Discovering the Secondary Metabolite Potential Encoded within Entomopathogenic Fungi

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Natural Product Reports</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>NP-HIG-04-2014-000054.R2</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Highlight</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>30-Jul-2014</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Gibson, Donna M.; USDA, ARS, Biological Integrated Pest Management Research Unit  
Donzelli, Bruno; Cornell University, Dept. of Plant Pathology and Plant Microbe Biology  
Krasnoff, Stuart; USDA, ARS, Biological Integrated Pest Management Research Unit  
Keyhani, Nemat; University of Florida, Dept. of Microbiology and Cell Science |
Discovering the Secondary Metabolite Potential Encoded within Entomopathogenic Fungi

Donna M. Gibson, a Bruno G. G. Donzelli, b Stuart B. Krasnoff, a and Nemat O. Keyhani c

This highlight discusses the secondary metabolite potential of the insect pathogens Metarhizium and Beauveria, including a bioinformatics analysis of secondary metabolite genes for which no products are yet identified.

Notes and references

a USDA-ARS, Biological Integrated Pest Management Research Unit, Robert W. Holley Center for Agriculture and Health, 538 Tower Road, Ithaca, NY 14853 USA Email: Donna.Gibson@ars.usda.gov; Stuart.Krasnoff@ars.usda.gov
b Dept. of Plant Pathology and Plant Molecular Biology, Cornell University, Ithaca, NY 14853 USA Email: bdd1@cornell.edu
c Dept. of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611 USA Email: keyhani@ufl.edu

† ABBREVIATIONS: A, adenylation domain; FAS, fatty acid synthase; KO, knockout; KS, ketosynthase domain; NP, natural product; NRPS, nonribosomal peptide synthetase; NRP, nonribosomal peptide; NP, natural product; PKS, polyketide synthase; PK, polyketide; SM, secondary metabolite; TC, terpene cyclase; TS, terpene synthase; WT, wild type

I. Introduction

1.1. A brief review of insect pathogenic fungi

Arthropods constitute the most species-rich group of animals, and the most diverse arthropod subgroup, the insects, occupy innumerable niches and have immense impacts as pests, pollinators, and agents of control of injurious species. Aside from their biological roles and functions, their enormous biomass provides opportunities as hosts or nutrient sources for parasites, pathogens, and predators, ranging from bacteria and fungi to plants and animals. Of particular interest are the 700 species (from ~100 genera) of entomopathogenic (insect-pathogenic) fungi 1, 2 that constitute a unique, highly specialized trophic subgroup. Fungal pathogens of insects are found within every ecosystem and all major fungal lineages with the principal exception of the higher basidiomycetes 3-5. Most entomopathogenic species are in the order Hypocreales (class Sordariomycetes, phylum Ascomycota). They associate with plants, animals and other fungi, and may subsist as pathogens, parasites, saprophytes, or in commensal and mutualistic relationships.

Fungi in the Hypocreales are thought to derive from plant-associated fungi 2.3. The entomogenous habit likely arose and spread concomitantly with the diversification of phytophagous insects that took place during the Cretaceous period 6. These fungi have undergone repeated inter-kingdom host-shifts among plants, arthropods and fungal hosts, accompanied by switches in nutritional habits 7. Such inter-kingdom host switching may have been facilitated in the soil or on aerial plant surfaces by the close proximity with sedentary insects feeding on the plants and saprobiic fungi. The co-evolutionary arms race between the insects and their fungal pathogens (especially obligate pathogens) often results in fascinating physiological and behavioral changes in insects, including reduced feeding, diminished reproductive fitness, defensive behaviors such as heat seeking and increased social grooming, to more specific behavioral manipulations of the host by the fungus, such as elevation-seeking prior to death, which enhances spore dispersal from the cadaver 7, 8.

1.2. Scope of the highlight

For this highlight, we will primarily focus on the secondary metabolites and molecular genetics of biosynthesis in two major cosmopolitan insect pathogenic fungal species, Beauveria bassiana and Metarhizium robertsii (anisopliae), with references to congeners where relevant. Both species have broad host-ranges encompassing over 1,000 insect species from > 50 different insect families, including pests of agricultural, veterinary and medical significance 7. Consequently, both Metarhizium and Beauveria species have been

This journal is © The Royal Society of Chemistry 2013

J. Name., 2013, 00, 1-3 | 1
developed for use as biological control agents worldwide. Besides being an effective pathogen, *B. bassiana* can exist as a saprophyte and, as a plant endophyte, and, in the latter case, can confer to its host-plant resistance against foraging insects and plant pathogens. *M. anisopliae* can also colonize plant roots and protect against root-feeding insects, and, like *Beauveria*, produces a variety of cell types in culture, including conidia, hyphae, appressoria, unicellular blastospores, and multi-cellular hyphal bodies.

2. Current use of entomopathogenic fungi and the infection process

Interest in biological control agents has steadily grown due to concerns over environmental pollution and pest resistance resulting from the use of insecticides. Unlike bacteria and viruses, insect pathogenic fungi infect their hosts by penetrating the cuticle and thus, with rare exceptions, need not be ingested, making them particularly well suited for broad application as contact biocides against a wide array of arthropods. In many cases, these agents are also compatible with other control measures, including traditional pesticides, BT transgenic-plants, or other biological controls such as predators and parasitoids, and thus offer an alternative and complementary tool for use in integrated pest management programs. Although, at present, entomopathogens represent only a small niche in the pesticide market, market share concerns over environmental pollution and pest resistance resulting from the use of insecticides are just beginning to address.

One of the key discoveries in fungal genomics is the large number of gene clusters with predicted roles in biosynthesis of secondary metabolites (SM), relative to the comparatively small number of characterized products. The biosynthetic machinery used for their production is energetically expensive, suggesting that NP biosynthesis confers a greater fitness or survival advantage to the organism. The genomes of filamentous Ascomycete fungi are replete with SM gene clusters, and the entomopathogenic fungi sequenced to date rank highly for the number of predicted SM gene clusters. The majority of these SM gene clusters are silent under standard laboratory conditions and require triggers, such as stress environments or intimate exposure to other organisms, for some of their genes to be expressed. One interesting aspect observed in other Ascomycete genomes is that the core genes (NRPSs, PKSs) encoding fungal secondary metabolite production are often organized in clusters with associated specific transcription factors and modifying enzyme genes. Many of these clusters are located in the fast-evolving, variable subtelomeric regions of chromosomes, which in some cases have been shown to contain species-specific genes correlated with adaptability to particular environmental niches.

We will focus on the secondary metabolite products of nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs), and hybrids of the two (NRPS-PKSs) since these chemistries are abundant in *Metarhizium* and *Beauveria* and their overall biosynthetic pathways are well characterized.

3.1. The *Metarhizium* genomes and known SMs.

3.1.1. Genome comparisons among *Metarhizium* species.

*Metarhizium robertsi* (formerly *anisopliae*) (ARS Collection of Entomopathogenic Fungi) Isolate 2575 has been sequenced (Genbank accession PRJNA230500), and there are two other *Metarhizium* genomes available - *M. anisopliae* and *M. acridum* (PRJNA38717; www.ncbi.nlm.nih.gov/assembly/243998/). Genome sizes are approximately 40 MB; predicted gene count for *M. robertsi*, *M. anisopliae*, and *M. acridum* is approximately 12,300, 10,600, and 9,900, respectively. In *M. robertsi* (ARSEF 2575), there are 85 core genes putatively involved in SM biosynthesis, encoding 16 NRPS, 24 PKS, 9 NRPS-like, 7 PKS-NRPS hybrids, and 28 involved in FAS/terpene or steroid-like biosynthesis. *M. anisopliae* was reported to have comparable numbers of SM gene clusters as *M. robertsi* as these genomes are highly similar.

3. Secondary metabolism in entomopathogenic fungi

Entomopathogenic fungi are a rich source of secondary metabolites, as reviewed previously. Yet recent genomic analyses with *Metarhizium* and *Beauveria* indicate that over 80% of the putative secondary metabolite-associated genes have no identified specific products and have sequences that are unique to this group of organisms. Indeed, in a comparative study of sequenced fungi, secondary metabolite core clusters are better represented in the genomes of entomopathogenic species than in genomes of fungi with other trophic associations. In addition to the obvious antibiotic service that fungal metabolites may perform in entomopathogens, they can play a role in pathogenicity and other interactions between a fungus and its insect host, or they may mediate an inter- or intra-specific communicative function, or aid in mitigating abiotic and biotic stresses. Also, surface host molecules, i.e. epicuticular waxes and lipids, may serve as substrates for enzymes, e.g. cytochrome P450s, involved in secondary metabolite biosynthesis and/or xenobiotic transformations.

3.1. The *Metarhizium* genomes and known SMs.

3.1.1. Genome comparisons among *Metarhizium* species.

*Metarhizium robertsi* (formerly *anisopliae* - ARSEF (ARS Collection of Entomopathogenic Fungi) Isolate 2575 has been sequenced (Genbank accession PRJNA230500), and there are two other *Metarhizium* genomes available - *M. anisopliae* and *M. acridum* (PRJNA38717; www.ncbi.nlm.nih.gov/assembly/243998/). Genome sizes are approximately 40 MB; predicted gene count for *M. robertsi*, *M. anisopliae*, and *M. acridum* is approximately 12,300, 10,600, and 9,900, respectively. In *M. robertsi* (ARSEF 2575), there are 85 core genes putatively involved in SM biosynthesis, encoding 16 NRPS, 24 PKS, 9 NRPS-like, 7 PKS-NRPS hybrids, and 28 involved in FAS/terpene or steroid-like biosynthesis. *M. anisopliae* was reported to have comparable numbers of SM gene clusters as *M. robertsi* as these genomes are highly similar. *M. acridum* has been reported to contain a total of 57 SM core gene clusters, including 13 NRPS, 13 PKS, 8 NRPS-like, 4 PKS-like, 1 NRPS-PKS, and 16 involved in FAS/terpene or steroid-like biosynthesis. Of the SM gene clusters, 3 NPs (destruxins (1), NG391/393 (2), serinocyclin (3)) have been functionally linked to a specific gene cluster in *M. robertsi* and *M. anisopliae*, using gene deletion studies.
precursor forms pipecolic acid that is then derivatized with acetate or produced during conidiation production, is produced by some isolates of potato beetle (coleopteran), and no differences in morphological or its expression in the later stages of infection to stationary phases of liquid culture, and in vivo analyses revealed expression patterns of destruxins suggests developmental regulation of a secondary metabolite gene from genome of ARSEF 2575 were used to test the role of serinocyclin (NG synthetase) was disrupted. No significant change in virulence against larvae of beet armyworm were observed, nor were any significant morphological changes seen, including responses to diffusion, it transports divalent cations across biological membranes and acts as an ionophore; beavercin increases cytoplasmic Ca$^{2+}$ concentration, causes ATP depletion, and activates calcium-sensitive cell apoptotic pathways. In in vitro studies, beavercin reverses the multidrug-resistance (MDR) phenotype in yeast and potentiates the fungicidal activity of fluconazole against fluconazole-resistant Candida albicans at sub-cytotoxic concentrations and known cytotoxic agents against multidrug-resistant (MDR) cancer cell lines. Beavercin hinders directional cell motility of cancer cells. The gene cluster containing the beavercin synthase NPS has been identified and analyzed. KO mutants showed small but measurable effects on virulence in comparison to WT strains when tested against Spodoptera frugiperda, Helicoverpa zea, and Galleria mellonella larvae. The octodepsipeptide bassianolide is a cyclic tetrameric ester of the dipetidopeptide bassianolide (10) are the best studied. The cyclocylogomer dipetidopeptide beavercin, an acyclic trimer of the dipetidopeptide monomer D-hydroxyisovaleric acid (D-Hiv) - N-methyl-L-phenylalanine (N-Me-Phe), is synthesized by multiple Beauveria spp., Paecilomyces and plant pathogenic Fusaria. Beavercin has moderate antibacterial, antifungal, and insecticidal activities as well as potent cytotoxic activity against human cell lines. Through diffusion, it transports divalent cations across biological membranes and acts as an ionophore; beavercin increases cytoplasmic Ca$^{2+}$ concentration, causes ATP depletion, and activates calcium-sensitive cell apoptotic pathways. In in vitro studies, beavercin reverses the multidrug-resistance (MDR) phenotype in yeast and potentiates the fungicidal activity of fluconazole against fluconazole-resistant Candida albicans at sub-cytotoxic concentrations and known cytotoxic agents against multidrug-resistant (MDR) cancer cell lines. Beavercin hinders directional cell motility of cancer cells. The gene cluster containing the beavercin synthase NPS has been identified and analyzed. KO mutants showed small but measurable effects on virulence in comparison to WT strains when tested against Spodoptera frugiperda, Helicoverpa zea, and Galleria mellonella larvae.

Several other SMs are reported from Metarhizium strains, but have yet to be linked to a specific SM gene cluster. Cytochalasins (C, D) (4a,b), macrocyclic lactone derivatives of perhydroindoles that inhibit actin and microtubule formation, were isolated from M. anisopliae, as well as other fungi; in Penicillium expansum, cytochalasin biosynthesis involves a hybrid NRPS-PKS, as suggested by RNA silencing. Swainsonine (5), an indolizidine alkaloid that inhibits α-mannosidases and blocks complex carbohydrate production, is produced by some isolates of M. anisopliae. A precursor forms piperacil acid that is then derivatized with acetate or malonate, although the biosynthetic genes have not been described. The dipetidopeptide tyrosine betaine (6) was identified from conidia of Metarhizium although its presumed NRPS has yet to be identified. The terpenoid helvolic acid (7), an antimicrobial terpenoid identified from Aspergillus fumigatus has also been isolated from M. anisopliae and a geranylgeranyl diposphate synthetase putatively involved in synthesis has been identified. Conidia of M. acridum contain metacraines (8), two novel 17-membered macrocycles consisting of a phenylalanine unit condensed with a nonaketide, suggestive of a hybrid NRPS-PKS biosynthetic route.

3.2. The Beauveria genome and known SMs.

3.2.1. Genome analysis of Beauveria Beauveria bassiana (ARSEF 2860) possesses a genome of approximately 34 MB, and contains approximately 10,400 protein-coding genes. Sequence identity between B. bassiana and M. anisopliae is about 58%, and blast analysis indicates that B. bassiana is more closely related to the Chinese medicinal fungus Cordyceps militaris than to Metarhizium sp. The B. bassiana genome contains 45 SM core genes, including 13 NRPS, 1 PKS, 7 NRPS-like, 1 PKS-like, 3 hybrid NRPS-PKS, and 12 genes related to FAS/terpene/steroid biosynthesis. Three of the SM genes are conserved among the 4 insect pathogenic genomes (1 NRPS, 1 PKS, and 1 terpene synthase (TS)), and 4 clusters are shared with Trichoderma reesei (2 NRPS-like, 1 NRPS, and 1 terpene cyclase (TC)).

3.2.2. Known SMs and SM genes of Beauveria species Beauveria species produce a number of cyclic peptides, of which the cyclocylogomer dipetidopeptide beavercin (9) and the octodepsipeptide bassianolide (10) are the best studied. The cyclocylogomer dipetidopeptide beavercin, an acyclic trimer of the dipetidopeptide monomer D-hydroxyisovaleric acid (D-Hiv) - N-methyl-L-phenylalanine (N-Me-Phe), is synthesized by multiple Beauveria spp., Paecilomyces and plant pathogenic Fusaria. Beavercin has moderate antibacterial, antifungal, and insecticidal activities as well as potent cytotoxic activity against human cell lines. Through diffusion, it transports divalent cations across biological membranes and acts as an ionophore; beavercin increases cytoplasmic Ca$^{2+}$ concentration, causes ATP depletion, and activates calcium-sensitive cell apoptotic pathways. In in vitro studies, beavercin reverses the multidrug-resistance (MDR) phenotype in yeast and potentiates the fungicidal activity of fluconazole against fluconazole-resistant Candida albicans at sub-cytotoxic concentrations and known cytotoxic agents against multidrug-resistant (MDR) cancer cell lines. Beavercin hinders directional cell motility of cancer cells. The gene cluster containing the beavercin synthase NPS has been identified and analyzed. KO mutants showed small but measurable effects on virulence in comparison to WT strains when tested against Spodoptera frugiperda, Helicoverpa zea, and Galleria mellonella larvae.
Tenellin (11a) and bassianin (11b) were first described as yellow pigments from B. bassiana 123, 124; bassianin differs from tenellin by 1 chain extension in the ketide moiety. Eley et al. 125 described the production of the acyltetramic acid tenellin by a hybrid PKS-NRPS gene cluster that shows much similarity to the gene cluster encoding NG39X in Metarhizium and fusarins in Fusarium spp. The tenellin knockout mutants showed no loss of virulence in assays against Galleria mellonella suggesting that tenellin is not involved in pathogenesis. A highly similar gene cluster involved in desmethylbassianin biosynthesis has also been identified from B. bassiana 126.

(Structures 12–14 here)

Several Beauveria SMs have no identified gene cluster as yet. The beauverolides (12) are a family of cyclic tetradepsipeptides that have been reported from B. bassiana 127, 128 as well as other fungal species 129, 130 for which the biosynthetic cluster has not been identified. Oosporein (13) is a red pigment found in both B. bassiana 131 and B. brongniartii 132, 133 that has antifungal activity 134. Earlier labelling studies with acetate and mevalonate suggest a PKS biosynthetic route 135, and it is chemically similar to orsellinic acid whose core PKS gene cluster has been identified in Aspergillus nidulans 136. Bassiatin (14) is a cyclized monomer unit of the trimer involved in beauvericin biosynthesis and thus it might be a shunt product of that pathway 137.

### 4. Comparison of predicted SM cluster genes of Metarhizium and Beauveria

NRPS and PKS genes from B. bassiana and M. robertsii were analyzed using two approaches. The first was a phylogenetic analysis in which the core protein in either NRPSs or PKSs from both fungi were compared to a collection of orthologous proteins of the same type which included proteins associated with a defined metabolite and proteins previously used in phylogenetic studies 138, 139. This analysis had the objective to classify B. bassiana and M. robertsii SM core genes and broadly infer the structure of their natural products. To this end, either adenylation (A) domains (NRPSs and hybrid PKS/NRPSs) or ketoacyl synthase (KS) domains (PKSs and hybrid PKS/NRPs) from target protein sequences were selected using their respective Hidden Markov Models obtained from PFAM (http://pfam.sanger.ac.uk/). Extracted fragments were aligned with Muscle 3.8.31 140, clustered with PhyML 3.1 141 and visualized with either Figtree (http://tree.bio.ed.ac.uk/software/figtree/) or TreeDyn 142 (Fig. 1 and 2). The second analysis was based on the web annotation service antiSMASH 143 and had the objective of identifying similarities among B. bassiana and M. robertsii gene clusters and those of other fungi or bacteria.

Overall, alignment of 81 reference sequences with 36 Metarhizium (EXU/EXV-series) and 18 Beauveria (EJP series) KS domains allowed for the segregation of PKSs into the broad structural classes of reducing, non-reducing and enzymes associated with bacterial type I PKSs as well as more narrowly defined groups (Fig. 1). While this analysis is speculative, it allows for classification and, to a certain extent, comparison of the PKSs found in B. bassiana and M. robertsii as well as very approximate predictions concerning putative backbone structure for the resultant products. This analysis indicates that both B. bassiana and M. robertsii have greater numbers of reducing (10 and 15, respectively, with hybrids not included) than non-reducing PKSs (2 and 10, respectively).

M. robertsii and B. bassiana were found to be each other's closest matches in 7 instances. Eight M. robertsii core PKS genes displayed greater similarity to sequences found in other fungi besides B. bassiana whereas three B. bassiana sequences displayed greater similarity to sequences outside those of M. robertsii. For M. robertsii, these included PKS genes involved in the syntheses of: (1) Aspergillus spp. ochratoxin A production, (EXV03186.1), (2) A. nidulans aspidorine synthesis (EXU98291.1), (3) A. claus clavosynthetic E synthesis (EXU97533.1), (4) Alternaria solani aslanipryone synthesis (EXU95862.1), (5) A. fumigatus pseudotaxin synthesis (EXU98505.1), (6) A. nidulans pkb4 (EXU95974.1), (7) A. alternata AF-toxin synthesis (EXU98661.1), and (8) Penicillium aethiopicum viridicatumtoxin synthesis (EXU97310.1). Similarly, B. bassiana PKSs bear similar to proteins involved in syntheses of: (1) A. terreus lovastatin nonaketide synthesis (QYAS5.1, although in this case a closely related Metarhizium gene also clustered with this group, EXU96139.1), (2) A. nidulans emericellamide synthesis (EJP62832.1), and (3) Cochliobolus heterosporus T-toxin synthesis (EJP70141.1). Neither B. bassiana nor M. robertsii appeared to contain any bacterial Type I or 6-MSAS-type PKSs. M. robertsii genome contains one NRPS-PKS (EXU97632.1), which is structurally similar to C. heterosporus PKS24 138, 144, while B. bassiana lacks this gene (Fig. 1, Bacterial PKS-associated clades). Additionally, within the non-reducing group, neither B. bassiana nor M. robertsii has PKSs clustering with either 1,3,6,8-tetrahydroxynaphthalene (T4HN) synthases or aflatoxin-like PKSs. Phylogenetic analysis show the presence of three M. robertsii and two B. bassiana PKSs divergent enough to be excluded from all the identified clades (Fig. 1).

Due to the modular nature of NRPSs, the phylogenetic analyses of these proteins from B. bassiana and M. robertsii are significantly more complex. Each NRPS module carries information concerning both function and evolutionary history; however neither aspect is well understood. In particular, the “nonribosomal code” of amino acid selection has not been precisely identified for fungal genomes 145,146. The evolutionary history of these modules as well as each NRPS as a whole appear to be a layering of vertical, possibly horizontal transmission events, and intricate genetic rearrangements which have led to a staggering complexity of module assortments 139. Despite these limitations, phylogenetic analysis based on 187 A domains identified from predicted M. robertsii and B. bassiana NRPS, NRPS-like and other adenylation enzymes and 188 reference A domains from other organisms provides several useful clues (Fig. 2). In contrast to PKS phylogeny, direct grouping of B. bassiana and M. robertsii A domains occurs frequently for NRPS-like proteins but only in 7 instances for NRPSs, of which three were from siderophore biosynthesis NRPSs. Candidate proteins involved in intracellular and extracellular (ferriricin and coprogens, respectively) siderophore biosynthesis are found in distinct clades as well as the respective γ-aminoadipate reductases (AARs), involved in lycine biosynthesis. The additional A domains found in these clades belong to proteins that differ in their structure from canonical siderophore or AARs NRPSs and their possible role in either iron metabolism or lycine biosynthesis is difficult to speculate. Characterized M. robertsii NRPS DXS (destruxin) and NPS1 (serinocyclin) have a complex clustering pattern. The first 4 DXS A domains group with both A domains found in perA (peramine biosynthesis NRPS). The last 2 A domains group with the second A domain of both C. heterosporus NPS3 and NPS1 and all are adjacent to a N-methyltransferase domain. For the serinocyclin synthetase NPS1, 6 out of 7 A domains are found in the ergot clade and the A domain from module #2 is clustered with EXV04526.1, an uncharacterized M. robertsii monomodular NRPS. Characterized B. bassiana bassianolide and beauvericin NRPSs have a more straightforward clustering pattern with each one of their A domains falling within either enniatin module 1 or cyclosporine/enniatin module 2 groups (Fig. 2). No B. bassiana or M. robertsii NRPS is included in well-defined clades.
such as penicillin, peptabols, or echinocandins. On the other hand, a clade composed of A domains from *B. bassiana* NRPS EJP64345.1 and two *M. robertsii* NRPSs (EXU95985.1, 8 modules; EXU97071, 8 modules) form a separate group. This clade coincides with the “Insect Pathogen Expanded Clade” identified by others.\(^49\)

Besides differing greatly in the number of predicted core genes, *M. robertsii* and *B. bassiana* secondary metabolism appears highly specific in terms of the nature of the natural products potentially produced (see Supplemental Tables 1, 2, 3, 4). Some PKSs from *B. bassiana* and *M. robertsii* share significant overall structural similarities (examples are EJP68806.1 and EXU96285.1; EJP67836.1 and EXU97187.1), but the respective predicted gene clusters are quite divergent, hinting at dissimilar natural products. Similarities between *M. robertsii* and *B. bassiana* NRPS metabolism are even less pronounced. The overlapping set includes essentially only siderophore and conidial pigment biosynthesis, both of which tend to be conserved among Ascomycota. Thus, whether or not a specific secondary metabolite has a role in pathogenicity or any interaction with other organisms, our analyses indicate that these two entomopathogens rely on a quite different set of natural products, a likely reflection of their different life styles and functional environments, and consistent with their entomopathogenicity being an example of convergent evolution in which one could predict a similar overall process mediated by unique mechanisms.

### 5. Predicting novel pathways

Phylogenetic analysis coupled with comparison of entire gene clusters to those characterized in other fungi can be used to approximately predict the structure of natural products associated with some of the *B. bassiana* and *M. robertsii* PKSs and NRPSs. Here, we give 4 examples.

*B. bassiana* harbours 2 nonreducing PKSs (EJP64619.1 and EJP62792.1) structurally similar to those involved in conidial pigment production in other fungal species. Together with *M. robertsii* EXU96629.1, EJP64619.1 clusters with YWA-like PKSs and is co-localized with genes similar to those involved in melanin biosynthesis (EJP64622.1; tetrahydroxynaphthalene reductase; EJP64623.1: scytalone dehydratase), which makes it a likely heptaketide naphthopyrone synthase. Despite its structural similarity to EJP64619.1 and other YWA-like enzymes, the second nonreducing PKS (EJP62792.1) does not group with either (Fig. 1). EJP62792.1 gene is co-localized with a transporter (EJP62793.1) and a laccase (EJP62796.1), which suggests secretion and dimerization of its product, respectively (Fig. 3A). This scenario is similar to that described for the biosynthesis of aurofusarin\(^65\) and consistent with a pathway leading to oosporein production.

A second prediction can be proposed for a mixed PKS-NRPS pathway. Both phylogenetic and AntiSMASH analyses indicate that the PKS EJP62832.1 and the NRPS EJP62835.1 may participate in the biosynthesis of a compound similar to emericellamide.\(^66\) In *A. nidulans* four proteins (NRPS, PKS, aetyltransferase, acyl-CoA ligase) are required for that pathway and all are found in close proximity in *B. bassiana* (Fig. 3B). This gene cluster is conserved in *C. militaris*, but appears to be absent in *M. robertsii* and *M. acridum*. *M. robertsii* also contains several recognizable pathways beyond those dedicated to the biosynthesis of melanin (EXV02536.1, ferricrocin (EXV05490.1) and a coprogen-type siderophore (EXV04699.1). One includes two NRPSs (EXU97504.1 and EXU97505.1) in which phylogenetic analyses associate with NRPSs involved in alkaloid biosynthesis and co-localize with genes also similar to those involved in ergot alkaloid biosynthesis in *C. purpurea* (Fig. 3C).\(^152\) This cluster is conserved in *M. acridum*. No clear ergot-like cluster was identified in *B. bassiana* although the NRPSs EJP6634.1, EJP67097.1, and EJP68775.1, respectively, have 2 out of 3, 3 out of 3 and 1 out of 1 adenylation domains within the ergot clade (Fig. 3). A further predicted pathway in *M. robertsii* is that of the PKS EXU97310.1. This non-reducing PKS is structurally similar to that responsible for the biosynthesis of the tetracycline-like antibiotic viridicat toxin in *P. aethiopicum*.\(^153\) EXU97310.1 is flanked by homologs responsible for the majority of the tailoring steps and secretion of this antibiotic, except for genes corresponding to VrtC, VrtE and VrtK (Fig. 3D). These proteins are presumably responsible for the addition and modification of a geranyl group to the tetracyclic polyketide in *P. aethiopicum*, which is likely to be absent in the corresponding metabolite in *M. robertsii*.

### 6. Specific biochemical transformations mediated by entomopathogenic fungi

The above sections summarize aspects of the phylogenetics of some of the secondary metabolite biosynthetic clusters present in *M. anisopliae* and *B. bassiana*. However, these organisms (*Beauveria* spp. in particular) have been extensively used in chemical compound transformations. While little is known about the specific enzymes involved in these reactions, select genes found in secondary metabolite clusters (e.g. especially cytochrome P450s) are likely to be involved in modification of exogenously supplied chemical compounds. Biotransformation involves the use of the enzymatic repertoire of an organism to catalyse the chemical modification of a compound. Reactions involving biotransformation can be separated into two broad categories: Xenobiotic transformations occur when an organism acts on a completely unknown (i.e. does not occur endogenously) substrate. In contrast, biosynthetically patterned transformations involve modification(s) of substrates related to compounds found in intrinsic biochemical pathways present in the catalysing organism. In the case of microbial agents, the use of the whole-cells as biocatalytic units has been considered a sustainable, “green chemistry” approach for large scale biosynthesis, novel compound discovery, biosynthetic pathway probing, as well as models that can be used to examine the metabolic fate of drug candidates.\(^154\)

Although an a number of entomopathogenic fungi have been used as whole cell biocatalysts for various modification of chemical substrates, often to isolate new bioactive compounds, *Beauveria* sp have by far been the most frequently utilized organisms for such purposes.\(^155\) Reactions catalyzed in such usages have included oxidations (hydroxylations), reductions, transglycosidations, and hydrolytic transformations. The major impetus for such conversions has been to produce (novel) compounds that display enhanced biological activities centered on providing human health benefits. Thus the use of chemical compounds typically start by being considered to possess various anti-microbial, anti-inflammatory, immune-stimulatory, anti-proliferative, and/or tissue-protective activities, often with a particular compound or various derivatives professed to display a remarkably wide range of the activities listed above. Here we will summarize the main chemical transformation reactions catalyzed by *B. bassiana* whole cells by providing recent examples of such work.

(Structures 15–15a here)

**Hydroxylation**

*Silybin* (15) is a flavonolignin that is the major constituent of herbal preparations derived from the seed extract of the milk thistle (*Silybum marianum*) used in pharmaceutical products for promotion of liver health. In addition to its hepatoprotective effects, silybin has been reported to display antioxidant (perhaps the mechanism behind
its liver health benefits?) and antipro liferative effects, the latter against a wide range of human cancer cell types including prostate, breast, cervix, lung, and liver. Silybin may also act as a neuroprotective, and offer treatment for heart and gastrointestinal problems. Silybin was transformed to 8-hydroxysilybin (15a) by B. bassiana, and the resultant product displayed 8-9 fold greater free radical scavenging activity than that the parent compound (silybin) 156.

Lactone derivatives of steroids are of interest due to their potential for displaying anticancer, antiandrogenic, and/or cholesterol-reducing activities. A number of microorganisms including B. bassiana are capable of catalyzing Baeyer–Villiger (BV) oxidations that result in the conversion of ketones to lactones or esters. The steroids, epi- and dehydroepiandrosterone (16, 17), androstenedione (18), androstenediol (19), and progesterone (20) were first transformed to 11-α-hydroxy-derivatives (16-20a, 18b) which then subsequently underwent BV oxidation reactions to form a series of 11α-hydroxy ring D δ-lactones. For some substrates, prolonged incubation with the fungal cells also resulted in reductions to corresponding (17β)-alcohols 157. The consequences of these modifications on the activities of the various steroids have yet to be reported.

(Structures 16–20 here (Biotransformations_1Final))

Lignans represent a class of structurally diverse plant-derived polypropenoids with a wide range of bio-therapeutic applications. Aryltetralins and aryltetralones, in particular, are of interest as precursors for the synthesis of the anticancer drugs etopoide and teniposide. Incubation of the aryltetralin lignan (−)-isosalbin (21) with B. bassiana resulted in the formation of (−)-8-hydroxyisosalbin (21a) as the only isolatable product 159.

Biotransformation of various steroid compounds by entomopathogenic fungi has also been investigated. The cardiovascular drug, mekrenone (22), was modified to two new derivatives, 11α- and 12β-hydroxymekrenone (22a, 22b), as well as the known metabolite 6β-hydroxymekrenone (22c) by B. bassiana 160.

(Structures 21–22 here)

Glycosidation
Flavonoids represent a class of pigmented ketone-containing plant metabolites characterized by a three-cycle (ring) backbone. These compounds are of particular interest due to potential beneficial effects as phytoestrogens, as part of hormone replacement therapies, as cancer chemo-preventive agents, as well as due to variously described antimicrobial, antiviral, and antioxidant activities.

A number of hop (beer) derived flavonoids have been transformed to glycosylated products after incubation with B. bassiana whole cells, including xanthohumol (23), isoxanthohumol (24), and 8-prenylaragavrin (25). Xanthohumol was transformed by B. bassiana via regioselective C-4′-glycosylation to xanthohumol 4′-O-β-D-(4′″-O-methyl) glucopyranoside (162)(23a), isoxanthohumol was converted to isoxanthohumol 7-O-β-D-(4″′-O-methyl)glucopyranoside (162)(24a), and 8-prenylaragavrin was converted to 8-prenylaragavin 7-O-β-D-(4″′-O-methyl)glucopyranoside (162)(25a) and 8-prenylaragavin 7-O-β-D-glucopyranoside (162)(25b) 167. The former product (24a) was reported to display higher anti-proliferative activity (against a human colon cancer cell line) than the parent chemical, and the latter two compounds (24a, 25a) were considered novel, reported for the first time by the authors. These data indicate the high probability of novel compound production, with flavonoids of tricyclic structure undergoing the highly selective 4′-O-methyl- β-D-glucosidation at the C7′-OH and for chalcones at C4′′-OH.

(Structures 23–25 here)

Selective glycosidation of anthraquinones, polyketide natural products being developed as leads for novel antimicrobial, antiproliferative, and anti-inflammatory compounds discovery, has also been reported 164. In addition, B. bassiana has been shown to transform 1-aminooctracone (26), a potent carcinogen, to various metabolites via acetylation (26a), oxidation (26b,c), and hydroxylation and O-methylglucosylation reactions (26c,d,e) 165.

(Structure 26 reaction scheme here (Biotransformations_2Final))

Demethylation
Donepezil (2,3-dihydro-5,6-dimethoxy-2-[[1-(phenylmethyl)-4-piperidinyl][methyl]-1H-inden-1-one (27), is among a small handful of chemical compounds used in the treatment of severe Alzheimer’s disease, acting as a reversible inhibitor of acetylcholinesterase. Donepezil transformation by B. bassiana produced trace amounts of 6-O-desmethyl donepezil (5-ODD) (27a) while the predominant product was 5-O-desmethyl donepezil (5-ODD) (27b), representing O-demethylation products 166.

(Structure 27 reaction scheme here (Biotransformations_2Final))

Phenylurea derivatives have been extensively used as herbicides for prevention of the growth of undesirable plants and the environmental fate and/or microbial transformation of these products is of significant interest. Diuron, N(3,4-dichlorophenyl)N′,N′-dimethyleurea (28), is employed for total weed control particularly on non-cultivated areas, e.g. roads, where it can accumulate in the soil. Incubation of diuron in the presence of B. bassiana resulted in the formation of two terminal nitrogen atom demethylated products, 3,4-dichlorohydroxyurea (28a) and N-3,4-dichlorophenyl-N′-methylurea (28b). Of note, chemical syntheses and testing of both (N-demethylated) degradative products revealed that they possessed greater toxicity in some assays than the parent compound 167.

Redirection of intrinsic biosynthetic pathways
Fungi produce a host of small peptides synthesized using NRPS enzymatic mechanisms. This class of compounds includes the β-lactams and in B. bassiana, beauvericins (cyclooligomeric depsipeptides) that have various biological functions. Strategies that have yielded combinatorial biosynthesis of novel compounds via redirection of the beauvericin biosynthesis pathway have recently been reviewed 168. Briefly, using both wild type and mutants deficient in the supply of the precursor compound, D-hydroxyisovalerate, feeding of substrate mimics, i.e. replacement of D-hydroxyisovalerate or L-phenylalanine with 2-hydroxybutyrate, 2-hydroxy-3-methylvalerate, or 3-fluorophenylalanine, resulted in novel compounds that displayed altered biological activities 169, 170.

Use of biotransformations in general
Whole-cell biotransformations seek to exploit enzymatic regio- and stereo- specificity to selectively modify chemical compounds. In many instances, use of B. bassiana as a whole cell catalyst has resulted in regioselective transformation of substrates into a small number of defined products. However, it should be noted that depending upon the substrate sequential reactions, i.e. demethylation, hydroxylation, and glycosylations could occur. Thus attempts in the
synthesis of the alkaloid epibatidine, a potent analgesic \(^{171}\), and derivatives using \textit{B. bassiana} resulted in the formation of O-demethylated compounds, which were further metabolized to β-4-methylglycosides, as well as their corresponding C-hydroxylated products \(^{172}\).

Most chemical transformation studies have focused on \textit{B. bassiana}, although several \textit{Isaria} strains have also been used. Aside from direct biotransformations, these systems have provided approaches for biomimetic syntheses. It is intriguing to speculate that the diverse range of chemical transformations already demonstrated, presumably mediated by an arsenal of fungal enzymes, is a consequence of the interactions that entomopathogens have with plants, insects, and other organisms, and hence the need for detoxification and/or assimilation of xenobiotic compounds. Given that there are several other tractable entomopathogenic fungi, i.e. \textit{Metarhizium} and \textit{Cordyceps sp} for which genomes and molecular tools are available, there is likely a significant potential for further exploitation of these organisms as whole cell biocatalysts.

Conclusions
Genomic data show that entomopathogenic fungi are rich in SM gene clusters, revealing the potential to yield a wealth of chemical compounds. Comparative genomics has allowed for the prediction of several of these orphan pathways, pointing to likely scaffolds for the final products, although these predictions await experimental validation. A number of SM pathways in entomopathogenic fungi have been genetically characterized; however, despite these initial findings, and as seen in most fungi, the vast majority of SM gene clusters, their biological role(s), and the compounds produced remain to be uncovered. Specific genes in SM clusters are also likely to be responsible for the wide range of biotransformations mediated by entomopathogenic fungi in drug discovery and remediation research. With the availability of sequenced genomes, future research in the field is now poised to unravel the SMs produced by these gene clusters and the specific enzymes involved in transformations of chemical compounds.

References


Figure legends

Figure 1. Phylogenetic/structural analysis of polyketide synthases (PKSs) and hybrid polyketide synthases-nonribosomal peptide synthetases (hybrid PKS-NRPSs) from *Beauveria bassiana* ARSEF 2860 and *Metarhizium robertsi* ARSEF 2575 inferred using Maximum-Likelihood (PhyML 3.1). Fifty three ketoacyl synthase (KS) domains identified in proteins from these two fungi were aligned with 82 KS domains from PKSs and hybrid PKS-NRPSs for which the natural product is known. Animal fatty acid synthases were used as the outgroup. Clades all have a bootstrap support >50%. T4HN: 1,3,6,8-tetrahydroxynaphthalene. Orph_00006 indicates a sequence found in a *M. robertsi* orphan contig not deposited in GenBank. Reference proteins used for the analysis were: *Alternaria alternata* AFT9-1; *A. solani* at5; *Aspergillus fumigatus* YWA1; *A. oryzae* PksL1; *A. terreus* LovF, LovB, AtX; *A. clavatus* CcsA, MsaS; *A. fumigatus* fma-PKS, psoA; *A. nidulans* stcA, WA, afoE, afoG, phka, phkB, easB, pkfA, AptA, cicF, pkgA, stcJ, orsA, ausA, apdA, pkdA, mdpG; *A. niger* An15g07920, YWA1, An14g04850, FUM1-like; *A. nomius* PksA; *A. ochraceoroseus* At1C; *A. ochraceus* LC35-12; *A. parasiticus* PksL1; *A. sojae* PksL1; *Bombyx mori* FAS, P270; *Cercospora nicotianae* CTB1; *Cochliobolus carbonum* PKS24; *C. heterostrophus* PKS1; PKS2; *Colletotrichum lagenarium* PKS1; *Elsinoe fawcettii* PKS1; *Fusarium heterosporum* eqxS, fisd; *F. oxysporum* FUM1; *F. pseudogroenlandicum* PKS10; *G. fujikuroi* Mon4, FusA; *G. moniliformis* FusS, FUM1; *G. zeae* PKS12, PKS13, PKS4; *Glairea lozoyensis* pks1; *Hypomyces subiculosus* hpm3; *Monascus purpureus* PkSCT; *Mucor musculus* FAS, *Mycobacterium sp.* NC_009077.1; *Mycobacterium ulcerans* PpsB, *Mycosphaerella pini* PksA; *M. zeae-maydis* PKS1; *Nodulisporium sp.* PKS1; *Nostoc sp.* all1648; *Penicillium aethiopicum* gsfA, VrtA; *P. brevicaespactum* mpatC; *P. expansum* cheA; *P. graveolulum* 6-MSAS; *P. nordicum* otapksPN; *Saccharomyces cerevisiae* fis2; *Saccharopolyspora erythraea* AM420293.1; *Salinispora arenicola* pks3a; *Sordaria macrospora* pks; *Streptomyces avermitilis* AVES3; *Wangiella dermatidis* PKS1; *M. robertsi* accessions are reported in green color; *B. bassiana* accessions are reported in red color. Included are examples of metabolites produced by a PKS found within a specific clade.

Figure 2. Phylogenetic/structural analysis of nonribosomal peptide synthetases (NRPSs) and hybrid polyketide synthases-nonribosomal peptide synthetases (hybrid PKS-NRPSs) from *Beauveria bassiana* ARSEF 2860 and *Metarhizium robertsi* ARSEF 2575 inferred using Maximum-Likelihood (PhyML 3.1). One hundred eighty seven adenylation domains (As) identified in proteins from these two fungi were aligned with 375 AMP domains from NRPSs, NRPS-like, acyl-CoA ligases and hybrid PKS-NRPSs for which respective products are known. Acyl-CoA ligases were used as the outgroup. Clades all have a bootstrap support >50%. T4HN: 1,3,6,8-tetrahydroxynaphthalene. Orph_00006 indicates a sequence found in a *M. robertsi* orphan contig not deposited in GenBank. Reference proteins used for the analysis were: *Alternaria alternata* AMT; *A. brassicaceae* Nps1; *Aspergillus clavatus* ccsA; *A. fumigatus* aarA, Afu5g10120, CPS1, FacA, gliP, nps1, pes1, psoA, sidC, sidD, sidB; *A. nidulans* acveA, cicB, easA, micA, TdiA; *A. terreus* apvA, btyA; *Claviceps purpurea* CPPS, CPPS2, CPPS3, CPPS4; *Cochliobolus carbonum* HTS1; *C. heterostrophus* CPS1, NPS1, NPS2, NPS3, NPS4, NPS5, NPS6, NPS7, NPS8, NPS9, NPS10, NPS11, NPS12; *Emericella rugulosa* EcdA; *Epicoccum festucae* PerA; *Fusarium equiseti* ESYN1; *F. graminearum* CPS1; *F. pseudogroenlandicum* CPS1; *Gibberella moniliformis* FusS; *G. zeae* AAR, FG03589.1, FG01671.7, NPS6, NPS10; *Glairea lozoyensis* GLNRPS4; *Leptosphaeria maculans* maal1, SirP, *Magnaporthe oryzae* SSMM1, SSM2, NSP6, syn2; *Metarhizium robertsi* PseA; *Neurospora crassa* AAR, NCU00608.7, NPS6; *Omphalotus olearius* Fso1; *Penicillium chrysogenum* ACVS1; *P. expansum* cheA; *Pyrenophora tritici-repentis* PTRG_06150; *Saccharomyces cerevisiae* ACS2, CPS1-like, LYS2; *Schizosaccharomyces pombe* aar, sid1; *Tolypocladium inflatum* SimA; *Trichoderma viride* tex1; *Ustilago maydis* ler3, sid2, UM01434.1, UM04803.1. *M. robertsi* accessions are reported in green color; *B. bassiana* accessions are reported in red color. Included are examples of metabolites produced by an NRPS found within a specific clade.

Figure 3. Comparison of known gene clusters to similar clusters in *Beauveria bassiana* and *Metarhizium robertsi*. A. Organization of the predicted oosporein biosynthesis gene cluster in *Beauveria bassiana*; B. Comparison of the emericellamide synthase gene cluster in *Aspergillus nidulans* to those found in *Beauveria bassiana* and *Cordyceps militaris*; C. Comparison of the ergot biosynthesis gene cluster in *Claviceps purpurea* to the ergot-like gene cluster identified in *Metarhizium robertsi*; D. Comparison of the viridiamide biosynthesis gene cluster in *Penicillium aethiopicum* to a similar gene cluster identified in *Metarhizium robertsi*. This journal is © The Royal Society of Chemistry 2012.
Figure 1.

**Examples**

- fumonisins
- pyranonigrin
- tenellin
- pseurotin
- fusarin C, NG-391
- lovastatin lovB
cytochalasin CcsA
- zearalenone PKS4
  - T toxin
- emericellamide
  - AF toxin
- asperfuranone afuG
  - lovastatin lovF
- aflatoxin
cercosporin
- YWA1
  - bikaverin
- asperlidiin
  - griseofulvin
- T4H synthases
- orsellinic acid
  - tryptophan
- asperfuranone afuE
  - aspermidine
- 6-MSAS
- NRPS-PKSs

Reducing PKSs and PKS-NRPS hybrids

Non-reducing PKSs

Clades associated to bacterial PKSs

This journal is © The Royal Society of Chemistry 2012
Figure 2.
Figure 3.