Eukaryot Cell* **5**: 1585-1595.

Maschmeyer, G., Haas, A., and Cornely, O.A. (2007) Invasive aspergillosis: epidemiology, diagnosis and management in immunocompromised patients. *Drugs* **67**: 1567-1601.

Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A. et al. (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **39**: W339-346.

Meyers, D.M., Obrian, G., Du, W.L., Bhatnagar, D., and Payne, G.A. (1998) Characterization of *aflJ*, a gene required for conversion of pathway intermediates to aflatoxin. *Appl Environ Microbiol* **64**: 3713-3717.

Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., Larsen, T.O., and Mortensen, U.H. (2013) Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. *PLoS One* **8**: e72871.

Nybakken, L., Solhaug, K.A., Bilger, W., and Gauslaa, Y. (2004) The lichens *Xanthoria elegans* and *Cetraria islandica* maintain a high protection against UV-B radiation in Arctic habitats. *Oecologia* **140**: 211-216.

Ohashi, H., Ueno, A., Nakao, T., Ito, J., Kimura, K., Ishikawa, M. et al. (1999) Effects of ortho-substituent groups of sulochrin on inhibitory activity to eosinophil degranulation. *Bioorg Med Chem Lett* **9**: 1945-1948.

Osherov, N., Kontoyiannis, D.P., Romans, A., and May, G.S. (2001) Resistance to itraconazole in *Aspergillus nidulans* and *Aspergillus fumigatus* is conferred by extra copies of the *A. nidulans* P-450 14alpha-demethylase gene, *pdmA*. *J Antimicrob Chemother* **48**: 75-81.

Parker, G.F., and Jenner, P.C. (1968) Distribution of trypacidin in cultures of *Aspergillus fumigatus*. *Appl Microbiol* **16**: 1251-1252.

Perrin, R.M., Fedorova, N.D., Bok, J.W., Cramer, R.A., Wortman, J.R., Kim, H.S. et al. (2007) Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. *PLoS Pathog* **3**: e50.

Rambaut, A. (2007). FigTree. URL http://tree.bio.ed.ac.uk/software/figtree/

Rohlfs, M., and Churchill, A.C. (2011) Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genet Biol* **48**: 23-34.

Roze, L.V., Arthur, A.E., Hong, S.Y., Chanda, A., and Linz, J.E. (2007) The initiation and pattern of spread of histone H4 acetylation parallel the order of transcriptional activation of genes in the aflatoxin cluster. *Mol Microbiol* **66**: 713-726.

Räisänen, R., Björk, H., and Hynninen, P.H. (2000) Two-dimensional TLC separation and mass spectrometric identification of anthraquinones isolated from the fungus *Dermocybe sanguinea*. *Z Naturforsch C* **55**: 195-202.

Sanchez, J.F., Entwistle, R., Hung, J.H., Yaegashi, J., Jain, S., Chiang, Y.M. et al. (2011) Genome-based deletion analysis reveals the prenyl xanthone biosynthesis pathway in Aspergillus nidulans. *J Am Chem Soc* **133**: 4010-4017.

Scherlach, K., Graupner, K., and Hertweck, C. (2013) Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu Rev Microbiol* **67**: 375-397.

Sekonyela, R., Palmer, J.M., Bok, J.W., Jain, S., Berthier, E., Forseth, R. et al. (2013) RsmA regulates *Aspergillus fumigatus* gliotoxin cluster metabolites including cyclo(L-Phe-L-Ser), a potential new diagnostic marker for invasive aspergillosis. *PLoS One* **8**: e62591.

Shimizu, K., and Keller, N.P. (2001) Genetic involvement of a cAMP-dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in *Aspergillus nidulans*. *Genetics* **157**: 591-600.

Shrimali, D., Shanmugam, M.K., Kumar, A.P., Zhang, J., Tan, B.K., Ahn, K.S., and Sethi, G. (2013) Targeted abrogation of diverse signal transduction cascades by emodin for the treatment of inflammatory disorders and cancer. *Cancer Lett* **341**: 139-149.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W. et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Simpson, T.J. (2012) Genetic and biosynthetic studies of the fungal prenylated xanthone shamixanthone and related metabolites in *Aspergillus* spp. revisited. *Chembiochem* **13**: 1680-1688.

Spikes, S., Xu, R., Nguyen, C.K., Chamilos, G., Kontoyiannis, D.P., Jacobson, R.H. et al. (2008) Gliotoxin production in *Aspergillus fumigatus* contributes to host-specific differences in virulence. *J Infect Dis* **197**: 479-486.

Sugui, J.A., Pardo, J., Chang, Y.C., Müllbacher, A., Zarember, K.A., Galvez, E.M. et al. (2007) Role of *laeA* in the regulation of *alb1*, *gliP*, conidial morphology, and virulence in *Aspergillus fumigatus*. *Eukaryot Cell* **6**: 1552-1561.

Szewczyk, E., Nayak, T., Oakley, C.E., Edgerton, H., Xiong, Y., Taheri-Talesh, N. et al. (2006) Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nat Protoc* **1**: 3111-3120.

Turner, W.B. (1965) The production of trypacidin and monomethylsulochrin by *Aspergillus fumigatus*. *J Chem Soc Perkin 1*: 6658-6659.

Twumasi-Boateng, K., Yu, Y., Chen, D., Gravelat, F.N., Nierman, W.C., and Sheppard, D.C. (2009) Transcriptional profiling identifies a role for BrlA in the response to nitrogen depletion and for StuA in the regulation of secondary metabolite clusters in *Aspergillus fumigatus*. *Eukaryot Cell* **8**: 104-115.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.

Upadhyay, S., Torres, G., and Lin, X. (2013) Laccases involved in 1,8-dihydroxynaphthalene melanin biosynthesis in *Aspergillus fumigatus* are regulated by developmental factors and copper homeostasis. *Eukaryot Cell* **12**: 1641-1652.

Van Waeyenberghe, L., Baré, J., Pasmans, F., Claeys, M., Bert, W., Haesebrouck, F. et al. (2013) Interaction of *Aspergillus fumigatus* conidia with *Acanthamoeba castellanii* parallels macrophage-fungus interactions. *Environ Microbiol Rep* **5**: 819-824.

Wiemann, P., Guo, C.J., Palmer, J.M., Sekonyela, R., Wang, C.C., and Keller, N.P. (2013) Prototype of an intertwined secondary-metabolite supercluster. *Proc Natl Acad Sci U S A* **110**: 17065-17070.

Woloshuk, C.P., Foutz, K.R., Brewer, J.F., Bhatnagar, D., Cleveland, T.E., and Payne, G.A. (1994) Molecular characterization of *aflR*, a regulatory locus for aflatoxin biosynthesis. *Appl Environ Microbiol* **60**: 2408-2414.

Xue, T., Nguyen, C.K., Romans, A., Kontoyiannis, D.P., and May, G.S. (2004) Isogenic auxotrophic mutant strains in the *Aspergillus fumigatus* genome reference strain AF293. *Arch Microbiol* **182**: 346-353.

Yang, X.-L., Awakawa, T., Wakimoto, T., and Abe, I. (2013) Induced production of novel prenyldepside and coumarins in endophytic fungi *Pestalotiopsis acaciae*. *Tetrahedron Letters* **54**: 5814-5817.

Yin, W.B., Baccile, J.A., Bok, J.W., Chen, Y., Keller, N.P., and Schroeder, F.C. (2013) A nonribosomal peptide synthetase-derived iron(III) complex from the pathogenic fungus *Aspergillus fumigatus*. *J Am Chem Soc* **135**: 2064-2067.

Appendix

Supplemental tables

Table S1: Product template domains used in the construction of the group V1 phylogenetic tree.

Gene	PT domain sequence
An_OrsA	RQMDVLLDELEASAAQVPFHAPTLPVASTVLGRIVRPGEQGVFDAN YLRRHTREPVAFLDAVRACETEGLIPDRSFAVEIGPHPICISLMATCLQ SAKINAWPSLRRGGDDWQSVSSTLAAAHSAQLPVAWSEFHKDHLDT VRLISDLPTYAFDLKTFWHSYKTPAAAVSAASATPSTTGLSRLASTTL HAVEKLQREEGKILGTFTVDLSDPKLAKAICGHVVDESAICPASIFID MAYTAAVFLEQENGAGAALNTYELSSLEMHSPLVL
An_AptA	EQVDPVLDQLTRVAETVHFKAPSIPIISPLLRSVVFDGKTINSSYLIRAT REPVHFAGAIEAAQDLGMVNDKTVWVDVGPHPICASFVRSLIPKAR VASSCRRNEDNYATMAKNLVALHLAGCTPVWDEYFRANEKAYNLL TLPKYAWNDVNYWIQYIGTWTLDKAHLKYTGTNGPPQVKPSSSALR TSLIHEIIEETIGEETATLKTVSDLQHPEFLEAVHGHRMNNCGVATSSI WTDMSLTVGEYLYNKLAPGSKVHMNVGELEVLHATVA
An_MdpG	SQTEAILDDFEEAAKDGVLFRAPNMPVISPLLGKVIFDDKTITAKYMR RATRETVNFLSALEMAQTFSTVDEETVWVEIGPHPVCMGFVNATLP AVNETVASMKRGEDNWVTLCNSLTALHCAGVPIEWNEYQRPFEKG LRLLDLPTYAWNDKTYWIQYNGDWALTKGNTFYDAEKSLKAQQTG QLASVPSGLRTSTVQQIIEESFNGSAGKVVMQSDMMQPDFLDAAHG HKMNGCGVVTSSIHGDIGFTLGGYLYKNLVKGGKAPDMNMANLVV LRGLVA
Ed_WdPks1	SQVDPILADFEKLASSVNYHPPRVPVISPLLSDVVSVGGVFDAFYLSR HCRKTVDFVGGLSAGMSTATISDTSLWLEVGGHPLCASMIKSCLSVP TLATMRRDEDPWKIISASMAGLYTAGKSLNWDAFHKENESLRVLND LPFYGFDEKNYWLQYTGDWLLYKGDYPKAIAPAPAAAAAARPAKA RKYLSTSVQGIVSEEVKGKTVTIVAESDFAHPKLFPVIAGHLVNGSGL CPSTLYADMAYTLGQLGVGLLKPGEKVDINIGTMDNPAPLLL
Gl_Pks1	AQVDAILDDFEKLAASVRFGSANIPLISPLLGRPLSEGESIDPTYLRNH AREAVNFLAGLTSSQELGIIDEKTVWIEVGPHSVCSGMVKSTFGATTI TAPSLRRNEEAYKTISASLCALQNAGMNIDWNEYHRDFSDSVRLLSL PTYSYDDKNYWIQYEGDWSLNKGRKAAGALPAPEIVKSSLSTTSVH EVVKETISDHVATVVTETDIARPDLRPLVTGHVVNGSYLCPSSLYGD MAMTVCEYGYKLIDPDVKDLVMDITHMEVPKSLIA

Cl_Pks1	AQVDPILDDFETLARAVSFERPQVPIISPLLGKMVESEPVNAAYLRNH ARDAVNFLGGLVHAQQSGSIDEKTVWLEVGPHPVCANFVKSSFGISS VAVPSLRRNEPTYKILSSTLCTLHTAGVNLDWDEFHRDFTDCTRLLD LPTYSFDEKNYWLQYTGDWCLTKNRGPSAVKAPLQIEPARPKLATTS IHAITNEDIKDDIAIIDTETNLSRPDTRPLVEGHLCNGTPLCPSTLYAD MAMTVADYAYKTLRPGTENIGLNVANVEVPKTLIF
Nsp_Pks1	AQVEPILDDFEKLARSVRFYEPKIPVISPLHGKLLEGEPINPAYLRNHA RDAVNFLGGLVSAQSSGVIDEKTAWLEVGPHPVCANMVKAAFGAT TIAVPTLRRNEATYKTLSATLCTLHSAGLNIDWNEFHRDFDASVRLL DLPAYSFDLKNYWLQYTGEWCLTKNRAALPAPTKAIEPAKPKLSTTT VQKITKEEVKGEIAILETESDLAEENLRKICLGHRCNGTTLTPSTLYAD MAETICEYGYKLLRPETEDIGMNVAHMEVTKTLIF
Fs_PksN	AQVEPILEEFESKAQAVRFYSPSIPYVSPLQSEVISKGCVLNGAYLAK ACRHTVNFQAALEAATNSGICSEKTIWLEVGPHPACSGMVKGTLGS NSATTSTLRKDTDTWKVLSTTLEQLHLAGIQIDWNEYHRGFEKFNRV LELPRYSWDLKNYWIMYRNDFCLTKGEGTQPATAVSDVPKPIKYISP SLQRIIEEEHSNESSSILVESEVHDPKLVSVFTGHRVNGAQLCPSSLWG DIAMQLTHYMMEKRSLSTKNTGMDVTRMQIQKSLIL
Fg_Aur1	AQVDPILSDLDTAASRVTFHSPQIPVLCALDSSVISPGNHGVIGPLHLQ RHCRETVNFEGALHAAEREKIINKTSTLWIEIGPHVVCSTFLKSSLGPS TPTIASLRRNDDCWKVLADGLSSLYSSGLTIDWNEYHRDFKASHQVL RLPCYSWEHKNYWIQYKYDWSLTKGDPPIAPNSSVEAVSALSTPSVQ KILQETSLDQVLTIVAETDLASPLLSEVAQGHRVNGVKVCTSSVYAD VGLTLGKYILDNYRTDLEGYAVDVHGIEVHKPLLL
An_WA	AQVEPILEDLEKALQGITFNKPSVPFVSALLGEVITEAGSNILNAEYLV RHCRETVNFLSAFEAVRNAKLGGDQTLWLEVGPHTVCSGMVKATL GPQTTTMASLRRDEDTWKVLSNSLSSLYLAGVDINWKQYHQDFSSS HRVLPLPTYKWDLKNYWIPYRNNFCLTKGSSMSAASASLQPTFLTTS AQRVVESRDDGLTATVVVHNDIADPDLNRVIQGHKVNGAALCPSSL YADSAQTLAEYLIEKYKPELKGSGLDVCNVTVPKPLIA
Af_PksP	AQVDPILESFEESAQGVIFHEPAVPFVSALNGEVITESNYSVLGPTYM VKHCREAVNFLGALEATRHAKLMDDATLWVEVGSHPICSGMIKSTF GPQATTVPSLRRDDDPWKILSNSLSTLHLAGVELNWKEFHQDFSSAH EVLELPRYGWDLKNYWIPYTNNFCLTKGGPVTAEVSAPKSTFLTTAA QKIVECREDGNTATLVVENNIAEPELNRVIQGHKVNGVALTPSSLYA DIAQTLVDHLITKYKPEYQGLGLDVCDMTVPKPLIA
Ff_Bik1	SQVDPILDDLEELASQVGFHEPKLPIVSPLLRTLLTGDTLGPQYIRRHC RETVDFLGAIKMAESQGIMDRSGMCIEIGAHPILTRMVKSIIGQDFRC LASLRRKEDHFKTLADSLCALHLAGFSVNWDEYHRDFASSRNVLQL PKYSWQLANYWMQYKYSWCLTKGDAPVENGPVGAVVQARALRLS DSVHNVIEQVHGDKRSSITVESDMHDPSLLAIAQNHRVNGLTMAPST LFADIAFTLAKHLIQNHGLDTHTNLPSINNMAVEKALIV

Fv_Bik1	SQVDPILDDLEELASQIEFHEPKLPIVSPLLCTLLTGDTLGPQYIRRHCR ETVDFLGAIKMAEAQGIMDRSGMCIEIGAHPILTRMVKSIIGQEFRCL ASLRRKEDHFKTLADSLCALHLAGFSVNWEEYHRDFASSRNVLPLPK YSWQLANYWMQYKYSWCLTKGDEPVENTAVGVPVHTRALRLSDS VHNVIEQAHGDKRSSITAESDMHDPSLLAIAQNHRVNNLTMAPSTLF ADIAFTLAKHLIENHGLDTQTNLPSINNMAVEKALIV
Ap_PksA	AQVQPILDDFEALAQGATFAKPQLLILSPLLRTEIHEQGVVTPSYVAQ HCRHTVDMAQALRSAREKGLIDDKTLVIELGPKPLISGMVKMTLGD KISTLPTLAPNKAIWPSLQKILTSVYTGGWDINWKKYHAPFASSQKV VDLPSYGWDLKDYYIPYQGDWCLHRHQQDCKCAAPGHEIKTADYQ VPPESTPHRPSKLDPSKEAFPEIKTTTTLHRVVEETTKPLGATLVVETD ISRKDVNGLARGHLVDGIPLCTPSFYADIAMQVGQYSMQRLRAGHP GAGAIDGLVDVSDMVVDKALVP
Ds_PksA	AQVQPILEDFEELAAGATFEKPKLAVISPLLGSVVEDEGVVGPNYLA RHCREAVGMVKALGVAKEKGIINEKTIVIEIGPKPLLCGMIKNILGQN IVALPTLKDKGPDVWQNLSNIFTTLYTGGLDINWTAFHAPFEPAKKV LQLPDYGWDLKDYFIQYEGDWVLHRHKIHCNCADAGKDVHNTSHY CPGKHTFAENVVVPGGAQKAVQEAPAAKTETKKMSKLDPTKEAYP GIPLTTTVHKVIEEKTEPLGAQFTVETDISRKDVNSIAQGHTVDSIPLC TPSFYADIALQVGKYAMDRIRAGHPGAGAIDGRVDVTDLVVDKALIP
An_StcA	AQVQPILEEFKNVARGVTFHKPQIPVLSPLLVKVIDEKGTVDPVYLAR HCREPVKMVSVLEHARDQHIITDRTIVIDVGPKALMAGMIKTTLDKD TSSALPTLGPSLDVWKSLTNILGTLYSRGLDINWVAYHEPFGSAKKVI ELPSYGWDLKDYFIPYKGEWCLHRHEIRCSCATPGKETATSDYQLPS DEQVAAKRPSKQDESKEAYPEIVATTTVHRVVEEKTEPLGATLVVET DISRPDVNQIAQGHLVDGIPLCTPSVYADIALHVGRYSMNRLRASHP GAMDGVVDVADMVIDKALIP
Cn_Ctb1	GQVDPILPELLQVAAACSIQDPQIPVISPAYGKVIRSAKDFQPEYFTHH CRSSVNMVDALQSAVEEGLLDKNVIGLEIGPGPVVTQFVKEAVGTT MQTFASINKDKDTWQLMTQALAKFYLAGASVEWSRYHEDFPGAQK VLELPAYGWALKNYWLQYVNDWSLRKGDPAVVVAASNLELSSSIH KVITNTITANSDGELVVDADLSREDLHPMVQGHQVYGVPLCTPSVY ADIALTLGEYIRQVIKPGEVAQTSVEVAEMNIQSALVA
Fs_Pgl1	SQLDPVLSEFQEIATGVTFSKPSVPVLRPLDGTVVDHCDSFGPEYLAN HSRQSVNMLGALSTAYRDHVITDKSMILELGPHPAITGMVKAVLGQ QVICIASLQRARQPWDVLCAALKVLYDAGANISWAEYQRDFGGFHS VVPLPAYSWDLKDYWIQYVNDWSLRKGDAPIVINNAPRLESTTIHSV VEESGDSKKTHMIVEADISRKDMSPLVQGHEVDGIPLCTPSVYADMA LTLGRYLLERYQPQKKENLIDVSDMTISKALIL

Fg_Pgl1	SQLEPVVAEIEMLASKVSFSAPSIPILCPLDGTVVEDDDVFGASYLAK HSRQPVNMLSALATAYREGTISDRHMMLEVGPHPAVTGMVKPTLG QQITCIASLQRGRAPWEMLSAALKTLYDAGSSINWAEYQSSFPGSHS VVALPAYSWELKDYWIKYVNDWSLRKGDPPLVINNAPKLESTTIHR VVEEEGDSNKIHIIVEADIARKDLSPLVQGHEVDGIPLCTPSVYADIGL TLGKYLLEKYQPQNRDNMVVVSDMTVSKALIL
Ff_Fsr1	SQLEPMVSDIEKLAGKVTFSDPKIPILCPLEGTVIENANPFNASYLARH SRQPVNMLTALTTAYRDGYLSDRSMVLEVGPHPAVSGMVKPTLGQ QITCVASLQRRRAPWDMLSAALKSLYDAGASINWVDYQSNFPGAHT VVDLPAYSWDLKEYWIQYVNDWSLRKGDPPLVINNVSKLESTTIHS VVEESGDSEKTGIVVEADIARKDLSPLVQGHEVDGIPLCTPSVYADIA LTLGKYLLERYQPQQKDDMVVVSDMTVSKALIL
Fv_Pgl1	SQLEPVVFDIEKLAKKVAFSEPRIPVLCPLEGTVIENENPFNASYLARH SRQPVNMLTALTTAYRDGYLSDRSMVLEVGPHPAVSGMVKPTLGQ QITCVASLQRSRAPWDMLSAALKSLYDAGASINWADYQSNFPGAHT VVDLPAYSWDLKEYWIQYVNDWSLRKGDPPLVINNVSKLESTTIHS VVEESGNSKKTHIVVEADIARKDLSPLVQGHEVDGIPLCTPSVYADIA LSLGKYLLERYQPQQKDDMVVVSDMTVYKALIL
Af_EncA	AQMDPLLGPFEHIARGVTFKAPNIPVMSPSLGDCVFDGKTINASYMC NVTRNPVKFVDALETARGMDLVDAKTVWVEIGPHASYSRFVGSAM PPGTATIASLNRNEDNWSTFARSMAQLHNLGVDLNWHEWHAPFESE LRLLTDLPAYQWNMKNYWIQYNGDWMLRKDGKSSAAAASHPHQA IPPALRTSLVHRLVCESVQETRVEVIVESDILHPDFFEAMNGHRMNGC AVATTAIHADIAFTLAKYLYSSIMPNSTDAPAINVKNMQVQHGLVA
Af_TpcC	AQMDPILDEFEALAASGVVFQAPNLPVISPLLSKVVFDEHTIDSVYMR RATRETVHFLSAMKMAHKISTIDDATVWVEIGPHPVCVNFVRSSLPS TSVTVPSFRRGEDNWVTLTSSLGILHCAGVPVDWNEFHQPFERALRL LDLPTYSWNEKTYWIQYQGNWALTKGNTFYDDEAPQTKALAGLAS ELRTSTVQQIIHEQYDGAAGSVVMQSDLMQPDFLAAAYGHKMNGR GVVTSSIHADIAFTLGEYLYKKLNPNQEPHMNIANLEVVKALVA
Af_NscA	AQMDPILDSLETLATPITFKAPSIPVLSPLLGSVVFDRKSIHAQYLRRA TRETVDFVAAIEAAQDFGLVDAKTIWIDVGPHPICASLVRGIDSSASVI SSCRRNEDNLATMSKSLVTLHLAGLTPCWAEYFRPREQEYSLLKLPT YSWNETDYWIPYIGTWTLDKALLKYGEKKAPLSLSMSRPSALRTSLV HQITTETVEATTATLHVLSDMQHPDFLEALHGHRMNNCGVATSSIW SDMAFTVGEYLYRRLVPQAKDVHMNLSDLEVLHAQVA
An_AdaA	AQMDAVRERLAKAVAAVPFKTPSVPVLSPLLGSVVFDGKSINPEYIV RATREPVQFATAIDAAQELGIVNSQTLWVDIGPHPICASFVRSLVPGA RIVSSCRRNEDNFATMAKSLCTLHLAGRTPSWAEYFRPDEQAYSLLR LPKYRWNEVNYWIQYLGTWTLDKAHLKNGGSQKRAITDVPSISSLR TSLIHQVTEETVDKTTATLKAISDIQHPDFLEAVHGHTMNNCGVATS SIWTDMAMTVGEHLYRRLVPGTDHVLMDLCDFEVQHAQVA

At_GedC	AQLDPILDTYEQIATKGAIFHPPNLPIISPLLGKVIFDDKTVNATYMRR ASRETVNFHAALETAQRISTVDDTTAWVEIGPHPVCMGFIRSTLQSTA LTVPSLRRGEESWVTITRSLSSLHCAGVEVHWNEFHRPFEQALRLLD LPTYSWNDKNYWIQYNGDWALTKGNTFYSSQQQNSAAVDELPSGP RTSTVQKIVEESFDGRAARVVMQSDLMQSDLLEAAYGHKMNGCGV VTSSIHADVGFTLGQYVYKKLNPNTKVPAMNMASLEVLKGLVA
Pa_VrtA	SQMDVVVDELEEIAKNIPFKAPSIPVLSTMLGTVVFDGKTINPTYLRR QTRGTVKFVAAVETARDLGLIDEKTVWVDLGPHPVCVGFIRKLSPES RIAASCRRNEENLSTITKSLVTLHLAGATPLWNEFFRPNEQVYRLLNL PKYSWNETNYWIPYLGTWALDKALLKYGITPVGAKAPATLPAAGLR TSTIHQTTLETIDSMTATLHVLSDMQAPEFRAAVYGHTMNNCGVATS SIWTDMALAVGEYLYRKLVPQAKEVHMNVCDLEVLHAQVI
Pa_GsfA	AQMDNVVNELIRQSQGIAYNTPKIPIMSPRDSSVIETGANIDSSYLPTS LKKAVDFAGALNAAWEAGVVSKSTVWLELSHHPVCSGFINRTLPNT SLTCSTLHRDSDNWTSLLKTLSSLYEVGLNIDWNEYHRPFEHALRLV SAPTYAWNNKDYWIQYRGDWNLTKGQVLPEAELPAVSGFRTSSIHR LYSENYDSSTAHLLGECNMTDLSLKGVIEGHAMNGYGVASSFLHAE MAFTLARRIQEKASLSTFTGMGINVTNFEYHDPVVK
Pf_PtaA	AQTDPILDDFESVSRTGVLFQAPNLPVISPLLGKVVFNDKTINANYVR RATRESVDFLSALEAAQKISIIDESTTWIEIGPHPVCMGFIRSAVPSIKV ASPSIRRGENNWQTLVQTLGALHLAGIPVDWNEYHRPFEQALRLLDL PTYSWNDKTYWIQYNGDWALTKGNTFYDAEKAAKAPRVGGDLPPS PISTSTVHRVIGETFDGTAGTVDIQSDLMQQDFHDAAYGHKMNNCG VVTSSIHADIVYTIGRYLHTKLKPGVKDIHMNISNLEVVKGLVA
Nf_101810	AQMDPILDEFEALAASGVVFQAPNLPVISPLLSKVVFDEHTIDSVYMR RATRETVNFLSAMEMAHKISTIDDATVCVEIGPHPVCVNFVRSSLPTT SVTVPSFRRGEDNWVTLTNSLGVLHCAGVPVDWNEFHQPFERALRL LDLPTYSWNEKTYWIQYNGNWALTKGNTFYDDEAAQSNALAGLAS ELRTSTVQQIIHEQFDGTAGSVVMQSDLMQPDFLAAAYGHKMNGRG VVTSSIHADIAFTLGEYLYKKLYPNQEPHMNIANLEVLKALVA
Nf_101660	AQTDPILHELEETAQTGVLFQPPRLPIISPLLGKVIFDEKTVNAKYICR ATRETVNFAAALEKALAMSTVDETMVWIEIGPHPVCLGFVRSIMSTV NVAVPSFRRGENNWQTLSQSLAAVHAAGVEVDWNEFHRPFEHGLR LLDIPTYAWNNKTYWHQYNGDWALTKGNNFYDSKKRSAAAGSPLA AAPVSSLRTSLVHRVIEESFSSTAGKVIVQSDMMQADFLAAAWGHQ MNGAGVVTSSIHADIAWTLGKYLLDSLRPNNKKGVVDMEISHLVVR EGLVA
Nf_045430	IQMDPLLGPFEQISRGVTFKAPNIPVISPSLGDCVFDGKTINASYMCNV TRNPVKFVDALEAARGMDLIDAKTVWVEIGPHASYSRFVGSAMPPG TTTIASLNRNENNWSTFARSMAQLHNLGVDLNWHEWHAPFESELRL LVDLPAYQWNTKNYWIQYNGDWMLRKDGKSLASSANRPHQSIPPA LRTSLVHRLVHESVQDNRAEVIVESDILHPDFFEAMNGHRMNGCAV ATTAIHADIAFTLAKYLSSSIMPNSTIAPAINVKNMQVQHGLVA

At_02434	AQMDPILDQLEEIAEGVTFHDPSIPVISPVEAACVFDGKSLGPSYIRGV TRRPVQFVDALKAAQDLGLVDEMTIWVEIGPRASYSHFVRSVMAPG TVTVPSLKKGEDNWLTLARGMAQMYSLGVPINWQEWHAPFESQLR LLELPSYQWNAKNHWIPYTGNWRLYKGDVPPAVVPSPTPSALRTSL VHSLVKESIKENEGEVVIQSDILQPEFLEEMKGHLMNGHPVATMTIHS DIAFTIATYLFSRLRPRTTIPGIDVRNMHFDHGLIA
Nf_112240	AQMDPILDSLETLATPIAFKAPSIPVLSPLLGSVVFDRKSIHAQYLRRA TREAVDFVAAIEAAQDFGLVDAKTIWIDVGPHPICAGLVRGIDSSASV ISSCRRNEDNLATMSKSLVTLHLAGLTPCWAEYFRPREREYSLLKLPT YSWNETDYWIPYIGTWTLDKALLKYGEKKAPLSLAMSRPSALRTSL VHQITAETVEATTATLHVLSDMQHPDFLEALHGHRMNNCGVATSSI WSDMAFTVGEYLYRRLVPQVKDVHMNLSDFEVLHAQVA
At_00145	AQIQVIVEQYRKVASSVHFGPPNVPVISPLLGDVVTDGNVFGPDYLC RQAREAVNFMGALKAAESKGVVDNNVIWLEIGPAPVCSAFVKSSLG SKALTLPSLRKQEDVWKTLSGTLSNLYSKGLTIEWEEVHREYEASHT VLALPSYCFEEKNYWLDYHNNWCLTKGQKLVESAAPKRRGRHLLT PSVQKVIKEDFGQTKITVVAESDLSDPDLNHAVTGHLVNGSALCPAG VYAESALTLAGYIYHRVKKTEDIGMDVRALEIVKPLIA
At_07500	AQVEPILTDFEASACHVTFHPPTIPFLSPLLGRAISVGDIGALASTYVS AACRGTVNFVKAVGVAADVVNIQDTIWLEIGAHPLCSGMVKGTLGP QIRTIATLRQNVEPYKTIVSGLQTLYLAGVEINWNEYHRHFPSSQTVI ELPLYSWDLKNYWIQYRNDFLLTKGEQPLPVAPPSRPLRKHLSPTAQ CIVEESHGTEKSSMVVESDIFDEKLLPILQGHLVNGAALAPSSMYADL ALTVAVYLISQSPNKLLVDTTGLDVANVRVDNPLIA

Table S1: Product template domains extracted from various NR-PKSs using Clustal Omega and UGene and used to create the phylogenetic tree (Fig. 1A).

Table S2: Comparison of the tpc and ged gene clusters.

Gene	Gene Identifier	Conserved function/domain	*A. terreus* best hit	Identity/Similarity (%)
	AFUA_4G14590	Zn_2Cys_6 transcription factor	ATEG_06771	62/73
tpcA	AFUA_4G14580	Emodin O-methyltransferase	*gedA*	73/83
tpcB	AFUA_4G14570	Metallo-β-lactamase	*gedB*	71/79
tpcC	AFUA_4G14560	NR-PKS	*gedC*	67/80
tpcD	AFUA_4G14550	Transcriptional coactivator	*gedD*	42/58
tpcE	AFUA_4G14540	Zn_2Cys_6 transcription factor	*gedR*	39/52
tpcF	AFUA_4G14530	Glutathione S-transferase	*gedE*	68/82
tpcG	AFUA_4G14520	NADH-dependent oxidoreductase	*gedF*	76/86
tpcH	AFUA_4G14510	Sulochrin O-methyltransferase	ATEG_01465	45/61
tpcI	AFUA_4G14500	Questin oxygenase	*gedK*	60/74
tpcJ	AFUA_4G14490	Dihydrogeodin oxidase-like	*gedJ*	63/77
tpcK	AFUA_4G14470	Dehydratase	*gedI*	75/88
tpcL	AFUA_4G14480	Emodin anthrone oxygenase	*gedH*	53/68
tpcM	AFUA_4G14460	Desmethylsulochrin O-methyltransferase	*gedG*	61/74
	AFUA_4G14450	Mannitol 2-dehydrogenase	ATEG_01190	86/93

Table S2: Comparison between the trypacidin and geodin biosynthetic clusters.

Table S3: SNPs identified in the CEA10 tpc cluster.

Non-conservative* SNPs		
Gene	Effect of mutation	Note
tpcJ	A16P	Last residue in SignalP prediction, corresponding residue in GedJ is threonine
tpcD	T330A	Nothing known about this part of the protein
tpcD	H222Y	Nothing known about this part of the protein
tpcC	D1248E, premature termination	Cuts off the very end of the AT domain and all of the ACP domain
tpcB	T353D	Nothing known about this part of the protein

Table S3: Non-conservative* SNPs in the trypacidin biosynthetic cluster in the CEA10 strain. *The amino acid change caused by this mutation is conservative, however, it also results in a frameshift and premature termination codon.

Table S4: Strains used in this study.

Strain	Genotype	Source	Parental strain
AF293		Osherov et al., 2001	NA
AF293.1		Osherov et al., 2001	AF293
AF293.6		Xue et al., 2004	AF293.1
Δ*laeA*	Δ*laeA::parapyrG; pyrG1*	Bok, Keller, 2005	AF293.1
Δ*brlA*	Δ*brlA::fumipyrG; pyrG1*	Mah and Yu, 2006	AF293.1
TFYL43.2	Δ*akuA::parapyrG; pyrG1; argB1*	This study	AF293.6
TFYL44.1	Δ*akuA; pyrG1; argB1*	This study	TFYL43.2
TFYL45.1	Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL43.2
TFYL81.1	*fumipyrG; fumiargB;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL45.1
TKOT458	Δ*tpcA::parapyrG; pyrG1*	This study	AF293.1
TWFZ57.1	Δ*tpcB::parapyrG; pyrG1*	This study	AF293.1
TKOT456	Δ*tpcC::parapyrG; pyrG1*	This study	AF293.1
TKOT455	Δ*tpcD::parapyrG; pyrG1*	This study	AF293.1
TKOT454	Δ*tpcE::parapyrG; pyrG1*	This study	AF293.1
TKOT459	ΔAfu4g14590*::parapyrG; pyrG1*	This study	AF293.1
TKOT3447	Δ*tpcK::fumiargB;* Δ*akuA::parapyrG; pyrG1; argB1*	This study	TFYL43.2
TKOT3448	Δ*tpcL::fumiargB;* Δ*akuA::parapyrG; pyrG1; argB1*	This study	TFYL43.2
TKOT3478	Δ*tpcK/L::fumiargB;* Δ*akuA::parapyrG; pyrG1; argB1*	This study	TFYL43.2
TKOT3925	ΔAfu4g09250*::fumiargB;* Δ*akuA::parapyrG; pyrG1; argB1*	This study	TFYL43.2
TKOT9447	Δ*tpcK::fumiargB;* Δ*encC::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL69.2
TKOT9448	Δ*tpcL::fumiargB;* Δ*encC::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL69.2
TKOT9478	Δ*tpcK/L::fumiargB;* Δ*encC::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL69.2
TKOT9925	Δ*Afu4g09250::fumiargB;* Δ*encC::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL69.2
TFYL68.1	Δ*encA::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL45.1
TFYL69.1	Δ*encC::parapyrG;* Δ*akuA::mluc; pyrG1; argB1*	This study	TFYL45.1

TFYL71.1	Δ*tpcC::fumiargB; ΔakuA::mluc; pyrG1; argB1*	This study	TFYL45.1
TFYL72.1	Δ*encA::parapyrG; ΔtpcC::fumiargB; ΔakuA::mluc; pyrG1; argB1*	This study	TFYL68.1
TFYL73.1	Δ*encA::parapyrG; fumiargB; ΔakuA::mluc; pyrG1; argB1*	This study	TFYL68.1
TFYL74.1	Δ*encC::parapyrG; fumiargB; ΔakuA::mluc; pyrG1; argB1*	This study	TFYL69.1
TFYL76.1	Δ*tpcC::fumiargB; fumipyrG; ΔakuA::mluc; pyrG1; argB1*	This study	TFYL71.1

Table S4: Strains used in this study.

Table S5: Primers used in this study.

Name	Sequence	Purpose
KTpyrGF	CGTAATACGACTCACTATAGGG	*pyrG* deletion cassette
KTpyrGR	ATTCGACAATCGGAGAGGCTGC	
KTargB5F	TTTCCTTGAATTCTGGTTCG	*argB* deletion cassette
KTargB3R	GAAGGAGAGACCCATACATCC	
KTpyrGF	CGTAATACGACTCACTATAGGG	*parapyrG*::*nidugpdA(p)* overexpression cassette
KTgpdAR	GTGATGTCTGCTCAAGCG	
A. para pyrG 2kb FOR	CATGTTTGACAGCTTATCATCG	*A. parasiticus pyrG* deletion cassette, TFYL strains
A. para pyrG 2kb REV	GACTATTCCGAGGGTGTGCTAT	
A. fumi argB FOR	GAACGCGGTCTGCATCCAAG	*A. fumigatus argB* deletion cassette, TFYL strains
A. fumi argB REV	ATCTGCAGAATTCGCCCTTG	
KT145905F	ACTACCCGAATACACGCATC	Afu4g14590 *pyrG* deletion construct 5' flank
KT145905R	CCCTATAGTGAGTCGTATTACGC GGTTGCTGACCAATAAATTAC	
KT145903F	GCAGCCTCTCCGATTGTCGAATT GAGGGAGTATTTCCAATGC	Afu4g14590 *pyrG* deletion construct 3' flank
KT145903R	CGGTTCATTGTGCAGGAC	
KT145605F	CGTCTGAGATAGGCTTCTGG	*tpcC pyrG* deletion construct 5' flank
KT145605R	CCCTATAGTGAGTCGTATTACGA AGTCGGAGAGGGATCTTTG	
KT145603F	GCAGCCTCTCCGATTGTCGAATT GGAGTATCCGACTGTAGGG	*tpcC pyrG* deletion construct 3' flank
KT145603R	GACGAACAAGCTTGACAGC	
KT145505F	TGGGTGACGTCTACATCCTC	*tpcD pyrG* deletion construct 5' flank
KT145505R	CCCTATAGTGAGTCGTATTACGT GTCCAGACCGTCTCTCTTC	
KT145503F	GCAGCCTCTCCGATTGTCGAATA TGGGTGGTGTTTCACAGAC	*tpcD pyrG* deletion construct 3' flank
KT145503R	CAGTCTCCGGATTGCTAATG	
KT145405F	CTTGGTAGTGGACGTCTTCG	*tpcE pyrG* deletion construct 5' flank
KT145405R	CCCTATAGTGAGTCGTATTACGC AATTCCGTTCGAGAACAAG	
KT145403F	GCAGCCTCTCCGATTGTCGAATA TCAGACTTTTACGCGAAGC	*tpcE pyrG* deletion construct 3' flank
KT145403R	TTGGTGTCATCCTGTACGTG	
Af-*akuA* 5' FOR	GAATGGGTCACCTCGTTGAC	*akuA* deletion construct 5'

Af-*akuA* 5' REV	CCAATTCGCCCTATAGTGAGTCG TATTACGGGTATGGATTGTCATC AGCC	flank
Af-*akuA* 3' FOR	CTGTCGCTGCAGCCTCTCCGATT GTCGAATGCATGCCGACTGTCT GAATG	*akuA* deletion construct 3' flank
Af-*akuA* 3'REV	ACGTTAGTAAGTAGCGTGGC	
4G00210 5'F FOR	CGCTGAACTCAGGCCATAGAC	*encA* deletion construct 5' flank
4G00210 5'F REV	CTGAGATCCATAGGATCAGCTT ATCGATGGGTCAACTCAATGAA TGCCCATCC	
4G00210 3'F FOR	CGTGTTGATAGCACACCCTCGG AATAGTCCTGTCCCACCATTGGT GATCTC	*encA* deletion construct 3' flank
4G00210 3'F REV	CCAAATGTGCAAAGCGCGG	
4G00225 5'F FOR	ATCACAGATCGCCACTGC	*encC* deletion construct 5' flank
4G00225 5'F REV	CTGAGATCCATAGGATCAGCTT ATCGATGACATGACGAACAACC TACGGC	
4G00225 3'F FOR	CGTGTTGATAGCACACCCTCGG AATAGTCCGCTAGACATCGAGG TAGTGGT	*encC* deletion construct 3' flank
4G00225 3'F REV	CCAGTATGACCAAGCCAC	
tpcC 5'F FOR	GTGTACGGCACGGACAGTAG	*tpcC* deletion construct 5' flank, TFYL strains
tpcC 5'F REV	AAATTTGTCTTGGATGCAGACC GCGTTCACCTCAGTCTCTCAACA GATACG	
tpcC 3'F FOR	CTCTCCTTCAAGGGCGAATTCTG CAGATCTTGGAGTATCCGACTGT AGGGG	*tpcC* deletion construct 3' flank, TFYL strains
tpcC 3'F REV	CCGCGTCTCGTAGGGATG	
KT144705F	CCTGAGAATCTGCCTCAAAC	*tpcK argB* deletion construct 5' flank
KT14470A5R	CGAACCAGAATTCAAGGAAATC ATGCAGATTAGAGGAGTCG	
KT14470A3F	GGATGTATGGGTCTCTCCTTCAC CGCTCTTGTTCTGTTCTG	*tpcK argB* deletion construct 3' flank
KT144703R	AATGACAGGTCAGTCGCTTC	
KT144805F	CGGTTAGTTATCCATCCCATC	*tpcL argB* deletion construct 5' flank
KT14480A5R	CGAACCAGAATTCAAGGAAAGA TAGAAGTGCAGTTCGAAGG	
KT14480A3F	GGATGTATGGGTCTCTCCTTCAC AGAATGATCGGTGGAGAG	*tpcL argB* deletion construct 3' flank

KT144803R	GGATATCAGGGTCATCAAAGC	
KT144805F	CGGTTAGTTATCCATCCCATC	*tpcK/L argB* deletion construct 5' flank
KT4748A5R	CGAACCAGAATTCAAGGAAAAC AGAATGATCGGTGGAGAG	
KT14470A3F	GGATGTATGGGTCTCTCCTTCAC CGCTCTTGTTCTGTTCTG	*tpcK/L argB* deletion construct 3' flank
KT144703R	AATGACAGGTCAGTCGCTTC	
KT49255F	CCTTTTACCCCATTCTGAGG	Afu4g09250 *argB* deletion construct 5' flank
KT4925A5R	CGAACCAGAATTCAAGGAAAGA ACGGGGAGTTGTGATATG	
KT4925A3F	GGATGTATGGGTCTCTCCTTCCG GCCTGGTTGTATTGTATC	Afu4g09250 *argB* deletion construct 3' flank
KT49253R	GGACCACAAATGGTGAACTC	
Af-*akuA* NEST FOR	TCTTACAGCCTTGACAGCGC	PCR amplification of *akuA* deletion construct
Af-*akuA* NEST REV	CTACCTAACACTGATCGCCC	
5-FOA *akuA-gpdA(p)* 5'F REV	CCCCAAGCCTTCGACATCCGGA TGGAATTGGTATGGATTGTCATC AGCC	PCR amplification of *pyrG* recycling construct
5-FOA *akuA-trpC(t)* 3'F FOR	TGAGGAATCCGCTCTTGGCTCCA CGCGGGGCATGCCGACTGTCTG AATG	
4G00210 NEST FOR	CGAAGATTACCTCGGCGC	PCR amplification of *encA* deletion construct
4G00210 NEST REV	GGATGACAGAGCGTCCAG	
tpcC NEST FOR	CCGTCTGAGATAGGCTTCTGG	PCR amplification of *tpcC* deletion construct, TFYL strains
tpcC NEST REV	CTGACGAACAAGCTTGACAGC	
KTgpdAintF	GAAGGGTGGTGCCAAGAAG	Positive control for PCR confirmation of deletants
KTgpdAintR	CAACGGAGACGTTGGAGGT	
KT14590DF	GCCATTTCCAATCTATCACG	PCR confirmation of Afu4g14590 deletion
KT14590DR	CAATGGAAGGGCTACATGAC	
KT14560DF	CGGGATCATAGTGACTGTCC	PCR confirmation of *tpcC* deletion
KT14560DR	CTCATGGGCCTATGTGATTC	
KT14550DF	CCATGGACCATCTCTACCAG	PCR confirmation of *tpcD* deletion
KT14550DR	AATGGCTAGGGGAGTTTCAG	
KT14540DF	ATTCATCACCCGCTGGATAC	PCR confirmation of *tpcE* deletion
KT14540DR	CTGTCCAAGGTCAGATGCTC	
mluc FOR	ATGGTCACCGACGCCAAG	PCR confirmation of firefly luciferase insertion
mluc REV	ACACGGCGATCTTTCCGC	
akuA Int FOR	AGCAGTAAGGGACGCTGTCC	PCR confirmation of *akuA* deletion
akuA Int REV	CGTCGACCTTCTCTTCACCC	
4G00210 scrn FOR	CGTATCGGGACCTTCTATGGC	PCR confirmation of *encA* deletion
4G00210 scrn REV	CGGTATCACAGTCGAAGCTGG	

4G00225 scrn FOR	CGAATCTATGAGCAGGGACAC	PCR confirmation of *encC* deletion
4G00225 scrn REV	CTCCAGTGACTGAGCAGCTC	
KT14470DF	GACTACCGCCGCTATATGAC	PCR confirmation of *tpcK* deletion
KT14470DR	TCATGATCGTGGAACAAATG	
KT14480DF	ACCCTCGCTTAATCCTCTTG	PCR confirmation of *tpcL* deletion
KT14480DR	ACAACCTGTCTGCTCTACGG	
KT48RTF2	CGTGCTGCTCGCAGTATAG	PCR confirmation of *tpcK/L* deletion
KT14470A5R	CGAACCAGAATTCAAGGAAATC ATGCAGATTAGAGGAGTCG	
KT9250DF	CCCTCACAGCATTCAACAC	PCR confirmation of Afu4g09250 deletion
KT9250DR	ATGAGCGATTTCCCAACC	
KT14560DF	CGGGATCATAGTGACTGTCC	PCR confirmation of *tpcC* overexpression
KTKSgpdAF	ATTCATCTTCCCATCCAAGAACC	
KT14530DF	ACATCCAACCCATCACAGTC	PCR confirmation of *tpcE* overexpression
KTKSgpdAF	ATTCATCTTCCCATCCAAGAACC	
KT145903R	CGGTTCATTGTGCAGGAC	Southern confirmation of Afu4g14590 deletion
KT145905F	ACTACCCGAATACACGCATC	
KT145603R	GACGAACAAGCTTGACAGC	Southern confirmation of *tpcC* deletion
KT145605F	CGTCTGAGATAGGCTTCTGG	
KT145503R	CAGTCTCCGGATTGCTAATG	Southern confirmation of *tpcD* deletion
KT145505F	TGGGTGACGTCTACATCCTC	
KT145403R	TTGGTGTCATCCTGTACGTG	Southern confirmation of *tpcE* deletion
KT145405F	CTTGGTAGTGGACGTCTTCG	
KT14470A3F	GGATGTATGGGTCTCTCCTTCAC CGCTCTTGTTCTGTTCTG	Southern confirmation of *tpcK* deletion
KT144703R	AATGACAGGTCAGTCGCTTC	
KT14480A3F	GGATGTATGGGTCTCTCCTTCAC AGAATGATCGGTGGAGAG	Southern confirmation of *tpcL* and *tpcK/L* deletion
KT144803R	GGATATCAGGGTCATCAAAGC	
KT4925A3F	GGATGTATGGGTCTCTCCTTCCG GCCTGGTTGTATTGTATC	Southern confirmation of Afu4g09250 deletion
KT49253R	GGACCACAAATGGTGAACTC	

KT56OE5F	TCTGTCCGTCTGAGATAGGC	*tpcC parapyrG::nidugpdA(p)* overexpression construct 5' flank
KT56OE5R	CCCTATAGTGAGTCGTATTACGC GTCGCTGCTATAGACCTCAG	
KT56OE3F	CGCTTGAGCAGACATCACATGA GGCCAGTAGACTTTACC	*tpcC parapyrG::nidugpdA(p)* overexpression construct 3' flank
KT56OE3R	AGTCGAGAGGATCGGGTATC	
KT54OE5F	CTTGGTAGTGGACGTCTTCG	*tpcE parapyrG::nidugpdA(p)* overexpression construct 5' flank
KT54OE5R	CCCTATAGTGAGTCGTATTACGC CAATTCCGTTCGAGAACAAG	

KT54OE3F	CGCTTGAGCAGACATCACATGC GGTCCCCAGCGTCTTC	*tpcE parapyrG::nidugpdA(p)* overexpression construct 3' flank
KT54OE3R2	AGACAATCAAGGCCAGCAG	

KT56OE5F	TCTGTCCGTCTGAGATAGGC	Southern confirmation of *tpcC* overexpression
KT56OE5R	CCCTATAGTGAGTCGTATTACGC GTCGCTGCTATAGACCTCAG	
KT54OE5F	CTTGGTAGTGGACGTCTTCG	Southern confirmation of *tpcE* overexpression
KT54OE5R	CCCTATAGTGAGTCGTATTACGC CAATTCCGTTCGAGAACAAG	
KT14560DF	CGGGATCATAGTGACTGTCC	Northern confirmation of *tpcC* overexpression
KT14560DR	CTCATGGGCCTATGTGATTC	
KT14540DF	ATTCATCACCCGCTGGATAC	Northern confirmation of *tpcE* overexpression
KT14540DR	CTGTCCAAGGTCAGATGCTC	
KT210RTF2	TGCGTCAGTGGAATTAGGC	Semi-quantitative RT-PCR of *encA* expression
KT210RTR2	ATTGTCGAGTCCGACATCC	
KT220RTF2	TTGTGCTCTGCTGCAAGG	Semi-quantitative RT-PCR of *encB* expression
KT220RTR2	CCGGAAGTCGATGTCCTAAG	
KT225RTF2	AAATCGTCGCAATCACTGG	Semi-quantitative RT-PCR of *encC* expression
KT225RTR2	AATCTGCGGGCTATGTCTG	
KT230RTF2	ACCACCCGATGTAGAGCAG	Semi-quantitative RT-PCR of *encD* expression
KT230RTR2	CTCCTGCAAGACGGAACTG	
KT57RTF2	CCCACTCTTGAGGATGTCG	Semi-quantitative RT-PCR of *tpcB* expression
KT57RTR2	ACTCGGGAATCCCCTGTC	
KT56RTF2	CCAGAGCCTTGACGACTTC	Semi-quantitative RT-PCR of *tpcC* expression
KT56RTR2	ATGGCCACAAGATGAATGG	
KT55RTF2	CGCTTCCTGGACGATAAAG	Semi-quantitative RT-PCR of *tpcD* expression
KT55RTR2	AACATACAGGAGGCGATGC	
KT540RTF	GGCTTCGCGTAAAAGTCTG	Semi-quantitative RT-PCR of *tpcE* expression
KT540RTR	GGGTGATGAATCAGCTTGG	
KT51RTF2	CCCTGCCTGCTCAAATTC	Semi-quantitative RT-PCR of *tpcH* expression
KT51RTR2	CGACCTCAACCTGGAAGAC	
KT470RTF	TCAGGTCGTCTTCGAGAGC	Semi-quantitative RT-PCR of *tpcK* expression
KT470RTR	TATCCATCCCATCCACAGC	
KTact1RTF1	CTTCCAGCCTAGCGTTCTG	Semi-quantitative RT-PCR of *act1* expression
KTact1RTR1	CATACGGTCGGAGATACCG	

Table S5: Primers used in this study.

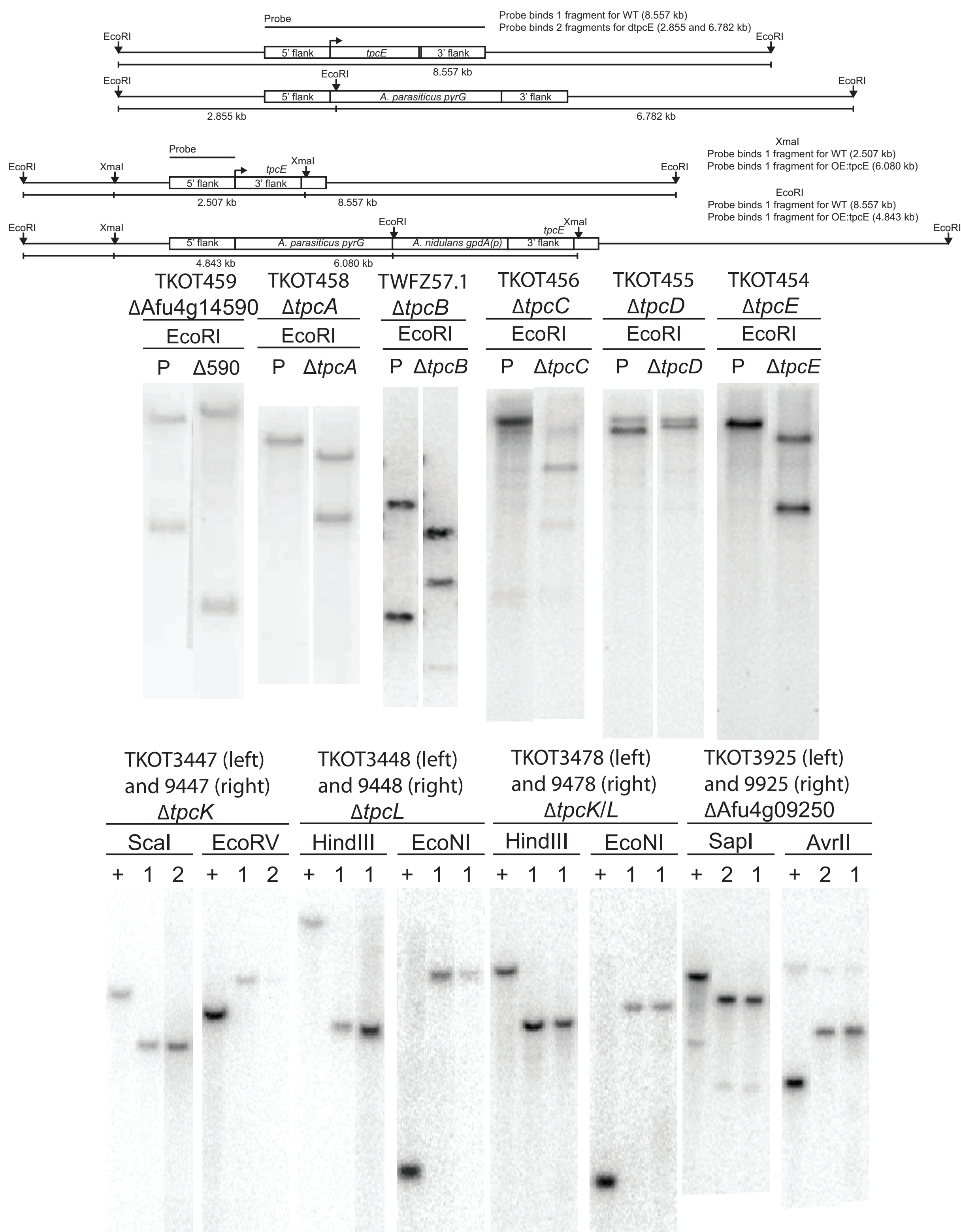
Table S6: Plasmids used in this study.

Plasmid	Genotype	Source
pJW24	*A. parasiticus pyrG*	Calvo et al., 2004
pJMP4	*A. fumigatus argB*	Sekonyela et al., 2013
pJMP9.1	*A. parasiticus pyrG*::*A. nidulans gpdA(p)*	Lim et al., 2012
pJMP10.1	*A. fumigatus argB*::*A. nidulans gpdA(p)*	Palmer and Keller, unpublished
pJMP147	*A. nidulans gpdA(p)*::*mluc*::*trpC(t)*	Palmer and Keller, unpublished
pKJA12.1	*A. fumigatus pyrG*	Affeldt and Keller, unpublished

Table S6: Plasmids used in this study.

Supplemental figures

Figure S1: Confirmation of mutants by Southern blot.



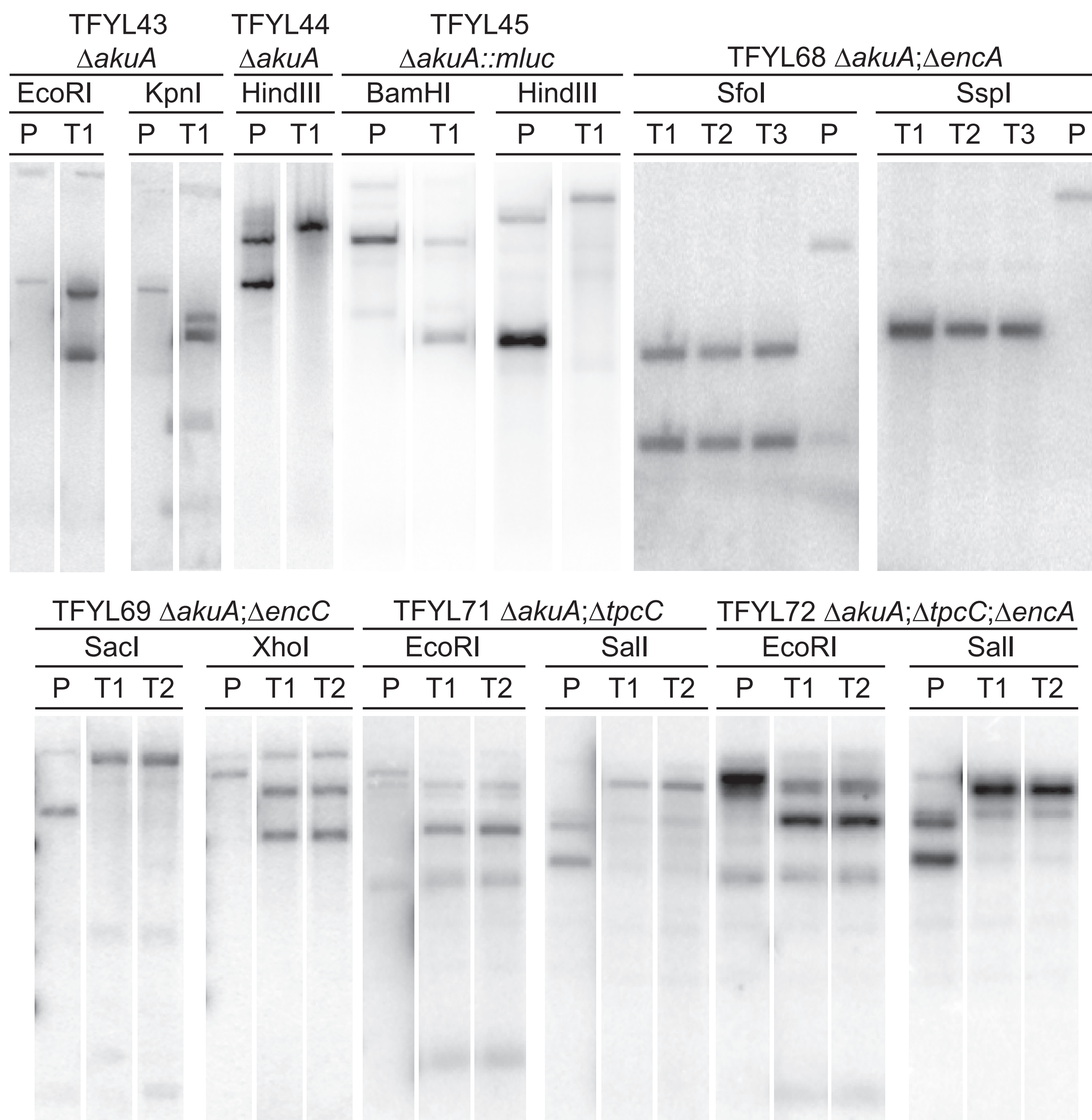


Figure S1: Confirmation of mutant strains with Southern blot analysis. Example overexpression and deletion constructs and probes are shown for *tpcE* mutants. The parental strain and successful transformants are shown. For Δ*tpcK*, Δ*tpcL*, Δ*tpcK*/*L*, and ΔAfu4g09250, the left transformant is in a WT *encC* background, whereas the right transformant is in a Δ*encC* background.

Figure S2: TLC analysis of standards used in this study.

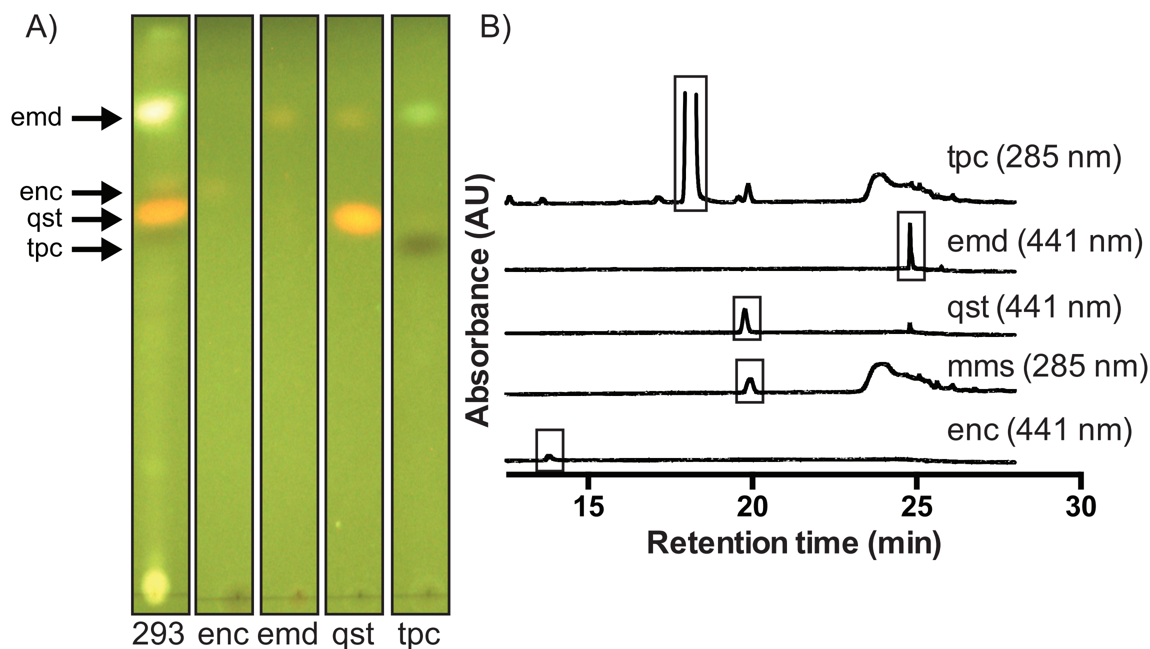


Figure S2: Thin layer chromatography (TLC) analysis of standards of endocrocin and trypacidin intermediates used in this study showing their positions relative to a wildtype extract (A) and High Performance Liquid Chromatography (HPLC) analysis of these standards showing only the relevant range of retention times and their absorbance at 285 and 441 nm (B). Black boxes surround the peak of interest to delineate it from other contaminating peaks. These peaks were confirmed by comparison of their UV absorption spectra to literature. These wavelengths were chosen as they are specific for the compound of interest at the relevant retention times and have relatively high absorbance. An exception to this is monomethylsulochrin, which could not be specifically detected as it co-eluted with questin, and so was excluded from presentation of results. 293=AF293, enc=endocrocin, emd=emodin, qst=questin, mms=monomethylsulochrin, and tpc=trypacidin.

Figure S3: TLC analysis of tpc cluster mutants.

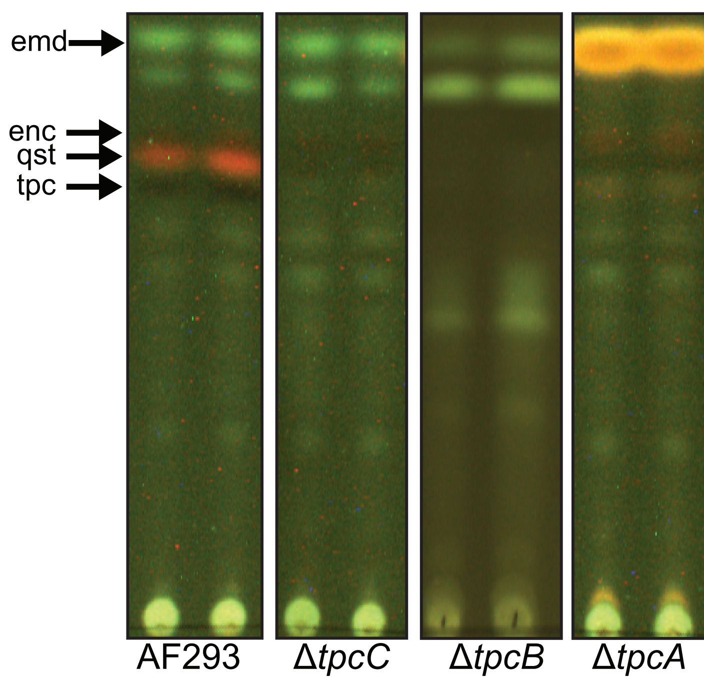


Figure S3: TLC analysis of *tpc* cluster mutants.

Figure S4: TLC analysis of tpc cluster regulators.

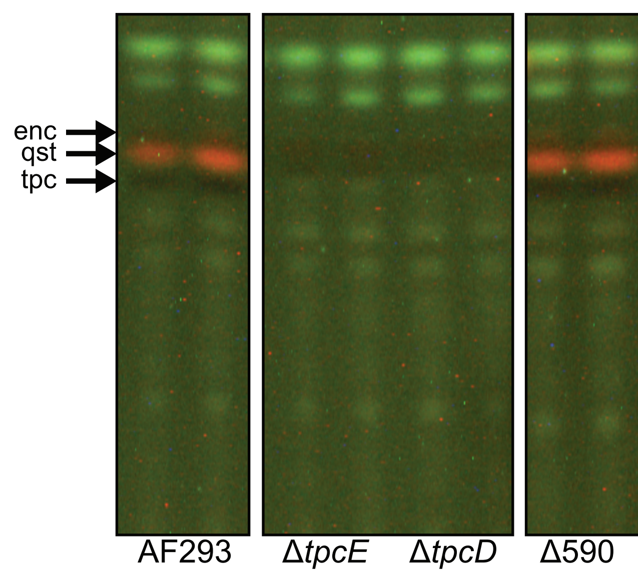


Figure S4: TLC analysis of *tpc* cluster mutants.

Figure S5: Semi-qRT-PCR analysis of LaeA and BrlA regulation of the enc and tpc gene clusters.

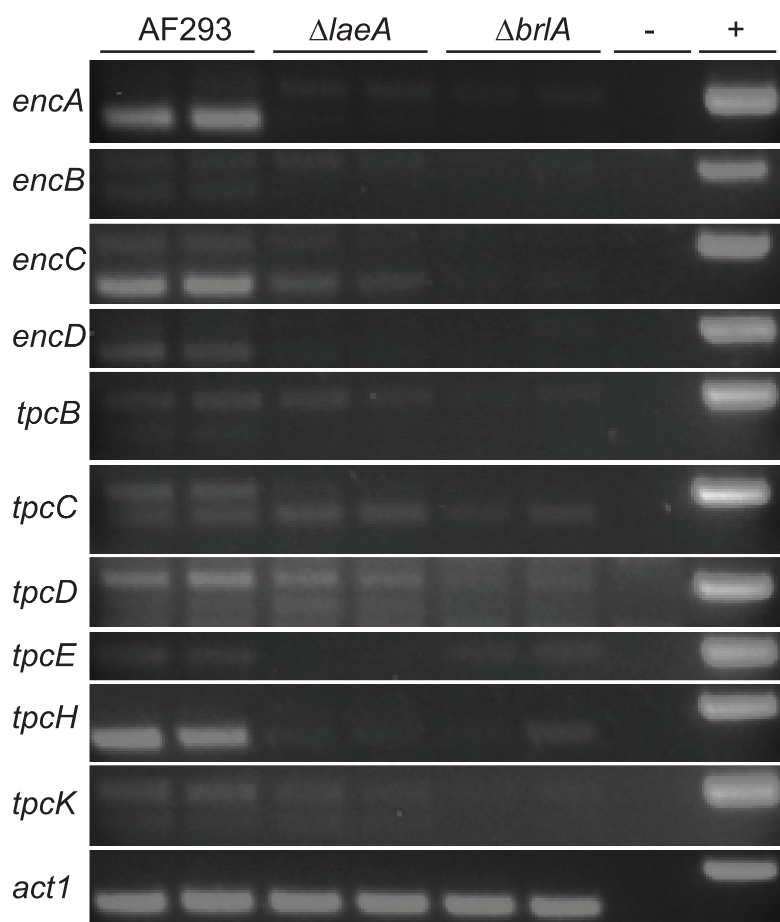


Figure S5: Semi-quantitative RT-PCR analysis of genes from the endocrocin and trypacidin gene clusters in WT (AF293), Δ*laeA*, and Δ*brlA*. - = water control, + = gDNA control. Primer pairs span an intron where possible (all cases but *tpcE*). 32 cycles were used for all genes.

Figure S6: TLC analysis of enc and tpc PKS deletants.

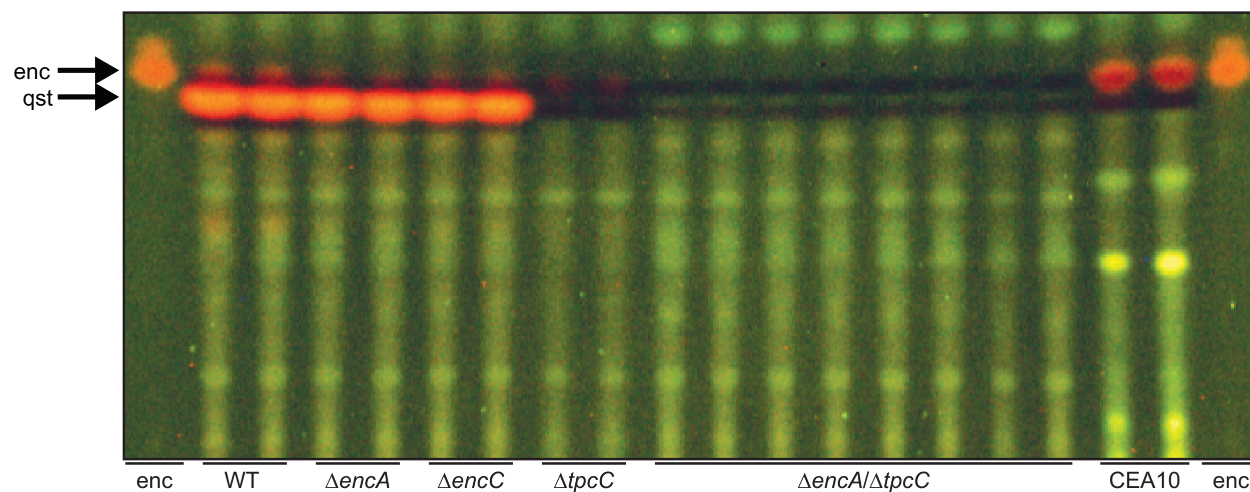


Figure S6: TLC analysis of *enc* and *tpc* cluster mutants.

Figure S7: TLC analysis of mdpH homolog mutants in the enc and tpc gene clusters.

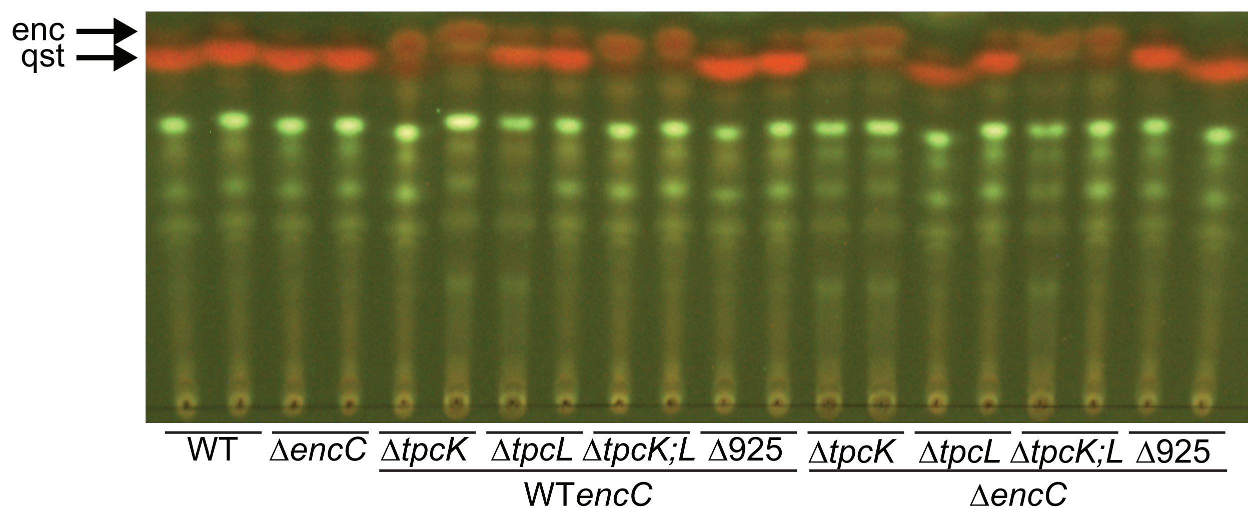


Figure S7: TLC analysis of *enc* and *tpc* early pathway genes.

Figure S8: Assessment of virulence of enc and tpc cluster mutants in Toll-deficient Drosophila.

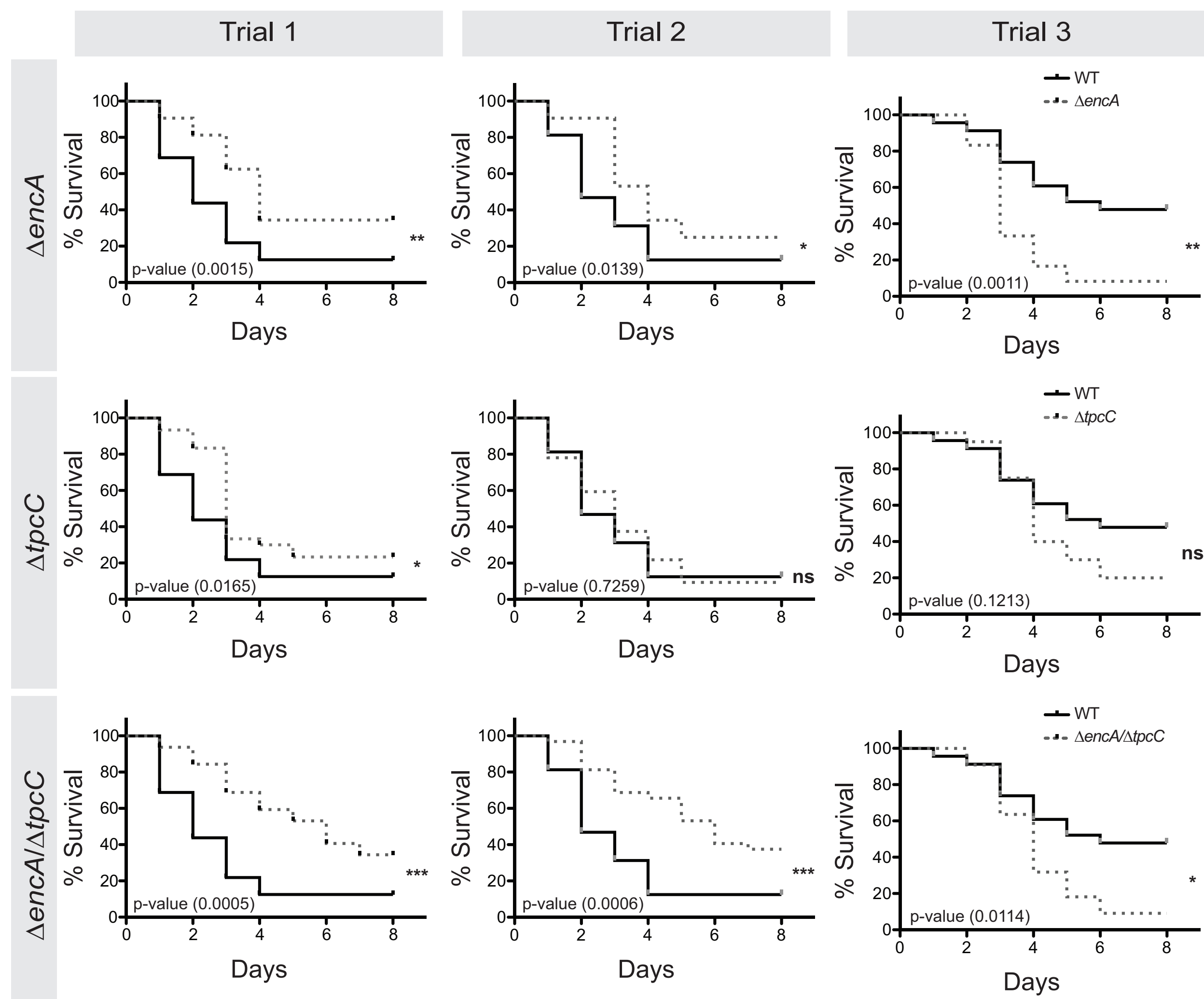


Figure S8: Survival curves of *Toll*-deficient *Drosophila* exposed to spores of the polyketide synthase mutants, Δ*encA*, Δ*tpcC*, and Δ*encA*/Δ*tpcC*.

CHAPTER 4

Concluding remarks and future directions

Group V NR-PKSs

In recent years, numerous studies have characterized SM gene clusters associated with group V NR-PKSs. In addition to the *ged* cluster which has been characterized piecemeal over the last two decades (Huang et al., 1995; Couch and Gaucher, 2004; Awakawa et al., 2009; Nielsen et al., 2013), twelve other group V gene clusters have been identified and characterized in the last seven years (Szewczyk et al., 2008; Chiang et al., 2010; Chooi et al., 2010; Li et al., 2011; Ahuja et al., 2012; Lim et al., 2012; Saha et al., 2012; Chooi et al., 2013; König et al., 2013; Xu et al., 2014; Chooi et al., 2015; Throckmorton et al., 2015). Previous studies have defined the subdivision of group V NR-PKSs into two groups, called V1 and V2. The MβLs associated with these PKS were also noted to have correspondingly diverged into two groups, indicating their co-evolution with the PKSs (Li et al., 2011). Group V1 and V2 PKSs generally produce metabolites with an anthracene (tricyclic) or naphthacene (tetracyclic) backbone, respectively, though these distinctions do not strictly hold true. The production of tetracyclic backbones was demonstrated to depend on multiple factors including the presence of a particular flavin-monooxygenase, an MβL with Claisen cyclase activity, and a PKS the produces a long enough polyketide chain. This characteristic ensemble of three enzymes was recognized and noted as a potentially useful signal in mining genomes for similar SM gene clusters and metabolites (Li et al., 2011). This ensemble was subsequently found in association with polycyclic prenyltransferases in a group of *nsc* and *vrt*-like clusters (Chooi et al., 2012; Chooi et al., 2013). We have expanded this concept to the level of whole SM gene clusters; PKSs from

different subgroups of group V are associated with distinct complements of decorating enzymes that can be used to search for and differentiate between them.

Product prediction directed by phylogenetics

Using only the sequence of the PKS as a lead, we identified 188 group V PKSs from public databases and, using MultiGeneBLAST analysis (Medema et al., 2013), identified their associated SM gene clusters. Information on the gene clusters associated with these PKSs was readily attainable for about two-thirds of these. It is possible that cluster information for the other one-third was not retrievable due to incomplete or insufficient sequence annotation, that these PKSs are not associated with gene clusters (i.e. they are orphaned PKSs), or that they or their associated clusters are so dissimilar as to have fallen below our thresholds of detection. We have no indication of which of these possibilities is true, but we expect that a combination of these reasons explain most, if not all of the missing information. For some PKSs represented in our study, data are available on the metabolite profile of the corresponding species. This was demonstrated for *T. cellulolyticus*, and we suggest that this PKS (GAM37897) would be a likely candidate for the production of purpactins in this species (Yilmaz et al., 2014), though this would, of course, need to be experimentally validated. Data on the metabolite profiles of other species harboring PKSs examined in our study are likely available, but our search was not exhaustive.

To paraphrase a recent perspective, drawing connections between metabolites and the PKSs (and associated clusters) that produce them ‘illuminates the nearby chemical space achievable by related PKSs as a lighthouse and contributes to the goal of prediction of metabolite

structures from primary sequences' (Chooi and Tang, 2012). In the context of this portion of the microcosm of secondary metabolism, we have utilized the 'lighthouses' built by our group and others over the last decade to make sense of the large amount of data available from fungal genome sequencing efforts and, hopefully, guide others in choosing directions for future research. Our analysis of uncharacterized group V NR-PKSs yielded a mass of information, including many previously unidentified SM gene clusters that might be interesting targets for future studies. These include numerous SM gene clusters in pathogens of plants, animals, and fungi, as well as some mutualistic fungi. Similar insight should be attainable at least in the other well-characterized groups of NR-PKSs, i.e. groups I-IV and VI-VII. This approach can also be applied at the species level, as demonstrated recently for *A. ustus* (Pi et al., 2015).

As new group V PKSs are identified in newly sequenced genomes, their relationship to characterized PKSs can be assessed to determine their position in the phylogenetic tree we have created. This will inform the researcher's expectations as to the type of SM gene cluster, if any, these PKSs are likely to belong to and the kind of metabolite these clusters might produce. This depends on two things – the existence of well-characterized SM gene clusters with which to compare and contrast, and large enough number of sequences to distinguish clades – this information can be used to predict the products of uncharacterized SM gene clusters. In our analysis, some areas of the phylogenetic tree meet these criteria better than others. Group V2 seems to have a high number of characterized clusters relative to the diversity of sequences, with nearly two-thirds of clusters identified being clearly *nsc*- or *vrt*-like. In contrast, group V3 seems relatively unsaturated, with most of the identified clusters being poor matches to the *pkg* cluster and none similar to the *gsf* cluster. This low-quality is complicated by the fact that all

characterized group V3 clusters besides the *gsf* cluster were not thoroughly delineated and consist of only a couple genes each. As suggested by a recent study characterizing one of the alternariol-producing clusters from group V3, (PKS=SNOG_15829), phylogenetics can be used to differentiate between clusters more similar to this one and *gsf*-like clusters. However, additional PKS sequences and information on their associated clusters are required in this portion of the phylogenetic tree before this can be implemented. Lastly, most of the characterized clusters from group V1, including the *mdp*, *ged*, *tpc*, and *pta* clusters, are very similar, with eight genes in common between them. Thus, the uncharacterized clusters we identified in this group did not clearly resolve into groups with a single most-similar characterized cluster. More in-depth analysis of the genes in these clusters was required to glean information from this analysis. This was possible however, because many of the enzymes typically encoded by group V1 cluster genes have previously defined roles. Most importantly, this includes the key scaffold-altering enzymes such as the anthraquinone-ring opening enzymes (Baeyer-Villiger oxidase and NADH-dependent oxidoreductase) and ring-closing multicopper oxidases (Huang et al., 1995; Ehrlich et al., 2005; Cary et al., 2006).

Toward rational design of SMs

The interplay between domains of PKSs has been demonstrated to be an important determinant of the activity of chimeric PKSs constructed through recombination of non-cognate domains. Most significantly, the product template (PT) and TE domains constrain the length and cyclization register of the possible products (Newman et al., 2014). These factors affect the efficient production of non-derailment metabolites from these PKSs. Thus, true prediction of the products of from sequence data and/or rational design of novel products by PKS domain

swapping, combinatorial construction of new SM gene clusters, or both will likely require more advanced knowledge of the dynamics of PKSs and their domains. This may be a longer term goal; in the immediate future, obtaining a thorough sampling of the diversity of polyketides produced in nature in terms of, for instance, their range of product lengths, registers of cyclization, constituent domains, and associated enzymatic activities will lay the groundwork for this long term goal by effectively creating a map with which to navigate this portion of the landscape of fungal secondary metabolism. Additionally, description of the substrate specificities of various decorating enzymes will be instrumental in the construction of novel SM gene clusters. To a certain extent, the substrate specificities can be inferred from the co-occurrence of certain well-characterized genes and the relative positions of the enzymes they encode in the likely biosynthetic pathway. A better understanding of the mechanisms of other types of PKSs, i.e. HR-PKSs, PR-PKSs, and hybrid PKS/NRPSs, will be required before these systems are as tractable, and other types of synthases, e.g. NRPSs, are similarly enigmatic at present.

Endocrocin, trypacidin, and group V1

In total, we have deleted and/or overexpressed seven of the thirteen genes predicted to be in the *tpc* cluster, with a preference toward those encoding enzymes involved in the early steps of the trypacidin biosynthetic pathway. These were chosen to address and confirm the possibility of production of endocrocin as a shunt product from the trypacidin biosynthetic pathway. This adds to previous studies partially characterizing other group V1 clusters including the endocrocin, monodictyphenone, geodin, and pestheic acid producing clusters in *A. fumigatus*, *A. nidulans*, *A. terreus*, and *P. fici*, respectively (Chiang et al., 2010; Lim et al., 2012; Nielsen et al., 2013; Xu et al., 2014). Five of the remaining six uncharacterized genes have homologs in the *mdp*, *ged*, or

pta clusters, and so would be of interest to delete in future studies. This would provide insight into the biosynthetic steps catalyzed by these enzymes, which would likely apply to all of these clusters.

In particular, the enzymes that catalyze the opening of the central anthraquinone ring to form benzophenones have been a point of interest in the study of the geodin, trypacidin, monodictyphenone, pestheic acid, and aflatoxin biosynthetic pathways. The enzymes reported to be required for this transition in the geodin, trypacidin, and monodictyphenone pathways include a Baeyer-Villiger oxidase, an NADH-dependent oxidoreductase, and a glutathione *S*-transferase (Chiang et al., 2010; Nielsen et al., 2013; Throckmorton et al., 2015). However, a similar transition is proposed to be catalyzed by enzymes encoded in the *pta* cluster which encodes no glutathione *S*-transferases (Xu et al., 2014). The mechanism of this transition in the group V1 biosynthetic pathway in which it has been most thoroughly addressed, the monodictyphenone pathway, has been a subject of some contention (Sanchez et al., 2011; Simpson, 2012). Examination of this transition in the trypacidin and other group V1 pathways might shed light on this debate as well. Additionally, the presumed functions of these enzymes are based in part on comparison to homologous biochemically characterized aflatoxin pathway enzymes (Skory et al., 1992; Ehrlich et al., 2005; Henry and Townsend, 2005b, a; Cary et al., 2006). In the aflatoxin biosynthetic pathway, this anthraquinone to benzophenone transition is catalyzed by a set of four enzymes, including a Baeyer-Villiger oxidase and an NADH-dependent oxidoreductase. Examination of this transition in group V1 biosynthetic pathways might provide an interesting contrast and potential insight into the shared evolutionary history of these gene clusters with the *afl* gene cluster.

In the course of the characterization of the trypacidin gene cluster and its redundant production of endocrocin as a shunt product, we examined the effect of single and double PKS-encoding gene deletions on the virulence of *A. fumigatus* in a *Toll*-deficient *Drosophila* model of Invasive Aspergillosis (IA) (Lionakis et al., 2005; Throckmorton et al., 2015). Though deletion of the gene encoding the endocrocin-producing PKS, *encA*, had a significant effect on the *Drosophila* survival in two of three trials, no consistent, significant, additive effect was observed for the gene encoding the trypacidin-producing PKS, *tpcC*, or the double PKS deletant. However, these experiments employed a growth medium other than the glucose minimal media that we usually use, and we observed a decrease in the levels of trypacidin on this medium. Thus, it would be useful to repeat these experiments using the usual medium. Additionally, it would be interesting to include some of our other mutants, specifically those that exhibit dramatic accumulation of endocrocin, i.e. the Δ*tpcK* strains, in these assays, with the expectation that they would be hyper-virulent.

References

Ahuja, M., Chiang, Y.M., Chang, S.L., Praseuth, M.B., Entwistle, R., Sanchez, J.F. et al. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J Am Chem Soc* **134**: 8212-8221.

Awakawa, T., Yokota, K., Funa, N., Doi, F., Mori, N., Watanabe, H., and Horinouchi, S. (2009) Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem Biol* **16**: 613-623.

Cary, J.W., Ehrlich, K.C., Bland, J.M., and Montalbano, B.G. (2006) The aflatoxin biosynthesis cluster gene, *aflX*, encodes an oxidoreductase involved in conversion of versicolorin A to demethylsterigmatocystin. *Appl Environ Microbiol* **72**: 1096-1101.

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microbiol* **76**: 2067-2074.

Chooi, Y.H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: from genes to molecules. *J Org Chem* **77**: 9933-9953.

Chooi, Y.H., Cacho, R., and Tang, Y. (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* **17**: 483-494.

Chooi, Y.H., Muria-Gonzalez, J.M., Mead, O.L., and Solomon, P.S. (2015) SnPKS19 encodes the polyketide synthase for alternariol mycotoxin biosynthesis in the wheat pathogen *Parastagonospora nodorum*. *Appl Environ Microbiol*.

Chooi, Y.H., Wang, P., Fang, J., Li, Y., Wu, K., and Tang, Y. (2012) Discovery and characterization of a group of fungal polycyclic polyketide prenyltransferases. *J Am Chem Soc* **134**: 9428-9437.

Chooi, Y.H., Fang, J., Liu, H., Filler, S.G., Wang, P., and Tang, Y. (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **15**: 780-783.

Couch, R.D., and Gaucher, G.M. (2004) Rational elimination of *Aspergillus terreus* sulochrin production. *J Biotechnol* **108**: 171-178.

Ehrlich, K.C., Montalbano, B., Boué, S.M., and Bhatnagar, D. (2005) An aflatoxin biosynthesis cluster gene encodes a novel oxidase required for conversion of versicolorin a to sterigmatocystin. *Appl Environ Microbiol* **71**: 8963-8965.

Henry, K.M., and Townsend, C.A. (2005a) Synthesis and fate of o-carboxybenzophenones in the biosynthesis of aflatoxin. *J Am Chem Soc* **127**: 3300-3309.

Henry, K.M., and Townsend, C.A. (2005b) Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. *J Am Chem Soc* **127**: 3724-3733.

Huang, K.X., Fujii, I., Ebizuka, Y., Gomi, K., and Sankawa, U. (1995) Molecular cloning and heterologous expression of the gene encoding dihydrogeodin oxidase, a multicopper blue enzyme from *Aspergillus terreus*. *J Biol Chem* **270**: 21495-21502.

König, C.C., Scherlach, K., Schroeckh, V., Horn, F., Nietzsche, S., Brakhage, A.A., and Hertweck, C. (2013) Bacterium induces cryptic meroterpenoid pathway in the pathogenic fungus *Aspergillus fumigatus*. *Chembiochem* **14**: 938-942.

Li, Y., Chooi, Y.H., Sheng, Y., Valentine, J.S., and Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**: 15773-15785.

Lim, F.Y., Hou, Y., Chen, Y., Oh, J.H., Lee, I., Bugni, T.S., and Keller, N.P. (2012) Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* **78**: 4117-4125.

Lionakis, M.S., Lewis, R.E., May, G.S., Wiederhold, N.P., Albert, N.D., Halder, G., and Kontoyiannis, D.P. (2005) *Toll*-deficient *Drosophila* flies as a fast, high-throughput model for the study of antifungal drug efficacy against invasive aspergillosis and *Aspergillus* virulence. *J Infect Dis* **191**: 1188-1195.

Medema, M.H., Takano, E., and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* **30**: 1218-1223.

Newman, A.G., Vagstad, A.L., Storm, P.A., and Townsend, C.A. (2014) Systematic domain swaps of iterative, nonreducing polyketide synthases provide a mechanistic understanding and rationale for catalytic reprogramming. *J Am Chem Soc* **136**: 7348-7362.

Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., Larsen, T.O., and Mortensen, U.H. (2013) Heterologous reconstitution of the intact geodin gene cluster in

Aspergillus nidulans through a simple and versatile PCR based approach. *PLoS One* **8**: e72871.

Pi, B., Yu, D., Dai, F., Song, X., Zhu, C., Li, H., and Yu, Y. (2015) A genomics based discovery of secondary metabolite biosynthetic gene clusters in *Aspergillus ustus*. *PLoS One* **10**: e0116089.

Saha, D., Fetzner, R., Burkhardt, B., Podlech, J., Metzler, M., Dang, H. et al. (2012) Identification of a polyketide synthase required for alternariol (AOH) and alternariol-9-methyl ether (AME) formation in *Alternaria alternata*. *PLoS One* **7**: e40564.

Sanchez, J.F., Entwistle, R., Hung, J.H., Yaegashi, J., Jain, S., Chiang, Y.M. et al. (2011) Genome-based deletion analysis reveals the prenyl xanthone biosynthesis pathway in *Aspergillus nidulans*. *J Am Chem Soc* **133**: 4010-4017.

Simpson, T.J. (2012) Genetic and biosynthetic studies of the fungal prenylated xanthone shamixanthone and related metabolites in *Aspergillus* spp. revisited. *Chembiochem* **13**: 1680-1688.

Skory, C.D., Chang, P.K., Cary, J., and Linz, J.E. (1992) Isolation and characterization of a gene from *Aspergillus parasiticus* associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. *Appl Environ Microbiol* **58**: 3527-3537.

Szewczyk, E., Chiang, Y.M., Oakley, C.E., Davidson, A.D., Wang, C.C., and Oakley, B.R. (2008) Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microbiol* **74**: 7607-7612.

Throckmorton, K., Lim, F.Y., Kontoyiannis, D., Zheng, W., and Keller, N.P. (2015) Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*. In *Environ Microbiol*. Madison, U.o.W.-. (ed).

Xu, X., Liu, L., Zhang, F., Wang, W., Li, J., Guo, L. et al. (2014) Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte *Pestalotiopsis fici*. *Chembiochem* **15**: 284-292.

Yilmaz, N., Visagie, C.M., Houbraken, J., Frisvad, J.C., and Samson, R.A. (2014) Polyphasic taxonomy of the genus *Talaromyces*. *Stud Mycol* **78**: 175-341.