ABSTRACT

NOAR, ROSLYN DANIELLE. Pathogenicity Mechanisms of *Mycosphaerella fijiensis*, Causal Agent of the Black Sigatoka Disease of Banana. (Under the direction of Dr. Margaret Daub).

Black Sigatoka, caused by the fungus *Mycosphaerella fijiensis*, is considered the most economically damaging banana disease. Despite its importance, little is known about the molecular basis of pathogenicity. Some of its close relatives produce polyketide toxins that kill host tissue and enable fungal nutrient acquisition. The goal of this research was to identify putative pathogenicity genes in this fungus, with focus on possible polyketide toxins. Bioinformatics analyses predicted seven polyketide synthases (PKS) and one hybrid PKS/non-ribosomal peptide synthase (NRPS) in the *M. fijiensis* genome. Putative biosynthetic gene clusters were also identified. The phylogenetics analyses and putative biosynthetic cluster comparisons predict that the *M. fijiensis PKS10-1* cluster encodes proteins involved in DHN melanin synthesis, and also show that the *PKS2-1*, *PKS8-2*, and *PKS10-2* gene clusters are similar to clusters in other fungi that produce alternapyrone, fumonisin, and solanapyrone, respectively. RNA-Seq analysis of *M. fijiensis*-infected leaf tissue in the necrotrophic phase showed that *PKS7-1*, *PKS8-2*, *Hybrid8-3*, *PKS8-4*, and *PKS10-2* are more highly expressed in infected leaf tissue than in culture medium.

Two polyketide synthases were selected for further analysis. Promoter:GFP fusion experiments suggested that PKS8-4 is expressed in developing pseudothecia in the sexual reproductive cycle. Phylogenetic analysis indicated that it is homologous to PKS enzymes in other fungi that are required for female fertility. A *pks8-4* disruption mutant was generated but was not altered in pathogenicity. GC-MS profiling identified significant changes in non-polar metabolites between the wild type and *pks8-4* mutant. Phylogenetic analysis of the

PKS8-1 cluster showed similarity to monodictyphenone-producing clusters in other fungi. *PKS8-1* was expressed in cultures producing a red-fluorescent compound with properties similar to perylenequinone toxins in other fungi, and expression of the MFS transporter gene in the PKS8-1 cluster correlated with production of the compound. From extracts of the red-fluorescing *M. fijiensis* hyphae, fourteen compounds that had not previously been reported from *M. fijiensis* were identified, including an idebenone-like metabolite, 1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene, and pulverochromenol, which are similar to molecules produced by polyketide biosynthetic pathways. Overexpression of a transcription factor gene from the *PKS8-1* gene cluster resulted in increased expression of many, but not all, genes in this cluster.

Transcriptome analysis enabled the identification of additional putative pathogenicity genes beyond those encoding polyketides that may be involved in the necrotrophic stage of disease. Genes with higher expression in infected leaf tissue include: genes encoding salicylate hydroxylase-like proteins, CFEM domain-containing proteins, hydrophobic surface binding proteins, proteins with characteristics common in effectors, proteins with Domain of Unknown Function 3328, and enzymes commonly associated with secondary metabolism. Analysis of differentially expressed gene clusters identified clusters for a NRPS and a novel fusicoccane. Transcriptome analysis also confirmed previous reports that 14 scaffolds in the *M. fijiensis* genome have characteristics common in dispensable chromosomes, as 14 putative dispensable scaffolds were identified with a distinctly different expression pattern than ‘core’ scaffolds. These included two putative dispensable scaffolds with a high percentage of genes with higher expression in infected leaf tissue, suggesting that those scaffolds may be involved in pathogenicity. No transcripts were mapped to two other

scaffolds, and PCR analysis suggested that these scaffolds are largely absent in the *M. fijiensis* isolate used for RNA-Seq. Overall, this research provides important insight into the pathogenicity mechanisms for an under-studied pathogen, and provides many exciting new avenues for future research.

Pathogenicity Mechanisms of *Mycosphaerella fijiensis*, Causal Agent of the Black Sigatoka Disease of Banana

by
Roslyn Noar

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Plant Pathology

Raleigh, North Carolina

2016

APPROVED BY:

_____	_____
Dr. Margaret Daub Committee Chair	Dr. Jean Ristaino
_____	_____
Dr. Gary Payne	Dr. Peter Balint-Kurti

DEDICATION

This work is dedicated to my loving husband Jesse Noar, and my supportive parents Cliff and Elaine Duvall, and in-laws Scott and Pat Noar.

BIOGRAPHY

Roslyn Noar was born in Ashland, Kentucky to her parents Cliff and Elaine Duvall. Throughout her childhood, she was fascinated by science, especially biology. While at Russell High School, she competed in science on the academic team and in Science Olympiad. Partly because of this experience, she decided to pursue research opportunities in college.

While at Cornell University, she gained research experience in several labs and realized that she particularly loved plant-microbe interactions. In 2007, she graduated with a Bachelor of Science degree in Biology with a Genetics concentration. After graduating, she worked as a research assistant in Maria Harrison's lab at Boyce-Thompson Institute for two years, and then decided to pursue her PhD. In 2010, she began her PhD in Plant Pathology at North Carolina State University in Margaret Daub's lab. She was funded for three years by a National Science Foundation Graduate Research Fellowship, and another two years by a Molecular Biotechnology Traineeship from the National Institutes of Health. She hopes that this and future research will contribute toward more profitable, efficient, and sustainable crop production.

ACKNOWLEDGMENTS

First, I would like to thank my advisor Dr. Margaret Daub, for the opportunity to work on this project and her countless hours of support, instruction, and guidance. Despite her busy schedule as department head, she was always available when I needed advice. I would also like to thank my other committee members, Drs. Jean Ristaino, Gary Payne, and Peter Balint-Kurti, for their valuable comments and suggestions. Thanks to Drs. Gary Payne and Marc Cubeta for the opportunity to serve as a teaching assistant for their classes, and thanks to all my other professors of my classes for sharing their knowledge which was so important for success in my research.

I would like to thank all of the Daub lab members for their support, friendship, advice, and assistance over the years: Elizabeth Thomas, Sonia Herrero, Morgan Carter, Dan DiCorpo, Aydin Beseli, Rika Judd, Hayde Eng, Marilia Goulart da Silva, and Nafisa Gomez. Thanks to Dr. Xie, Dongming Ma, and Seyit Yüzüak for their assistance and valuable comments and instruction for analyzing polyketides. Thanks to Jessica Pope, Erica Lassiter, and Greg O'Brian for their help while I was a rotation student in the Ristaino and Payne labs. Thanks to my classmates in the Department of Entomology and Plant Pathology and the Department of Plant and Microbial Biology for their friendship and encouragement.

Thanks to the National Science Foundation, the National Institutes of Health, and Dole Food Company for funding my education and research, without which this work would not have been possible.

I would also like to thank my family for believing in and supporting me. I would especially like to thank my husband Jesse, because without his support and encouragement, this would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 5: A Novel Polyketide Synthase Gene Cluster in the Plant Pathogenic Fungus *Mycosphaerella fijiensis*

LIST OF FIGURES

CHAPTER 1: Literature Review

CHAPTER 2: Bioinformatics Prediction of Polyketide Synthase Gene Clusters from *Mycosphaerella fijiensis*

CHAPTER 3: Transcriptome Sequencing of *Mycosphaerella fijiensis* During Association with *Musa acuminata* Reveals Candidate Pathogenicity Genes

CHAPTER 4: A Polyketide Synthase Gene Cluster Associated with the Sexual Reproductive Cycle of the Banana Pathogen, *Mycosphaerella fijiensis*.

CHAPTER 5: A Novel Polyketide Synthase Gene Cluster in the Plant Pathogenic Fungus *Mycosphaerella fijiensis*

CHAPTER 1

Literature Review

Introduction

Bananas and plantains are one of the most important crops in developing countries, in terms of gross production value[1]. Bananas are important both as a subsistence crop and as an export crop: about 87% of bananas grown are consumed locally[1], and in 2013, banana exports exceeded 17 million tons[2]. Current edible banana cultivars include diploids, triploids, and tetraploids of two banana species: *Musa acuminata* and *Musa balbisiana*[3]. Cavendish cultivars make up the vast majority of bananas grown for export; these are triploid "dessert" banana varieties of *M. acuminata*[4,5].

Black Sigatoka, caused by the ascomycete fungus *Mycosphaerella fijiensis*, is considered the most important banana disease[6]. This disease results in necrotic streaks on the leaves, loss of photosynthetic capacity, and reduced yield[5]. It also causes premature fruit ripening, which can be problematic for export companies, since the fruit can become over-ripe in transit[7]. Yield losses can exceed 50% in the absence of treatment[8]. Black Sigatoka was first described in Sigatoka Valley in Fiji in 1963, though analyses of herbarium samples suggest that it was present in Taiwan by 1927[5,6,9]. Since then it has spread to most banana-growing regions around the world, including Latin America, Asia, Africa, and throughout the Pacific[5].

M. fijiensis reproduces both sexually and asexually through ascospores and conidia, respectively. Ascospores are considered the primary means of disease spread since they are smaller, more abundant, and can travel several kilometers via wind currents[6,10]. Once conidia

or ascospores land on new leaves in wet or high humidity conditions, they germinate and grow epiphytically on the leaf surface before forming stomatopodia (hyphal swellings) and infecting via the stomata[5,11,12]. There, the hyphae encircle the substomatal chamber and then ramify intercellularly through the leaf tissue[5,13]. Although no haustoria form, there is an extended period of biotrophy that can last for several weeks before the fungus switches to necrotrophy, at which point necrotic leaf streaks can be observed[5,14]. Hyphae in the substomatal chamber give rise to both conidiophores producing conidia and spermagonia producing the male gametes called spermatia[5,15]. Pseudothecia with ascospores are produced in necrotic leaf tissue, and ascospores are discharged upon leaf wetting, continuing the cycle[16,17].

Disease control is accomplished on commercial plantations by fungicide treatments, which are applied up to 60 times per year and are estimated to comprise about a third of the total production cost of banana[5–7,18]. Although the need for fungicides is an expensive problem for commercial plantations, the disease is especially devastating to subsistence farmers, who may not be able to afford the fungicides used to treat the disease[6]. Chemical control is achieved via a combination of site-specific and protectant/multi-site fungicides. Eight main site-specific classes of fungicides with different modes of action are commonly used in mixtures or alternating patterns: QoIs (quinone outside inhibitors), DMIs (demethylation inhibitors), SDHIs (succinate dehydrogenase inhibitors), BCMs (benzimidazoles), N-phenylcarbamates, anilinopyrimidines, amines and guanidines[19]. High levels of resistance to BCMs has become widespread, and resistance has been reported for QoIs, DMIs and SDHIs[5,7,19,20]. Multi-site fungicides used for black Sigatoka control include

mancozeb, chlorothalonil, metiram, captan, thiram, and propineb; these are considered fungicides of low resistance risk[19].

In addition to chemical control, cultural practices can aid in the management of black Sigatoka disease. Growing plants in good, fertile soil results in more vigorous plants with reduced disease severity[21–23]. Pruning necrotic leaves or leaf areas greatly reduces the amount of inoculum and therefore reduces disease spread[21]. Pruned leaves should ideally be removed from the plantation and burned or buried to eliminate the inoculum, though in practice this can be prohibitively labor-intensive[21]. More commonly, pruned leaves are left lying against the plantation soil, which reduces inoculum but not as effectively[21]. Urea may also be applied to the leaf debris to accelerate decomposition[24]. Another important cultural practice is the use of clean planting material. While bananas have traditionally been planted from suckers, this practice can spread a variety of diseases from field to field[25]. Therefore, large companies have turned to tissue culture to propagate vigorous plants that are free from many important pathogens[26]. Though wind dispersal of *M. fijiensis* ascospores is likely to infect new plantings, tissue culture can help ensure that the plants are as free from the stress of other diseases as possible. It has long been observed that symptoms of Sigatoka diseases are much less severe when plants are grown in partial shade[5]. Yellow Sigatoka is a disease caused by *Mycosphaerella musicola*, a very close relative of *M. fijiensis*[5]. An early management recommendation for yellow Sigatoka was to plant in the shade of coffee trees to mitigate symptoms. This was shown to increase marketable banana yields by 50%[27]. For black Sigatoka, though the incubation period is the same regardless of lighting conditions,

development of further symptoms is slowed in shaded conditions, leading to fewer necrotic leaves[28].

The use of resistant banana cultivars could also provide an environmentally-friendly means of black Sigatoka control. Black Sigatoka resistance was identified in certain banana varieties, such as Yangambi Km 5, Calcutta 4, and P. Lilin[29]. Ortiz and Vuylsteke studied the genetic inheritance of resistance by crossing the diploid resistant banana 'Calcutta 4' with susceptible triploid plantain cultivars[30]. Segregation ratios for resistance were consistent with one major recessive allele named bs_1 and two independent alleles with additive effects, named bsr_2 and bsr_3[30]. This is similar to the genetic basis of resistance to *M. musicola*, which is thought to depend in part on multiple recessive alleles[30].

Although conventional breeding for resistance is difficult and labor intensive due to sterility of edible banana cultivars, successful crosses have been done by Fundación Hondureña de Investigación Agrícola (FHIA) and International Institute of Tropical Agriculture (IITA), and black Sigatoka-resistant dessert bananas, cooking bananas, and plantain varieties have been developed[31]. Of the hybrids developed, FHIA-01, or 'Goldfinger' was one of the most promising varieties, with the main flaw of this variety being its softness compared to Cavendish when ripe[31]. Unfortunately, breakdown of resistance to black Sigatoka has already been observed for this variety in Samoa[32]. Development of resistant cultivars with different resistance genes will therefore be needed to help manage black Sigatoka disease.

Genomic and Transcriptomic Analyses

In recent years, a number of genetic resources have become available and have facilitated research on the *Musa acuminata-M. fijiensis* pathosystem. A genetic linkage map based on AFLP, SSR markers, and the MAT locus has been developed for *M. fijiensis*[33], and the genome of *M. fijiensis* has been sequenced through the Joint Genome Institute's (JGI) Community Sequencing Program. The genome sequence of the banana host *M. acuminata* subsp. *malaccensis* doubled-haploid Pahang is publicly available[34]; this variety was chosen for sequencing because it contributed one of the three *M. acuminata* genomes to Cavendish[35]. A BAC library has been created of the wild diploid banana 'Calcutta 4' clone of *Musa acuminata* subsp. *burmannicoides*, which is resistant to black and yellow Sigatoka, Panama disease, burrowing nematodes, and banana weevils[36,37]. In addition to the genetic tools available for *Musa acuminata*, a draft genome sequence is available for *Musa balbisiana* doubled-haploid variety Pisang Klutuk Wulung, which is one of the possible ancestral parents that provided the B genome that is present in today's plantain and cooking banana varieties[38].

Having the genome sequences of both banana and fungus enables better analysis of transcriptomics data, and therefore predictions can be made of genes involved in pathogenicity and resistance to disease. Several studies have analyzed expression of banana genes during infection with *M. fijiensis*. Rodriguez et al used microarray analysis to compare genes expressed in the susceptible ‘Williams’ variety to the resistant ‘Calcutta 4’ variety, when exposed to *M. fijiensis*[39]. ‘Calcutta 4’ showed higher expression of genes encoding phenylalanine ammonia lyase, peroxidase, disease resistance response 1, PR-4 and PR-10,

when compared to 'Williams,' between 6 and 24 hours post-inoculation[39]. Timm et al generated a suppression-subtractive cDNA library of 'Calcutta-4' at different time points in the biotrophic phase compared to uninoculated 'Calcutta-4'[40]. Genes up-regulated after being challenged with *M. fijiensis* include ones encoding putative transcription factors, the NBS-LRR protein RGA1, and cytochrome P450 pathways[40]. They also detected effects on genes involved in the glycolysis/gluconeogenesis pathway and scavenging of reactive oxygen species (ROS)[40]. Portal et al generated a suppression subtractive hybridization cDNA library to identify transcripts that are upregulated late in the infection process of the susceptible 'Grand Nain' banana cultivar by *M. fijiensis* [41]. Banana genes with antifungal action, genes for phenylpropanoid synthesis, detoxification, flavonoid and isoflavone-related genes, and genes for synthesis of ethylene and activation of jasmonic acid were expressed[41]. Unfortunately, because the RNA extraction protocol was not well suited to *M. fijiensis*, the only fungal transcript identified was encoding UDP-glucose pyrophosphorylase enzyme, which is involved in carbohydrate metabolism[41].

Transcriptome sequencing studies have also been used to characterize expression of banana genes after challenge by close relatives of *M. fijiensis*. Passos et al used 454 sequencing to compare transcripts between the resistant 'Calcutta 4' and susceptible 'Grand Nain' varieties of banana after being challenged with *M. musicola*[42]. Genes upregulated in 'Calcutta 4' compared to 'Grand Nain' included numerous transcription factors as well as genes potentially related to detoxification such as superoxide dismutases, glutathione-S-transferases and metallothionen-like proteins, which protect cells against damage from ROS[42]. Gene transcripts for phenylalanine ammonia-lyase (PAL) were abundant, and these

promote synthesis of phytoalexins and salicylic acid[42]. An enzyme involved in production of flavonoids, 4-coumarate-CoA ligase, was also upregulated in Calcutta 4[42]. Uma et al used Illumina transcriptome sequencing of the susceptible 'Grand Nain' versus resistant 'Manoranjitham' banana varieties after being challenged with another close relative of *M. fijiensis*, *Mycosphaerella eumusae*[43]. They observed very similar expression patterns as those found in other studies. For example, they observed higher expression of flavonoid-producing genes and genes involved in scavenging ROS (glutathione S-transferase and peroxidase) in the resistant variety[43]. However, they also observed expression of genes in the resistant variety that may be involved in production of ROS (polyphenol oxidase and hexokinase-3). Furthermore, they showed that genes in the phenylpropanoid and other secondary metabolite pathways had higher expression in the resistant banana variety[43].

Studies of *M. fijiensis* gene expression have been more limited. Cho et al identified expressed sequence tags in *M. fijiensis* grown in different types of culture medium: Fries liquid medium, Fries liquid medium with banana leaf extract, and Potato Dextrose Agar (PDA)[44]. In the Fries liquid medium and modified Fries medium, a homolog of the Ecp2 effector from *Cladosporium fulvum* was identified, and in all three libraries, a homolog of the Avr4 effector from *C. fulvum* was identified[44].

We used transcriptome sequencing to compare *M. fijiensis* gene expression in infected leaf tissue from the necrotrophic stage and Potato Dextrose Broth culture medium (Chapters 2-5)[45,46]. Several types of genes with pathogenicity-related homologs were identified as having higher expression in infected leaf tissue, including: genes encoding CFEM domain-containing proteins, hydrophobic surface binding proteins, salicylate

hydroxylase-like proteins, and proteins with characteristics common in effectors (Chapter 3)[46]. Eleven genes encoding Domain of Unknown Function (DUF) 3328 proteins had higher expression in infected leaf tissue, whereas none had higher expression in medium, suggesting that these may also be important for pathogenicity (Chapter 3)[46]. We showed that genes commonly associated with secondary metabolism[47], such as short-chain dehydrogenases, oxidoreductases in the 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily, and cytochrome P450s had higher expression in infected leaf tissue (Chapter 3)[46]. Secondary metabolism genes are typically clustered together in fungal genomes[47], so we identified clusters of genes with similar differential expression. Sixteen gene clusters were identified with higher expression in leaf tissue, including clusters for the synthesis of polyketide synthases, a non-ribosomal peptide synthase, and a novel fusicoccane (Chapters 2-3)[45,46].

In addition to our prediction of putative pathogenicity genes, our transcriptome analysis provided support for the hypothesis that *M. fijiensis* has dispensable chromosomes[46]. These have been reported in many other fungi[48]. In many cases, their function is cryptic, but in some fungi, genes encoded on the dispensable chromosomes play roles in pathogenicity. For example, genes for AM-toxin, HC-toxin production and phytoalexin detoxification are encoded on dispensable chromosomes in *Alternaria alternata*, *Cochliobolus carbonum* and *Nectria haematococca*, respectively[49–51].

Most dispensable chromosomes in filamentous fungi are small and rich in repetitive sequences[48]. These chromosomes are thought to have originated at least in part by horizontal gene transfer, because codon usage and G+C content often differ from the core chromosomes[52]. The related pathogen *Mycosphaerella graminicola* has eight chromosomes

that are dispensable, and are easily lost during sexual reproduction[53]. Fourteen scaffolds from the *M. fijiensis* genome share several characteristics with the dispensable chromosomes from *M. graminicola*: they are small, have low G+C content, different codon usage, highest proportion of repetitive DNA, lowest gene density, and lowest proportion of genes encoding proteins with PFAM domains[54,55]. Therefore, it has been hypothesized that the 14 scaffolds from *M. fijiensis* may also represent dispensable chromosomes[54,55]. If these *M. fijiensis* chromosomes are indeed dispensable, they have a different evolutionary origin from those in *M. graminicola*, since the genes on each species' dispensable chromosomes or putative dispensable chromosomes are absent in the genome of the other species[55].

We analyzed the transcriptome of leaf tissue infected with *M. fijiensis* isolate 14H1-11A compared to *M. fijiensis* grown in liquid medium, mapping the transcripts to both the banana genome and the *M. fijiensis* genome of isolate CIRAD86. This analysis revealed that the 14 putative dispensable scaffolds from *M. fijiensis* have a very different pattern of gene expression than the putative 'core' scaffolds (Chapter 3)[46]. Each of the putative 'core' scaffolds had a small percentage of genes with higher expression in infected leaf tissue and a small percentage of genes with higher expression in media[46]. On two of the putative dispensable scaffolds, a very high percentage (>30% and >50%) of the genes had higher expression in infected leaf tissue, whereas no genes on those scaffolds had higher expression in medium, which suggests that these two scaffolds may be important for pathogenicity[46]. For the remaining putative dispensable scaffolds, some had a small percentage of genes with higher expression in one condition but not the other, and some putative dispensable scaffolds had no differentially expressed genes[46]. Further analysis of scaffolds with no differentially

expressed genes revealed that for two of them, no transcripts were detected at all[46]. Using PCR assays, we were unable to amplify any of three genes tested from one of the scaffolds from which no transcripts were detected, and we were only able to amplify one of three genes from the other scaffold from which no transcripts were detected (Chapter 3)[46]. These results suggest that these scaffolds or large parts of them may be absent in isolate 14H1-11A, which is consistent with the hypothesis that these scaffolds represent dispensable chromosomes.

Secondary Metabolites Produced by *M. fijiensis*

Many pathogens produce secondary metabolites such as polyketides and non-ribosomal peptides as important pathogenicity factors. These can have a wide variety of modes of action, but often they cause toxicity to plant cells during a necrotrophic stage of infection by the pathogen. *M. fijiensis* is closely related to species of fungi that are known to produce polyketide light-activated toxins, such as *Cercospora* spp. which produce cercosporin, and *Ramularia* spp. which produce rubellin[56–59]. As mentioned previously, it has long been observed that Sigatoka disease symptoms are mitigated by growing plants in partial shade[5]. Based on these observations, it has been suspected that *M. fijiensis* may also produce light-activated toxins. Therefore, a number of experiments have been done to identify toxic secondary metabolites from *M. fijiensis*.

The first phytotoxic secondary metabolite from *M. fijiensis* to be reported, fijiensin (Figure 1), is produced in modified M-1-D liquid medium[60]. It was found to induce necrotic lesions on both resistant and susceptible banana cultivars but not non-host species[60,61].

Fijiensin, also called vermistatin, is produced by many other fungi, including *Talaromyces flavus*[62], *Talaromyces thailandiasis*[63], *Penicillium vermiculatum*[64,65], *Penicillium verruculosum*[66], *Penicillium simplicissimum*[67], *Eurotium rubrum*[68], and *Guignardia* sp[69]. No fungicidal activity of fijiensin could be observed toward *Saccharomyces cerevisiae* or *Aspergillus niger*[69]. However, it does potentiate the effects of the fungicide miconazole against *Candida albicans*[62]. Compounds related to fijiensin have been shown to slightly stimulate rooting of lettuce, and Chinese cabbage seedlings[70]. In murine leukemia cells and Ehrlich ascites, it acts as an RNA synthesis inhibitor, and it has shown some promise as an antitumor compound[71].

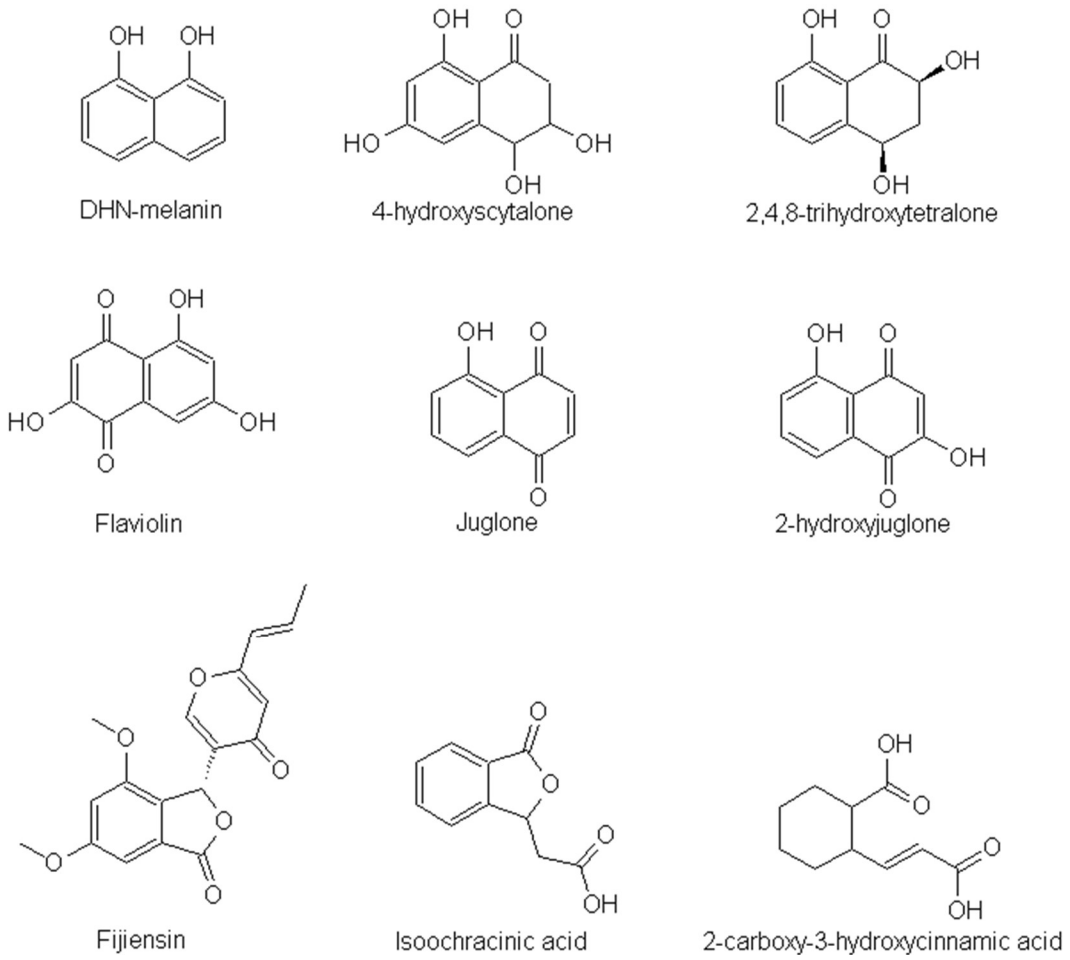


Figure 1. Phytotoxic metabolites identified from *M. fijiensis* cultures.
The figure shows the structures of secondary metabolites identified from *M. fijiensis* that are known to be toxic to banana tissue, as well as the structure of DHN-melanin, from which pathway some of the phytotoxic shunt metabolites are derived.

Additional phytotoxic metabolites of *M. fijiensis* were identified by Stierle and colleagues from cultures grown in M-1-D medium and extracted in ethyl acetate: 2,4,8-trihydroxytetralone (2,4,8-THT), juglone, isoochracinic acid, 2-carboxy-3-hydroxycinnamic acid, and 4-hydroxyscytalone (Figure 1)[61]. Of these, 2,4,8-trihydroxytetralone, juglone, and 4-hydroxyscytalone are melanin shunt pathway metabolites[72,73]. Juglone and 2,4,8-THT were

only found in *M. fijiensis* cultures older than 20 days, at which point 2,4,8-THT was produced at 30 times the rate of the other toxins. Toxicity of these metabolites was compared for resistant and susceptible banana cultivars, and only 2,4,8-THT exhibited host selectivity with this assay[61].

The fungicide tricyclazole blocks melanin biosynthesis and leads to accumulation of melanin shunt pathway metabolites. For *M. fijiensis*, addition of tricyclazole to the culture medium results in increased production of 2,4,8-THT and flaviolin, but not 2-hydroxyjuglone or juglone. Tricyclazole treatment of *M. fijiensis*-infected banana plants resulted in large necrotic leaf spots[74]. 2,4,8-THT production by *M. fijiensis* in culture was shown to be increased by banana leaf extracts in the medium, especially extracts from resistant compared to susceptible banana cultivars[74]. Based on these data, 2,4,8-THT was proposed as an important pathogenicity factor for black Sigatoka, and it was recommended as a screening tool to identify resistant plants in tissue culture systems.

Other studies have examined the potential role in the disease process of the melanin shunt metabolite juglone. In the study by Stierle and colleagues, juglone was isolated at lower concentrations but was more phytotoxic than 2,4,8-THT[61]. Juglone has previously been reported to be produced by plants of the family *Juglandaceae*[75]. Busogoro and colleagues extracted fungal metabolites from culture filtrates using ethyl acetate, and injected these metabolites into banana leaves. They showed toxicity that was light dependent and that resulted in swelling of the chloroplasts[76]. Juglone inhibits electron transport in banana chloroplasts, and this inhibition is greater for the black Sigatoka-susceptible cultivar Grand Nain compared to the partially resistant cultivar Fougamou[77]. This mode of action for juglone

toxicity provides one possible explanation for the light dependence for symptoms of black Sigatoka. However, symptoms of yellow Sigatoka are also alleviated by growing the banana plants under shade, and Stierle and colleagues were unable to show production of juglone or 2,4,8-THT by *M. musicola* in culture[61].

Melanin itself is an important pathogenicity factor for many fungi (Figure 1). It is crucial for penetration of cells by certain appressoria-producing fungi, including *Magnaporthe grisea*[78] and *Venturia inaequalis*[79]. For *M. grisea*, melanized appressorial cell walls allow accumulation of glycerol and high turgor pressure in the appressoria[78]. In the human pathogen *Wangiella dermatitidis*, melanin-deficient mutants are less pathogenic, and grow much more slowly through agar medium, an effect which is thought in part to be due to reduced ability for hyphal tip protrusion[80]. Melanin also reduces cell wall porosity, which can protect the fungus against fungicides and other toxic compounds, perhaps ones produced by the host plant[81].

Melanin can scavenge and protect against ROS. For example, melanized cells of *Cryptococcus neoformans* survive better than non-melanized cells in the presence of oxygen and nitrogen free radicals[82]. This is important for a pathogen, since oxidative burst is a common host defense in both plants and animals[83,84]. In other cases, such as under UV light, melanin can actually generate ROS, including singlet oxygen, hydroxyl radicals, and hydrogen peroxide[85,86]. Isolated melanin as well as whole mycelium from *M. fijiensis* has been shown to produce singlet oxygen when irradiated with 532 nm light[87]. Production of ROS could be used to kill host tissue to obtain nutrients, a strategy reported for many necrotrophic and hemibiotrophic pathogens.

More recently, the polyketide synthase gene for melanin production has been knocked out, and this knockout is reported to be as virulent as the wild-type strain[5]. Since no alternative pathways to melanin shunt metabolites have been reported, this knockout should be unable to produce melanin or melanin shunt-pathway metabolites. This study suggests that melanin and the shunt metabolites 2,4,8-THT, juglone, and 4-hydroxyscytalone may not play an important role in the disease process. However, chemical analyses of the mutant are required to confirm its inability to produce 2,4,8-THT and the other shunt metabolites.

While no evidence of phytotoxins was found in the aqueous phase of M-1-D culture filtrates[61], multiple studies have found evidence of hydrophilic phytotoxins when *M. fijiensis* is grown in liquid V-8 medium[88,89]. For these assays, a pigment-free fraction was obtained using activated charcoal, and this fraction was shown to be more phytotoxic than the original crude filtrate[88]. Although the identity of these phytotoxins has not yet been published, assays have been done for the biological activity of one of them. It had similar activity in both light and dark, and caused an increase in ROS and increased membrane permeability[89]. This phytotoxin was toxic to non-host species and to resistant and susceptible banana varieties, indicating that it is not a host-selective toxin[89].

Genome sequencing of *M. fijiensis* has also enabled bioinformatics analyses about the secondary metabolites that may be produced. We showed that one non-ribosomal peptide synthase gene cluster and one NRPS-like gene cluster had higher expression in infected leaf tissue compared to medium (Chapter 3)[46]. The closest characterized homolog to the NRPS from *M. fijiensis* is a destruxin synthetase; destruxins have insecticidal and phytotoxic properties and are produced by fungal pathogens of both insects and plants[90–96]. We also

identified a diterpenoid biosynthetic gene cluster from *M. fijiensis* that is very similar to a gene cluster producing a fusicoccane called brassicicene C from *Alternaria brassicicola* (Chapter 3)[46]. Fusicoccanes affect plant physiology in a variety of ways, including opening of stomata[97].

We did an extensive analysis of polyketide synthases from *M. fijiensis*. The Secondary Metabolites Unique Regions Finder (SMURF) program was used to predict seven polyketide synthases and one hybrid polyketide synthase/non-ribosomal peptide synthase from the *M. fijiensis* genome (Chapter 2)[45]. An analysis of conserved domains predicted that PKS7-1, PKS8-1, and PKS10-1 are non-reducing enzymes; PKS2-1 and PKS8-2 are highly reducing; and Hybrid8-3, PKS8-4 and PKS10-2 are partially reducing. Secondary metabolite biosynthetic genes are typically clustered together in fungal genomes, and the biosynthetic clusters typically encode transcription factors, transporters, cytochrome P450s, short-chain dehydrogenases, and others[47]. Using this information, we predicted which genes may be part of each biosynthetic cluster.

Through phylogenetics analysis and putative biosynthetic cluster comparisons, we predicted that *PKS10-1* encodes DHN melanin synthesis, and the *PKS2-1*, *PKS8-2*, and *PKS10-2* gene clusters are similar to alternapyrone, fumonisin, and solanapyrone-producing clusters, respectively. *PKS7-1*, *PKS8-2*, *Hybrid8-3*, *PKS8-4*, and *PKS10-2* have higher expression in infected leaf tissue than in medium (Chapters 2 and 4)[45]. We generated a disruption mutant of *pks8-4* and showed that this mutant is still pathogenic. PKS8-4 is expressed in developing pseudothecia, which form in infected leaf tissue (Chapter 4), and

PKS8-4 is related to polyketide synthases from *Neurospora crassa* and *Sordaria macrospora* that are important for female fertility and perithecia development (Chapter 4)[98,99].

Cercospora spp. and *Ramularia collo-cygni* are closely related to *M. fijiensis* and produce red-fluorescent polyketide toxins[59,100,101]. We showed that *M. fijiensis* hyphae have similar fluorescence emission spectra to cercosporin-producing hyphae of *Cercospora nicotianae* (Chapter 5). From extracts of red-fluorescing *M. fijiensis* hyphae, we identified fourteen compounds that have not previously been reported from *M. fijiensis*, including pulverochromenol, an idebenone-like metabolite and 1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene, which are all similar to molecules produced by polyketide biosynthetic pathways (Chapter 5). We showed that production of the red-fluorescent compound correlates to expression of the MFS transporter gene from the *PKS8-1* gene cluster, and that this cluster is likely to produce a polyketide similar to monodictyphenone (Chapter 5). Since pulverochromenol is similar in structure to monodictyphenone, we hypothesize that this compound may be produced by the PKS8-1 biosynthetic cluster (Chapter 5). Overexpression of an *aflR*-like transcription factor gene from the *PKS8-1* gene cluster resulted in increased expression of many, but not all, genes in the *PKS8-1* gene cluster, and further research is necessary to clarify whether disease symptoms may be more severe in plants inoculated with the transcription factor over-expressers compared to the wild-type (Chapter 5).

Effectors Produced by *M. fijiensis*

During the biotrophic phase of its life cycle, *M. fijiensis* would be expected to produce effectors to inhibit host defenses and prevent the death of host cells. To date, very few effectors have been characterized from *M. fijiensis*. However, it has been shown that *M. fijiensis* shares a few conserved effectors with *Cladosporium fulvum*, including Avr4, Ecp2, and Ecp6. In *C. fulvum*, Avr4 functions by binding to chitin and protecting against plant chitinases. *M. fijiensis* Avr4 is similar enough to *C. fulvum* Avr4 to trigger an HR response in tomato plants with the cognate R protein[102]. *C. fulvum* effector Ecp6 functions in a similar manner to Avr4: it binds to chitin, outcompeting the host receptor for chitin, and thereby prevents an immune response[103]. Ecp6 has homologs in *M. fijiensis* and many other fungi[104]. *M. fijiensis* has three homologs of *C. fulvum* effector Ecp2, and one of these triggers an HR response in tomato plants with the cognate R protein for *C. fulvum* Ecp2. Expression of *M. fijiensis* Ecp2 also triggers some necrosis in tomato that is independent of the presence of the R gene[102].

Because pathogenic fungi secrete effector proteins into the apoplast to modulate host physiology, effectors tend to be small secreted proteins that are cysteine-rich to provide greater stability in the apoplast through the disulfide bonds[105–107]. The host often evolves to recognize pathogen effectors and trigger stronger host defenses, so genes encoding effectors are typically rapidly evolving and highly variable between different species and even between strains of the same species[108]. We showed that *M. fijiensis* genes encoding 30 short, cysteine-rich, secreted proteins have higher expression in infected leaf tissue compared to culture medium (Chapter 3)[46]. For approximately half of these putative effector genes,

homologs were restricted to species within the *Mycosphaerellaceae*, which is consistent with the observation that effectors often have a restricted phylogenetic distribution (Chapter 3)[108].

Other Putative Pathogenicity Genes from *M. fijiensis*

The fungal cell wall is an attractive target site for developing fungicidal compounds. Whereas plant cell walls are rich in cellulose, fungal cell walls are unique because they are rich in chitin, mannans, glucans, and glycoproteins[109]. Kantun-Moreno et al have identified and analyzed the expression of a class of mannoproteins in *M. fijiensis*: the glucosyl phosphatidylinositol (GPI) proteins[110]. GPI proteins have been reported to function in cell wall remodeling, adhesion to host tissue, biofilm formation, and as virulence factors[111–113]. *MfGas1* and *MfGas2* are homologous to Gas proteins (glycophospholipid surface proteins) from pathogenic *Fusarium* species[110]. In *Fusarium* spp., these Gas proteins are involved in pathogenesis, cell wall formation, morphology, and conidiation[112,114]. Expression of *MfGas1* is nearly constant throughout different culture conditions and the infection process, whereas *MfGas2* is more highly expressed in conidia compared to mycelium, and during the infection process it is most highly expressed at the speck stage[110].

An ABC transporter, MgAtr4, was identified in the closely related fungus *M. graminicola*. This transporter was shown to contribute to virulence and to aid in colonization of the substomatal cavities[115]. More recently, a homolog of *MgAtr4* was identified in the *M. fijiensis* genome. This homolog, *MfAtr4*, was most highly expressed in early stages of infection, and was not expressed at the latest necrotrophic phase of infection[116]. The identity of the substance being transported has not yet been determined, but ABC transporters may

play several roles in virulence. They can protect the fungus against plant defense compounds, or they may transport effectors or secondary metabolites.

Potential for Transgenic Banana Plants Resistant to Black Sigatoka

Transformation protocols for banana have been developed[117], and a better understanding of *M. fijiensis* pathogenicity genes may help identify potential transgenes to engineer black Sigatoka resistance in banana. Genetic modification can be done in an agronomically desirable genetic background, which would minimize problems that could otherwise arise from conventional breeding with resistant but agronomically undesirable banana varieties. Unlike some other crops, concerns of traits spreading to other farmed bananas or their wild counterparts should be minimal due to the sterility of cultivated banana.

There are several promising options available for choice of transgene. One option is antimicrobial peptides (AMPs), which are small, cysteine-rich peptides with broad toxicity against fungi and bacteria. Transgenic banana plants producing an analogue of the AMP magainin had resistance to both *M. fijiensis* and *F. oxysporum* f.sp. *cubense*[118]. Hundreds of transgenic *Musa* sp. lines have been developed to express different AMPs, and some of these lines showed resistance under in vitro conditions to another banana pathogen, *Colletotrichum musae*, but further research is needed to show whether these transgenic lines are resistant to black Sigatoka and other diseases[119–121]. Genes encoding hydrolytic enzymes such as chitinases have also been used. Transgenic 'Gros Michel' banana plants expressing a rice chitinase gene developed lesions from black Sigatoka more slowly than the non-transformed control plants[122]. One interesting strategy that has been tested in tobacco is to use a pathogen-

inducible promoter to drive expression of an elicitor. The elicitor is then recognized by the plant and triggers HR, leading to resistance[123].

Host-induced gene silencing (HIGS) is another potential strategy for engineering disease resistance. For this strategy, an RNAi silencing construct is designed against an important fungal gene, and this construct is transformed into the plant host. When the plant expresses this construct, dsRNAs are transported from the plant to the pathogen, where they silence the pathogen genes. HIGS is reported to be effective not only for biotrophic pathogens such as powdery mildew[124] and rust fungi[125], but also hemibiotrophic pathogens such as *Fusarium verticillioides*[126], and necrotrophic pathogens such as *Sclerotinia sclerotiorum*[127].

Use of HIGS does require an intact silencing pathway in the fungus being targeted, and this pathway is not ubiquitous among fungi. For example, *Saccharomyces cerevisiae*, *Ustilago maydis*, and *Cryptococcus gattii* lack intact silencing pathways[128–130], indicating that this pathway has been lost multiple times during fungal evolution[128]. Fortunately, the silencing pathway appears to be intact in *M. fijiensis*, because dsRNAs for adenylate cyclase and DNA polymerase were able to inhibit *M. fijiensis* spore germination in vitro[131]. Therefore, HIGS is a plausible control strategy for black Sigatoka.

Choice of transgene will depend on several factors in addition to the gene's ability to confer resistance. Any transgenes resulting in agronomic off-types must be discarded. For example, the AMP MsDEF1 was transformed into potato plants and conferred strong field resistance to *Verticillium dahliae*[132]. However, the transgenic plants produced smaller tubers

and had reduced yields compared to the non-transformed control[133]. This problem may be mitigated by using tissue-specific promoters.

There are several major obstacles to development and adoption of transgenic resistant banana plants. Development and regulatory testing of a new genetically modified (GM) crop is very expensive. From 2008 to 2012, the mean cost of bringing a new GM trait to market was $136 million[134]. Furthermore, there are significant hurdles in public opinion where GM crops are concerned. Much of the export banana crop is exported to Europe, and consumer acceptance of GM crops in Europe is particularly low[2,135].

One non-transgenic approach with much of the functionality of HIGS is to use dsRNA sprays as a pesticide to target and knock down expression of important pathogen genes. It has been shown that dsRNAs sprayed on leaf surfaces can retain biological activity for over 4 weeks[136]. The technology has shown some promise for insect pests as well as viral and fungal pathogens, though research with fungal pathogens is still in preliminary stages[136–138].

Regardless of the type of control methods chosen, some thought should be given to likelihood of the fungus evolving to overcome the methods used. Achieving durable control depends on understanding the biology of the pathogen and its potential to evolve. A pathogen has higher evolutionary potential if it has a high mutation rate, a large population size, both asexual and sexual reproduction, and propagules that can spread long distances[139]. Unfortunately, *M. fijiensis* has many of these risk factors for high evolutionary potential. It has a greatly expanded genome even compared to other *Mycosphaerella* species, due to transposable elements[140]. The abundance of transposable elements would contribute to a

higher mutation rate[141]. It produces abundant spores through both sexual and asexual reproduction, and the sexual spores can be dispersed by wind for long distances[5,10]. There are a few strategies available to mitigate these risk factors. If overcoming host resistance requires a large fitness penalty onto the fungus, then evolution of new virulent races is less likely[142]. Furthermore, multiple banana varieties with different resistance genes could be mixed in the same field, or different resistance genes could be stacked in the same variety, so that there isn't one large, single selection pressure on the fungus to evolve. Since *M. fijiensis* has such a high evolutionary potential, many control measures should be used in parallel, ideally combining resistant banana varieties, chemical control methods, and cultural control methods in a way that is as economical and environmentally friendly as possible.

Other Prospects for Disease Control

Disease forecasting is one strategy for reducing fungicide applications on susceptible hosts. Several studies have investigated the relationship between rainfall and humidity and disease severity, and in many cases fungicide applications could be reduced during the dry season[143–145]. Delays in needed fungicide applications can result in unsatisfactory control[145], so effective use of forecasting would require an accurate model that is accessible and easy to use so that farmers can quickly make fungicide application decisions.

Aside from weather conditions, another factor to consider is the severity of existing infections in the field. Easily scored indicators such as number of disease-free leaves per plant or disease score on particular leaves could be incorporated into a forecasting model[143]. However, these symptoms occur late during the infection process, so it could be useful to

have an earlier indicator of infection. PCR-based assays are available and are the most sensitive means of detecting infection at an early stage[146]. However, these assays are expensive and not practical for routine monitoring of infection for a forecasting system.

ELISA tests are cheaper and more practical in a field situation, and a few have been developed to detect *Mycosphaerella* pathogens in infected leaves. One such test is able to identify infected leaves two weeks earlier than symptom-based assays, but it does not distinguish between the different *Mycosphaerella* species[147,148]. A polyclonal antibody-based ELISA test 'Insight' from Novartis Crop Protection has also been developed which quantifies *Mycosphaerella* spp. proteins during early infection, but it has limited independent use[149]. A triple antibody system (TAS-ELISA) test has also been developed which is quantitative and can distinguish between *M. fijiensis* and *M. musicola*[150].

Resistance of Sigatoka pathogens to available fungicides must be well-monitored so that farmers can choose fungicides that are more likely to be effective. Fungicide sensitivity testing is typically done using a microtiter plate assay recommended by the Fungicide Resistance Action Committee (FRAC), but in some cases there are PCR-based assays to detect particular mutations that are associated with fungicide resistance[10,151].

Although models exist for helping to predict disease severity, all of the relevant data needs to be integrated into one system that is easy for farmers to access and understand. Ideally, such a system would be accessible not only for large export companies, but also for smaller-scale farmers. Partnerships need to be developed between researchers, industry, and smaller-scale growers. Affordability of fungicides is an issue for small-scale growers, so a forecasting system would not help poor farmers with fungicide applications. However, there

are a number of cultural control methods that help with disease control, and educational programs are needed to teach them to farmers.

Conclusion

Mycosphaerella fijiensis is a major threat to banana production, but little is known about the molecular basis by which this fungus causes disease. Close relatives of *M. fijiensis* produce polyketides that generate reactive oxygen species in the light and kill the host plant tissue. Because symptoms of black Sigatoka are mitigated by partial shade, it has long been suspected that *M. fijiensis* may produce similar light-activated toxins as its close relatives. The following chapters provide details of experiments used to identify putative pathogenicity genes in this fungus, with special focus on the role of secondary metabolism and polyketide production in pathogenesis. This research generates many hypotheses about pathways that could be targeted for new black Sigatoka control methods.

CHAPTER 2

Bioinformatics Prediction of Polyketide Synthase Gene Clusters from *Mycosphaerella fijiensis*

Roslyn D. Noar and Margaret E. Daub

Noar, R. D., and Daub, M. E. (2016). Bioinformatics Prediction of Polyketide Synthase Gene Clusters from *Mycosphaerella fijiensis*. *PLoS ONE*, *11*(7), e0158471. http://doi.org/10.1371/journal.pone.0158471

Abstract

Mycosphaerella fijiensis, causal agent of black Sigatoka disease of banana, is a Dothideomycete fungus closely related to fungi that produce polyketides important for plant pathogenicity. We utilized the *M. fijiensis* genome sequence to predict PKS genes and their gene clusters and make bioinformatics predictions about the types of compounds produced by these clusters. Eight PKS gene clusters were identified in the *M. fijiensis* genome, placing *M. fijiensis* into the 23rd percentile for the number of PKS genes compared to other Dothideomycetes. Analysis of the PKS domains identified three of the PKS enzymes as non-reducing and two as highly reducing. Gene clusters contained types of genes frequently found in PKS clusters including genes encoding transporters, oxidoreductases, methyltransferases, and non-ribosomal peptide synthases. Phylogenetic analysis identified a putative PKS cluster encoding melanin biosynthesis. None of the other clusters were closely

aligned with genes encoding known polyketides, however three of the PKS genes fell into clades with clusters encoding alternapyrone, fumonisin, and solanapyrone produced by *Alternaria* and *Fusarium* species. A search for homologs among available genomic sequences from 103 Dothideomycetes identified close homologs (>80% similarity) for six of the PKS sequences. One of the PKS sequences was not similar (< 60% similarity) to sequences in any of the 103 genomes, suggesting that it encodes a unique compound. Comparison of the *M. fijiensis* PKS sequences with those of two other banana pathogens, *M. musicola* and *M. eumusae*, showed that these two species have close homologs to five of the *M. fijiensis* PKS sequences, but three others were not found in either species. RT-PCR and RNA-Seq analysis showed that the melanin PKS cluster was down-regulated in infected banana as compared to growth in culture. Three other clusters, however were strongly upregulated during disease development in banana, suggesting that they may encode polyketides important in pathogenicity.

Introduction

Mycosphaerella fijiensis is the causal agent of black Sigatoka, also known as black leaf streak disease of banana and plantain. Black Sigatoka was first described in Fiji in 1963, and since then it has spread to most banana-growing regions around the world, including Latin America, Asia, Africa, and throughout the Pacific[5]. It has become one of the most important diseases of banana world-wide, causing up to 50% yield loss[6]. Infection of banana plants by *M. fijiensis* leads to necrotic streaking on leaves and loss of photosynthetic capacity[7]. The disease also causes premature fruit ripening, which is problematic for banana

export companies, since bananas may over-ripen in transit[6]. Control of black Sigatoka is primarily through the extensive application of fungicides, which are estimated to account for 25-30 % of the cost of production[5–7,18]. Development of fungicide-resistant strains is an on-going problem[7], and in developing countries, fungicide costs are prohibitive, and the disease results in severe yield losses[6].

Initial symptoms of the disease are chlorotic specks, followed by red-brown streaking, followed by the development of necrotic lesions surrounded by chlorotic halos and extensive blighting of the leaf tissue (Figure S1)[5–7,12]. These symptoms (chlorosis, necrosis, streaking) as well as histological observations of symptoms ahead of hyphal colonization argue for the importance of toxins in disease development[12]. *M. fijiensis* is closely related to several fungi in the family Mycosphaerellaceae that are known to produce polyketide toxins important in pathogenicity. For example, *Cercospora* spp. produce the light-activated perylenequinone polyketide toxin, cercosporin, which kills host plant tissue through the production of reactive oxygen species; mutants deficient for the CTB1 polyketide synthase in the cercosporin biosynthetic pathway are reduced in both number and size of lesions on host plants[57]. *Ramularia collo-cygni* produces the anthraquinone toxin rubellin D. Although the role rubellin plays in the virulence of *R. collo-cygni* is not fully understood, it has been shown to generate reactive oxygen species in the light, resulting in fatty acid peroxidation and toxicity[56,152]. Also, *Dothistroma septosporum* has been shown to produce the polyketide dothistromin, which is structurally similar to a precursor of aflatoxin[153]. Dothistromin is important both for pathogenicity and for sporulation of the fungus[154].

The important role of polyketide toxins in related fungal pathogens has led to the investigation of possible production of polyketide toxins by *M. fijiensis.* A number of phytotoxic compounds have been identified from *M. fijiensis* culture filtrates, including several polyketides. These include fijiensin, 2,4,8-trihydroxytetralone, juglone, 4-hydroxyscytalone, and isoochracinic acid[60,61,71,155]. The compounds 2,4,8-trihydroxytetralone, 4-hydroxyscytalone and juglone are melanin shunt pathway metabolites[72,73,156,157]. These metabolites are toxic to banana tissue, and 2,4,8-trihydroxytetralone was shown to have host selectivity when comparing resistant versus susceptible banana varieties[61]. These results have implicated the melanin polyketide pathway as an important pathway for production of pathogenicity factors involved in disease development. In addition to toxic shunt metabolites, the melanin polyketide itself is known in many fungi to play important roles in pathogenicity, including appressorial function and entry into host tissue[78,79], hyphal tip protrusion[80,158], protection against plant defense compounds[82,159], and in some cases, production of reactive oxygen species[87] which may contribute to killing of host tissue later in the infection.

Recent sequencing of the *M. fijiensis* genome[55,160] allows for a bioinformatics analysis of PKS genes and clusters in the genome. Polyketide synthases in fungi are large, multi-domain enzymes that act iteratively to produce the polyketide chain. An iterative PKS enzyme must contain a ketosynthase domain (KS), an acyltransferase (AT) domain, and an acyl carrier protein domain (ACP)[161]. The reduction of the resulting polyketide is determined by three optional domains: ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER)[161]. These three domains act sequentially to reduce a keto to a hydroxyl group, dehydrate

a hydroxyl to an enoyl group, and reduce an enoyl to an alkyl group, respectively[162–164]. PKS enzymes lacking some or all of these domains will produce partially reduced or non-reduced polyketides.

The initial polyketide product may then be modified by other enzymes, typically encoded by genes that cluster together in the genome with the PKS gene[47,165]. These genes include methyltransferases, oxidoreductases, cytochrome P450s, transporters, and transcription factors[47]. While experimental data is necessary to confirm the involvement of a gene in the biosynthesis of a polyketide, predictions of genes in the cluster can be made based on proximity to the adjacent PKS and whether a gene is homologous to common types of secondary metabolite biosynthetic genes. Polyketide biosynthetic clusters can be small, as in the case of bikaverin from *Fusarium fujikuroi*, with an 18 kb cluster[166], or much larger, as for the aflatoxin biosynthetic cluster from *Aspergillus parasiticus* which spans 82 kb[167].

The main objectives of the work reported here were to predict PKS genes and their gene clusters in the published *M. fijiensis* genome, make bioinformatics predictions about the types of compounds that may be produced by these biosynthetic clusters, and identify PKS clusters with a possible role in pathogenicity by assaying expression in infected leaf tissue compared to expression during saprophytic growth in medium.

Results

Prediction of Polyketide Synthase Gene Clusters

Polyketide synthase genes from the *M. fijiensis* genome were identified using the program SMURF (Secondary Metabolite Unique Regions Finder), which is designed to

identify secondary metabolite gene clusters from fungi[47]. SMURF identified one PKS gene on *M. fijiensis* scaffold 2 (*PKS2-1*), one on scaffold 7 (*PKS7-1*), a hybrid PKS-NRPS gene and three PKS genes on scaffold 8 (*PKS8-1*, *PKS8-2*, *Hybrid8-3*, and *PKS8-4*), and two PKS genes on scaffold 10 (*PKS10-1*, *PKS10-2*). PKS protein ID numbers are indicated in Table S1. In addition to these PKS and hybrid PKS-NRPS genes, SMURF identified one PKS-like gene on scaffold 3.

Comparison of Number of PKS Genes for Dothideomycete Fungi

Since production of polyketides is a common strategy for pathogenesis among necrotrophic and hemibiotrophic fungi, the repertoire of PKS genes in *M. fijiensis* was compared to that of other Dothideomycete fungi. SMURF was used to predict PKS genes from 74 additional Dothideomycete genomes available from Joint Genome Institute (JGI) and NCBI. This analysis revealed that the *M. fijiensis* genome falls into the 23rd percentile for number of PKS genes compared to the other Dothideomycete genomes and into the 34th percentile for the number of PKS genes in the order Capnodiales, of which *M. fijiensis* is a member (Figure 1, Table S2).

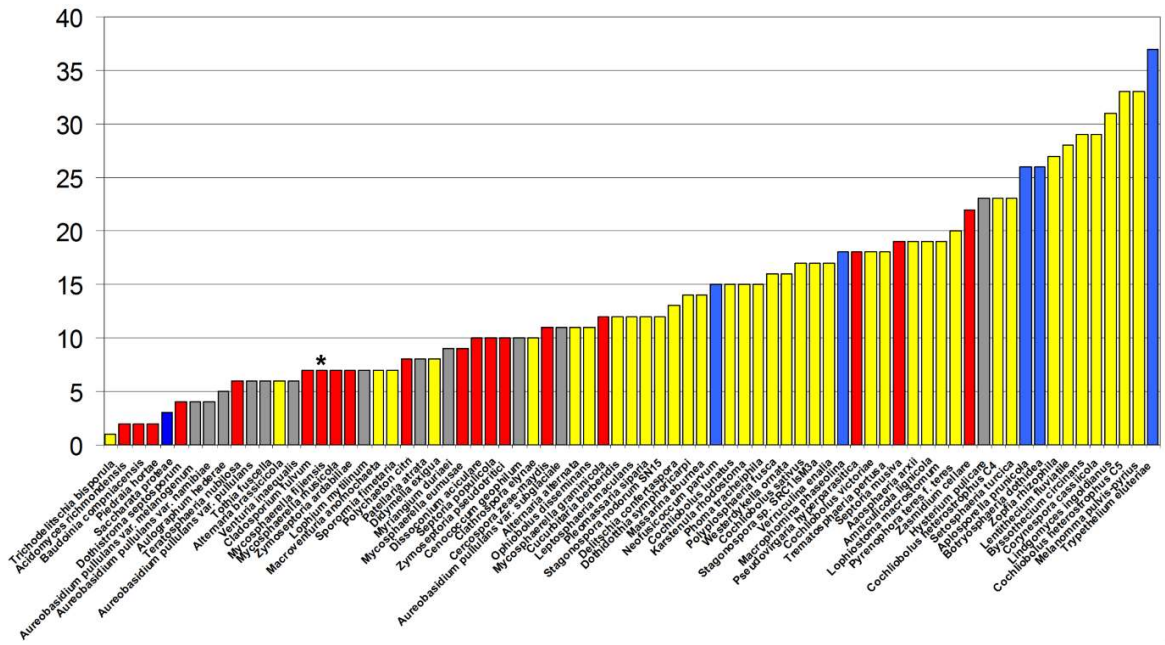


Figure 1. **Number of PKS genes predicted from 75 Dothideomycete fungal genomes.** The number of PKS genes predicted by the web tool SMURF (Secondary Metabolite Unique Regions Finder) from 75 Dothideomycete fungal genomes from JGI are shown. Red bars = Capnodiales; Yellow bars = Pleosporales; Blue bars = Botryosphaeriales; Gray bars = other orders within class Dothideomycetes. *M. fijiensis* is indicated with an asterisk above its bar.

Prediction of *M. fijiensis* PKS Domains

Prediction of which domains are encoded in each PKS gene can provide clues as to whether the PKS enzyme has the necessary domains to be functional, and whether the PKS enzyme produces a reduced or a non-reduced product. Blastp analysis was done for all of the predicted PKS and PKS-like protein sequences, using the NCBI database. NCBI's Conserved Domain Database[168] predicted that the PKS-like protein encoded on scaffold 3 only contains a 3-oxoacyl-(acyl-carrier-protein) synthase domain, which confirms that it is unlikely to be a true PKS. Ketosynthase, acyltransferase, and phosphopantetheine attachment site (characteristic of acyl carrier protein and peptidyl carrier protein[169]) domains were predicted

by the Conserved Domain Database to be included in each *M. fijiensis* PKS enzyme (Figure 2). Based on the optional domains predicted to be present in each *M. fijiensis* PKS enzyme, PKS7-1, PKS8-1, and PKS10-1 are likely to be non-reducing, since they do not contain KR, DH, or ER domains. PKS2-1 and PKS8-2 are likely to be highly reducing, since they contain all three reducing domains. Hybrid8-3, PKS8-4, and PKS10-2 are likely to be partially reducing, since they each contain KR and DH domains, but not complete ER domains. PKS10-2 does have a region with homology to an ER domain, but with a poor E-value when compared to E-values for ER domains in the other PKS enzymes (Table S3).

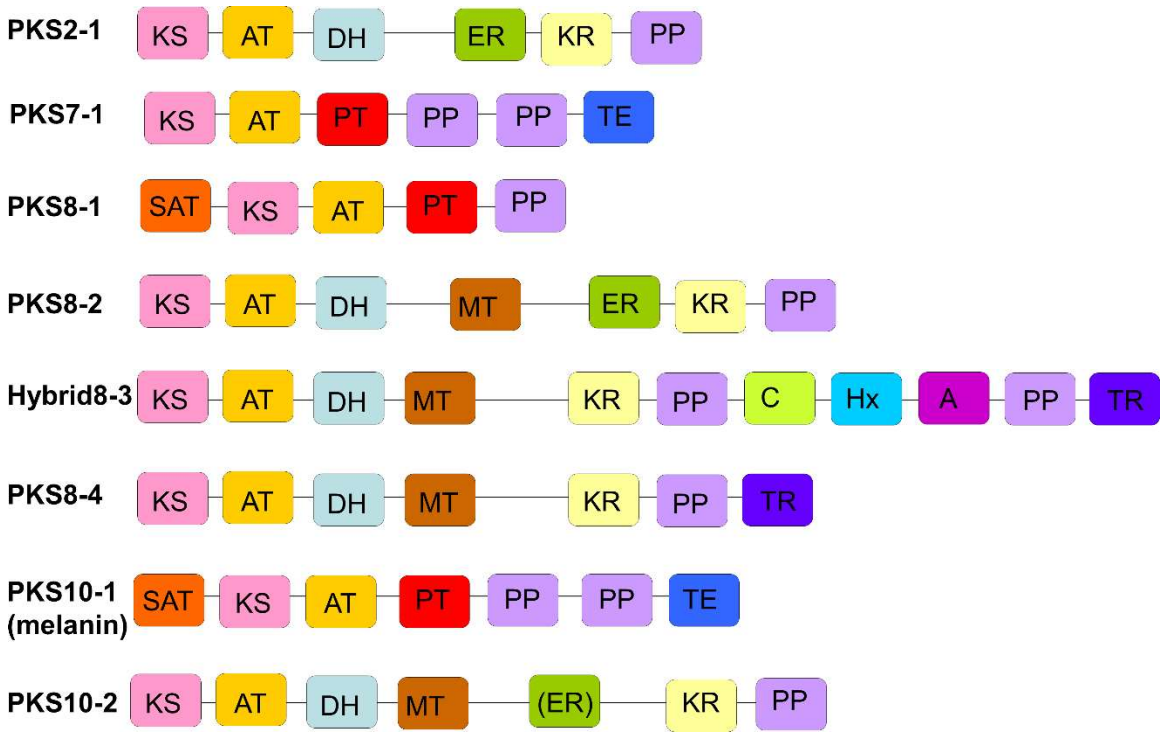


Figure 2. Domains predicted to be present in each *M. fijiensis* PKS or PKS-NRPS hybrid enzyme.
PKS domains are: SAT = starter unit acyltransferase; KS = ketosynthase; AT = acyltransferase; DH = dehydratase; MT = methyltransferase; ER = enoyl reductase; KR = ketoreductase; PT = product template; PP = phosphopantetheine attachment site; TE = thioesterase; TR = thioester reductase. NRPS domains are: C = condensation; Hx = HxxPF repeat domain; A = adenylation. Each type of domain is shown in a different color, to better display the similarities and differences between the domain organizations of different PKS enzymes. The ER domain of PKS10-2 is shown in parenthesis to indicate that it may or may not be functional, since the E-value for this domain is much lower than for ER domains of other PKS protein sequences (Table S3).

Prediction of Polyketide Biosynthetic Cluster Genes

For each locus with a PKS gene in the *M. fijiensis* genome, BLAST searches were done for genes adjacent to the PKS to determine if they are homologous to genes common in secondary metabolite clusters. Starting at the PKS and working outward, genes were proposed to be part of the cluster if BLAST results indicated that the type of gene is

commonly found in secondary metabolite clusters. Descriptions of the genes proposed for each of the eight PKS clusters are provided by Noar and Daub[45], including the GO, InterPro, and KOG descriptions of each gene, functional domains, homologs in other species, and links to NCBI and JGI descriptions. Once a region was found containing gene types not commonly found in clusters, these regions were proposed to flank the biosynthetic cluster. Genes and sizes of putative biosynthetic clusters and their flanking regions are shown graphically in Figure 3. A summary of the genes found in each proposed cluster is shown in Table 1.

Figure 3. Proposed *M. fijiensis* PKS and PKS-NRPS gene clusters.
The PKS or PKS-NRPS hybrid are shown, surrounded by the neighboring genes in the genome. Genes are labeled with putative functions of the corresponding proteins as determined by BLAST[45]. PKS or PKS-NRPS genes are shown in dark blue, and NRPS genes are shown in purple, with other proposed biosynthetic cluster genes shown in light blue. Genes proposed to not be part of the biosynthetic cluster are shown in gray. For the melanin biosynthetic cluster, genes known to be part of melanin biosynthesis are shown in red. A) *PKS2-1* cluster; B) *PKS7-1* cluster; C) *PKS8-1* cluster; D) *PKS8-2* cluster; E) *Hybrid8-3* cluster; F) *PKS8-4* cluster; G) *PKS10-1* (melanin) cluster; H) *PKS10-2* cluster. No genes are shown to one side of the putative biosynthetic clusters for *PKS2-1* and *PKS7-1* because no gene models are found within 95 kb of the ones shown.

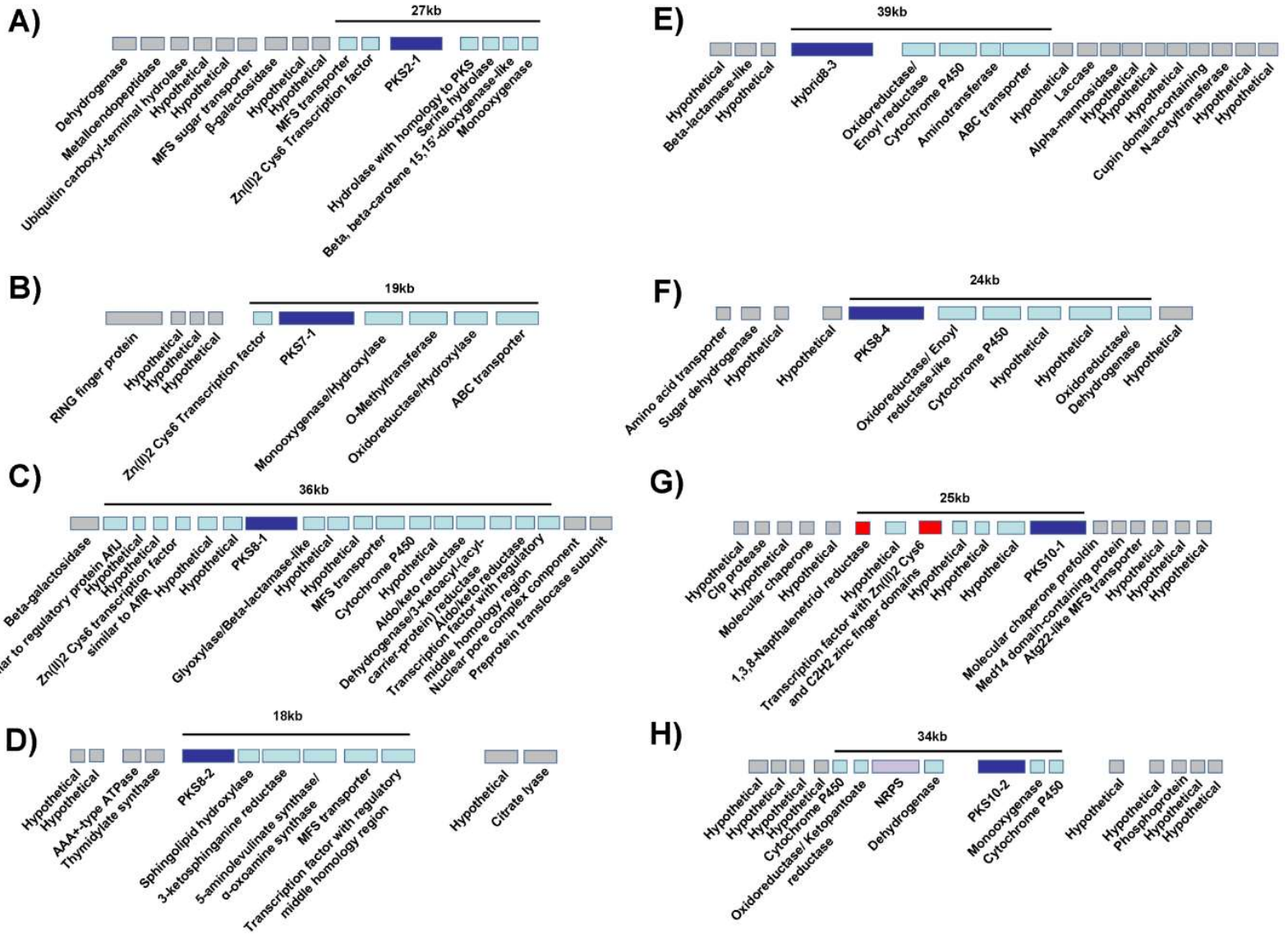

A)
27kb
Dehydrogenase
Metalloendopeptidase
Ubiquitin carboxyl-terminal hydrolase
Hypothetical
Hypothetical
MFS sugar transporter
β-galactosidase
Hypothetical
Hypothetical
MFS transporter
Zn(II)2 Cys6 Transcription factor
PKS2-1
Hydrolase with homology to PKS
Serine hydrolase
Beta, beta-carotene 15,15'-dioxygenase-like
Monooxygenase
B)
19kb
RING finger protein
Hypothetical
Hypothetical
Hypothetical
Zn(II)2 Cys6 Transcription factor
PKS7-1
Monooxygenase/Hydroxylase
O-Methyltransferase
Oxidoreductase/Hydroxylase
ABC transporter
C)
36kb
Beta-galactosidase
Similar to regulatory protein AflJ
Hypothetical
Hypothetical
Zn(II)2 Cys6 transcription factor
similar to AflR
Hypothetical
Hypothetical
PKS8-1
Glyoxylase/Beta-lactamase-like
Hypothetical
Hypothetical
MFS transporter
Cytochrome P450
Hypothetical
Aldo/keto reductase
Dehydrogenase/3-ketoacyl-(acyl-carrier-protein) reductase
Transcription factor with regulatory
Aldo/keto reductase
middle homology region
Nuclear pore complex component
Preprotein translocase subunit
D)
18kb
Hypothetical
Hypothetical
AAA+-type ATPase
Thymidylate synthase
PKS8-2
Sphingolipid hydroxylase
3-ketosphinganine reductase
5-aminolevulinate synthase/
α-oxoamine synthase
MFS transporter
Transcription factor with regulatory
middle homology region
Hypothetical
Citrate lyase
E)
39kb
Hypothetical
Beta-lactamase-like
Hypothetical
Hybrid8-3
Oxidoreductase/
Enoyl reductase
Cytochrome P450
Aminotransferase
ABC transporter
Hypothetical
Laccase
Alpha-mannosidase
Hypothetical
Hypothetical
Cupin domain-containing
N-acetyltransferase
Hypothetical
Hypothetical
F)
24kb
Amino acid transporter
Sugar dehydrogenase
Hypothetical
Hypothetical
PKS8-4
Oxidoreductase/ Enoyl
reductase-like
Cytochrome P450
Hypothetical
Hypothetical
Oxidoreductase/
Dehydrogenase
Hypothetical
G)
25kb
Hypothetical
Clp protease
Hypothetical
Molecular chaperone
Hypothetical
1,3,8-Naphthalenetriol reductase
Hypothetical
Transcription factor with Zn(II)2 Cys6
and C2H2 zinc finger domains
Hypothetical
Hypothetical
PKS10-1
Molecular chaperone prefoldin
Med14 domain-containing protein
Atg22-like MFS transporter
Hypothetical
Hypothetical
Hypothetical
H)
34kb
Hypothetical
Hypothetical
Hypothetical
Hypothetical
Cytochrome P450
Oxidoreductase/ Ketopantoate
reductase
NRPS
Dehydrogenase
PKS10-2
Monooxygenase
Cytochrome P450
Hypothetical
Hypothetical
Phosphoprotein
Hypothetical
Hypothetical

Table 1. Predicted PKS and hybrid PKS-NRPS biosynthetic clusters.
For each PKS and hybrid PKS-NRPS gene predicted from *M. fijiensis* genome, presence or absence of nearby commonly found polyketide biosynthetic genes is indicated.

	PKS 2-1 cluster	**PKS 7-1 cluster**	**PKS 8-1 cluster**	**PKS 8-2 cluster**	**Hybrid 8-3 cluster**	**PKS 8-4 cluster**	**PKS 10-1 cluster**	**PKS 10-2 cluster**
PKS	**X**	**X**	**X**	**X**		**X**	**X**	**X**
NRPS								**X**
Hybrid PKS-NRPS					**X**			
Methyltransferase		**X**						
Transporter	**X**	**X**	**X**	**X**	**X**			
Oxidoreductase	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
Cytochrome P450			**X**		**X**	**X**		**X**
Transcription Factor	**X**	**X**	**X**	**X**			**X**	

Identifying Homologs of *M. fijiensis* PKS Genes

To determine if any of the PKS genes in *M. fijiensis* have close, well-characterized homologs, predicted protein sequences were aligned from *M. fijiensis* PKS genes and PKS genes from other fungi with well-characterized products. RAxML was used to generate a phylogenetic tree of the predicted PKS protein sequences (Figure 4). This analysis revealed that *M. fijiensis* PKS10-1 is in the same clade as PKS enzymes that produce DHN melanin from other fungi. None of the other PKS proteins from *M. fijiensis* were found to have close homology to characterized PKS proteins, but three fall into clades with excellent bootstrap values. PKS2-1 and the alternapyrone and T-toxin-producing PKS sequences form a clade with a bootstrap value of 98%. The alternapyrone PKS from *Alternaria solani* is the most similar to PKS2-1, with a sequence similarity of 58% when aligned using BLAST. PKS8-2 forms a clade with the fumonisin PKS from *Fusarium verticillioides*, with a bootstrap value

of 100%, and has 58% similarity when aligned using BLAST. PKS10-2 forms a clade with the solanapyrone PKS from *Alternaria solani* with a bootstrap value of 100%, and has 52% similarity when aligned using BLAST. Percent similarity values between these three *M. fijiensis* PKS sequences and the characterized PKS sequences are lower than those between known orthologs (for example the cercosporin biosynthetic PKS is 91% similar between *C. nicotianae* and *C. zeae-maydis*, and the aflatoxin PKS from *Aspergillus parasiticus* and sterigmatocystin PKS from *A. nidulans* have 77% similarity), suggesting that they encode different products.

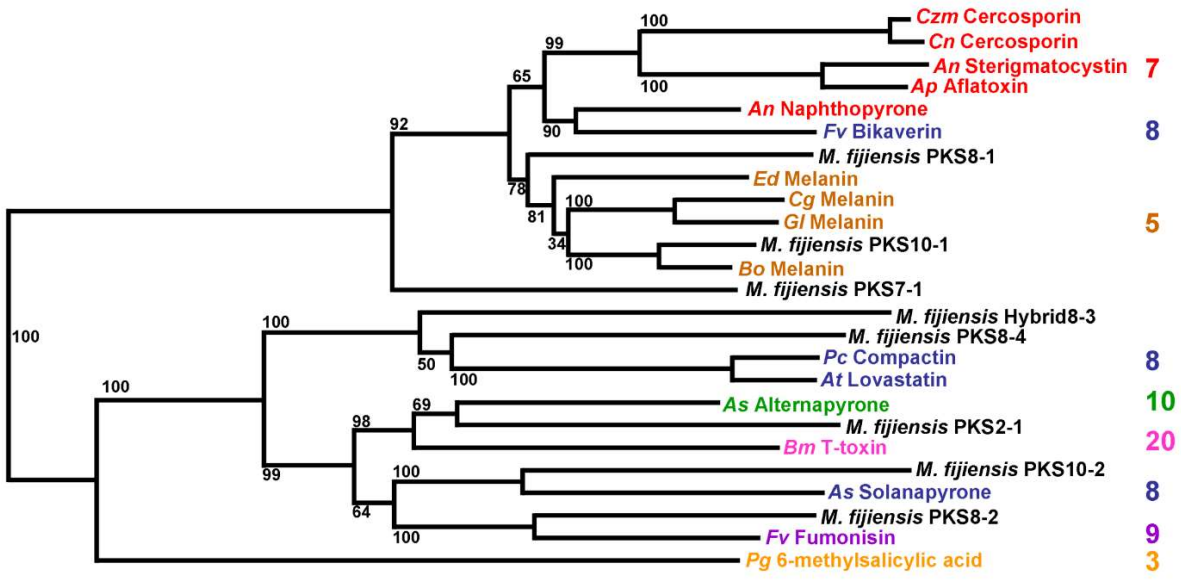


Figure 4. Maximum likelihood phylogenetic tree of PKS protein sequences. RAxML was used to generate a maximum likelihood phylogenetic tree of *M. fijiensis* PKS protein sequences as well as protein sequences from well-characterized PKS genes from other species. Products of each well-characterized PKS enzyme are indicated on the tree, along with abbreviations for species names. Numbers on branches indicate bootstrap support for the clade. Scale bar of branch length indicates substitutions per site. Number of iterations catalyzed by each PKS enzyme is indicated next to the name of the polyketide product of the PKS, in a matching color font. Orange = 3 ketide subunits; Brown = 5 ketide subunits; Red = 7 ketide subunits; Blue = 8 ketide subunits; Purple = 9 ketide subunits; Green = 10 ketide subunits; Pink = 20 ketide subunits. *An = Aspergillus nidulans*; *Ap = Aspergillus parasiticus*; *As = Alternaria solani*; *At = Aspergillus terreus*; *Bm = Bipolaris maydis*; *Bo = Bipolaris oryzae*; *Cg = Colletotrichum graminicola*; *Cn = Cercospora nicotianae*; *Czm = Cercospora zeae-maydis*; *Ed = Exophiala dermatitidis*; *Fv = Fusarium verticillioides; Gl = Glarea lozoyensis*; *Pc = Penicillium citrinum; Pg = Penicillium griseofulvum*. Accession numbers for each PKS sequence are available in Table S1.

In order to further characterize the similarities between the *M. fijiensis* PKS sequences and those of the characterized polyketides, we looked at domain structure as well as organization of the cluster. PKS2-1 formed a clade with a good bootstrap value to the alternapyrone and T-toxin-producing PKS sequences, with greatest similarity to the

alternapyrone-producing PKSN protein from *Alternaria solani* (Figure 4). The *A. solani* PKSN has the following domain organization: KS-AT-DH-MT-ER-KR-ACP[170]. The *M. fijiensis* PKS2-1 has the same domain organization except that it lacks a MT domain (Figure 5A). Therefore, while alternapyrone is a highly methylated decaketide[170], the product of PKS2-1 may lack these methyl groups. Alternapyrone was identified as the product of PKSN by cloning the PKSN-encoding gene *alt5* from *A. solani* and expressing it in *Aspergillus oryzae*[170]. Since only the PKS was expressed in a heterologous system, the final product of the biosynthetic cluster in *A. solani* is currently unknown. There are genes for three cytochrome P450s and an oxidase adjacent to the *alt5* gene in the *A. solani* genome[170], which may modify alternapyrone to create a different metabolite. In the *M. fijiensis PKS2-1* biosynthetic cluster, *PKS2-1* is adjacent to several putative cluster genes such as a transporter and a transcription factor with a fungal Zn(2)-Cys(6) binuclear cluster domain, but not the cytochrome P450s or the oxidase that may be part of the *alt5* biosynthetic cluster (Figure 5B).

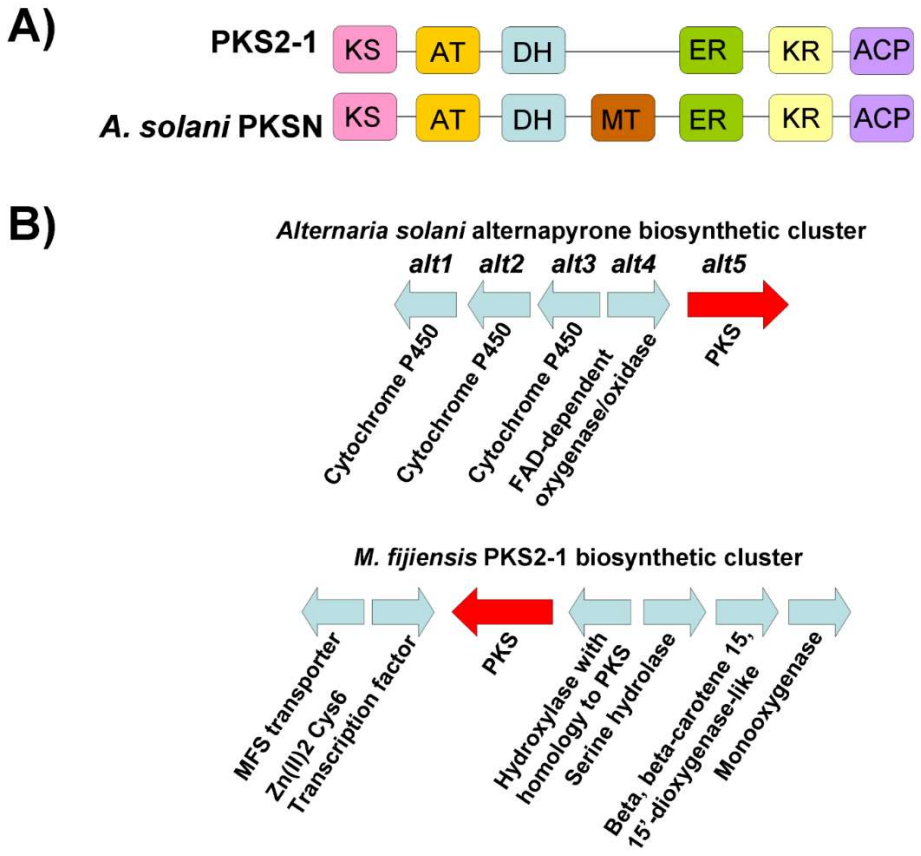


Figure 5. Comparison of domains and putative cluster genes for *M. fijiensis PKS2-1* and the alternapyrone-producing PKS.
A) Domain organization of *M. fijiensis* PKS2-1 protein sequence, compared to domain organization of *A. solani* PKSN, which is encoded by the gene *alt5*. Each domain is shown in a different color. KS = ketosynthase (pink); AT = acyltransferase (orange); DH = dehydratase (blue); MT = methyltransferase (brown); ER = enoyl reductase (green); KR = ketoreductase (yellow); ACP = acyl carrier protein domain (purple). B) Putative biosynthetic cluster for *M. fijiensis PKS2-1* gene compared to alternapyrone biosynthetic cluster genes. Types of genes common between the two clusters being compared are shown in the same color. Types of genes that are different between the two clusters are shown in light blue.

PKS8-2 forms a clade with the Fum1p fumonisin PKS protein sequence from *Fusarium verticillioides* (Figure 4). Fum1p has a protein sequence similarity of 72% with the AAL-toxin PKS Alt1p as well as an identical sequence of PKS domains: KS-AT-DH-

MT-ER-KR-ACP[171]. The two PKS genes are also clustered with similar genes. The PKS enzymes produce similar, dimethylated, highly reduced polyketides, with the carbon backbone for AAL-toxin being 2 carbons shorter. Polyketide products for many PKS enzymes are released from the PKS via the thioesterase domain[172], however Fum1p and Alt1p do not contain thioesterase domains[171]. Instead, release of the 18-carbon chain for fumonisin and the 16-carbon chain for AAL-toxin is catalyzed via a α-oxoamine synthase, which condenses the polyketide chain with the α-carbon of L-alanine or glycine, respectively[171,173–177].

Based on the domain analysis, PKS8-2 has the same domain organization as Fum1p and Alt1p (Figure 6A). Gene clusters were also compared for *PKS8-2* and *FUM1* (Figure 6B). Sixteen genes have been described from the fumonisin biosynthetic cluster, versus 6 genes predicted from the *M. fijiensis PKS8-2* cluster. Both clusters were predicted to contain a PKS, a transcription factor, a transporter, and an α-oxoamine synthase. For the genes with similar functions between these clusters, protein sequence similarity was determined by alignment with BLAST. While genes with similar predicted functions had significant alignment hits, sequence similarity was not especially high, with only the transcription factor having >60% sequence similarity (Table S4).

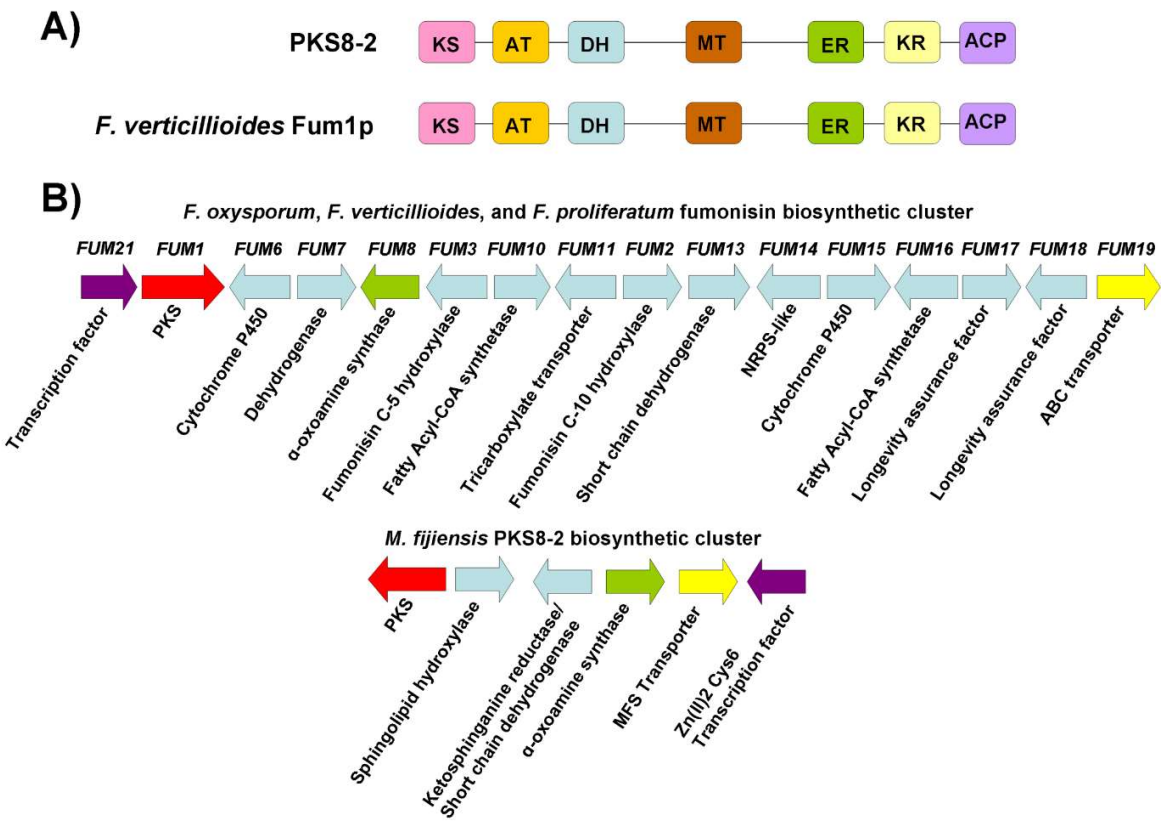


Figure 6. Comparison of domains and putative cluster genes for *M. fijiensis PKS8-2* and the fumonisin-producing PKS.
A) Domain organization of *M. fijiensis* PKS8-2 protein sequence, compared to domain organization of *F. verticillioides* Fum1p. Each domain is shown in a different color. KS = ketosynthase (pink); AT = acyltransferase (orange); DH = dehydratase (blue); MT = methyltransferase (brown); ER = enoyl reductase (green); KR = ketoreductase (yellow); ACP = acyl carrier protein domain (purple). B) Putative biosynthetic cluster for *M. fijiensis* PKS8-2 gene compared to fumonisin biosynthetic cluster. Types of genes common between the two clusters being compared are shown in the same color. Types of genes that are different between the two clusters are shown in light blue.

PKS10-2 formed a clade with good bootstrap support with the solanapyrone PKS Sol1 from *Alternaria solani*, which produces solanapyrone (Figure 4). Sol1 has a domain organization of: KS-AT-DH-MT-ER-KR-ACP[178]. PKS10-2 has a similar domain organization, except that while the E-value for the Sol1 ER domain is 3.69e-94, the E-value

for the PKS10-2 ER domain is only 2.22e-05, and the Conserved Domain Database does not recognize the PKS10-2 ER active site (Figure 7A, Table S3). Therefore, it is not clear if the PKS10-2 ER domain is functional. Since the role of the ER domain in a polyketide synthase is to reduce a double bond to a single bond, the initial polyketide product of PKS10-2 may have an additional double bond that solanapyrone lacks. The solanapyrone cluster contains a PKS, an O-methyltransferase, a dehydrogenase, a transcription factor, a oxidase, and a cytochrome P450[178]. The *PKS10-2* cluster is predicted to contain a PKS, a mono-oxygenase, two cytochrome P450s, an oxidoreductase with homology to ketopantoate reductase, a dehydrogenase, and a non-ribosomal peptide synthase (Figure 7B). While both clusters contain a cytochrome P450 and a dehydrogenase, they are otherwise very different, especially considering the PKS10-2 cluster is proposed to include a non-ribosomal peptide synthase (Figure 7B). Furthermore, while genes with similar predicted functions in the two clusters had significant alignment hits, sequence similarity was low, at 40% and 45% for the cytochrome P450 and the dehydrogenase, respectively (Table S4).

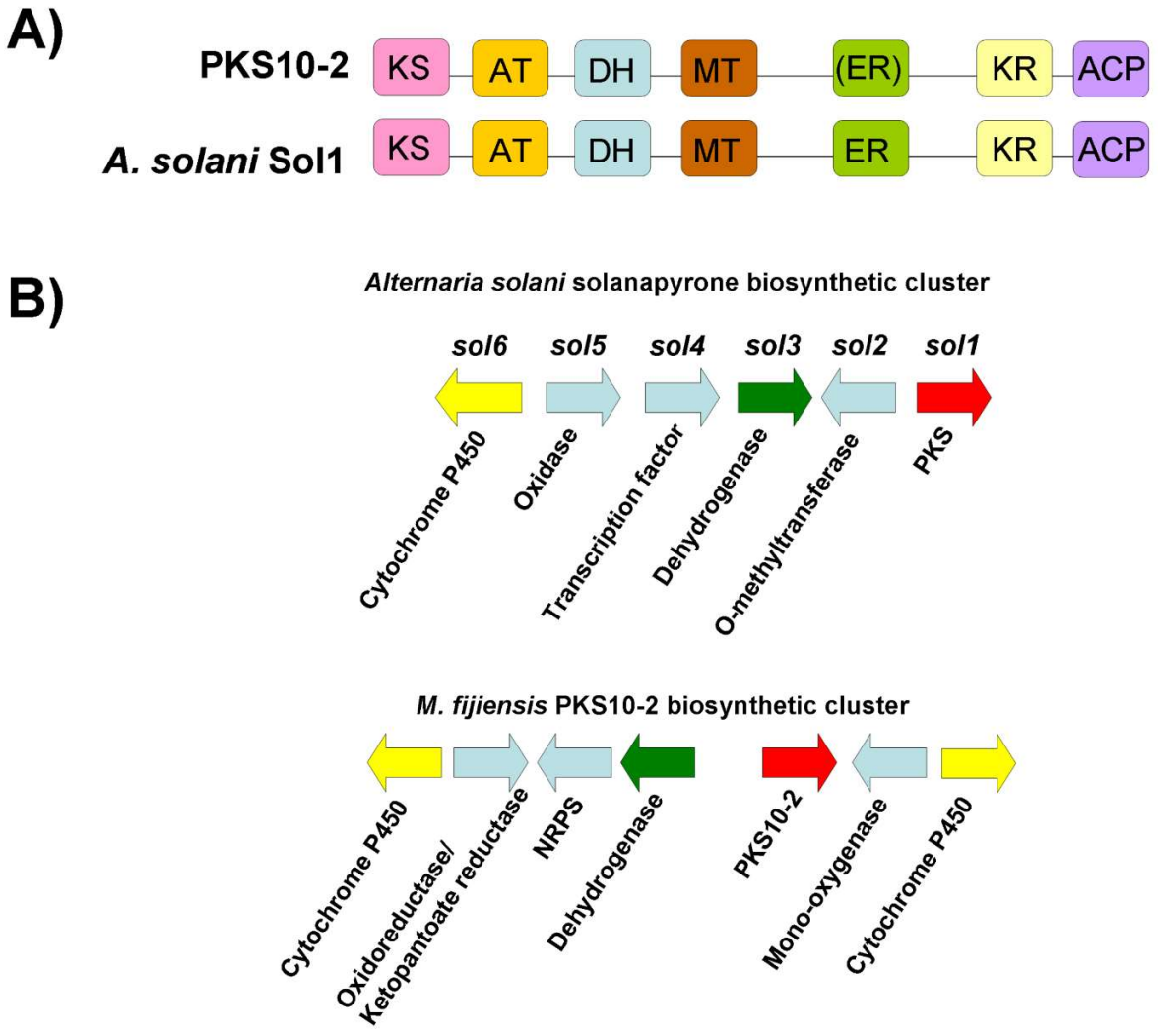


Figure 7. Comparison of domains and putative cluster genes for *M. fijiensis PKS10-2* and the solanapyrone-producing PKS.
A) Domain organization of *M. fijiensis* PKS10-2 protein sequence, compared to domain organization of *A. solani* Sol1. Each domain is shown in a different color. KS = ketosynthase (pink); AT = acyltransferase (orange); DH = dehydratase (blue); MT = methyltransferase (brown); ER = enoyl reductase (green); KR = ketoreductase (yellow); ACP = acyl carrier protein domain (purple). The ER domain of PKS10-2 is shown with parentheses because while this domain was predicted using NCBI's Conserved Domain Database, the E-value was very poor compared to E-values predicted for ER domains in other PKS sequences (Table S3), and therefore it is uncertain whether the domain is functional. B) Putative biosynthetic cluster for *M. fijiensis PKS10-2* gene compared to the solanapyrone biosynthetic cluster. Types of genes common between the two clusters being compared are shown in the same color. Types of genes that are different between the two clusters are shown in light blue.

KS Modeling of Tertiary Structure

It has been shown that size of the substrate-binding cavity in the ketosynthase (KS) domain is important for determining the number of iterations[179]. Larger substrate binding cavities correlate with more iterations to form the polyketide product. KS domain cavities fall into three main size categories: small, medium, and large. For example, 6-methylsalicylic acid synthase (MSAS) and related PKS enzymes have a small substrate binding cavity of about 300 cubic angstroms, and they perform three iterations. Naphthopyrone (NAP) and related PKS enzymes have intermediate sized binding cavities of about 800 cubic angstroms, and perform five to eight iterations. The T-toxin PKS has the largest sized binding cavity, and it performs 20 iterations[179] (Figure 4).

Polyketide synthase and fatty acid synthase genes share similar domains and have a common evolutionary origin[180], and a crystal structure of the fatty acid synthesis enzyme 1kas is available[181]. Tertiary structure models comparing MSAS- and NAP-type PKS enzymes with the 1kas crystal structure reveals very similar structures. In each, two amino acid residues (positions 229 and 400 in 1kas) protrude into the substrate binding cavity, affecting its size. In MSAS-type PKS enzymes carrying out three iterations, these two amino acid residues are completely conserved (Y,Y). NAP-type PKS enzymes have Tyr and Ala residues in these positions, allowing their substrate binding cavity to be larger than that of MSAS[179].

Using the PKS/NRPS Analysis Website[182], KS domain sequences were predicted for each of the *M. fijiensis* PKS protein sequences, as well as for the other PKS sequences used for generation of the phylogenetic tree in Figure 4. These sequences were then aligned with

the fatty acid synthesis protein sequence for 1kas. Using this alignment, amino acid residues corresponding to position 229 and 400 in 1kas were identified. Since PKS7-1, PKS8-1, and PKS10-1 were most closely related to NAP-type rather than MSAS-type or T-toxin-type PKS proteins (Figure 4), it was expected that they would have Tyr and Ala for the residues protruding into the active site. Indeed, Tyr and Ala were found to be in those two positions (Table S5). The observation that PKS enzymes in the NAP-type clade have Tyr and Ala in these positions and all catalyze 5-8 iterations[179] is consistent with our observation that PKS7-1, PKS8-1, and PKS10-1 also have Tyr and Ala in these positions.

Since PKS2-1, PKS8-2, Hybrid8-3, PKS8-4, and PKS10-2 were more closely related to the T-toxin PKS than to the MSAS- or NAP-type PKS proteins (Figure 4), it is uncertain whether these two residues are informative. It is possible that the tertiary structure of their KS domains is different such that the two amino acid residues no longer line the cavity in the same way. However, PKS2-1 and PKS10-2 have the same two residues as the T-toxin and the solanapyrone PKS enzymes which catalyze 20 and 8 iterations respectively; PKS8-2 has the same two residues as the fumonisin PKS which catalyzes 9 iterations; and Hybrid8-3 and PKS8-4 have the same two residues as the compactin and lovastatin PKS enzymes, which both catalyze 8 iterations (Table S5).

Uncharacterized PKS Homologs in Dothideomycete Genomes

The phylogenetic analysis we conducted (Figure 4) only included sequences of characterized polyketide synthases with known products. In recent years, several genomes or partial genomes within the Mycosphaerellaceae have been sequenced and are publicly

available including those of *Mycosphaerella musicola, Mycosphaerella eumusae, Mycosphaerella graminicola*, *Dothistroma septosporum*, *Septoria musiva*, *Pseudocercospora pini-densiflorae*, *Cercospora zeae-maydis*, *Mycosphaerella laricina*, and *Mycosphaerella arachidis*. Many of these are hemibiotrophic plant pathogens like *M. fijiensis* that are hypothesized to use toxic metabolites to facilitate pathogenesis. Since the range of polyketides produced by these fungi has not yet been fully characterized, it is possible that *M. fijiensis* produces similar metabolites as some of its close relatives.

Tblastn searches were done for each *M. fijiensis* PKS sequence against 103 Dothideomycete fungal genomes available from JGI and NCBI[45], and hits were arranged by bitscore to identify the closest homologs for each *M. fijiensis* PKS. Results are shown in Table 2. This analysis revealed that PKS2-1, PKS8-1, Hybrid8-3, PKS8-4, PKS10-1, and PKS10-2 all have close homologs with >80% sequence similarity in Mycosphaerellaceae genomes. By contrast, the best homologs of PKS7-1 have <60% sequence similarity, suggesting that this cluster may produce a unique product. Comparison of the *M. fijiensis* PKS sequences with those of the other two banana pathogens, *M. musicola* and *M. eumusae,* showed that these two species have close homologs to PKS2-1, Hybrid8-3, PKS10-1, and PKS10-2. By contrast, a PKS8-4 homolog was only found in *M. musicola*, and PKS7-1, PKS8-1, and PKS8-2 homologs were not found in either of the other species that infect banana.

Table 2. Top 10 tblastn hits by bitscore, for each *M. fijiensis* PKS.
Table indicates the *M. fijiensis* PKS, species where the blast hit was found, percent sequence identity, percent sequence similarity, bitscore, and E-value. Hits are arranged in the table by bitscore, and are color-coded by percent sequence similarity: pink, >80% similarity; orange, >70-80% similarity; yellow, >60-70% similarity; green,>50-60% similarity; gray, >40-50% similarity. A) PKS2-1; B) PKS7-1; C) PKS8-1; D) PKS8-2; E) Hybrid 8-3; F) PKS8-4; G) PKS10-1; H) PKS10-2.

Table 2A

***M. fijiensis* PKS**	**Species of BLAST hit**	**% identity**	**% similarity**	**bitscore**	**E-value**
PKS2-1	*Pseudocercospora pini-densiflorae*	92.51	95.29	4074	0
PKS2-1	*Mycosphaerella musicola*	92.53	95.72	4010	0
PKS2-1	*Mycosphaerella eumusae*	91.35	94.39	4008	0
PKS2-1	*Zymoseptoria passerinii*	70.26	80.76	2777	0
PKS2-1	*Mycosphaerella graminicola*	66.91	79.18	2140	0
PKS2-1	*Zymoseptoria pseudotritici*	67.6	79.46	2136	0
PKS2-1	*Zymoseptoria brevis*	66.91	79	2133	0
PKS2-1	*Zymoseptoria ardabiliae*	66.73	79.24	2120	0
PKS2-1	*Rhytidhysteron rufulum*	50.28	66.12	1475	0
PKS2-1	*Zasmidium cellare*	47.71	64.59	1043	0

Table 2B

***M. fijiensis* PKS**	**Species of BLAST hit**	**% identity**	**% similarity**	**bitscore**	**E-value**
PKS7-1	*Macroventuria anomochaeta*	36.02	55.48	1107	0
PKS7-1	*Lizonia empirigonia*	35.83	55.59	1107	0
PKS7-1	*Delitschia confertaspora*	36.05	54.69	1101	0
PKS7-1	*Didymella exigua*	36.34	55.2	1097	0
PKS7-1	*Hysterium pulicare*	36.25	54.67	1095	0
PKS7-1	*Pleomassaria siparia*	35.88	55.11	1095	0
PKS7-1	*Didymella zeae-maydis*	35.76	54.8	1094	0
PKS7-1	*Sporormia fimetaria*	37.51	55.03	1092	0
PKS7-1	*Elsinoe ampelina*	36.19	54.18	1091	0
PKS7-1	*Melanomma pulvis-pyrius*	35.98	54.82	1090	0

Table 2C

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
PKS8-1	*Pseudocercospora pini-densiflorae*	89.78	92.44	2572	0
PKS8-1	*Byssothecium circinans*	69.53	81.47	2416	0
PKS8-1	*Passalora fulva*	61.84	76	2283	0
PKS8-1	*Cladosporium fulvum*	61.84	76	2283	0
PKS8-1	*Pyrenophora tritici-repentis*	65.36	77.94	2222	0
PKS8-1	*Pyrenophora teres f. teres*	64.95	77.98	2217	0
PKS8-1	*Septoria populicola*	66.13	78.08	2083	0
PKS8-1	*Septoria musiva*	65.88	77.89	2083	0
PKS8-1	*Trematosphaeria pertusa*	60.92	75.53	2038	0
PKS8-1	*Cercospora canescens*	66.5	78.04	1942	0

Table 2D

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
PKS8-2	*Zymoseptoria pseudotritici*	45.14	63.39	2043	0
PKS8-2	*Mycosphaerella graminicola*	45.3	63.19	2035	0
PKS8-2	*Zymoseptoria passerinii*	45.21	63.3	2030	0
PKS8-2	*Aulographum hederae*	51.35	68.61	1892	0
PKS8-2	*Lindgomyces ingoldianus*	43.28	61.17	1828	0
PKS8-2	*Decorospora gaudefroyi*	45.09	62.41	1698	0
PKS8-2	*Zasmidium cellare*	40.6	58.46	1622	0
PKS8-2	*Polyplosphaeria fusca*	40.05	58.29	1620	0
PKS8-2	*Pleomassaria siparia*	42.61	62.21	1585	0
PKS8-2	*Amniculicola lignicola*	39.29	57	1584	0

Table 2E

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
Hybrid8-3	*Pseudocercospora pini-densiflorae*	83.96	89.59	6688	0
Hybrid8-3	*Mycosphaerella eumusae*	86.18	91.32	4383	0
Hybrid8-3	*Mycosphaerella musicola*	85.07	91.06	4315	0
Hybrid8-3	*Cercospora zeae-maydis*	69.61	81.3	3565	0
Hybrid8-3	*Cercospora canescens*	68.93	80.92	3553	0
Hybrid8-3	*Septoria populicola*	67.38	78.76	3441	0
Hybrid8-3	*Passalora fulva*	66.15	78.59	3388	0
Hybrid8-3	*Cladosporium fulvum*	66.15	78.59	3388	0
Hybrid8-3	*Mysosphaerella sp. Ston1*	67.5	79.59	3302	0
Hybrid8-3	*Mycosphaerella arachidis*	65.39	77.82	3267	0

Table 2F

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
PKS8-4	*Mycosphaerella musicola*	80.84	88.21	2907	0
PKS8-4	*Pyrenophora teres f. teres*	47.77	63.93	2460	0
PKS8-4	*Lizonia empirigonia*	42.25	60.09	2120	0
PKS8-4	*Alternaria alternata*	45.2	63.2	2028	0
PKS8-4	*Glonium stellatum*	46.56	64.8	1823	0
PKS8-4	*Pyrenophora tritici-repentis*	49.92	65.89	1746	0
PKS8-4	*Setosphaeria turcica*	49.77	66.6	1723	0
PKS8-4	*Pseudovirgaria hyperparasitica*	38.77	56.7	1558	0
PKS8-4	*Stagonospora sp.* SRC1lsM3a	38.86	57.64	1553	0
PKS8-4	*Cenococcum geophilum*	39.29	57.45	1523	0

Table 2G

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
PKS10-1	*Pseudocercospora pini-densiflorae*	97.65	98.64	4080	0
PKS10-1	*Mycosphaerella musicola*	95.48	97.65	4012	0
PKS10-1	*Mycosphaerella eumusae*	95.75	97.06	4011	0
PKS10-1	*Mycosphaerella laricina*	83.71	90.6	3604	0
PKS10-1	*Passalora fulva*	83.94	90.47	3575	0
PKS10-1	*Cladosporium fulvum*	83.94	90.47	3575	0
PKS10-1	*Mycosphaerella arachidis*	82.82	89.68	3565	0
PKS10-1	*Cercospora zeae-maydis*	83.21	89.85	3551	0
PKS10-1	*Dothistroma septosporum*	82.68	90.15	3549	0
PKS10-1	*Mysosphaerella sp. Ston1*	81.83	88.25	3534	0

Table 2H

M. fijiensis PKS	Species of BLAST hit	% identity	% similarity	bitscore	E-value
PKS10-2	*Mycosphaerella eumusae*	80.87	89.18	1412	0
PKS10-2	*Mycosphaerella musicola*	80.48	88.01	1398	0
PKS10-2	*Mycosphaerella sp. Ston1*	43.58	59.15	1282	0
PKS10-2	*Melanomma pulvis-pyrius*	33.02	49.34	1163	0
PKS10-2	*Pyrenochaeta sp.* DS3sAY3a	32.95	50.28	1160	0
PKS10-2	*Cochliobolus carbonum*	35.72	51.89	1035	0
PKS10-2	*Stagonospora sp.* SRC1lsM3a	33.55	50.45	1034	0
PKS10-2	*Setosphaeria turcica*	32.09	48.07	984	0
PKS10-2	*Cucurbitaria berberidis*	35.08	52.77	984	0
PKS10-2	*Aureobasidium pullulans* var. *subglaciale*	42.68	58.86	937	0

Since the polyketide synthases PKS2-1, PKS8-4, and PKS10-2 each had close homologs in the related banana pathogens *M. musicola* and *M. eumusae* (Table 2), we compared the biosynthetic clusters for each of the corresponding PKS genes in the different species. Conserved domains were predicted for proteins encoded by genes neighboring the PKS gene in each species[45]. For sequences with similar predicted functions between species, a blastp search was done for each *M. fijiensis* protein sequence to determine percent identity and similarity with its ortholog in the other species[45]. This analysis revealed that the biosynthetic clusters are extremely similar for *PKS2-1*. Clusters for all three species share genes encoding a PKS, an MFS transporter, a transcription factor, a monooxygenase, a dioxygenase, and at least one serine hydrolase (Figure 8), and most of the orthologous protein sequences have >90% similarity[45]. For the *PKS8-4* gene cluster, both *M. fijiensis* and *M. musicola* have genes encoding a PKS, a cytochrome P450, and a dehydrogenase (Figure 9). However, the *M. fijiensis* gene cluster contains a gene encoding an enoyl reductase-like protein which *M. musicola* lacks, and the *M. musicola* gene cluster contains a second dehydrogenase gene that *M. fijiensis* lacks (Figure 9). The *PKS8-4* cluster proteins that the two species share all have at least 83% sequence similarity[45]. All three species have nearly identical biosynthetic clusters for *PKS10-2* (Figure 10). All three species have genes encoding a PKS, an NRPS, a dehydrogenase, a monooxygenase, a cytochrome P450, and an oxidoreductase with similarity to ketopantoate reductase. In *M. fijiensis* and *M. eumusae*, the dehydrogenase is predicted to be encoded by a stand-alone gene, whereas in *M. musicola*, it is predicted to be a part of the NRPS (Figure 10). *M. fijiensis* and *M. eumusae* have a second cytochrome P450 gene distal to the monooxygenase. The *M. musicola* scaffold ends just after

the monooxygenase gene, and a close homolog of this cytochrome P450 is found at the end of another scaffold. It may be that greater sequencing depth of the *M. musicola* genome would reveal that these two scaffolds are in fact part of the same chromosome. Regardless, all proteins encoded by the *M. fijiensis PKS10-2* biosynthetic cluster have orthologs with high sequence similarity (at least 70%) in these two other banana pathogens.

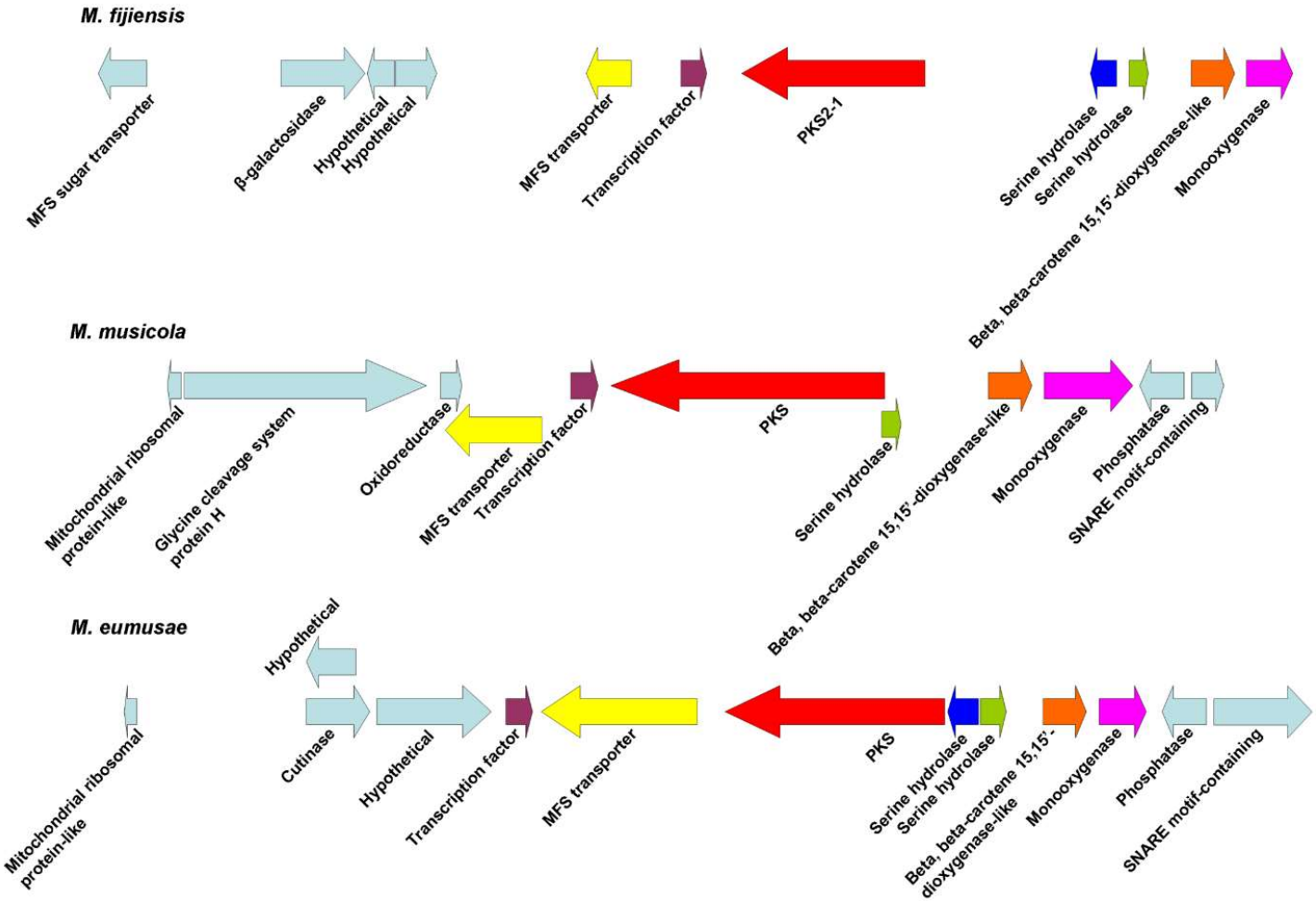


Figure 8. Comparison of *PKS2-1* gene clusters from *Mycosphaerella* species.
Putative biosynthetic cluster for *M. fijiensis PKS2-1* gene compared to the orthologous cluster in *M. musicola* and *M. eumusae*. Putative orthologous genes are shown in the same color. Genes flanking the putative biosynthetic cluster are shown in light blue.

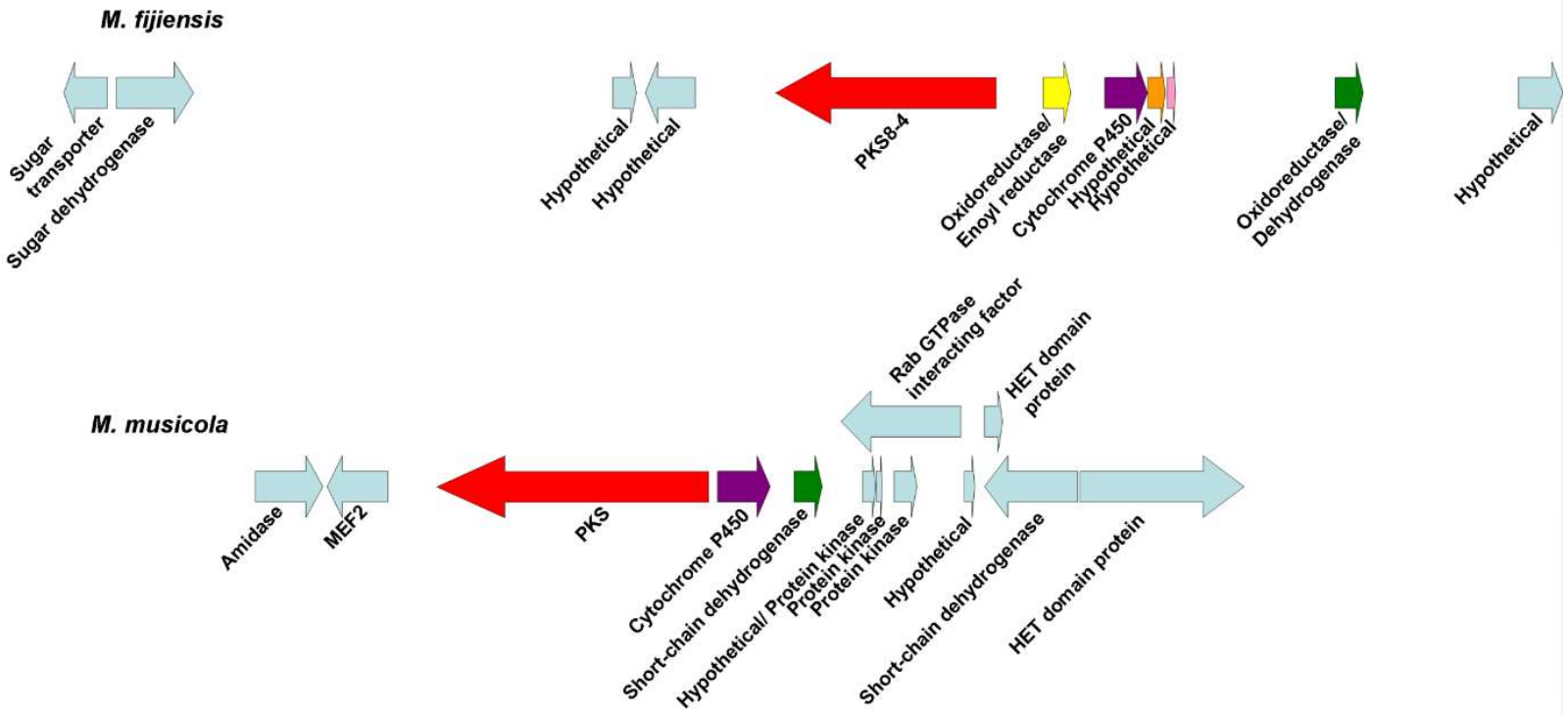


Figure 9. Comparison of *PKS8-4* gene clusters from *M. fijiensis* and *M. musicola*.
Putative biosynthetic cluster for *M. fijiensis PKS8-4* gene compared to the orthologous cluster in *M. musicola*. Putative orthologous genes are shown in the same color. Genes flanking the putative biosynthetic cluster are shown in light blue.

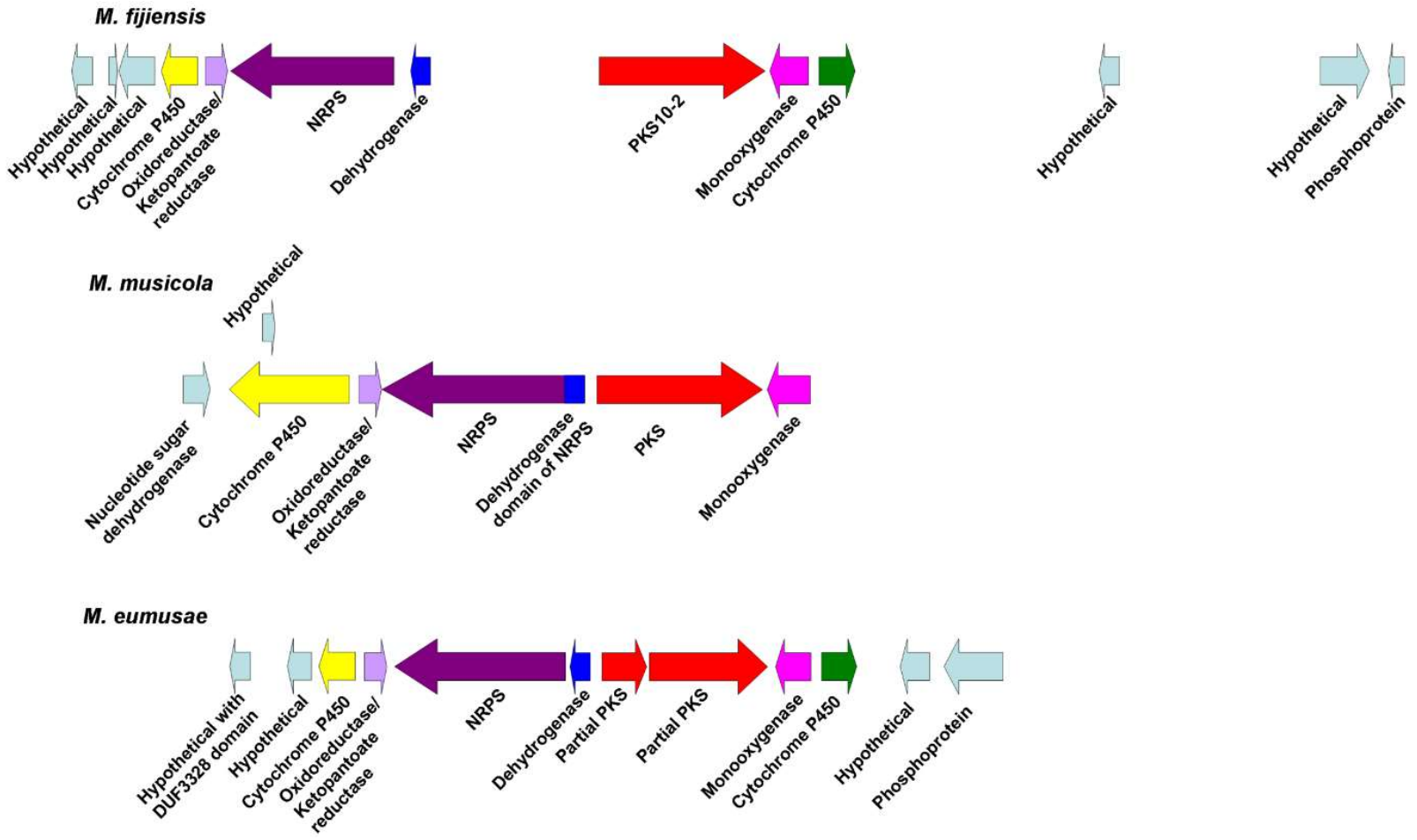


Figure 10. Comparison of *PKS10-2* gene clusters from *Mycosphaerella* species.
Putative biosynthetic cluster for *M. fijiensis PKS10-2* gene compared to the orthologous cluster in *M. musicola* and *M. eumusae*. Putative orthologous genes are shown in the same color. Genes flanking the putative biosynthetic cluster are shown in light blue.

Expression of *M. fijiensis* PKS Genes During Colonization of Banana

We hypothesized that PKS genes important for pathogenicity should be expressed during infection of host banana tissue. To identify PKS genes that are expressed during infection, we first compared expression of the *M. fijiensis* PKS genes in inoculated leaves of in vitro-cultured banana plants as compared to mycelial cultures using RT-PCR analysis. Isolate 10CR1-24 was used to inoculate leaves of tissue cultured plants, and leaves were harvested at 5 weeks post-inoculation, once they had become symptomatic (Figure S1C). As a control, isolate 10CR1-24 was grown in liquid PDB medium, and mycelium was harvested after 1 week incubation. RT-PCR assays were performed using tissue from both conditions (Figure 11). This analysis revealed that *Hybrid8-3*, *PKS2-1*, and *PKS8-1* were most strongly expressed in culture, with less expression in infected leaf tissue. *PKS8-2* and *PKS8-4* were expressed more often in infected leaves than in culture, and *PKS10-2* was strongly expressed in the infected leaf samples, but not expressed in the flask samples. *PKS10-1* was expressed under both conditions, and *PKS7-1* expression was not detected in any samples in either condition.

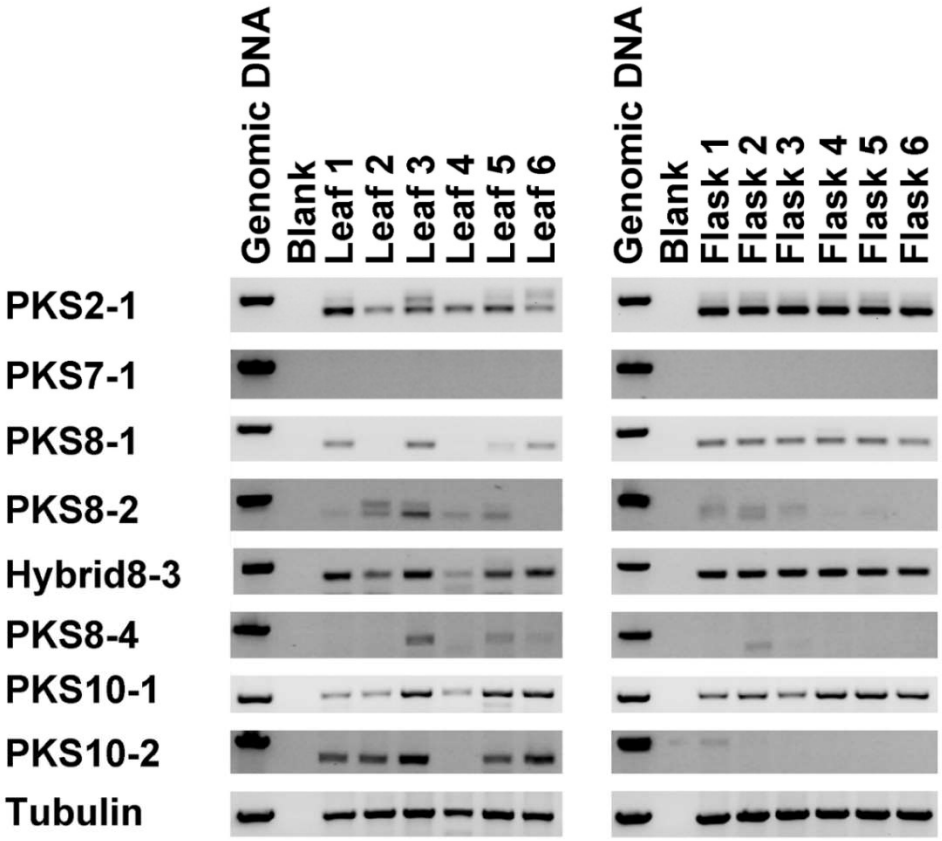
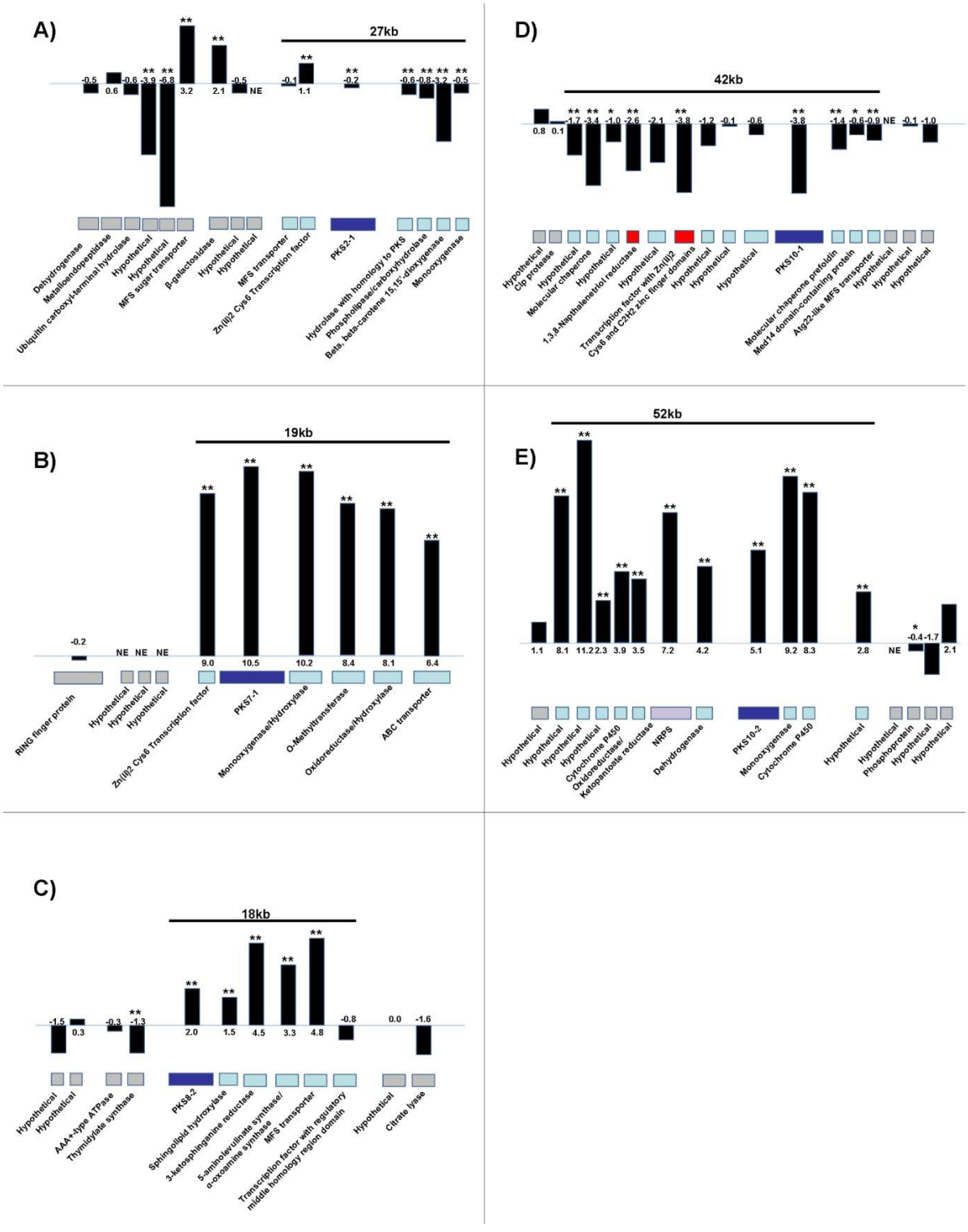


Figure 11. RT-PCR analysis of PKS gene expression in infected leaves and mycelial cultures.
M. fijiensis isolate 10CR1-24 was used to infect banana tissue-culture plants, as well as to inoculate PDB medium. RT-PCR assays for each PKS or hybrid PKS-NRPS were performed on the resulting leaf lesions or mycelium from the PDB liquid culture. Where possible (for all genes except beta-tubulin and *PKS10-1*), primers were designed to span an intron, to distinguish cDNA from any gDNA contamination. Genomic DNA isolated from 10CR1-24 was used as a positive control, and reactions were also done with no added template, as a negative control. Each lane represents RNA isolated from individual infected leaves (left) or mycelium grown in flask culture (right).

We then conducted RNA-Seq analysis to further characterize expression of PKS genes as well as the other genes proposed to be in each cluster. We inoculated potted 'Grand Nain' banana plants grown in the greenhouse or PDB flasks with *M. fijiensis* isolate 14H1-11A. Mycelium from PDB flasks was harvested after 1 week, and infected leaves were harvested at 6 weeks post-inoculation. RNA was isolated from these tissue samples and sequenced using the Illumina HiSeq platform. We focused our analysis on the PKS genes and clusters identified as being of interest through our phylogenetic, PKS domain, and cluster analysis (Figures 4-7). Results are shown in Figure 12.

Figure 12. RNA-Seq results of PKS cluster gene expression in infected leaves relative to mycelial culture.
Potted banana plants or PDB medium were inoculated with conidia of *M. fijiensis* isolate 14H1-11A, and tissue was harvested for RNA-Seq analysis. The PKS is shown, surrounded by its neighboring genes in the genome. Each gene is labeled with its function as predicted by BLAST. PKS genes are colored in dark blue, NRPS genes are colored in purple, genes known to be involved in melanin biosynthesis are colored in red, putative biosynthetic cluster genes based on a combination of bioinformatics predictions and fold changes are colored in light blue, and putative flanking genes are colored in gray. For each gene, $\log_2$ fold change values (log2FC) are shown for expression in the infected tissue versus expression in liquid culture. Expression data shown as black bars; actual log2FC values are shown at the base of the bars. Gene expression differences that are significant at $p<0.05$ are indicated by a single asterisk, and gene expression differences that are significant at $p<0.01$ are indicated by two asterisks. NE = no expression detected. A) *PKS2-1* cluster; B) *PKS7-1* cluster; C) *PKS8-2* cluster; D) *PKS10-1* (melanin) cluster; E) *PKS10-2* cluster.


A)
27kb
-0.5 0.6 -0.6 -3.9 -6.8 3.2 2.1 -0.5 NE -0.1 1.1 -0.2 -0.6 -0.8 -3.2 -0.5
Dehydrogenase
Metalloendopeptidase
Ubiquitin carboxyl-terminal hydrolase
Hypothetical
Hypothetical
MFS sugar transporter
β-galactosidase
Hypothetical
Hypothetical
MFS transporter
Zn(II)2 Cys6 Transcription factor
PKS2-1
Hydrolase with homology to PKS
Phospholipase/carboxyhydrolase
Beta, beta-carotene 15,15'-dioxygenase
Monooxygenase
D)
42kb
0.8 0.1 -1.7 -3.4 -1.0 -2.6 -2.1 -3.8 -1.2 -0.1 -0.6 -3.8 -1.4 -0.6 -0.9 NE -0.1 -1.0
Hypothetical
Clp protease
Hypothetical
Molecular chaperone
Hypothetical
1,3,8-Naphthalenetriol reductase
Hypothetical
Transcription factor with Zn(II)2 Cys6 and C2H2 zinc finger domains
Hypothetical
Hypothetical
Hypothetical
PKS10-1
Molecular chaperone prefoldin
Med14 domain-containing protein
Atg22-like MFS transporter
Hypothetical
Hypothetical
Hypothetical
B)
19kb
-0.2 NE NE NE 9.0 10.5 10.2 8.4 8.1 6.4
RING finger protein
Hypothetical
Hypothetical
Hypothetical
Zn(II)2 Cys6 Transcription factor
PKS7-1
Monooxygenase/Hydroxylase
O-Methyltransferase
Oxidoreductase/Hydroxylase
ABC transporter
E)
52kb
1.1 8.1 11.2 2.3 3.9 3.5 7.2 4.2 5.1 9.2 8.3 2.8 NE -0.4 -1.7 2.1
Hypothetical
Hypothetical
Hypothetical
Hypothetical
Cytochrome P450
Oxidoreductase/
Ketopentiotte reductase
NRPS
Dehydrogenase
PKS10-2
Monooxygenase
Cytochrome P450
Hypothetical
Hypothetical
Phosphoprotein
Hypothetical
Hypothetical
C)
18kb
-1.5 0.3 -0.3 -1.3 2.0 1.5 4.5 3.3 4.8 -0.8 0.0 -1.6
Hypothetical
Hypothetical
AAA+-type ATPase
Thymidylate synthase
PKS8-2
Sphingolipid hydroxylase
3-ketosphinganine reductase
5-aminolevulinate synthase/
α-oxoamine synthase
MFS transporter
Transcription factor with regulatory
middle homology region domain
Hypothetical
Citrate lyase

In this dataset, *PKS2-1* showed a small but statistically significant decreased expression in the infected leaf compared to when grown in PDB (Figure 12A). These results are consistent with the greater expression found from mycelial culture relative to infected tissue-cultured plants in the initial RT-PCR analysis (Figure 11). With the exception of the transcription factor, all of the genes predicted to be in the cluster (MFS transporter, hydrolase, phospholipase, dioxygenase, and monooxygenase) (Figures 3A and 5B) were also all unchanged or significantly decreased in expression in infected leaves relative to mycelial culture (Figure 12A).

In contrast to *PKS2-1*, RNA-Seq analysis showed *PKS7-1* to have a highly significant, large increase in expression in the infected leaf as compared to when grown in liquid medium (Figure 12B). All of the genes predicted to be in the cluster (transcription factor, monooxygenase, O-methyltransferase, oxidoreductase, and ABC transporter genes) (Figure 3B) showed similar increases in expression, while genes flanking those such as three hypothetical genes and a RING finger protein gene did not show such differences.

RNA-Seq analysis of *PKS8-2* showed a statistically significant increase in expression in the infected leaves relative to when grown in PDB (Figure 12C). The neighboring sphingolipid hydroxylase, ketosphinganine reductase, α-oxoamine synthase, and MFS transporter predicted to be part of the biosynthetic cluster (Figures 3D and 6B) also had statistically significant increases in expression in infected leaf tissue, whereas the neighboring transcription factor gene was not differentially expressed.

RNA-Seq analysis showed a large, statistically significant decrease in expression of *PKS10-1* in the infected leaf versus when grown in PDB (Figure 12D). Fungal melanin

biosynthesis is known to require a 1,3,8-trihydroxynaphthalene reductase[183,184], and a transcription factor[185], both of which are clustered with the melanin PKS. These two genes also had decreased expression in infected leaf tissue as compared to when grown in PDB (Figure 12D). In addition to the genes known or predicted to be involved in melanin biosynthesis (Figure 3G), six more flanking genes had significantly decreased expression in infected leaf tissue: two molecular chaperone genes, a gene for a Med14-domain containing protein, two genes for hypothetical proteins, and a vacuolar transporter-like gene (Figure 12D).

Finally, RNA-Seq analysis of *PKS10-2* showed a large, statistically significant increase in expression in infected leaf tissue relative to when grown in PDB (Figure 12E). Ten neighboring genes also had large, significant increases in expression, suggesting that the cluster is larger than predicted (Figures 3H and 7B) and may include four hypothetical proteins, two cytochrome P450s, an oxidoreductase with homology to ketopantoate reductase, a non-ribosomal peptide synthase, a dehydrogenase, and a monooxygenase. The RNA-Seq results are also consistent with the RT-PCR analysis from tissue-cultured plants showing expression only in infected leaf tissue and not in culture (Figure 11).

Discussion

While polyketide toxins have long been suspected as pathogenicity factors from *M. fijiensis*, this is the first study to investigate the repertoire of polyketide biosynthetic genes from this fungus. Using the publically available genome sequence, we have predicted that there are seven PKS genes and one hybrid PKS-NRPS gene in the genome of *M. fijiensis*.

When compared to the number of PKS gene clusters predicted from other fungi, this is relatively few (Figure 1). However, some fungal species with few PKS genes are known to produce polyketides as pathogenicity factors. For example, only four PKS genes were detected in the genome of *Dothistroma septosporum* (Figure 1, Table S2), yet it is known to produce the polyketide dothistromin, which is an important pathogenicity factor[154]. Only 6 PKS genes were detected in the genome of *Alternaria brassicicola* (Figure 1, Table S2), yet it is known to produce the polyketide depudecin, which acts as a histone deacetylase inhibitor and contributes to virulence[186]. Therefore, it is not possible to predict the importance of polyketides in virulence based on the number of PKS genes in the genome.

All of the *M. fijiensis* PKS and hybrid PKS-NRPS genes were predicted to encode enzymes with all of the necessary PKS domains (KS, AT, and ACP). They contain different combinations of optional domains, which are consistent with PKS7-1, PKS8-1, and PKS10-1 being non-reducing; Hybrid8-3, PKS8-4, and PKS10-2 being partially reducing; and PKS2-1 and PKS8-2 being highly reducing PKS enzymes.

In fungal genomes, secondary metabolite biosynthetic genes are typically clustered together, and genes such as those encoding transcription factors, transporters, oxidoreductases and dehydrogenases are known to commonly be found in biosynthetic clusters[47]. Therefore, we were able to predict which genes may be part of biosynthetic clusters with each *M. fijiensis* PKS or hybrid PKS-NRPS (Figure 3). The sizes of these predicted biosynthetic clusters ranged from 18 to 39 kb, which is very consistent with sizes of characterized polyketide biosynthetic clusters from other fungi. For example, the bikaverin, solanapyrone, cercosporin, fumonisin, and aflatoxin clusters are, respectively 18

kb, 21 kb, 36 kb, 46 kb, and 82 kb[166,167,178,187–189]. Our RNA-Seq experiment provided further clues as to which genes may be co-regulated and part of a biosynthetic cluster. These data largely confirmed our bioinformatics predictions. For example, we predicted that the *PKS7-1* cluster would include a transcription factor, a mono-oxygenase, an O-methyltransferase, an oxidoreductase, and an ABC transporter (Figure 3B). The RNA-Seq data showed that all of these genes had higher expression in infected leaf tissue, with $\log_2$ fold change values of at least 6.4, whereas other neighboring genes were not differentially expressed (Figure 12B). RNA-Seq analysis also confirmed cluster predictions for the *PKS2-1* and *8-2* (Figures 3, 5B, 6B, and 12). By contrast, RNA-Seq data enabled us to refine our prediction for the *PKS10-1* and *10-2* clusters: while all of the genes predicted to be in the clusters were differentially expressed, there were several neighboring hypothetical genes that had similar differential expression, suggesting that they are part of the clusters (Figures 3, 7, and 12).

Phylogenetic analysis of the *M. fijiensis* PKS protein sequences and well-characterized sequences from other species revealed that PKS10-1 is clearly part of the clade of melanin-producing PKS enzymes (Figure 4). *M. fijiensis* is highly melanized in culture, and melanin shunt metabolites such as 2,4,8-trihydroxytetralone, 4-hydroxyscytalone and juglone have been shown to be toxic to banana and have been implicated as toxins important in disease development[61,72,73,156,157]. Our expression analysis showed that the melanin *PKS10-1* is expressed in infected leaf tissue (Figure 11), however RNA-Seq analysis (Figure 12D) identified a significant down regulation of the cluster genes in leaf tissue relative to expression in culture. These results suggest that melanin shunt metabolites may not be involved in disease development, at least at the stages assayed in our studies (Figure S1). Our

expression studies do not address the possible role of melanin in other stages of pathogenicity such as entry into host tissue.

None of the other *M. fijiensis* PKS protein sequences had close, well-characterized homologs in our phylogenetic analysis (Figure 4). Further, protein sequences of genes in the predicted PKS clusters also had low sequence similarity to the types of genes clustered with the closest well-characterized PKS homolog (Tables S4)[45]. These results suggest that the PKS enzymes other than PKS10-1 are likely to produce novel compounds.

In spite of the lack of clear PKS homologs, some predictions can be made based on clades in the phylogenetic analysis, PKS domain organization, and modeling of the KS domain. For example, the closest well-characterized homologs of PKS2-1 are the alternapyrone and T-toxin PKS sequences (Figure 4). While PKS2-1 has a very different biosynthetic cluster (Figure 5B) and is unlikely to produce the same product, the alternapyrone and T-toxin PKS enzymes both produce highly saturated products, which supports our prediction from the domain analysis that PKS2-1 is a highly saturating PKS (Figure 2). Furthermore, the alternapyrone and T-toxin PKS enzymes catalyze 10 and 20 iterations respectively, which is more than the other well-characterized PKS enzymes in Figure 4, and suggests that the product of PKS2-1 is also large. PKS2-1 is highly homologous ($\geq$79 % sequence similarity) with as yet uncharacterized PKS proteins within several members of the Mycosphaerellaceae including *M. musicola*, *M. eumusae*, *M. graminicola, P. pini-densiflorae*, and several *Zymoseptoria* species (Table 2). Furthermore, the genes predicted to be in the PKS2-1 biosynthetic clusters of *M. fijiensis*, *M. musicola*, and *M. eumusae* are all very similar. Our results suggest that the product may be commonly

produced by this group of fungi. In our expression analysis, *PKS2-1* and its clustered genes were more highly expressed in culture than in leaf tissue (Figure 12A), suggesting a possible role in normal growth processes and not specifically in pathogenicity.

The *PKS10-2* cluster is of significant interest to us given its high expression in leaf tissue suggesting a possible role in disease development. Further, sequence similarity to PKS proteins from the related banana pathogens *M. musicola* and *M. eumusae* was >80%, whereas there was little similarity (<60%) to PKS proteins in other Mycosphaerellaceae (Table 2). Comparison of the *M. fijiensis PKS10-2* gene cluster with clusters from *M. musicola* and *M. eumusae* showed that the three species have nearly identical gene clusters. All of the genes predicted to be in the *M. fijiensis PKS10-2* gene cluster have close homologs in the other two fungi. Thus PKS10-2 may be of interest for further investigation for a possible role in pathogenicity of banana. PKS10-2 has a high bootstrap support value to a solanapyrone-producing PKS from *Alternaria solani* (Figure 4). Solanapyrone is a phytotoxic polyketide that inhibits DNA polymerase[190]. It has been thought to play a role in pathogenicity since it is less toxic to non-host than to host plants[191] although it is not essential for pathogenicity[192].

The solanapyrone and putative *PKS10-2* biosynthetic clusters also share some similar types of genes (PKS, cytochrome P450, and dehydrogenase). However, the putative *PKS10-2* cluster also contains an NRPS and other types of genes not found in the solanapyrone biosynthetic cluster (Figure 7B). Therefore, these are unlikely to produce the same compound. This conclusion is supported by the low sequence similarity between the clustered dehydrogenase and cytochrome P450 genes in the *PKS10-2* and solanapyrone clusters (Table S4B).

RNA-Seq analysis also showed the *M. fijiensis PKS8-2* cluster to be upregulated in leaf tissue relative to growth in culture (Figure 12C). Similarity to PKS proteins in other Mycospharellaceae was not strong (<70%) (Table 2), but PKS8-2 has a high bootstrap support value to the fumonisin-producing PKS (Figure 4). The fumonisin and *PKS8-2* biosynthetic clusters both contain genes homologous to α-oxoamine synthase (Figure 6B)[171]. α-oxoamine synthase is an enzyme involved in sphingolipid biosynthesis[193,194]. A homolog is used in fumonisin biosynthesis to produce a polyketide that can act as an analog of sphinganine and thereby inhibit subsequent steps in sphingolipid biosynthesis[171,194]. In addition to the α-oxoamine synthase homolog, the predicted *PKS8-2* cluster also includes genes homologous to ketosphinganine reductase and sphingolipid hydroxylase (Figure 6B), which further suggests that the polyketide produced by the *PKS8-2* cluster may perturb sphingolipid metabolism. Disruption of sphingolipid biosynthesis by fumonisin results in toxicity because sphingolipids are important components of eukaryotic cell membranes, and are involved in signal transduction for a variety of processes[195,196]. Fumonisin is an important pathogenicity factor, causing necrosis in sensitive plant hosts[197]. It is also cytotoxic and carcinogenic to animals[198].

Of all the clusters, the *PKS7-1* cluster remains a mystery. It is of interest due to its high expression in infected plant tissue in the RNA-Seq experiment. The genes clustered with *PKS7-1* (oxidoreductase, monooxygenase, O-methyltransferase) are also similar to genes in the biosynthetic cluster for the production of cercosporin, a light-activated toxin produced by the related *Cercospora* species that has been shown to be involved in disease development[57,187]. BLAST analysis, however, did not identify any close homologs (Figure 4,

Table 2), including homologs in the banana pathogens *M. musicola* and *M. eumusae.* Thus PKS7-1 seems to be unique among PKS enzymes identified to date.

In conclusion, we have identified eight PKS or hybrid PKS-NRPS biosynthetic gene clusters in the *M. fijiensis* genome and we have used bioinformatics to make predictions about the products synthesized by these gene clusters. Our data predict that *PKS10-1* is involved in melanin biosynthesis, and that three other PKS clusters (*PKS2-1*, *8-2*, and *10-2*) are similar to clusters that produce alternapyrone, fumonisin, and solanapyrone, respectively. Four of the clusters (*PKS2-1, Hybrid8-3, PKS10-1,* and *PKS10-2*) are found in both of the related banana pathogens *M. musicola* and *M. eumusae* (additionally, *PKS8-4* is found in *M. musicola* only), however three of the clusters (*PKS7-1, PKS8-1,* and *PKS8-2*) are not found in these related species. Three of the clusters (*PKS7-1*, *PKS8-2*, and *PKS10-2*) are highly expressed in infected leaf tissue and are thus potential targets for further characterization of these pathways and polyketide products for a role in *M. fijiensis* pathogenicity.

Methods

Prediction of Polyketide Synthase Gene Clusters

To identify polyketide synthase gene clusters from the *M. fijiensis* genome as well as genomes of 74 other Dothideomycete fungi (Table S2)[199,200], the SMURF (Secondary Metabolites Unique Region Finder) tool from J. Craig Venter Institute was used. SMURF identifies PKS and NRPS genes from user-provided protein sequences in FASTA format, and predicts which genes may be a part of the biosynthetic cluster based on genome coordinate information[47].

For each PKS or hybrid PKS-NRPS gene predicted in *M. fijiensis*, a blastp search was done against the non-redundant protein sequences in the NCBI database. NCBI's Conserved Domain Database[168] was used to predict the domains of each PKS or PKS-like protein sequence (Figure 2). Open reading frames near each PKS locus were identified using the JGI *M. fijiensis* genome browser. Blastp searches were done of each of these open reading frames to predict possible functions and compare them to types of genes commonly found in PKS gene clusters.

Identifying Homologs of *M. fijiensis* PKS Genes

A phylogenetic tree was created with all of the PKS protein sequences from *M. fijiensis*, as well as the protein sequences of some well-characterized PKS enzymes from other species. Sequences were aligned using the MUSCLE algorithm v3.8.31[201] in Mesquite v3.04[202] and ModelGenerator v0.85[203] was used to identify the best evolutionary model. The best evolutionary model was predicted by both the Akaike Information Criterion and the Bayesian Information Criterion to be LG+I+G+F[204]. Therefore, this model was used with RaxmlGUI v1.3.1[205] to generate the phylogenetic tree, using maximum likelihood with slow bootstrap, no outgroup, and the autoMRE function.

Repeat masked assembly scaffolds were downloaded for 93 Dothideomycete fungal genomes available on JGI, as well as 10 additional Mycosphaerellaceae genomes available on NCBI[45]. BLAST+[206] was used to create a local blast database of these Dothideomycete genome sequences, and this database was searched for hits of the *M. fijiensis* PKS sequences,

using the tblastn algorithm. Results were sorted by bitscore and color coded by percent sequence similarity (Table 2).

Ketosynthase Domain Alignment and Prediction of Iteration Number

KS domain sequences were identified using the University of Maryland PKS/NRPS Analysis Web Server[182], for each of the *M. fijiensis* PKS protein sequences and the other PKS protein sequences used for creating the RAxML phylogeny. These were aligned using MUSCLE v3.8.31 in Mesquite v3.04, along with the 1kas fatty acid synthesis enzyme sequence (Accession 13BN_A). Amino acid residues corresponding to positions 229 and 400 in 1kas were determined for each sequence.

Fungal Cultures

Isolate 10CR1-24 was obtained from infected banana leaves collected from a commercial banana plantation in Guapiles, Costa Rica. Single ascospore isolations were done based on the protocol kindly provided by Dr. Miguel Muñoz, Dole Food Company (Personal Communication). Regions of infected leaves containing pseudothecia were cut into 2 cm squares, stapled to paper and submerged in sterile water for 10 minutes at room temperature to allow pseudothecia to hydrate. Leaf squares were attached to the lids of petri dishes suspended above 1% water agar, and incubated for 60 minutes at room temperature to allow ascospores to be released. Presence of ascospores on the water agar surface was confirmed using a dissecting microscope, and ascospores were recovered by pipetting into sterile water. The resulting spore suspension was transferred to a new Potato Dextrose Agar (PDA) (BD

Difco) or Mycophil agar (BD Difco) plate, and a cell spreader was used to spread the ascospores across the plate. Plates were sealed with parafilm and incubated at 25 °C to allow colonies to grow. Isolate 14H1-11A was obtained using the same method and was kindly provided by Jean Ristaino (North Carolina State University). The species of each resulting colony was confirmed by ITS sequencing followed by BLAST: ITS sequences of 10CR1-24 and 14H1-11A have 100% identity with *M. fijiensis* isolate UQ H444 (Accession AY923762.1).

Conidia were obtained from isolates using the protocol of Peraza-Echeverria et al[207]. Briefly, mycelial cultures of *M. fijiensis* isolates 10CR1-24 and 14H1-11A grown on PDA medium were macerated in water, and 2 mL of the resulting suspension was pipetted onto plates of modified V8 medium (0.2g/L $CaCO_3$, 100 mL/L V8 juice and 20g/L Difco agar). Cultures were incubated at 18 °C under continuous, cool-white fluorescent and black light. After 5-6 days, sporulation plates were stimulated to produce conidia by adding 2 mL water and brushing the plates with a paint brush or cell spreader, and removing the resulting suspension. After another 5-6 days, conidia were harvested in the same way, adding 2 mL water or 0.5% Tween 20 solution, brushing the plates to dislodge spores, and removing the spore suspension by pipetting.

Banana Tissue Culture and Inoculation

Grand Nain banana tissue culture plants (kindly provided by Miguel Muñoz, Dole Food Company) were maintained on modified Murashige and Skoog medium[208]. Growth medium was prepared with 4.33 g/L Murashige and Skoog basal salts (Caisson labs), 30g/L

sucrose, 200 ug/L glycine, 50 ug/L niacin, 50 ug/L pyridoxine, 10 ug/L thiamin, 10 ug/L myo-inositol, 1 ug/L cysteine, 2 g/L Phytagel (Sigma-Aldrich), with or without 4.5 mg/L 6-benzylaminopurine (BAP), with the pH adjusted to 5.8. Medium with BAP is used for bud proliferation, while medium without BAP is used for rooting. Plants were maintained on an 18h light/6h dark photoperiod with cool white fluorescent light at 25-30 °C.

Growth Conditions for Semi-quantitative RT-PCR and RNA-Seq Assays

For semi-quantitative RT-PCR assays, in vitro-cultured banana plants grown on medium without BAP were used for inoculations. A mix of conidia and mycelial fragments harvested from V8 sporulation plates of isolate 10CR1-24 described above were applied as 10 µL droplets onto the leaves. Plants were maintained under an 18h light/6h dark photoperiod before harvesting lesions at 5 weeks post-inoculation by cutting out the lesions and flash-freezing them in liquid nitrogen. For growth in culture, macerated mycelium of 10CR1-24 harvested from PDA plates was grown in 100 mL of PDB in 250 mL flasks. Flasks were incubated at 28 °C in an incubator shaker at 250 rpm for 1 week in the dark. Tissue was harvested by filtering through Miracloth (Millipore), blotting the tissue dry, and flash-freezing in liquid nitrogen. For both infected leaf tissue and mycelial tissue, tissue was ground using a mortar and pestle, RNA was extracted using the Spectrum Plant Total RNA kit, and samples were DNase treated with TURBO DNase (Ambion). RNA quality was verified by gel electrophoresis with 300 ng RNA as estimated by Nanodrop (Thermo Scientific). cDNA was synthesized using iScript Select cDNA synthesis kit (Bio-Rad). PCR

reactions used ExTaq (TaKaRa) according to manufacturer's instructions, with primers, annealing temperatures, extension times, and cycle numbers as indicated in Table S6.

For RNA-Seq analysis, rooted in vitro-cultured banana plants were transferred to soil and grown in the greenhouse until plants were approximately 20 cm in height, after which they were moved to an incubator at 25 °C and under an18h light/6h dark photoperiod, with cool white light. Conidia of isolate 14H1-11A were diluted in sterile 0.5% Tween 20 to a final concentration of $5.2x10^4$/mL, and 25 mL of the conidia suspension was sprayed onto each plant. For the first week post-inoculation, plants were covered in clear plastic bags to maintain high humidity conditions conducive to infection. After 6 weeks post-inoculation, symptomatic leaves were harvested by cutting out leaf areas with lesions, and flash freezing in liquid nitrogen. For mycelial cultures for RNA-Seq, 10 µL of $1.3x10^6$/mL conidia of 14H1-11A were used to inoculate 50 mL PDB in 125 mL flasks. Flasks were incubated on a rotary shaker at 150 rpm at 25 °C in the dark. After 1 week, mycelium was harvested by filtering through Miracloth and flash freezing in liquid nitrogen. RNA was extracted from the infected leaf and mycelial tissue and purity was confirmed as described above for the semi-quantitative RT-PCR experiment except that RNA samples were DNAse treated using DNase I (Roche).

cDNA Library Construction and Illumina HiSeq Sequencing

Total RNA was submitted to the North Carolina State University Genomic Sciences Laboratory for sequencing. Quality of RNA was further confirmed using an Agilent Bioanalyzer, and then strand-specific libraries were created using the NEBNext Ultra

Directional library prep kit (New England BioLabs). Sequencing was done using an Illumina HiSeq 2500 platform to generate 125-base single-end reads. Sequencing resulted in an average yield of 32 million reads per sample. RNA-Seq data are available through NCBI through SRP075820.

Identification of Differentially Expressed Genes

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to verify quality of the RNA-Seq reads for each sample. Illumina Truseq adapter sequences and low-quality bases were trimmed using CutAdapt v1.7 with a quality cutoff of 20 and a minimum sequence length of 36[209].

Sequences from each sample were mapped to the banana genome, *Musa acuminata* subsp. *malaccensis* double-haploid Pahang[210], and to the *M. fijiensis* genome[55] using Tophat v2.0.9[211]. Gene expression levels were determined using HTSeq v0.6.0[212], with gene annotations available from JGI. DESeq2 v1.4.5 was used to identify differentially expressed genes[213].

Acknowledgements

We thank Dr. Sonia Herrero, Syngenta Biotechnology, for consultation on gene cluster identification, and Dr. Miguel Muñoz, Dole Food Company, for providing tissue-cultured banana plants and for the protocol for single ascospore isolations. We also thank Dr. Jean Ristaino from North Carolina State University for providing infected leaf material

collections, and Suchitra Chavan and Jessica Pope from North Carolina State University with their help with single ascospore isolations of *M. fijiensis*.

Supplemental Files

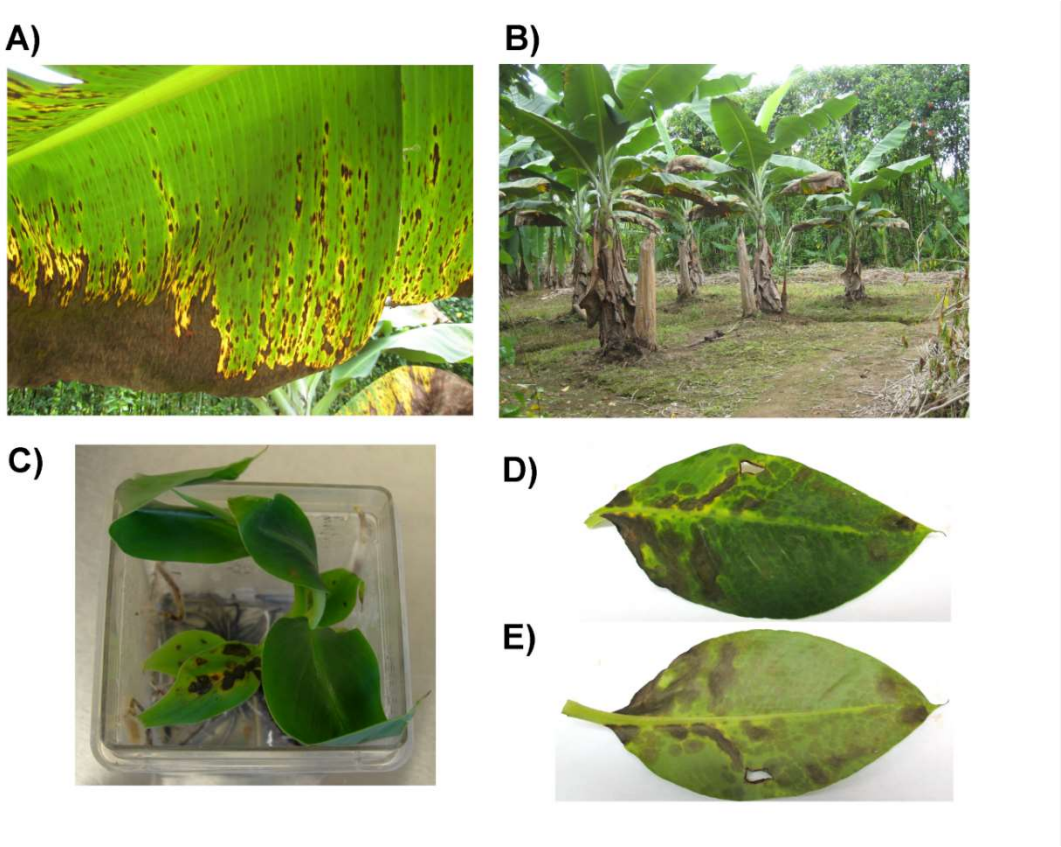


Figure S1. Examples of black Sigatoka symptoms.
A) Infected leaf on banana plant in Costa Rica B) Infected banana plants in Costa Rica banana plantation C) Example of infected tissue culture plant D) Example of infected leaf harvested for transcriptome analysis, adaxial side E) Abaxial side.

Table S1. Accession numbers for PKS sequences.
Accession numbers or JGI protein ID numbers for *M. fijiensis* PKS protein sequences, and well-characterized PKS protein sequences from other species. Names of PKS proteins are indicated as well as the polyketide product produced.

Species	Polyketide product	Protein name	Citation	Accession number	JGI protein ID
Mycosphaerella fijiensis	-	PKS2-1		XP_007922357.1	29867
Mycosphaerella fijiensis	-	PKS7-1		XP_007928748.1 (partial sequence)	141883 (full sequence)
Mycosphaerella fijiensis	-	PKS8-1		XP_007929626.1	34361
Mycosphaerella fijiensis	-	PKS8-2		XP_007929903.1	166988
Mycosphaerella fijiensis	-	Hybrid8-3		XP_007930181.1	93879
Mycosphaerella fijiensis	-	PKS8-4		XP_007930394.1	190533
Mycosphaerella fijiensis	-	PKS10-1		XP_007931490.1	216850
Mycosphaerella fijiensis	-	PKS10-2		XP_007931278.1	44966
Bipolaris oryzae	Melanin	PKS1	[214]	BAD22832.1	-
Colletotrichum graminicola	Melanin	PKS1	[215]	ACN32207.2	-
Exophiala dermatitidis	Melanin	PKS1	[216]	AAD31436.3	-
Glarea lozoyensis	Melanin	PKS1	[217]	AAN59953.1	-
Cercospora nicotianae	Cercosporin	CTB1	[57]	AAT69682.1	-
Cercospora zeae-maydis	Cercosporin	CTB1		-	42949
Aspergillus parasiticus	Aflatoxin	PKSL1	[218]	Q12053.1	-
Fusarium fujikuroi	Bikaverin	PKS4	[219]	CCT67991.1	-
Penicillium citrinum	Compactin	MlcA	[220]	BAC20564.1	-
Aspergillus terreus	Lovastatin	LNKS	[221]	Q9Y8A5.1	-
Aspergillus nidulans	Sterigmatocystin	PKSst	[222]	XP_681094.1	-
Fusarium verticillioides	Fumonisin	Fum1p	[173]	AAD43562.2	-
Alternaria solani	Alternapyrone	PKSN	[170]	BAD83684.1	-
Alternaria solani	Solanapyrone	Sol1	[178]	BAJ09789.1	-
Bipolaris maydis	T-toxin	PKS1	[223]	AAB08104.3	-
Penicillium griseofulvum	6-methylsalicylic acid	MSAS	[224]	CAA39295.1	-
Aspergillus nidulans	Naphthopyrone YWA1	WA	[225]	CAA46695.2	-

Table S2. Number of PKS genes identified in genomes of Dothideomycete fungi.
For each species, the number of PKS genes identified by SMURF is indicated along with the name of each species' taxonomic order.

Order	Species	Number of PKS genes predicted
Asterinales	*Aulographum hederae*	5
Botryosphaeriales	*Saccharata proteae*	3
Botryosphaeriales	*Neofusicoccum parvum*	15
Botryosphaeriales	*Macrophomina phaseolina*	18
Botryosphaeriales	*Aplosporella prunicola*	26
Botryosphaeriales	*Botryosphaeria dothidea*	26
Capnodiales	*Acidomyces richmondensis*	2
Capnodiales	*Baudoinia compniacensis*	2
Capnodiales	*Piedraia hortae*	2
Capnodiales	*Dothistroma septosporum*	4
Capnodiales	*Teratosphaeria nubilosa*	6
Capnodiales	*Cladosporium fulvum*	7
Capnodiales	*Mycosphaerella fijiensis*	7
Capnodiales	*Mycosphaerella musicola*	7
Capnodiales	*Zymoseptoria ardabiliae*	7
Capnodiales	*Polychaeton citri*	8
Capnodiales	*Mycosphaerella eumusae*	9
Capnodiales	*Dissoconium aciculare*	10
Capnodiales	*Septoria populicola*	10
Capnodiales	*Zymoseptoria pseudotritici*	10
Capnodiales	*Cercospora zeae-maydis*	11
Capnodiales	*Mycosphaerella graminicola*	12
Capnodiales	*Pseudovirgaria hyperparasitica*	18
Capnodiales	*Septoria musiva*	19
Capnodiales	*Zasmidium cellare*	22
Dothideales	*Aureobasidium pullulans var. melanogenum*	4
Dothideales	*Aureobasidium pullulans var. namibiae*	4
Dothideales	*Aureobasidium pullulans var. pullulans*	6
Dothideales	*Aureobasidium pullulans var. subglaciale*	11
Hysteriales	*Hysterium pulicare*	23
Microthyriales	*Tothia fuscella*	6
Myriangiales	*Myriangium duriaei*	9
Mytilinidiales	*Lophium mytilinum*	7
Mytilinidiales	*Cenococcum geophilum*	10
Patellariales	*Patellaria atrata*	8
Pleosporales	*Trichodelitschia bisporula*	1
Pleosporales	*Alternaria brassicicola*	6
Pleosporales	*Macroventuria anomochaeta*	7
Pleosporales	*Sporormia fimetaria*	7

Table S2 (continued)

Order	Species	Number of PKS genes predicted
Pleosporales	*Didymella exigua*	8
Pleosporales	*Clathrospora elynae*	10
Pleosporales	*Alternaria alternata*	11
Pleosporales	*Ophiobolus disseminans*	11
Pleosporales	*Cucurbitaria berberidis*	12
Pleosporales	*Leptosphaeria maculans*	12
Pleosporales	*Pleomassaria siparia*	12
Pleosporales	*Stagonospora nodorum SN15*	12
Pleosporales	*Delitschia confertaspora*	13
Pleosporales	*Dothidotthia symphoricarpi*	14
Pleosporales	*Massarina eburnea*	14
Pleosporales	*Cochliobolus lunatus*	15
Pleosporales	*Karstenula rhodostoma*	15
Pleosporales	*Phoma tracheiphila*	15
Pleosporales	*Polyplosphaeria fusca*	16
Pleosporales	*Westerdykella ornata*	16
Pleosporales	*Cochliobolus sativus*	17
Pleosporales	*Stagonospora sp. SRC1lsM3a*	17
Pleosporales	*Verruculina enalia*	17
Pleosporales	*Cochliobolus victoriae*	18
Pleosporales	*Trematosphaeria pertusa*	18
Pleosporales	*Aaosphaeria arxii*	19
Pleosporales	*Amniculicola lignicola*	19
Pleosporales	*Lophiostoma macrostomum*	19
Pleosporales	*Pyrenophora teres f. teres*	20
Pleosporales	*Cochliobolus heterostrophus C4*	23
Pleosporales	*Setosphaeria turcica*	23
Pleosporales	*Zopfia rhizophila*	27
Pleosporales	*Lentithecium fluviatile*	28
Pleosporales	*Byssothecium circinans*	29
Pleosporales	*Corynespora cassiicola*	29
Pleosporales	*Lindgomyces ingoldianus*	31
Pleosporales	*Cochliobolus heterostrophus C5*	33
Pleosporales	*Melanomma pulvis-pyrius*	33
Trypetheliales	*Trypethelium eluteriae*	37
Venturiales	*Venturia inaequalis*	6

Table S3. E-values for domains in each *M. fijiensis* PKS or hybrid enzyme.
Blastp with the Conserved Domain Database from NCBI was used to predict enzyme domains in each PKS or PKS-NRPS enzyme. E-values are shown for each domain predicted. A) *M. fijiensis* PKS enzymes. B) *M. fijiensis* hybrid PKS-NRPS enzyme. Abbreviations for domains: SAT = starter unit acyltransferase; KS = ketosynthase; AT = acyltransferase; PT = product template; DH = dehydratase; MT = methyltransferase; ER = enoyl reductase; KR = ketoreductase; PP 1 = first phosphopantetheine attachment site; PP 2 = second phosphopantetheine attachment site; TE/TR = thioesterase or thioester reductase; C = condensation; Hx = HxxPF repeat domain; A = adenylation. The Conserved Domain Database also identifies binding sites. Shown are the presence (NAD(P)+) or absence (NAD(P)-) of an NAD(P) binding site in the ER, KR, and TE/TR domains; presence (SAM+) or absence (SAM-) of a SAM binding site in the MT domain; and presence of an AMP binding site (AMP +) and acyl-activating enzyme consensus motif (Acyl-act+) in the A domain. Some PKS proteins such as those for melanin and cercosporin (CTB1) are known to have a SAT domain at the N terminus. Blastp searches were done against each *M. fijiensis* PKS or hybrid PKS-NRPS, using the region containing the SAT domain in CTB1. E-values for hits for this search are shown in red text.

A)

	SAT	KS	AT	PT	DH	MT	ER	KR	PP 1	PP 2	TE/TR
PKS2-1	N/A	1.27e-138 Active site+	1.30e-102	N/A	5.53e-42	N/A	3.81e-36 NAD(P) -	4.99e-66 Active site+ NAD(P) +	6.76e-07	N/A	N/A
PKS7-1	N/A	3.27e-110 Active site+	3.85e-62	4.21e-108	N/A	N/A	N/A	N/A	4.30e-11	3.26e-06	4.26e-22 Active site -
PKS8-1	7e-16	1.71e-117 Active site+	2.57e-96	1.11e-123	N/A	N/A	N/A	N/A	6.72e-11	N/A	N/A
PKS8-2	N/A	6.14e-124 Active site+	1.56e-88	N/A	7.93e-58	2.91e-04 SAM +	1.33e-140 NAD(P) +	7.09e-58	4.38e-06	N/A	N/A
PKS8-4	N/A	2.75e-140 Active site+	1.16e-82	N/A	8.47e-37	1.79e-07 SAM +	N/A	7.41e-52 Active site + NAD(P) +	6.5e-04	N/A	4.86e-24 Active site +
PKS10-1	1.41e-04 3e-29	2.56e-119 Active site+	1.18e-97	6.26e-136	N/A	N/A	N/A	N/A	9.33e-11	1.35e-13	1.75e-23
PKS10-2	N/A	2.56e-119 Active site+	1.64e-98	N/A	1.57e-49	5.42e-08 SAM +	2.22e-05 NAD(P) -	2.63e-56 Active site + NAD(P) -	9.69e-10	N/A	N/A

B)

	KS	AT	DH	MT	KR	PP 1	C	Hx	A	PP 2	TE/TR
Hybrid8-3	2.32e-141 Active site+	1.18e-106	5.23e-44	9.99e-04 SAM +	1.86e-49 Active site - NAD(P) -	8.98e-08	2.85e-43	3.8e-10	3.49e-112 AMP + Acyl-act +	8.54e-06	6.45e-34 Active site + NAD(P) +

Table S4. Sequence similarity for proteins with similar functions encoded by polyketide biosynthetic clusters.
A) *M. fijiensis* PKS8-2 cluster compared to fumonisin biosynthetic cluster from *F. verticillioides*; B) *M. fijiensis* PKS10-2 cluster compared to solanapyrone biosynthetic cluster from *A. solani*.

A)

Gene	***M. fijiensis* accession**	***F. verticillioides* accession**	***F. verticillioides* gene name**	**Protein sequence similarity**
PKS	XP_007929903.1	AAD43562.2	*FUM1*	59%
Transcription factor	XP_007929904.1	ABQ95367.1	*FUM21*	66%
Transporter	XP_007929664.1	AAN74822.1	*FUM19*	51%
α-oxoamine synthase	XP_007929637.1	ADQ39012.1	*FUM8*	43%

B)

Gene	***M. fijiensis* accession**	***A. solani* accession**	***A. solani* gene name**	**Protein sequence similarity**
PKS	XP_007931278.1	BAJ09789.1	*sol1*	52%
Dehydrogenase	XP_007928748.1	BAJ09787.1	*sol3*	45%
Cytochrome P450	XP_007931281.1	BAJ09784.1	*sol6*	40%
Cytochrome P450	XP_007931280.1	BAJ09784.1	*sol6*	No hits

Table S5. Amino acid residues for tertiary structure analysis of ketosynthase domains. Amino acid residues for *M. fijiensis* PKS sequences and sequences from well-characterized PKS proteins, corresponding to positions 229 and 400 in the fatty acid synthase 1kas. For well-characterized PKS enzymes, the number of iterations catalyzed is shown. Also indicated is whether each PKS is more closely related to the PKS producing 6-methylsalicylic acid (MSAS), naphthopyrone (NAP), or T-toxin (T-tox) (Figure 4), and the number of iterations catalyzed by members of this clade, if known. These are color coded yellow, blue, or pink, respectively. For *M. fijiensis* PKS sequences in the T-toxin PKS clade, well-characterized PKS enzymes with identical residues at positions 229 and 400 are indicated, along with the number of iterations catalyzed by those PKS enzymes.

PKS	229 (1kas)	400 (1kas)	Closest PKS by homology	Actual number of iterations
MSAS *P. griseofulvum*	Y	Y	MSAS (3 iterations)	3
NAP *A. nidulans*	Y	A	NAP (5-8 iterations)	5
Melanin *B. oryzae*	Y	A	NAP (5-8 iterations)	5
Melanin *C. graminicola*	Y	A	NAP (5-8 iterations)	5
Melanin *E. dermatitidis*	Y	A	NAP (5-8 iterations)	5
Melanin *G. lozoyensis*	Y	A	NAP (5-8 iterations)	5
Aflatoxin *A. parasiticus*	Y	A	NAP (5-8 iterations)	7
Sterigmatocystin *A. nidulans*	Y	A	NAP (5-8 iterations)	7
Cercosporin *C. nicotianae*	Y	A	NAP (5-8 iterations)	7
Cercosporin *C. zeae-maydis*	Y	A	NAP (5-8 iterations)	7
Bikaverin *F. fujikuroi*	Y	A	NAP (5-8 iterations)	8
PKS7-1	Y	A	NAP (5-8 iterations)	Unknown
PKS8-1	Y	A	NAP (5-8 iterations)	Unknown
PKS10-1	Y	A	NAP (5-8 iterations)	Unknown

Table S5 (continued)

PKS	229 (1kas)	400 (1kas)	Closest PKS by homology	Actual number of iterations
T-toxin *B. maydis*	Y	Y	T-tox	20
Compactin *P. citrinum*	Y	F	T-tox	8
Lovastatin *A. terreus*	Y	F	T-tox	8
Fumonisin *F. verticillioides*	Y	I	T-tox	9
Alternapyrone *A. solani*	F	F	T-tox	10
Solanapyrone *A. solani*	Y	Y	T-tox	8
PKS2-1	Y	Y	T-tox	Unknown Same residues as T-toxin and solanapyrone PKS proteins (20 and 8 iterations, respectively)
PKS8-2	Y	I	T-tox	Unknown Same residues as fumonisin PKS (9 iterations)
Hybrid8-3	Y	F	T-tox	Unknown Same residues as compactin and lovastatin PKS proteins (both catalyze 8 iterations)
PKS8-4	Y	F	T-tox	Unknown Same residues as compactin and lovastatin PKS proteins (both catalyze 8 iterations)
PKS10-2	Y	Y	T-tox	Unknown Same residues as T-toxin and solanapyrone PKS proteins (20 and 8 iterations, respectively)

Table S6. Primer sets and conditions for semi-quantitative RT-PCR assays.
For each RT-PCR assay, primer names, sequences, annealing temperature (ann temp), extension time (ext time), number of cycles (cyc), and expected product sizes are indicated. *Same product size expected for cDNA and gDNA products.

Gene	Forward primer sequence	Reverse primer sequence	Ann temp	Ext time	Cyc	Size - cDNA (bp)	Size - gDNA (bp)
Beta-tubulin	cagctcgagcgcatgaacg	ggtgcgaaaccgaccatgaag	59	1:00	40	745*	745*
PKS2-1	atggcagtcaataccgatga	gcttcgtatgagacttccag	53	0:45	45	344	450
PKS7-1	gccgtctatgactattgaca	tggctgttgattacgctctt	52	1:30	45	660	774
PKS8-1	caggacgcatcaactacttc	ctcggcggagtggttagttc	54	0:45	45	362	511
PKS8-2	caggaagattgacgaaaggc	catagtgttggatcatgtcg	52	0:45	45	524	574
Hybrid8-3	ctcgccgaacttgatggaga	tacaggcatcggaacgacgagg	59	1:30	45	495	585
PKS8-4	tatgctctcacctgacgg	agattcatagctctcgat	47	1:30	45	542	886
PKS10-1	ttctcatacacgaccactgc	gcaagcaatctcggtcatcttg	55	0:45	45	603*	603*
PKS10-2	cttcgtgtttcatagggaac	catcaacttcaatcggatcg	51	1:30	45	615	809

CHAPTER 3

Transcriptome Sequencing of *Mycosphaerella fijiensis* During Association with *Musa acuminata* Reveals Candidate Pathogenicity Genes

Roslyn D. Noar and Margaret E. Daub

Noar, R. D., & Daub, M.E. (2016). Transcriptome Sequencing of *Mycosphaerella fijiensis* during Association with *Musa acuminata* Reveals Candidate Pathogenicity Genes. BMC Genomics 17:690.

http://doi.org/10.1186/s12864-016-3031-5

Abstract

Background: *Mycosphaerella fijiensis*, causative agent of the black Sigatoka disease of banana, is considered the most economically damaging banana disease. Despite its importance, the genetics of pathogenicity are poorly understood. Previous studies have characterized polyketide pathways with possible roles in pathogenicity. To identify additional candidate pathogenicity genes, we compared the transcriptome of this fungus during the necrotrophic phase of infection with that during saprophytic growth in medium.

Results: Transcriptome analysis was conducted, and the functions of differentially expressed genes were predicted by identifying conserved domains, Gene Ontology (GO) annotation and GO enrichment analysis, Carbohydrate-Active EnZymes (CAZy) annotation, and identification of genes encoding effector-like proteins. The analysis showed that genes

commonly involved in secondary metabolism have higher expression in infected leaf tissue, including genes encoding cytochrome P450s, short-chain dehydrogenases, and oxidoreductases in the 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily. Other pathogenicity-related genes with higher expression in infected leaf tissue include genes encoding salicylate hydroxylase-like proteins, hydrophobic surface binding proteins, CFEM domain-containing proteins, and genes encoding secreted cysteine-rich proteins characteristic of effectors. More genes encoding amino acid transporters, oligopeptide transporters, peptidases, proteases, proteinases, sugar transporters, and proteins containing Domain of Unknown Function (DUF) 3328 had higher expression in infected leaf tissue, while more genes encoding inhibitors of peptidases and proteinases had higher expression in medium. Sixteen gene clusters with higher expression in leaf tissue were identified including clusters for the synthesis of a non-ribosomal peptide. A cluster encoding a novel fusicoccane was also identified. Two putative dispensable scaffolds were identified with a large proportion of genes with higher expression in infected leaf tissue, suggesting that they may play a role in pathogenicity. For two other scaffolds, no transcripts were detected in either condition, and PCR assays support the hypothesis that at least one of these scaffolds corresponds to a dispensable chromosome that is not required for survival or pathogenicity.

Conclusions: Our study revealed major changes in the transcriptome of *Mycosphaerella fijiensis*, when associating with its host compared to during saprophytic growth in medium. This analysis identified putative pathogenicity genes and also provides support for the existence of dispensable chromosomes in this fungus.

Keywords

Mycosphaerella fijiensis, black Sigatoka, transcriptome, effectors, secondary metabolism, non-ribosomal peptide synthase, fusicoccane, Domain of Unknown Function 3328, salicylate hydroxylase, dispensable chromosome

Background

Banana (including common "dessert" bananas as well as cooking and plantain types) is one of the world's most important food crops, grown in tropical and subtropical regions in over 120 countries[5]. Only about 10-15% of bananas are grown for export, with the rest serving as an important subsistence crop in many developing countries[6,7]. Black Sigatoka, caused by the ascomycete fungus *Mycosphaerella fijiensis*, is a major threat to banana production. It is found in almost all banana-growing countries, and can cause up to 50% yield loss as well as premature ripening of fruit[6]. Control of the disease is through frequent applications of fungicides, which are estimated to account for 25-30% of the total banana production cost[5–7,18]. Fungicide resistance is an ongoing problem that threatens the viability of this method of black Sigatoka control[7,226].

M. fijiensis produces both conidia and ascospores, both of which can infect banana leaves via the stomata[5]. The fungus is a hemibiotroph. The conidia or ascospores germinate, forming mycelium that initially grows epiphytically on the leaf surface prior to penetration through the stomata and into the leaf[5]. The fungus colonizes the intercellular spaces of the leaf during its biotrophic phase[5]. The fungus then switches to a necrotrophic stage, leading to the death of leaf cells and the formation of necrotic leaf lesions[5].

As a hemibiotroph, *M. fijiensis* would be expected to produce both effectors to suppress host defense responses and prevent death of host cells during biotrophy[108], and toxic secondary metabolites and proteins to kill host tissue during necrotrophy[227]. However, little is known about the repertoire of effectors and toxic metabolites produced by *M. fijiensis* during its association with banana. Homologs of the *Cladosporium fulvum* effectors Ecp2, Ecp6, and Avr4 have been identified from the *M. fijiensis* genome[102,104]. However, most fungal pathogens have a large repertoire of effectors, and most effectors have a restricted phylogenetic distribution, so *M. fijiensis* is likely to produce many other effectors[108]. In addition to effectors, several studies have been done to identify toxins secreted by *M. fijiensis*. Several phytotoxic metabolites have been identified from *M. fijiensis* including 2,4,8-trihydroxytetralone, which showed some host selectivity and was thought to be an important pathogenicity factor[60,61,228]. However, 2,4,8-trihydroxytetralone is a melanin shunt metabolite[156], and disruption of the melanin biosynthetic pathway was shown to have no effect on pathogenicity[5]. Phytotoxic activity has also been identified from the hydrophilic portion of culture filtrates, but the identity of these toxins is unknown[88,89]. All of these studies were done using mycelium grown in culture conditions, which may not fully reflect what is produced during the association of *M. fijiensis* with its host.

In previous work we used the publicly available *M. fijiensis* genome sequence (NCBI Genome ID 10962)[160], obtained from isolate CIRAD86, to predict the capacity of *M. fijiensis* to produce polyketides[45], an important class of secondary metabolites that are used as pathogenicity factors by closely related fungi[57,154]. In this study, seven putative polyketide synthase gene clusters and one hybrid polyketide synthase/non-ribosomal peptide synthase

gene cluster were identified[45]. Among the clusters were ones with similarity to clusters producing melanin, as well as the secondary metabolites fumonisin, solanapyrone, and alternapyrone produced by *Alternaria* and *Fusarium* species[45]. Melanin has been shown to play important roles in fungal pathogenicity of plants including penetration into host tissue[78,79]. Fumonisin promotes *Fusarium* spp. pathogenicity by perturbing sphingolipid biosynthesis in the host[194,197].

The publicly available *M. fijiensis* genome sequence has also been used to investigate possible dispensable chromosomes. Many fungi use genes located on conditionally dispensable chromosomes to assist in pathogenicity, host specificity, and other functions that are useful but not required for survival[48]. Ohm et al. observed that the CIRAD86 *M. fijiensis* genome contains 14 scaffolds that are very different from the rest of the scaffolds in the genome: they are small, have a low G+C content, have the lowest gene density and the lowest proportion of genes encoding proteins with PFAM domains, have the highest proportion of repetitive DNA, and have different codon usage[54]. Though it has not been proven that these 14 scaffolds represent dispensable chromosomes, they share their unusual characteristics with dispensable chromosomes from the related species *Mycosphaerella graminicola*[54,55].

Next-generation transcriptome sequencing has greatly improved our understanding of the genetic mechanisms of pathogenicity in other species[229,230]. For *M. fijiensis*, it can help identify genes encoding effectors, secondary metabolite pathways, and other proteins that may be important for pathogenicity. It can also identify portions of the genome with an abundance of genes that are expressed during association with banana, and thereby suggest

which putative dispensable chromosomes may be important for pathogenesis. To date, however, research on changes in gene expression during the *M. fijiensis*-*Musa* spp. (banana) interaction has largely been limited to the *Musa* spp. transcriptome. Portal et al. created suppression subtractive hybridization cDNA libraries from late stages of infection to identify expressed genes from banana and *M. fijiensis*[41]. They identified banana genes involved in biosynthesis of phenyl-propanoids, jasmonic acid and ethylene, genes encoding pathogenesis-related (PR) proteins, and genes involved in detoxification such as glutathione S-transferases[41]. Although many defense-related banana genes were identified, the inefficiency of CTAB-based RNA extraction protocols with *M. fijiensis* had not been reported at the time[41]. As a result, the only fungal gene identified from their libraries was a gene for UDP glucose pyrophosphorylase, which is involved in trehalose biosynthesis[41]. Another study used microarray analysis to compare genes expressed in the resistant banana variety Calcutta 4 versus the susceptible variety, Williams, when challenged with *M. fijiensis*[39]. Banana genes encoding peroxidase, PR-4, PR-10, phenylalanine ammonia lyase, and disease resistance response 1 showed higher expression in Calcutta 4 compared to Williams, between 6 and 24 hours after inoculation[39].

Studies of *M. fijiensis* gene expression have been more limited. Expressed sequence tags have been identified from *M. fijiensis* grown in three different culture media: Potato Dextrose Agar (PDA), Fries liquid medium, and Fries liquid medium with banana leaf extract added[44]. This analysis found a homolog of the *Cladosporium fulvum* effector gene Avr4 in all three libraries, and a homolog of the *C. fulvum* effector gene Ecp2 in both the Fries liquid medium and medium supplemented with leaf extract[44]. In our study on *M.*

fijiensis polyketide synthases, we used transcriptome sequencing to compare expression of the polyketide synthase gene clusters in infected banana leaf tissue relative to in culture medium[45]. The genes in the previously mentioned polyketide synthase clusters with similarities to fumonisin and solanapyrone clusters had increased expression in infected leaf tissue, suggesting that these gene clusters may produce polyketide products that are important for pathogenicity[45]. By contrast, the melanin cluster had lower expression in infected leaf tissue as compared to expression in culture medium[45]. This was the first study in which transcriptome sequencing was used to analyze expression of *M. fijiensis* during its association with banana. However, our analysis was limited to polyketide synthase gene clusters.

Other transcriptome sequencing studies have been done with banana infected with the related banana pathogens *Mycosphaerella musicola* and *Mycosphaerella eumusae*. One study compared the resistant banana variety Calcutta 4 to the susceptible Grand Nain variety during association with *M. musicola*[42]. A homolog of the *C. fulvum* effector Ecp6 gene[104] was shown to have higher expression in *M. musicola* during association with Calcutta 4 relative to Grand Nain, whereas genes encoding a SAP family cell cycle dependent phosphate-associated protein, two Hsp70 family proteins, an FAD binding domain protein, and a calcium channel all had higher expression in *M. musicola* in the association with Grand Nain compared to Calcutta 4[42]. Another study compared the transcriptomes of resistant (Manoranjitham) and susceptible (Grand Nain) banana varieties challenged or unchallenged with *M. eumusae*[43]. Banana genes with higher expression in the resistant compared to susceptible banana variety included those encoding enzymes involved in the

phenylpropanoid pathway, abscisic acid biosynthesis, alkaloid biosynthesis, and scavenging of reactive oxygen species[43].

The limited information on pathogenicity-related genes in *M. fijiensis* and other *Mycosphaerella* banana pathogens is a barrier to the development of new control methods for this devastating disease. Thus, the objectives of this paper were to expand beyond our focus on polyketides to use transcriptome sequencing data from symptomatic leaf tissue and mycelium growing saprophytically in medium to predict other *M. fijiensis* genes that may have roles in pathogenicity. In addition, we were interested to determine if putative pathogenicity-related genes are concentrated in particular regions of the genome, such as on scaffolds predicted by Ohm et al. to represent dispensable chromosomes[54].

Results

Identification of Differentially Expressed Genes

Banana plants and Potato Dextrose Broth (PDB) medium were inoculated with *M. fijiensis* conidia. Tissue harvested from the fungus grown in liquid medium as well as from symptomatic leaf tissue were used for RNA isolation and transcriptome sequencing. Principal component analysis showed that the samples from the same treatment (leaf tissue, liquid medium) clustered together, separately from samples in the other treatment group (Figure S1). A total of 802 differentially expressed genes were identified (Figure S2)[46]; of these, 483 genes were more highly expressed in infected leaf tissue, and 319 genes were more highly expressed in culture medium[46].

To identify the genes with the greatest differential expression, differentially expressed genes were sorted based on their $\log_2$ fold change (log2FC) values. Lists of the most differentially expressed genes (the 20 with highest and 20 with lowest expression in infected leaf tissue compared to culture medium) can be seen in Tables 1 and 2, respectively. The majority of the genes (13 of the top 20) with higher expression in infected leaf tissue are predicted to encode hypothetical proteins, 11 of which have no conserved domains and two of which have a DUF (Domain of Unknown Function) 3328 domain (Table 1). Other than genes encoding hypothetical proteins, two of the genes most highly expressed in the banana leaf relative to culture medium encode proteins in a polyketide synthase gene cluster (*PKS7-1*) previously described[45]. Other types of genes with higher expression in infected leaf tissue include those encoding an oxidoreductase, a 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily enzyme, a Major Facilitator Superfamily (MFS) multidrug transporter-like protein, a peptidase, and a transcription factor (Table 1). Of genes with lower expression in infected leaf tissue, a majority (11 out of 20) were also predicted to encode hypothetical proteins (Table 2). One of these hypothetical proteins contains a CFEM (Common in Fungal Extracellular Membrane) domain, a cysteine-rich domain present in some proteins that play important roles in pathogenicity[231]. Genes encoding a cupredoxin, a copper transporter, a cysteine synthase, a 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily enzyme, a cytochrome P450, a heme peroxidase, an α/β-hydroxylase, an oxidoreductase, and a glutamine amidotransferase were also among the top 20 differentially expressed genes with reduced expression in infected leaf tissue (Table 2).

Table 1. Genes with highest expression in infected leaf tissue compared to culture medium as determined by log2FC.
Table indicates the JGI gene and protein IDs, the predicted function of the encoded protein based on blast and conserved domains, and the log2FC value.

Gene ID	Protein ID	Predicted function of encoded protein	Log2FC
fgenesh1_kg.9_#_12_#_4417424:1	183842	Hypothetical; 44% identity to WI-1 adhesin from *Blastomyces dermatitidis*	13.3
estExt_Genewise1Plus.C_90054	157089	Hypothetical with Domain of Unknown Function (DUF) 3328	11.4
Mycfi1.estExt_fgenesh1_pg.C_120010	87989	Hypothetical, no conserved domains	11.3
Mycfi1.fgenesh1_pg.C_scaffold_29000110	84397	Hypothetical, no conserved domains	11.2
Mycfi1.e_gw1.43.35.1	46458	FAD-dependent oxidoreductase	11.1
e_gw1.2.477.1	132918	2-oxoglutarate and Fe(II)-dependent oxygenase superfamily	10.9
fgenesh1_pm.1_#_1217	185508	MFS multidrug transporter-like protein	10.8
fgenesh1_pg.3_#_601	195588	Hypothetical, no conserved domains	10.6
Genemark.4551_g	173539	Hypothetical, no conserved domains	10.6
Mycfi1.gw1.34.54.1	20039	PKS7-1	10.5
fgenesh1_pm.4_#_231	188143	Zinc peptidase	10.5
estExt_Genemark.C_30183	207097	Hypothetical, no conserved domains	10.3
fgenesh1_pm.7_#_421	190048	Monooxygenase in PKS7-1 cluster	10.2
fgenesh1_pg.7_#_244	198484	Hypothetical, no conserved domains	10.0
Genemark.2108_g	171096	Hypothetical, no conserved domains	9.96
Genemark.4552_g	173540	Hypothetical, no conserved domains	9.82
fgenesh1_pg.9_#_17	199506	Hypothetical with DUF3328	9.76
Mycfi1.fgenesh1_pm.C_scaffold_23000030	65382	Hypothetical protein in cupin superfamily	9.69
Mycfi1.e_gw1.22.234.1	43729	Transcription factor	9.55
Mycfi1.fgenesh1_pg.C_scaffold_1001762	76887	Hypothetical, no conserved domains	9.37

Table 2. Genes with lowest expression in infected leaf tissue compared to culture medium as determined by log2FC.
Table indicates the JGI gene and protein IDs, the predicted function of the encoded protein based on blast and conserved domains, and the log2FC value.

Gene ID	Protein ID	Predicted function	Log2FC
fgenesh1_pm.2_#_327	186279	Hypothetical, no conserved domains	-7.1
estExt_Genewise1Plus.C_90903	157466	Cysteine synthase	-6.8
estExt_fgenesh1_kg.C_20418	210593	Hypothetical, no conserved domains	-6.8
gw1.2.3679.1	120453	Heme peroxidase	-6.7
estExt_fgenesh1_pm.C_60071	204084	Hypothetical, no conserved domains	-6.7
fgenesh1_kg.1_#_500_#_4410656:1	181367	Hypothetical, with some homology to phosphate carrier protein	-6.7
fgenesh1_pg.1_#_1546	193207	Hypothetical, no conserved domains	-6.4
fgenesh1_pm.2_#_99	186051	α/β-hydroxylase	-6.4
estExt_fgenesh1_pg.C_80340	216449	Hypothetical, with CFEM domain	-6.3
Mycfi1.e_gw1.1.178.1	26018	Cytochrome P450	-6.2
Mycfi1.estExt_fgenesh1_pg.C_160286	88531	Hypothetical, no conserved domains	-6.2
Mycfi1.e_gw1.3.802.1	31919	FAD dependent oxidoreductase	-6.1
Genemark.2479_g	171467	Hypothetical, no conserved domains	-6.1
estExt_Genemark.C_80180	208708	2-oxoglutarate and Fe(II)-dependent oxygenase superfamily	-6.1
estExt_fgenesh1_kg.C_80203	212291	Glutamine amidotransferase	-6.0
estExt_fgenesh1_kg.C_30187	210875	Copper transporter	-5.9
estExt_fgenesh1_kg.C_50238	211639	Hypothetical, no conserved domains	-5.8
estExt_fgenesh1_kg.C_20200	210380	Hypothetical, no conserved domains	-5.8
estExt_fgenesh1_kg.C_20513	210685	Hypothetical, no conserved domains	-5.8
Mycfi1.estExt_Genewise1.C_12242	48579	Cupredoxin	-5.8

Validation of RNA-Seq Results by RT-qPCR

To validate the identification of differentially expressed genes from the RNA-Seq dataset, four genes were chosen for further expression analysis using RT-qPCR on the same RNA samples. These genes were: a cytochrome P450, a proteinase, an oligopeptide transporter, and a ferric-chelate reductase. RT-qPCR assays confirmed the RNA-Seq analysis: the cytochrome P450, proteinase, and oligopeptide transporter had higher

expression in infected leaf tissue, and the ferric-chelate reductase had lower expression in infected leaf tissue compared to culture medium (Figure 1).

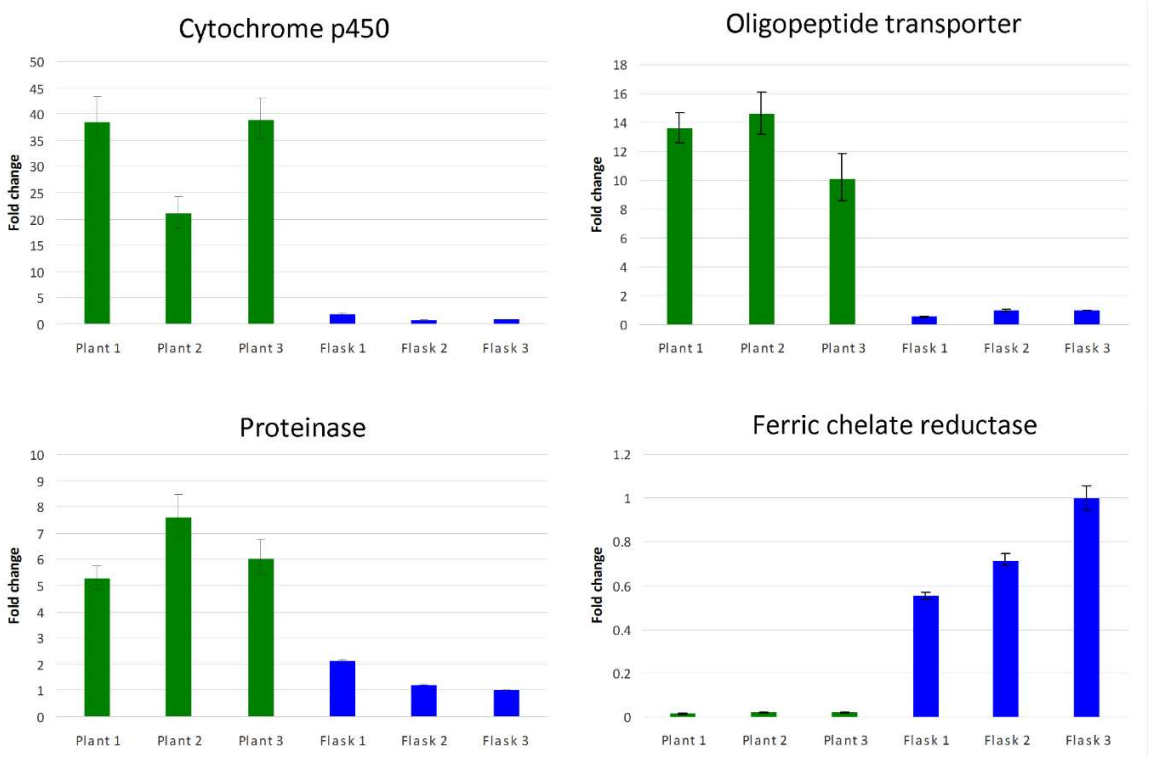


Figure 1. Validation of RNA-Seq results using RT-qPCR assays.
The cytochrome P450 and oligopeptide transporter were normalized against ubiquitin-conjugating enzyme E2 6, and the proteinase and ferric-chelate reductase were normalized against transcription initiation factor TFIID subunit 10. Fold change was calculated using the $2^{-\Delta\Delta C}{}_{T}$ method, relative to the Flask #3 sample. Error bars represent standard error from mean of technical replicates for each sample. Results confirmed the RNA-Seq data that showed higher expression of the cytochrome P450, proteinase, and oligopeptide transporter in infected leaf tissue, whereas the ferric-chelate reductase had lower expression.

Prediction of Differentially Expressed Gene Functions

Blast and Conserved Domain Analysis

The functions of differentially expressed genes based on conserved domains were predicted by identifying homologs using blastp on the National Center for Biotechnology Information's (NCBI) non-redundant protein sequences database and by identifying conserved domains using NCBI's Conserved Domain Database[46,168]. In cases for which multiple differentially expressed genes had the same conserved domains, the number of genes with those conserved domains with higher expression in infected leaf tissue or in culture was determined. This analysis identified genes involved in secondary metabolism, pathogenesis, and nutrient acquisition as well as other functions. For example, genes commonly involved in secondary metabolism[47], such as those encoding cytochrome P450s, short-chain dehydrogenases, methyltransferases, and 2-oxoglutarate and Fe(II)-dependent oxygenases all had many more genes more highly expressed in infected leaf tissue as compared to growth in culture (Figure 2). Twenty-three cytochrome P450 genes were more highly expressed in infected plant tissue, and only five had lower expression. Similarly, 11 short-chain dehydrogenase genes were more highly expressed in infected leaf tissue, and none had lower expression (Figure 2). These findings are consistent with the identification of secondary metabolite genes as among those most highly expressed in the infected plant tissue (Table 1).

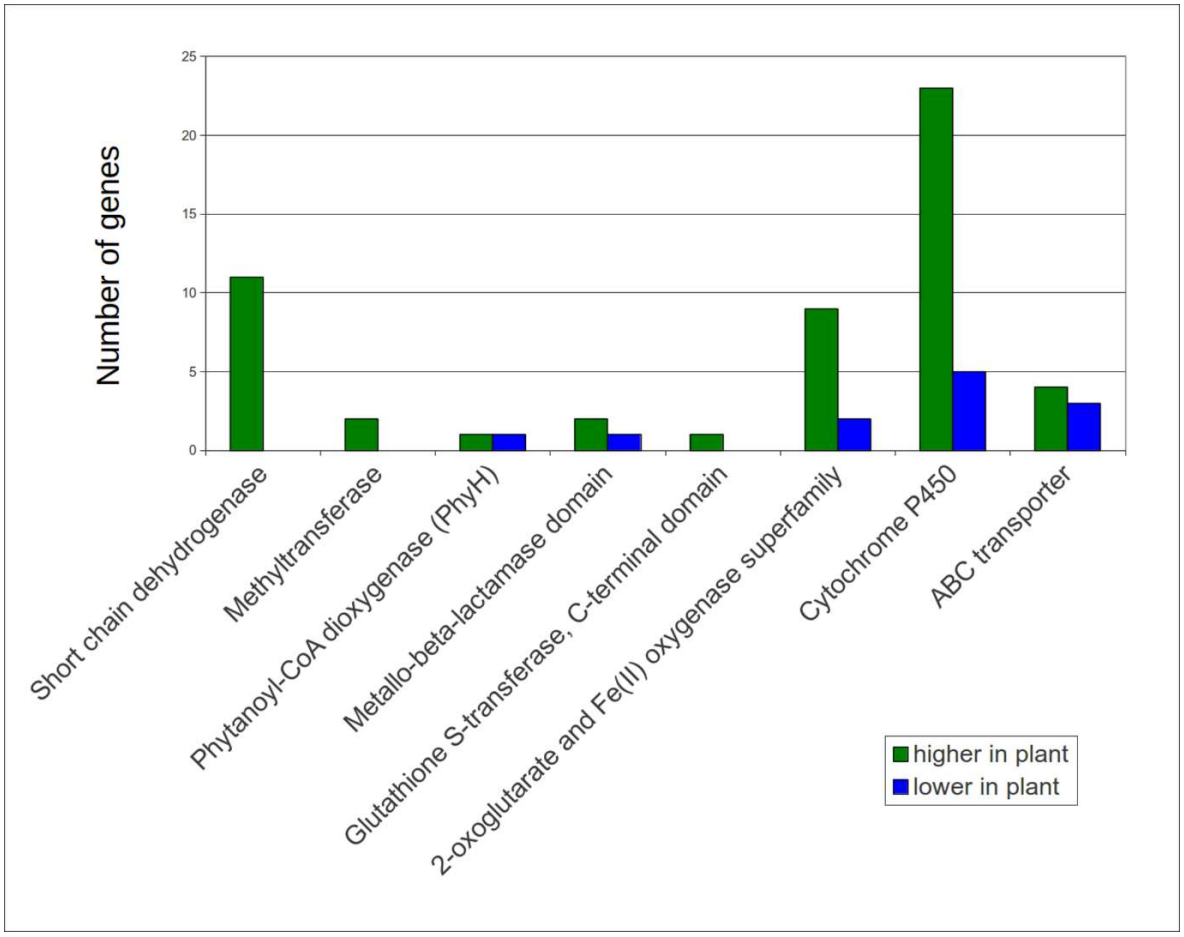


Figure 2. Differentially-regulated genes identified through domain analysis with homology to secondary metabolite genes.
For each of the 802 differentially expressed genes identified from RNA-Seq, a blastp search was done, and conserved protein domains were identified. The figure shows genes with homology to genes commonly involved in secondary metabolism[47]. Green bars = genes more highly expressed in infected leaf tissue relative to mycelium grown in medium; Blue bars = genes with lower expression in infected leaf tissue.

In addition to secondary metabolite genes, types of genes previously implicated in pathogenesis were also identified through the domain analysis (Figure 3). For example, although one of the most highly expressed genes in culture compared to infected leaf tissue encodes a CFEM domain-containing protein (Table 2), five genes encoding proteins with

CFEM domains were more highly expressed in infected leaf tissue, whereas only one had lower expression than in culture. Six genes encoding proteins with conserved Hydrophobic Surface Binding Protein A (HsbA) domains, found in proteins implicated in the recruitment of cutinases important for pathogenicity[232,233], were found to be more highly expressed in infected leaf tissue, and none were found with lower expression. Two differentially regulated transcripts were identified that encode proteins with homology to Pathogenesis-Related Protein 1 (PR-1), one with higher expression in the infected leaf and one with higher expression in culture (Figure 3). In plants, PR-1 proteins are synthesized in response to pathogen attack and play roles in defense[234,235]. In fungal pathogens, however, some PR-1-like proteins have roles in pathogenesis and are required for full virulence[236,237]. Finally, four salicylate hydroxylase-like genes were more highly expressed in the infected leaf tissue, and one had lower expression (Figure 3). Salicylic acid is important for plant defense responses[238], and salicylate hydroxylase interferes with defense by degrading salicylic acid to catechol[239]. To further characterize these sequences, blastp analysis was conducted using the Epichloë festucae salicylate hydroxylase sequence[46,240]. The homolog with the highest similarity (Accession XP_007932011.1) is among the sequences with higher expression in infected leaf tissue, with a log2FC of 3.9[46].

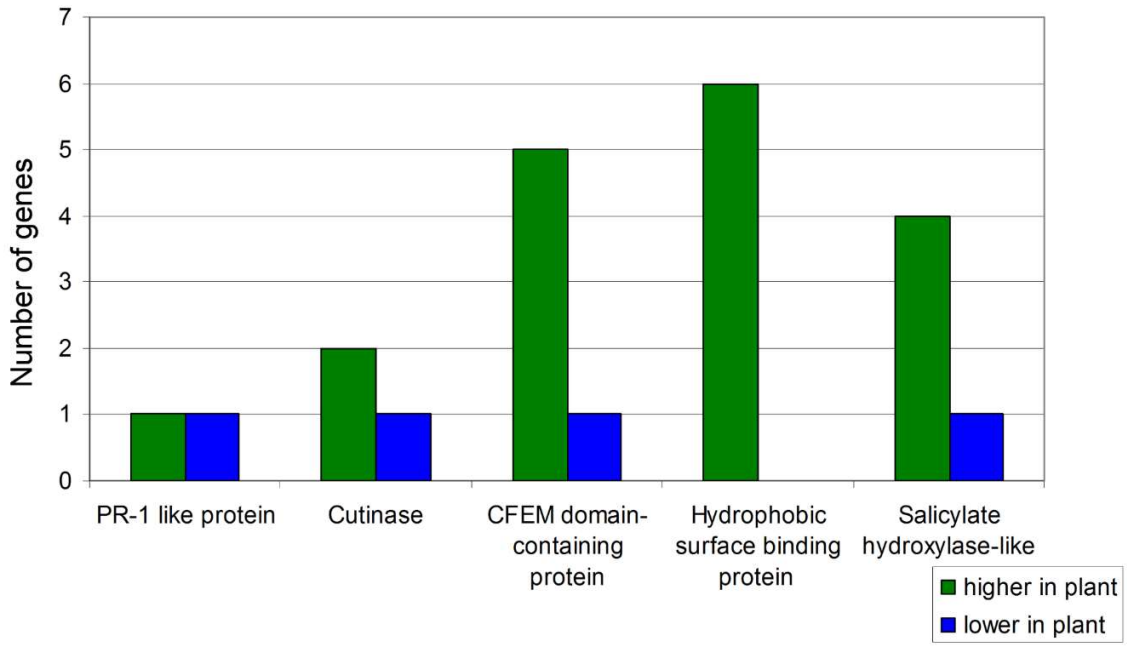


Figure 3. Differentially-regulated genes identified through domain analysis with homology to genes with roles in pathogenesis.
For each of the 802 differentially expressed genes identified from RNA-Seq, a blastp search was done, and conserved protein domains were identified. The number of genes with higher expression in the infected leaf tissue or in culture medium from each category of genes with roles in pathogenesis in other species was determined and is indicated in the bar chart. Green bars = genes more highly expressed in infected leaf tissue relative to mycelium grown in culture medium; Blue bars = genes with lower expression in infected leaf tissue.

Domain analysis also showed that differentially expressed genes have possible roles in acquiring nutrients from the environment. For example, 23 MFS sugar transporter genes were more highly expressed and only 4 had lower expression in infected leaf tissue relative to the fungus grown in medium (Figure 4). More genes encoding peptidases, proteases, proteinases, amino acid transporters, and oligopeptide transporters were more highly expressed in infected leaf tissue, whereas more genes encoding inhibitors of peptidases and proteinases were more highly expressed in culture (Figure 4). Several ferric-chelate reductase

and copper transporter genes were more highly expressed in culture, and none showed higher expression in infected leaf tissue (Figure 4).

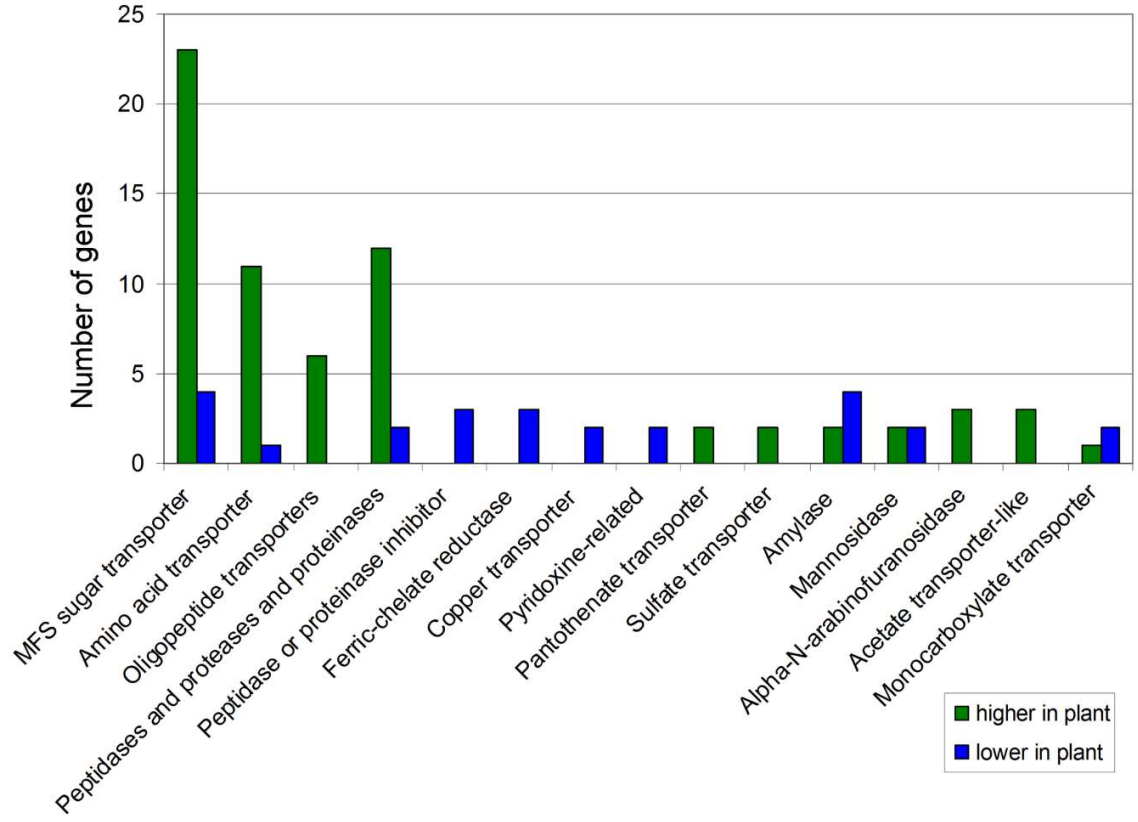


Figure 4. Differentially-regulated genes identified through domain analysis with homology to genes with roles in nutrition.
For each of the 802 differentially expressed genes identified from RNA-Seq, a blastp search was done, and conserved protein domains were identified. The number of genes with higher expression in the infected leaf vs. in culture medium from each category of genes with putative roles in response to nutrient levels in the environment was determined and is indicated in the bar chart[241–252]. Green bars = genes with higher expression in infected leaf tissue relative to mycelium grown in medium; Blue bars = genes with lower expression in infected leaf tissue.

Other genes for which a greater number were more highly expressed in the infected leaf tissue were ones that encode proteins with domains for pyridine nucleotide disulfide

oxidoreductases, α/β-hydrolases, chloroperoxidases, retrograde regulation proteins, ethyl tert-butyl ether degradation proteins, and proteins with DUF3328 domains (Figure 5). The difference in expression of transcripts encoding DUF3328-containing proteins was especially dramatic: eleven genes encoding proteins with DUF3328 domains were more highly expressed in infected leaf tissue and none were more highly expressed in culture (Figure 5). These results agree with the previous finding that two of the 20 genes with highest expression in the infected leaf tissue compared to in culture medium have DUF3328 domains (Table 1). Although the function of DUF3328 domain-containing proteins is unknown, some studies have suggested involvement in sexual reproduction in fungi[253,254]. In contrast, there were more transcripts encoding protein kinases, cupredoxins, ATP synthases, acetyltransferases, cytochrome b5-like proteins, glutathione-dependent formaldehyde-activating enzymes, heme peroxidases, and genes annotated as stress responsive that were more highly expressed in culture than in the infected leaf (Figure 5). Many genes encoding transcription factors were differentially expressed, but there were similar numbers of genes with higher expression in the infected leaf tissue (9 genes) versus in medium (10 genes) (Figure 5). Transcription factors were further characterized based on conserved domains[46]. Overall, the transcription factors encoded by genes more highly expressed in medium had a greater diversity of domains, including the helix-loop-helix, Zn2Cys6, jumonji, bZIP, fungal-specific transcription factor, and NDT80/PhoG-like transcription factor domains[46]. Transcription factors encoded by genes with higher expression in infected leaf tissue had fewer types of domains, and five of these nine transcription factors had only a fungal-specific transcription factor domain[46].

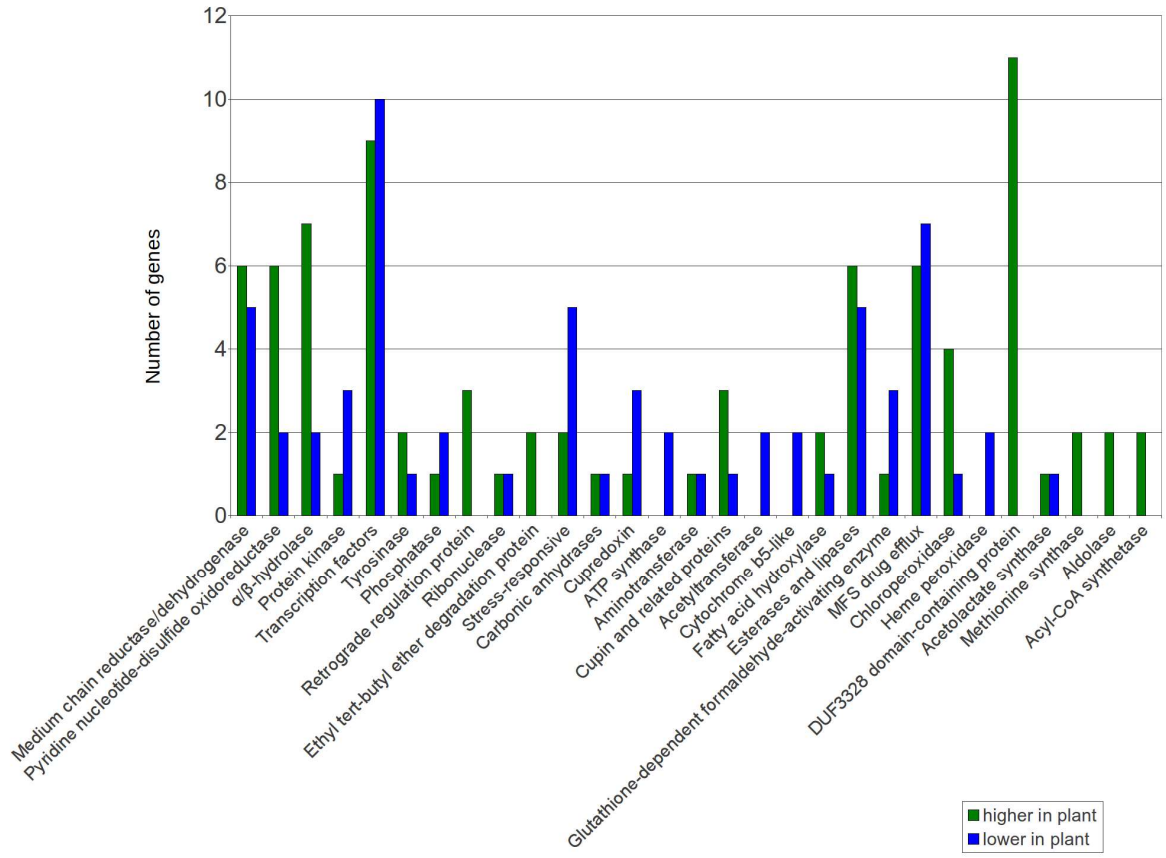


Figure 5. Differentially-regulated genes identified through domain analysis with homology to genes with miscellaneous biological roles.
For each of the 802 differentially expressed genes identified from RNA-Seq, a blastp search was done, and conserved protein domains were identified. The number of genes with higher expression in the infected leaf or in culture medium from each category of genes with miscellaneous biological roles was determined and is indicated in the bar chart. Green bars = genes with higher expression in infected leaf tissue relative to mycelium grown in medium; Blue bars = genes with lower expression in infected leaf tissue.

GO Annotation of Differentially Expressed Genes

Blast2GO is a tool for associating Gene Ontology (GO)[255] terms with sequences of interest[256,257]. This program uses blast to find homologs of the input sequences. InterProScan searches against all the European Bioinformatics Institute databases to find protein

signatures[258,259]. GO terms associated with the blast hits are mapped using annotation files from the GO Consortium. Depending on the similarity of the input sequence with the blast hit, as well as the quality of the evidence code, the input sequences are finally annotated with these GO terms[256,257].

Blast2GO was used with *M. fijiensis* sequences, resulting in 12598 sequences (96.1% of total sequences) with blast hits (Figure S3). Gene Ontology (GO) annotations were obtained for 6678 sequences (50.9% of total sequences) (Figure S3). After annotations were obtained for each gene, Blast2GO was used for GO enrichment analysis. This analysis identified GO terms that were significantly over-represented from sequences having higher expression in infected leaf tissue or in culture medium, compared to their representation in the total set of genes from the *M. fijiensis* genome. Several GO terms were significantly over-represented in sequences having higher expression in infected leaf tissue (Table 3), whereas no GO terms were found to be over-represented in sequences having higher expression in medium. Oxidoreductase, monooxygenase, dioxygenase, and O-methyltransferase activities were found to be significantly over-represented in sequences having higher expression in infected leaf tissue, which is consistent with our previous results (Figure 2). Carbohydrate transport was also found to be over-represented, which is consistent with our finding that more sugar transporters had higher than lower expression in infected leaf tissue (Figure 4). Of the 33 genes annotated by GO as encoding iron ion binding activities, 23 correspond to genes shown by blastp and conserved domain analysis to encode cytochrome P450s, which is consistent with the finding in Figure 2 that more genes encoding cytochrome P450s are more highly expressed in infected leaf tissue.

Table 3. Over-represented GO terms in sequences of transcripts having higher expression in infected leaf tissue than in culture medium.
Table indicates the over-represented GO term, its ID number, its category (Molecular Function, Molecular Process, or Cellular Component), false discovery rate (FDR), and p-value.

GO-ID	Term	Category	FDR	P-Value
GO:0016491	oxidoreductase activity	Molecular Function	1.31E-16	1.69E-20
GO:0055114	oxidation-reduction process	Molecular Process	1.49E-14	3.85E-18
GO:0004497	monooxygenase activity	Molecular Function	1.11E-13	4.29E-17
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	Molecular Function	1.84E-12	9.51E-16
GO:0005506	iron ion binding	Molecular Function	2.61E-12	1.68E-15
GO:0020037	heme binding	Molecular Function	2.35E-08	2.12E-11
GO:0046906	tetrapyrrole binding	Molecular Function	2.35E-08	2.12E-11
GO:0055085	transmembrane transport	Molecular Process	1.38E-04	1.42E-07
GO:0016021	integral component of membrane	Cellular Component	1.83E-04	2.12E-07
GO:0031224	intrinsic component of membrane	Cellular Component	2.71E-04	3.49E-07
GO:0003824	catalytic activity	Molecular Function	3.62E-03	5.12E-06
GO:0046914	transition metal ion binding	Molecular Function	1.29E-02	2.00E-05
GO:0051213	dioxygenase activity	Molecular Function	1.54E-02	2.58E-05
GO:0008643	carbohydrate transport	Molecular Process	4.86E-02	9.18E-05
GO:0008171	O-methyltransferase activity	Molecular Function	4.86E-02	9.38E-05

CAZy Annotation of Differentially Expressed Genes

In addition to the analysis by GO annotation done by Blast2GO, the differentially expressed genes were also analyzed by CAZy annotations. The CAZy database describes families of enzymes that create, degrade, or modify glycosidic bonds[260]. Genes with CAZy annotations were compared with the list of differentially expressed genes. The majority of genes encoding enzymes in the CAZy database were not differentially regulated (Figure 6). For those that were, the analysis revealed similar numbers of genes with higher or lower expression in plant tissue. These included genes encoding carbohydrate esterases, glycoside hydrolases, and polysaccharide lyases; slightly more glycosyl transferases were more highly expressed in culture than in infected leaf tissue. No sequences annotated as having carbohydrate binding modules were differentially expressed (Figure 6). Differentially expressed CAZymes were further analyzed[46]. This analysis revealed that differentially expressed CAZymes were distributed to two carbohydrate esterase families, sixteen glycoside hydrolase families, and four glycosyl transferase families[46], but no patterns in families with higher expression in either condition were apparent.

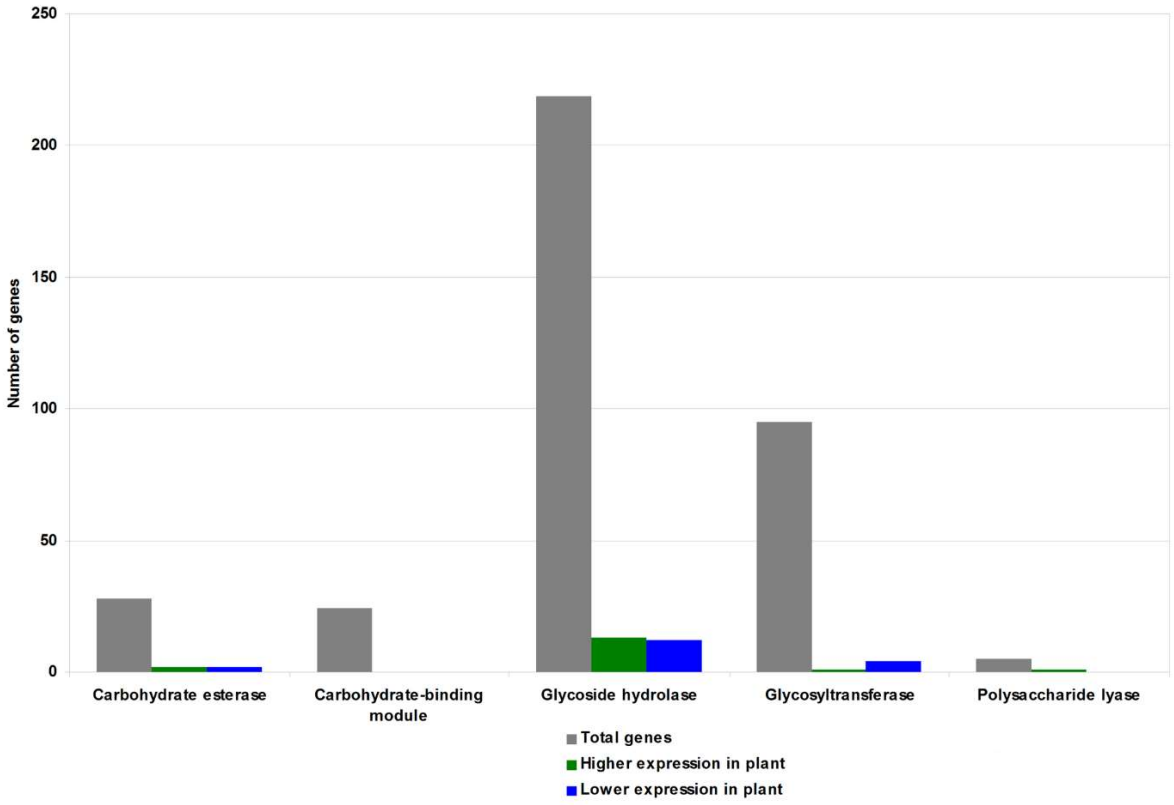


Figure 6. Number of genes with CAZy annotations.
Indicated are the number of genes with each CAZy annotation within: Gray = All *M. fijiensis* genes; Green = List of genes having higher expression in infected leaf tissue; Blue = List of genes having higher expression during growth in medium.

Effector Protein Predictions

Pathogenic fungi secrete effector proteins into the plant apoplast that modulate host physiology and suppress or otherwise protect the pathogen from host defenses[105]. Secreted proteins can be predicted by the presence of a signal peptide, which are N-terminal peptides that target proteins for translocation across the endoplasmic reticulum membrane and that are cleaved off during the translocation process[261]. The overwhelming majority of known fungal effectors are less than 300 amino acids in length after the signal peptide is cleaved[105].

Further, most avirulence effectors are cysteine-rich[106], because the disulfide bonds from the cysteines provide stability against plant proteases in the apoplast[107].

To identify small, cysteine-rich, secreted proteins, the program SignalP 4.1[262] was first used to predict *M. fijiensis* protein sequences that contain a signal peptide. From this analysis, 863 protein sequences (7% of total sequences) were predicted to contain a signal peptide. SignalP 4.1 was also used to generate predictions of the mature protein sequences once the signal peptides are cleaved. Of the 863 mature protein sequences, 394 were less than 300 amino acids in length, and 231 of these sequences were considered cysteine-rich, containing four or more cysteine residues.

Genes encoding 40 of the 231 short, cysteine-rich, secreted proteins were differentially expressed in our RNA-Seq analysis[46]. Thirty were more highly expressed in infected leaf tissue, whereas only 10 were more highly expressed in culture[46]. Of the protein sequences encoded by the 30 transcripts that were more highly expressed in infected leaf tissue, six had conserved domains: one had a PR-1-like protein domain, two were predicted to be cutinases, one had a CFEM domain, one had a DUF3328 domain, and one had a globin-like domain[46]. Two identified as hypothetical were among those with the highest expression in infected leaf tissue compared to culture medium (Table 1)[46]. For the 10 short, cysteine-rich, secreted proteins whose transcripts were more highly expressed in culture, two had conserved domains: one had a CFEM domain, and the other had a serine/threonine phosphatase domain[46]. The CFEM domain-encoding transcript was also identified in the list of genes with highest expression in culture medium compared to infected leaf tissue (Table 2)[46]. For the remaining 32 differentially expressed genes encoding short, cysteine-rich,

secreted proteins, no significant homology to characterized protein sequences and no predicted conserved domains could be identified[46]. Overall, 16 of the 30 differentially expressed putative effector genes with higher expression in the infected leaf tissue had homologs restricted to species within the Mycosphaerellaceae[46]; others had homologs outside this family. This result is consistent with the observation that many fungal effectors have a restricted phylogenetic distribution[108].

Homologs of the Ecp2, Ecp6 and Avr4 effectors of the tomato pathogen *Cladosporium fulvum* have been identified in *M. fijiensis*, and the *M. fijiensis* Ecp2 and Avr4 effectors are recognized by tomato R proteins that normally recognize the *C. fulvum* Ecp2 and Avr4 effectors[102,104]. In *C. fulvum*, Avr4 and Ecp6 bind chitin and protect fungal cell walls from plant chitinases[263]. The function of Ecp2 is not known[102]. In our dataset, none of the genes were strongly differentially expressed: MfAvr4 had a log2FC of -0.8 with an adjusted p-value of 0.05, and the expression of MfEcp2 and MfEcp6 was unchanged between the two conditions[46]. This result may be due to our focus on the necrotrophic phase in our transcriptome analysis.

Identification of Differentially Expressed Gene Clusters

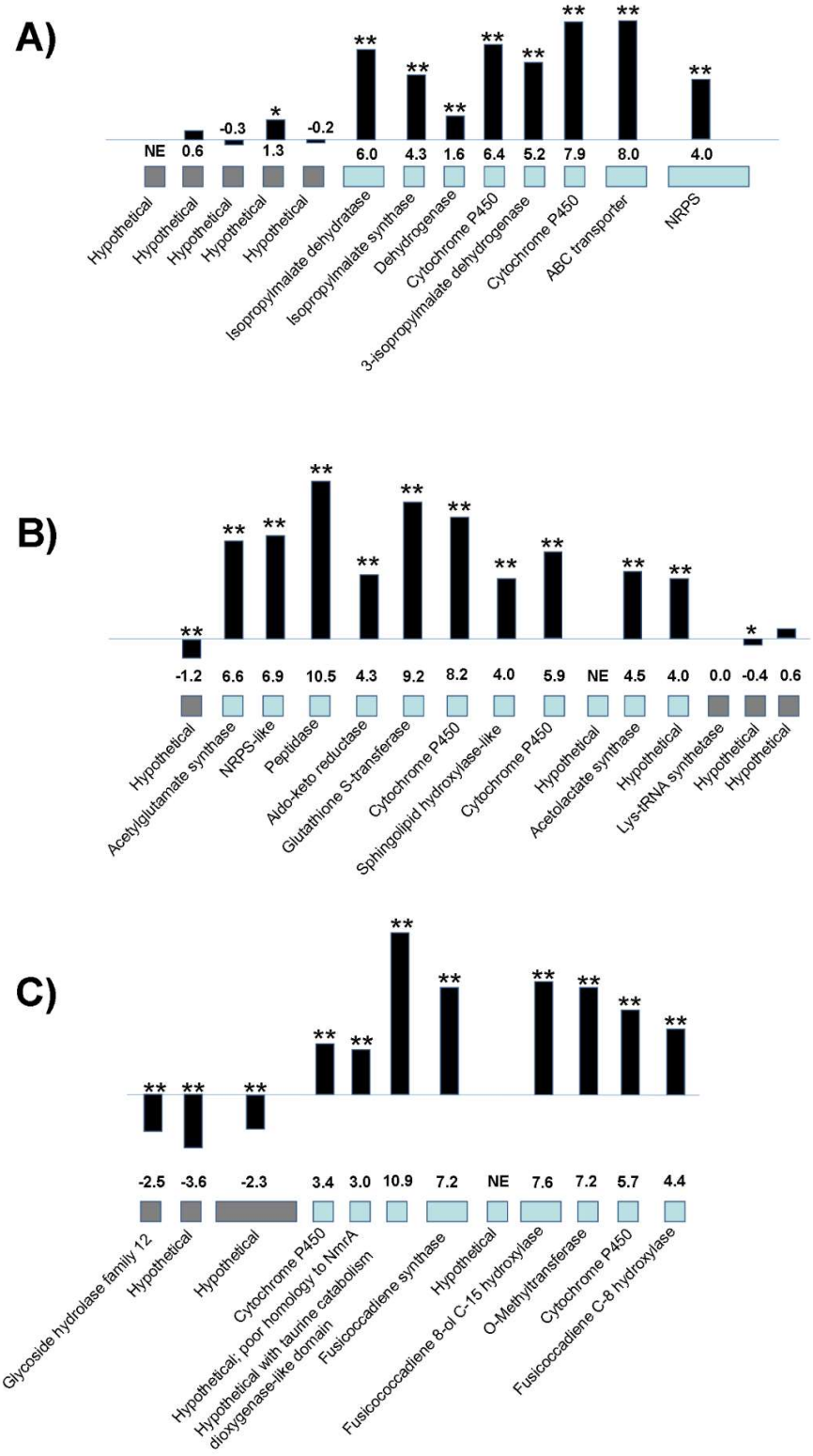
To identify potential gene clusters, we searched for loci in the genome with at least three adjacent genes that were similarly differentially expressed. Using this method, 16 gene clusters were identified with higher expression in infected leaf tissue, and 5 clusters were identified with lower expression in infected leaf tissue. Three of the putative clusters with

higher expression in infected leaf tissue encode polyketide pathways, and have been previously described[45]. The remaining 18 clusters are detailed by Noar and Daub[46].

Genes encoding secondary metabolite pathways are often clustered in fungal genomes[47], and our cluster analysis supports the importance of secondary metabolism in *M. fijiensis* disease development. Two of the genes shown in Table 1 with highest expression in infected leaf tissue compared to culture medium are part of a polyketide synthase cluster (*PKS7-1*) previously described, and two other PKS clusters were also previously shown to be more strongly expressed in infected leaf tissue[45]. Of the remaining 13 clusters identified with higher expression in infected leaf tissue, two have genes similar to non-ribosomal peptide synthases (NRPS). One of these is an NRPS on scaffold 7 (Figure 7A). Adjacent genes encoding an ATP-Binding Cassette (ABC) transporter, two cytochrome P450s, a 3-isopropylmalate dehydrogenase, a glyoxylate/hydroxypyruvate reductase, an α-isopropylmalate synthase, and a 3-isopropylmalate dehydratase also showed higher expression in infected leaf tissue (Figure 7A). A blastp search revealed that the closest homologs of this NRPS are in the related banana pathogens *Mycosphaerella musicola* and *Mycosphaerella eumusae*; both of these species have homologs with 60% sequence similarity[46]. Aside from the homologs in *M. musicola* and *M. eumusae*, none of the other top 15 homologs were found in Dothideomycete species[46]. The closest characterized homolog is a destruxin synthetase from *Metarhizium guizhouense*, with 53% sequence similarity to the *M. fijiensis* NRPS[46]. Destruxins are cyclic hexadepsipeptides produced by insect pathogens such as *Metarhizium* spp., *Aschersonia* spp., and *Beauveria felina*[90–92]. In these pathosystems, destruxins act insecticidally. Destruxins have also been identified from the

plant pathogens *Alternaria brassicae* and *Ophiosphaerella herpotricha*, and are phytotoxic to some plant species[93–96]. In addition to the NRPS gene cluster on scaffold 7, there is a gene cluster containing an NRPS-like gene on scaffold 4 (Figure 7B). While a true NRPS enzyme must contain a condensation domain, an adenylation domain, and a phosphopantetheine attachment site[264], the NRPS-like enzyme encoded on scaffold 4 was predicted to contain an adenylation domain, but no condensation domain or phosphopantetheine attachment site[46]. Other genes in this cluster that were similarly more highly expressed in infected leaf tissue include genes encoding an N-acetylglutamate synthase, a peptidase, an aldo/keto reductase, a glutathione S-transferase, two cytochrome P450s, and a sphingolipid hydroxylase-like protein (Figure 7B). A blastp search of the *M. fijiensis* NRPS-like sequence against the NCBI non-redundant protein sequence database revealed that *M. musicola* has a close homolog with 87% sequence similarity, but *M. eumusae* does not[46]. There were no close homologs of this NRPS-like sequence for which a function has been described[46].

Figure 7. Non-ribosomal peptide synthase (NRPS), NRPS-like, and fusicoccane clusters with higher expression in infected leaf tissue.
The NRPS or NRPS-like gene is shown along with its neighboring genes in the *M. fijiensis* genome. The description of each gene as determined by blastp of the corresponding protein is shown along with its log2FC value of expression in infected leaf tissue versus expression in liquid medium. Black bars are proportional to the log2FC value. Gene expression differences that are significant at $p < 0.01$ are shown with two asterisks above the corresponding bar, and those significant at $p < 0.05$ are shown with a single asterisk. NE = no expression detected. A) NRPS gene cluster on scaffold 7; B) NRPS-like gene cluster on scaffold 4; C) Fusicoccane gene cluster on scaffold 2.


A)
NE 0.6 -0.3 1.3 -0.2 6.0 4.3 1.6 6.4 5.2 7.9 8.0 4.0
Hypothetical
Hypothetical
Hypothetical
Hypothetical
Hypothetical
Isopropylmalate dehydratase
Isopropylmalate synthase
Dehydrogenase
Cytochrome P450
3-isopropylmalate dehydrogenase
Cytochrome P450
ABC transporter
NRPS
B)
-1.2 6.6 6.9 10.5 4.3 9.2 8.2 4.0 5.9 NE 4.5 4.0 0.0 -0.4 0.6
Hypothetical
Acetylglutamate synthase
NRPS-like
Peptidase
Aldo-keto reductase
Glutathione S-transferase
Cytochrome P450
Sphingolipid hydroxylase-like
Cytochrome P450
Hypothetical
Acetolactate synthase
Hypothetical
Lys-tRNA synthetase
Hypothetical
Hypothetical
C)
-2.5 -3.6 -2.3 3.4 3.0 10.9 7.2 NE 7.6 7.2 5.7 4.4
Glycoside hydrolase family 12
Hypothetical
Hypothetical
Cytochrome P450
Hypothetical; poor homology to NmrA
Hypothetical with taurine catabolism dioxygenase-like domain
Fusicoccadiene synthase
Hypothetical
Fusicoccadiene 8-ol C-15 hydroxylase
O-Methyltransferase
Cytochrome P450
Fusicoccadiene C-8 hydroxylase

Another gene cluster with higher expression in the infected leaf tissue is one with genes having homology to fusicoccane biosynthetic genes (Figure 7C). Fusicoccanes are diterpenoids with a variety of effects on plant physiology, including opening of stomata[97]. A blastp search was performed of the *M. fijiensis* fusicoccadiene synthase using the non-redundant protein sequence database through NCBI, which identified the best homolog as being from *Alternaria brassicicola* with 83% sequence similarity[46]; this enzyme catalyzes the first step in the synthesis of a fusicoccane called brassicicene C[265]. The top homologs[46] were used to create a phylogenetic tree, which showed a bootstrap value of 100 for the relationship between the *M. fijiensis* fusicoccadiene synthase and the *A. brassicicola* and *Bipolaris victoriae* fusicoccadiene synthases (Figure S4). None of the top homologs identified were from Mycosphaerellaceae species[46], even though genome sequences from several members of this family are publicly available on NCBI, including very close relatives of *M. fijiensis* such as *M. musicola* and *M. eumusae* (NCBI Genome IDs 43744 and 43743, respectively). The gene content and gene orientations for the *M. fijiensis* and *A. brassicicola* fusicoccane biosynthetic clusters were further compared, showing that both clusters contain genes encoding a fusicoccadiene synthase, an α-ketoglutarate-dependent dioxygenase, a hypothetical protein with similarity to nmrA, four cytochrome P450s, and an O-methyltransferase (Figure 8). The *M. fijiensis* cluster contains a gene encoding a hypothetical protein which the *A. brassicicola* cluster does not have, and the *A. brassicicola* cluster contains genes that the *M. fijiensis* cluster lacks, including: genes encoding a cytochrome P450, a short-chain dehydrogenase, and an acetyltransferase. Gene orientation is largely conserved, with only a cytochrome P450 on one end of the cluster being in different

orientations between the two clusters. These results suggest that *M. fijiensis* may produce a fusicoccane very similar, though not identical, to brassicicene C.

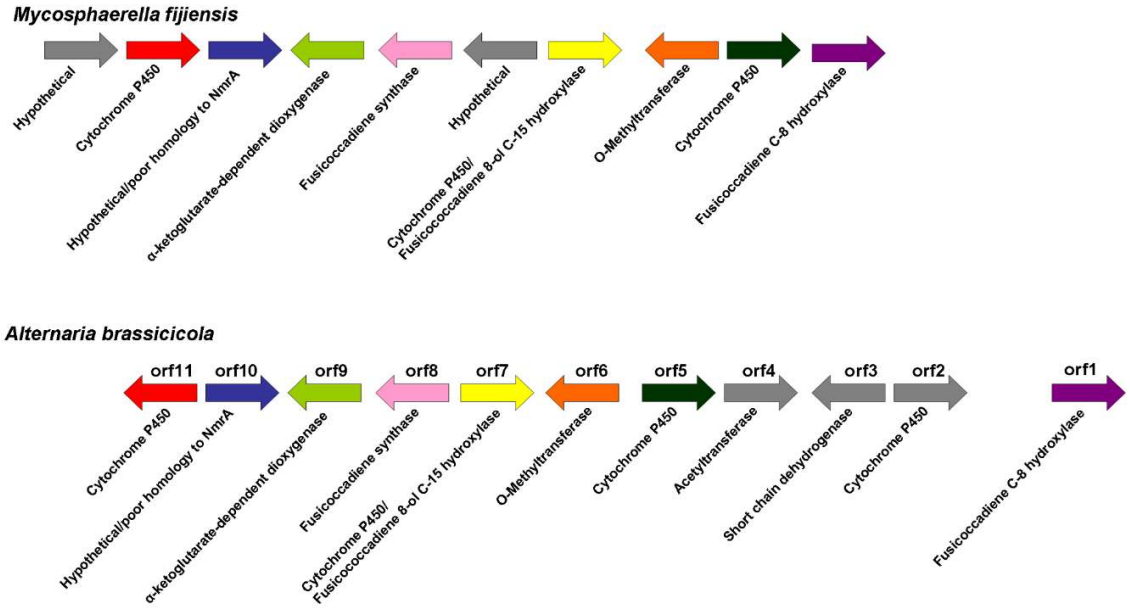
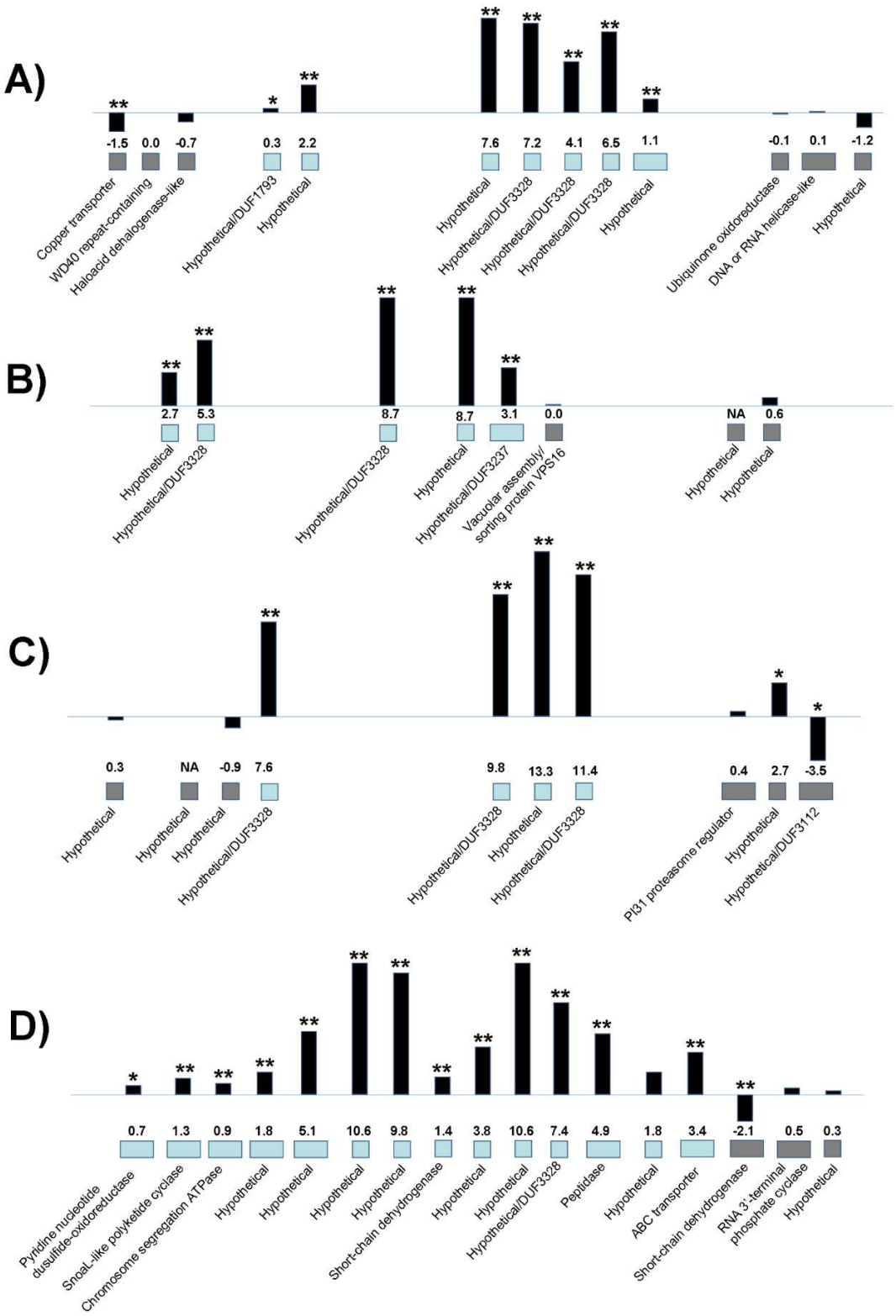


Figure 8. Comparison of *M. fijiensis* and *Alternaria brassicicola* fusicoccane biosynthetic clusters.
Genes in the *M. fijiensis* fusicoccane biosynthetic cluster are shown compared to those in the *A. brassicicola* cluster. Orientation of each gene is indicated by the direction of each arrow. Genes encoding proteins with similar functions are indicated by the same color, and genes not shared between the two clusters are shown in gray.

Interestingly, four gene clusters with higher expression in infected leaf tissue include genes encoding DUF3328-containing proteins (Figure 9)[46]. These gene clusters include from one to three DUF3328 genes, and considered together, these clusters contain 9 of the 11 DUF3328 genes with higher expression in infected leaf tissue (Figures 5 and 9)[46]. As noted

earlier, the function of DUF3328 domain-containing proteins is unknown, although some studies have suggested involvement in fungal sexual reproduction[253,254].

Figure 9. Gene clusters encoding Domain of Unknown Function (DUF) 3328-containing proteins.
Gene clusters with higher expression in infected tissue are shown, which contain genes encoding DUF3328 proteins. The description of each gene as determined by blastp of the corresponding protein is shown along with its log2FC value of expression in infected leaf tissue versus expression in liquid medium. Black bars are proportional to the log2FC value. Gene expression differences that are significant at $p < 0.01$ are shown with two asterisks above the corresponding bar, and those significant at $p < 0.05$ are shown with a single asterisk. NE = no expression detected. A) Gene cluster on scaffold 12; B) Gene cluster on scaffold 2; C) Gene cluster on scaffold 9; D) Gene cluster on scaffold 3.


A)
**
**
*
**
**
**
**
**
-1.5
0.0
-0.7
0.3
2.2
7.6
7.2
4.1
6.5
1.1
-0.1
0.1
-1.2
Copper transporter
WD40 repeat-containing
Haloacid dehalogenase-like
Hypothetical/DUF1793
Hypothetical
Hypothetical
Hypothetical/DUF3328
Hypothetical/DUF3328
Hypothetical/DUF3328
Hypothetical
Ubiquinone oxidoreductase
DNA or RNA helicase-like
Hypothetical
B)
**
**
**
**
**
2.7
5.3
8.7
8.7
3.1
0.0
NA
0.6
Hypothetical
Hypothetical/DUF3328
Hypothetical/DUF3328
Hypothetical
Hypothetical/DUF3237
Vacuolar assembly/
sorting protein VPS16
Hypothetical
Hypothetical
C)
**
**
**
**
**
*
*
0.3
NA
-0.9
7.6
9.8
13.3
11.4
0.4
2.7
-3.5
Hypothetical
Hypothetical
Hypothetical
Hypothetical/DUF3328
Hypothetical/DUF3328
Hypothetical
Hypothetical/DUF3328
PI31 proteasome regulator
Hypothetical
Hypothetical/DUF3112
D)
*
**
**
**
**
**
**
**
**
**
**
**
**
**
0.7
1.3
0.9
1.8
5.1
10.6
9.8
1.4
3.8
10.6
7.4
4.9
1.8
3.4
-2.1
0.5
0.3
Pyridine nucleotide
disulfide-oxidoreductase
SnoaL-like polyketide cyclase
Chromosome segregation ATPase
Hypothetical
Hypothetical
Hypothetical
Hypothetical
Short-chain dehydrogenase
Hypothetical
Hypothetical
Hypothetical/DUF3328
Peptidase
Hypothetical
ABC transporter
Short-chain dehydrogenase
RNA 3'-terminal
phosphate cyclase
Hypothetical

Two gene clusters with higher expression in infected leaf tissue, on scaffolds 7 and 21, consist almost entirely of genes encoding hypothetical proteins (Figure S5A-B)[46]. The other clusters contain genes with putative functions, but their roles as gene clusters are unclear (Figure S5C-F)[46]. For example, one cluster contains several genes commonly found in secondary metabolite gene clusters[47] including one encoding a 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily enzyme, two methyltransferases, an oxidoreductase, and a transporter (Figure S5C)[46]. Another cluster contains genes encoding two hydrophobic surface binding proteins with homology to HsbA and a pyridine nucleotide-disulfide oxidoreductase (Figure S5D)[46]. HsbA domains are found in proteins implicated in recruitment of cutinases[232,233], but the role of the oxidoreductase in this gene cluster is unclear.

The five gene clusters with lower expression in infected leaf tissue contain genes encoding proteins with putative functions, but the role of the genes together as a cluster is unclear (Figure S6)[46]. For example, one cluster contains genes encoding a putative cerato-platanin, two hypothetical proteins, and a C2 domain-containing protein (Figure S6B)[46]. Cerato-platanin was first described as a small, secreted fungal toxin from *Ceratocystis fimbriata*[266]. However, cerato-platanins are produced by both pathogenic and non-pathogenic fungi[267,268], and they are believed to play multiple roles in fungal biology including promotion of cell wall expansion and hyphal elongation[269,270]. Though the putative cerato-platanin encoded by this gene cluster may play a role in hyphal elongation in the culture medium, it is unclear whether cerato-platanin would interact with any of the other gene products in this cluster. Another example of a gene cluster with lower expression in the infected leaf tissue contains genes encoding an oxidoreductase, a short-chain dehydrogenase,

a choline dehydrogenase, an MFS transporter, an aminoacyl-tRNA ligase, a protein containing ankyrin repeats, two hypothetical proteins, and a sequence with homology to pyoverdine/dityrosine biosynthesis protein (Figure S6C)[46]. Dityrosine is a cross-linking agent that is present in the cell walls of some fungi and protects against adverse environmental conditions[271,272]. Dityrosine is present in the outermost wall layer of ascospores in many members of the family Saccharomycetaceae[273]. In *Candida albicans*, dityrosine is also produced along the surface of yeast bud scars and germinated cells[274]. Pyoverdine is a siderophore produced by *Pseudomonas* spp. which consists of a fluorescent chromophore, a peptide chain, and an acyl side chain[275]. While putative functions for many genes in this cluster can be predicted, it is unclear how proteins encoded by this cluster may work together to produce a product.

Distribution of Differentially Expressed Genes on *M. fijiensis* Genome Scaffolds

Many fungi have one or more dispensable chromosomes that assist pathogenicity on a particular host, or have other functions that are useful but not strictly necessary for survival[48]. The related species *Mycosphaerella graminicola* has eight dispensable chromosomes which are readily lost when the fungus undergoes meiosis[53]. These dispensable chromosomes from *M. graminicola* have characteristics that distinguish them from the core chromosomes: they are the smallest chromosomes, have the lowest G+C content, have the lowest gene density, have the lowest proportion of genes encoding proteins with PFAM domains, have the highest proportion of repetitive DNA, and have different codon usage[55,276]. There are 14 scaffolds from the *M. fijiensis* genome that were predicted to also be dispensable, since they have the

same characteristics as the dispensable chromosomes from *M. graminicola*: small scaffolds, low G+C content, low gene density, low proportion of genes encoding proteins with PFAM domains, a high proportion of repetitive DNA, and different codon usage[54,55]. If these scaffolds do correspond to dispensable chromosomes, some may play important roles in pathogenicity as is the case for some other pathogenic fungi[48,277–281].

Scaffolds which contain at least 25 genes and were predicted to be dispensable by Ohm et al[54] are indicated in Figure 10. To determine whether these putative dispensable scaffolds from *M. fijiensis* have a different percentage of differentially expressed genes compared to the core scaffolds, the differentially expressed genes were sorted based on their scaffold of origin (Figure 10). This analysis revealed that 31% (13 genes out of 42) and 52% (21 genes out of 40) of the genes on the predicted dispensable scaffolds 15 and 21, respectively, were more highly expressed in infected leaf tissue, and no genes on these scaffolds had lower expression (Figure 10)[46]. By contrast, less than 10% of the genes on the core scaffolds were more highly expressed in infected leaf tissue (Figure 10). These results suggest that scaffolds 15 and 21 may play roles in pathogenicity. About two-thirds of the predicted proteins encoded by genes on scaffolds 15 and 21 had no blast hits and no conserved domains (61% and 69%, respectively) (Figure S7)[46]. For those from scaffolds 15 and 21, respectively, that did have homologs, 80% and 100% were in the related banana pathogens *M. musicola* and *M. eumusae* (Figure S7)[46]. Two gene models from scaffold 15 and one from scaffold 21 had conserved domains: one had a fungal Zn(2)-Cys(6) binuclear cluster domain common in transcription factors, one had a chromosome segregation protein

domain, and one had a serine/threonine protein kinase domain. All three of these genes were more highly expressed in infected leaf tissue[46].

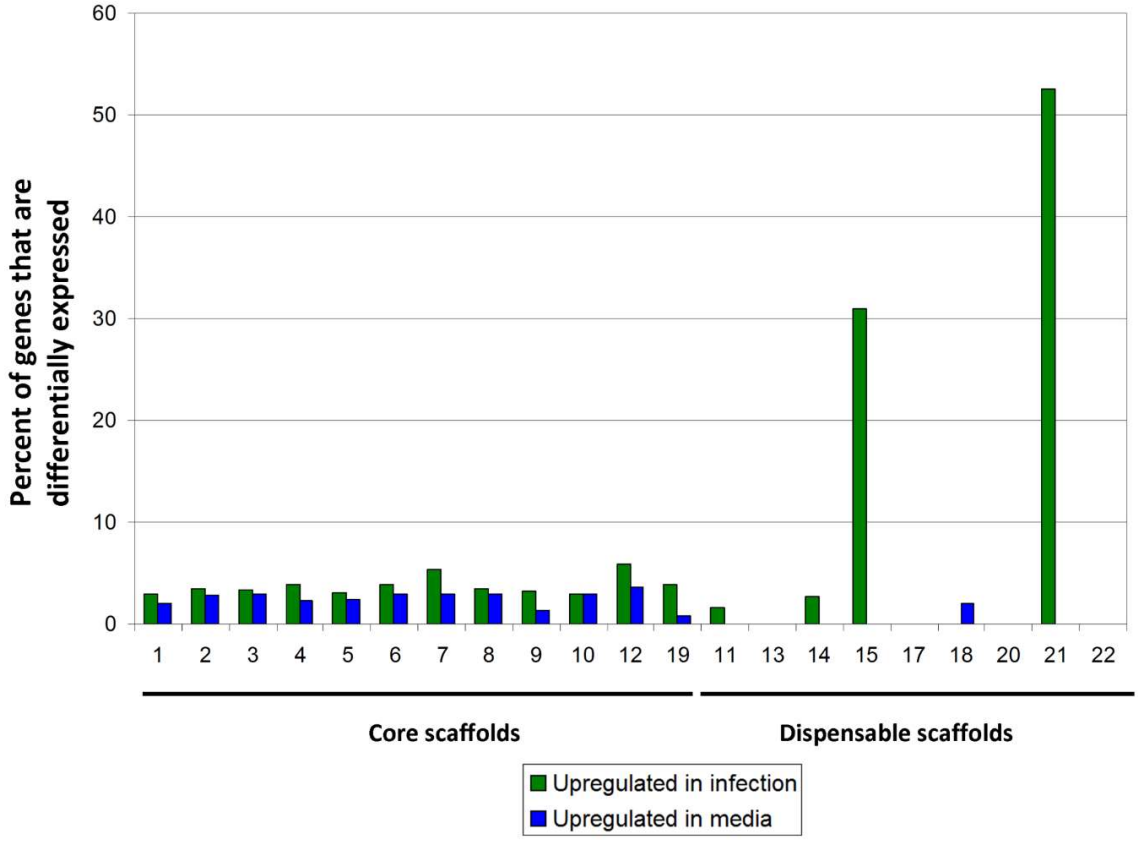


Figure 10. Percent of genes on each *M. fijiensis* scaffold that are differentially expressed.
For each *M. fijiensis* genome scaffold with at least 25 predicted genes, the number of genes with higher expression in the infected leaf or in culture is expressed as a percentage of the total number of genes on that scaffold. Predictions of whether the scaffold may represent a core or dispensable scaffold[54,55] are indicated underneath the scaffold number. Scaffolds were predicted to be dispensable based on the following characteristics: small scaffolds, low G+C content, low gene density, low proportion of genes encoding proteins with PFAM domains, high proportion of repetitive DNA, and different codon usage[54,55]. Green = Genes more highly expressed in infected leaf tissue; Blue = Genes with lower expression in infected leaf tissue.

While the other putative dispensable scaffolds did not have as large a proportion of differentially expressed genes as were found on scaffolds 15 and 21, they did have very different patterns of differential expression compared to the core scaffolds (Figure 10). Each of the core scaffolds had some genes that were more highly expressed in the infected leaf tissue and other genes that were more highly expressed in culture medium (Figure 10). By contrast, the scaffolds predicted to be dispensable had only genes that were more highly expressed in infected leaf tissue (scaffolds 11, 14, 15, and 21), more highly expressed in culture (scaffold 18), or no differentially expressed genes (scaffolds 13, 17, 20, and 22)[46].

Although none were differentially expressed between the two conditions, transcripts were detected for 82% of the 65 genes and 43% of the 30 genes on scaffolds 17 and 22[46]. In contrast, no transcripts were detected from genes on scaffolds 13 and 20, out of 51 and 29 total genes on these scaffolds, respectively[46]. Since no transcripts were detected from scaffolds 13 and 20, we used PCR to assay isolate 14H1-11A (used in the RNA-Seq analysis) for genes on these scaffolds in order to determine whether these scaffolds are present in this isolate. The PCR assays were done for three hypothetical genes on each of scaffolds 13 and 20, and β-tubulin as a positive control. Genomic DNA extracted from isolates 14H1-11A (the isolate used for RNA-Seq) and CIRAD86 (the isolate for which the reference genome is available) was used as a template. This analysis revealed that while the β-tubulin PCR assay resulted in equally strong bands for both isolates, PCR amplifications of all genes on scaffolds 13 and 20 resulted in strong bands for isolate CIRAD86 only, whereas only one gene on scaffold 20 yielded a comparably sized, though fainter, band for isolate 14H1-11A (Figure 11).

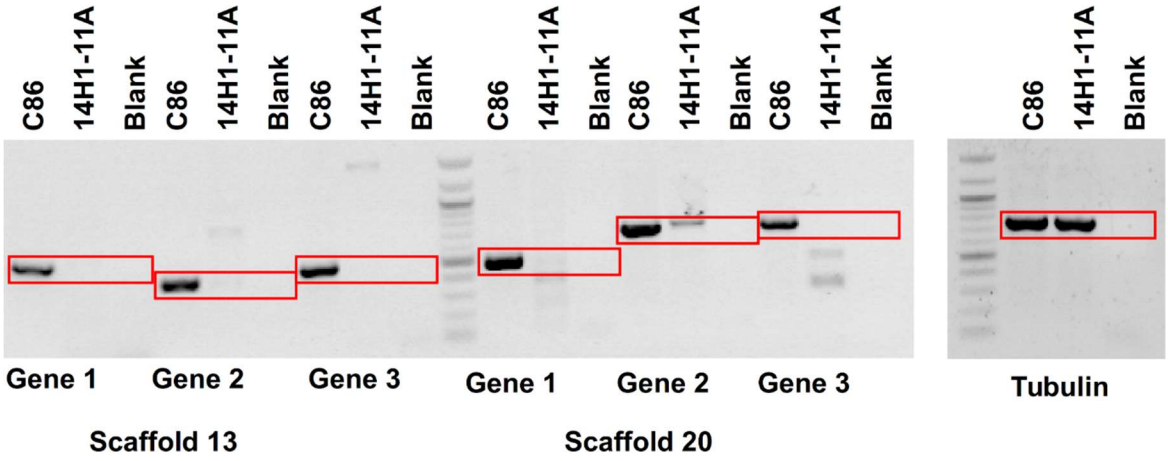


Figure 11. PCR amplification of genes from scaffolds 13 and 20 in isolates 14H1-11A and CIRAD86.
PCR amplification was done for three genes encoding hypothetical proteins on scaffolds 13 and 20, as well as β-tubulin as a positive control. Genomic DNA from isolates 14H1-11A (used in this RNA-Seq analysis) and CIRAD86 (genome sequence publicly available) was used as a template for PCR assays, with water used as a negative control. Quick-Load 100 bp DNA ladder (NEB) was used as a molecular weight marker. Red rectangles mark the expected product size for each assay.

From the lack of transcripts detected from scaffold 13 as well as the lack of amplification of scaffold 13 genes in our PCR amplification of genomic DNA from isolate 14H1-11A, we have no evidence that scaffold 13 is present in isolate 14H1-11A. Results are less clear for scaffold 20. As noted, no transcripts were identified for this scaffold in our transcriptome analysis, and two out of the three genes tested by PCR were not amplified. However, a faint band was amplified from 14H1-11A for gene 2 on this scaffold (Figure 11). This band was gel purified and sequenced for both isolates, and the sequences were used for a blastn search against the *M. fijiensis* genome. The blast search showed that the DNA sequence for isolate 14H1-11A has 83% identity with the published sequence from isolate

CIRAD86 (Table 4). A blastx search was then done of the isolate 14H1-11A translated nucleotide sequence to align this predicted protein sequence against the published sequence from isolate CIRAD86. This analysis revealed that the changes in nucleotide sequence result in several internal stop codons in isolate 14H1-11A (Figure S8), and therefore 14H1-11A is unlikely to produce a functional protein product. Scaffold 20 may be present but diverged in isolate 14H1-11A, or it may be that only part of that scaffold is present.

Table 4. Blast hits of scaffold 20 gene amplified from both isolates CIRAD86 and 14H1-11A (gene 2 in Figure 11).
Bands of the expected product size were gel purified for isolate CIRAD86 and 14H1-11A, and were sequenced. Blastn searches of both the resulting sequences were done against the *M. fijiensis* genome. The table indicates the isolate from which the sequence was obtained, the description of the hit, the accession of the hit, the bit score, E-value, percent identity, and percent of gaps.

Isolate of query sequence	Description of hit	Hit Accession	Bit score	E-value	Percent identity	Gaps
CIRAD86	*Pseudocercospora fijiensis* CIRAD86 hypothetical protein partial mRNA	XM_007934411.1	1129	0	611/611 (100%)	0/629 (0%)
14H1-11A	*Pseudocercospora fijiensis* CIRAD86 hypothetical protein partial mRNA	XM_007934411.1	538	5.00E-149	505/609 (83%)	12/609 (1%)

Discussion

Although several studies have investigated banana genes involved in defense against *M. fijiensis* and other *Mycosphaerella* pathogens[39,41–43], our work is the first to identify candidate pathogenicity genes based on analysis of the *M. fijiensis* transcriptome during its association with banana. Our analysis identified 802 genes that were differentially expressed in infected leaf tissue compared to culture medium. Of these, 483 genes had higher expression in infected leaf tissue, and 319 genes had higher expression in medium.

Secondary metabolic pathways have long been suspected as important during the interaction between this fungus and banana[5,61,74,76,228,282,283], and we found that two of the genes with the highest expression in infected leaf tissue compared to culture medium (Table 1) were from the previously described *PKS7-1* polyketide gene cluster[45], whose product and function are currently unknown. Many types of genes commonly involved in secondary metabolism were also found to have higher expression in infected leaf tissue compared to growth in medium. These genes include ones encoding cytochrome P450s, short-chain dehydrogenases, and oxidoreductases in the 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily (Figures 2 and 7)[47]. In addition to the three polyketide synthase gene clusters recently described[45], we showed that an NRPS, an NRPS-like, and a fusicoccane gene cluster had higher expression in infected leaf tissue compared to medium (Figure 7). The closest characterized homolog to the *M. fijiensis* NRPS protein sequence was a destruxin synthase. Some fungal pathogens produce destruxins which are toxic to their insect or plant hosts and are thought to be involved in virulence[90–96]. Further research is needed to identify the product of the NRPS and its possible role in virulence.

Another gene cluster identified with higher expression in the infected leaf tissue is predicted to encode a fusicoccane. Fusicoccanes are diterpenoids produced by a variety of organisms including fungi, plants, and liverworts. They share a common C20 core, but vary on stereochemistry, degree of saturation, and substitutions, which all affect their physiological activity[97]. Some fusicoccanes cause stomatal opening and wilting, affect seed germination and cell elongation, have antibacterial and antifungal activities, and cause inhibition of biological nitrification and of lysophospholipase[284]. The predicted fusicoccane gene cluster in *M. fijiensis* is very similar to a cluster producing brassicicene C from *Alternaria brassicicola*, though it lacks a cytochrome P450, a short-chain dehydrogenase, and an acetyltransferase that the *A. brassicicola* cluster contains (Figure 8). These differences suggest that the fusicoccane side groups may slightly differ, thus we hypothesize that the cluster encodes a novel fusicoccane. No significant antimicrobial activity has been detected from brassicicene, and its phytotoxicity and effect on the pathogenicity of *A. brassicicola* are unknown[285].

Other types of genes previously implicated in pathogenicity in other fungi also had higher expression in the infected leaf tissue. These include genes predicted to encode salicylate hydroxylase-like proteins, hydrophobic surface binding proteins, and CFEM domain-containing proteins (Figure 3). We showed that the *M. fijiensis* homolog with highest similarity to the *E. festucae* salicylate hydroxylase is among the genes with higher expression in infected leaf tissue[46]. Since salicylic acid is important for plant defense responses[238], and salicylate hydroxylase converts salicylate to catechol[239], production of salicylate hydroxylase by *M. fijiensis* could be a strategy to dampen the salicylic acid defense pathway. In

Aspergillus oryzae, the hydrophobic surface binding protein HsbA is secreted from fungal tissue and promotes degradation of the hydrophobic compound polybutylene succinate-co-adipate (PBSA) by recruiting cutinase[232]. PBSA is structurally similar to waxes in the plant cuticle[233,286]; therefore, it is thought that HsbA and related proteins may recruit cutinases important for plant pathogenicity[233]. CFEM domains are cysteine-rich domains that have been identified in proteins important for pathogenicity[231].

We observed that more genes encoding amino acid transporters, oligopeptide transporters, peptidases, proteases and proteinases had higher expression in infected leaf tissue compared to culture medium (Figure 4). One proteinase gene was among the list of 20 genes with highest expression in infected leaf tissue compared to culture medium (Table 1). Conversely, more peptidase or proteinase inhibitor genes had lower expression in infected leaf tissue (Figure 4). These results may be expected if the in planta environment is lower in nitrogen than the PDB culture medium. In other plant pathogenic fungi and bacteria, low nitrogen conditions are known to induce expression of virulence genes, and it is thought that this reflects nitrogen limitation for the pathogen during its interaction with the plant host[287,288]. Likewise, more sugar transporters had higher expression in infected leaf tissue compared to medium (Figure 4). This may reflect a higher sugar concentration in the PDB medium due to added dextrose, and thus a need for higher expression of sugar transporters in the plant environment compared to a relatively sugar-rich medium.

We identified four gene clusters with higher expression in the infected leaf tissue that encode one or more hypothetical proteins with Domain of Unknown Function (DUF) 3328 (Figure 9). Two DUF3328-encoding genes were among the list of genes with highest

expression in the infected leaf tissue compared to medium (Table 1). Although the function of this domain is not known, some studies have associated changes in DUF3328 gene expression with the fungal sexual cycle. In *Sordaria macrospora*, a DUF3328 gene cluster was identified as up-regulated in wild-type protoperithecia compared to protoperithecia of a mutant for sexual reproduction[253]. In *Podospora anserina*, genes encoding DUF3328 proteins were identified as having higher expression in mat- compared to mat+ strains[254]. In natural infections where both mating types of *M. fijiensis* are present, pseudothecia develop in infected banana tissue, and ascospores are considered the primary means of pathogen spread[5]. Therefore, although the role of DUF3328 proteins remains unclear, one hypothesis is that they may be important for pseudothecia development in this fungus.

We identified genes predicted to encode small, cysteine-rich, secreted proteins, since these are features common in fungal effectors[105,106]. We found 30 such genes with higher expression in infected leaf tissue, and 10 with higher expression in medium[46]. Two of these were among the list of genes with highest expression in infected leaf tissue compared to culture medium (Table 1), and one was among the list of genes with highest expression in culture medium compared to infected leaf tissue (Table 2). Among the proteins encoded by the 30 genes with higher expression in infected leaf tissue were proteins implicated in other species as being important for pathogenicity, including cutinases and proteins with PR-1-like and CFEM domains[231,236,237,289]. For 16 of the putative effector genes with higher expression in the infected leaf tissue, all homologs were restricted to species within the Mycosphaerellaceae. This result is consistent with the observation that many fungal effectors have a restricted phylogenetic distribution[108]. Although it has previously been shown that the

M. fijiensis genome encodes homologs of the *C. fulvum* effectors Avr4, Ecp2 and Ecp6[102,104], none of these genes had higher expression in infected leaf tissue compared to culture medium[46]. It is possible that these effectors are important for pathogenicity at different time points, such as the biotrophic phase, than for the necrotrophic phase assayed in this experiment.

Finally, we were able to provide support for the hypothesis that *M. fijiensis* may have dispensable chromosomes, and that some of these may be important for pathogenicity[54,55]. Ohm et al. showed that the *M. fijiensis* genome has 14 scaffolds with similar characteristics to the dispensable chromosomes from *M. graminicola*[54]. Compared to the rest of the genome, these 14 scaffolds are small, have a lower G+C content, different codon usage, a high proportion of repetitive DNA, a low gene density, and a lower proportion of genes encoding proteins with PFAM domains[54]. We identified two of these putative dispensable scaffolds (15 and 21) for which 31% and 52 % of the genes, respectively, on the scaffolds had higher expression in infected leaf tissue, and no genes on these scaffolds had higher expression in culture medium. It is still unknown whether these scaffolds correspond to dispensable chromosomes, but our data suggest that they may play a role in pathogenicity. We also identified two more scaffolds (13 and 20) for which no transcripts were detected. PCR assays were unable to amplify the three genes tested on scaffold 13, and two of three genes on scaffold 20 in our isolate 14H1-11A, whereas amplification was possible for all of these genes in isolate CIRAD86, for which the genome sequence is publicly available (NCBI Genome ID 10962)[160] (Figure 11). Considering our transcriptome and PCR data, we have no evidence that scaffold 13 is present in our isolate, though PCR data show that at least part of

scaffold 20 remains. Our data support the hypothesis that scaffold 13 corresponds to a dispensable chromosome that is not required for survival or pathogenicity.

Conclusions

This study is the first to identify candidate pathogenicity genes of the fungus *Mycosphaerella fijiensis*, based on transcriptome data from this fungus during its necrotrophic phase of infection of banana, compared to during saprophytic growth in culture medium. We showed that gene clusters predicted to synthesize a non-ribosomal peptide and a fusicoccane, as well as many types of genes encoding proteins commonly involved in secondary metabolism, such as cytochrome P450s and short-chain dehydrogenases, have higher expression in infected leaf tissue. We identified several other types of genes with higher expression in infected leaf tissue, including genes predicted to encode salicylate hydroxylase-like proteins, hydrophobic surface binding proteins, CFEM domain-containing proteins, amino acid and sugar transporters, and proteins with DUF3328 domains. Furthermore, we identified two putative dispensable scaffolds with a large proportion of genes with higher expression in infected leaf tissue, suggesting that these scaffolds may play a role in pathogenicity. We also identified two other scaffolds for which no transcripts were detected in either condition, and PCR assays support the hypothesis that at least one of these scaffolds corresponds to a dispensable chromosome that is not required for survival or pathogenicity. Together, these results suggest exciting avenues of further research for an important and understudied pathogen.

Methods

Fungal Cultures

M. fijiensis isolate 14H1-11A, isolated from the FHIA research station in La Lima, Honduras, was kindly provided by Dr. Jean Ristaino (North Carolina State University) and was routinely cultured on Potato Dextrose Agar (PDA) (Difco). Conidial production was induced as described[45,207].

Banana Tissue Culture

'Grand Nain' banana tissue culture plants were obtained from Miguel Muñoz (Dole Food Company) and were maintained on modified Murashige and Skoog medium as described[45].

Inoculation of Plants and Flasks

Rooted banana plants grown in modified Murashige and Skoog medium were transferred to potting mix under greenhouse conditions. When they were approximately 20 cm in height, plants were transferred to an incubator at 25 °C under cool-white fluorescent light on a 18h light/6h dark photoperiod. Plants were inoculated by atomizing with $5.2x10^4$/mL conidia in 0.5% Tween 20 as described[45]. Plants were covered with clear plastic bags to maintain high humidity conditions for 1 week. Symptomatic banana leaf tissue was harvested at 6 weeks post-inoculation by flash-freezing in liquid nitrogen.

For growth in culture, 50 mL flasks of Potato Dextrose Broth (PDB) (Difco) were inoculated with 10 µL of $1.3x10^6$/mL conidia and incubated in a rotary shaker at 150 rpm in

the dark at 25-30 °C. After one week, mycelium was harvested by filtering through Miracloth, blotting dry, and flash freezing in liquid nitrogen.

cDNA Library Construction and Illumina HiSeq Sequencing

RNA was isolated using the Spectrum Plant Total RNA kit (Sigma). Samples were DNase treated with DNase I (Roche). Total RNA sequencing was conducted at the Genomic Sciences Laboratory, North Carolina State University. RNA quality was confirmed by gel electrophoresis and an Agilent Bioanalyzer. Strand-specific libraries were created using the NEBNext Ultra Directional library prep kit (New England BioLabs). Single-end 125-base reads were generated using an Illumina HiSeq 2500 platform. Average sequencing yield was 32 million reads per sample.

Identification of Differentially Expressed Genes

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to verify the quality of RNA-Seq reads. Illumina Truseq adapter sequences and low-quality bases were trimmed using CutAdapt v1.7 with a quality cutoff of 20 and a minimum sequence length of 36[209].

Sequences were mapped from each sample to both the banana genome, *Musa acuminata* subsp. *malaccensis* double-haploid Pahang[34,210], and to the *M. fijiensis* genome[55,160] using Tophatv2.0.9[211]. Gene expression levels were determined using HTSeqv0.6.0[212] and the gene annotations available from the Joint Genome Institute (JGI). Differentially expressed genes were identified with an adjusted p-value < 0.01, a log2FC

value > 3, and a basemean (mean of normalized read counts) > 10, using the program DESeq2 v1.4.5[46,213]. Principal component analysis and a volcano plot were done to verify that biological replicates were similar in expression pattern to each other (Figure S1) and to visualize the distribution of differentially expressed transcripts (Figure S2).

Validation by RT-qPCR

Four differentially expressed genes were chosen for validation of RNA-Seq results. Genes were chosen with a basemean value of >150 in both conditions, a log2FC value >3, and an adjusted p-value of < 0.01, to select differentially expressed genes that were well-expressed in both conditions.

cDNA was synthesized with iScript Select reverse transcriptase (BioRad). qPCR reactions were performed with iQ SYBR Green SuperMix, in a CFX Connect Real-Time System machine. The program was done with an initial denaturation at 95°C for 2 minutes, followed by 50 cycles of 95°C for 10 seconds, 57°C for 30 seconds, 72°C for 30 seconds with a plate read, and then 78°C for 30 seconds with a plate read. A melt curve was also done to verify that a single product was formed for each reaction.

Each gene of interest was normalized against two *M. fijiensis*-specific reference genes having the same reaction efficiency as the gene of interest, and fold-change in infected leaf tissue was calculated compared to when grown in liquid medium, using the $2^{-\Delta\Delta C_T}$ method[290]. Primer sequences are provided by Noar and Daub[46].

Prediction of Differentially Expressed Gene Functions

Blast and Conserved Domain Analysis

For each of the 802 differentially expressed genes, a blastp search was performed using the non-redundant protein sequence database from NCBI (July 2015). NCBI's Conserved Domain Database[168] was used to predict domains from each protein sequence[46]. To identify the closest *M. fijiensis* homolog of salicylate hydroxylase, a blastp search was done with the *Epichloë festucae* salicylate hydroxylase sequence (Accession = AIY25489.1)[240] against all the *M. fijiensis* isolate CIRAD86 predicted protein sequences in NCBI's non-redundant protein sequence database (6/1/2016)[46].

CAZy and GO Annotations and Enrichment Analysis

Blast2GO Basic v3.1.3 was used to do blastx v2.2.32+ searches of *M. fijiensis* gene catalog coding sequences from JGI against the NCBI non-redundant protein sequences (nr database), using the default parameters. InterProScan, mapping, and annotation steps were all done with the default parameters. Using the resulting annotation data, Fisher's Exact Test was used to identify GO terms that were over-represented from sequences having higher expression in infected leaf tissue or in medium, compared to their representation in the total set of genes from the *M. fijiensis* genome.

Genome annotations, including those based on the CAZy database, were downloaded from NCBI for *M. fijiensis* (Genome ID: 10962). Genes with CAZy annotations were compared with the list of differentially expressed genes.

Effector Protein Predictions

M. fijiensis Gene Catalog protein models were downloaded from the published genome sequence from JGI, in FASTA format. Signal peptides were predicted from these sequences using SignalP 4.1[262], with the organism group set to Eukaryotes, and the D-cutoff value set to the default. SignalP 4.1 was also used to make predictions of the mature protein sequences once signal peptides are cleaved. Mature protein sequences were then sorted by length to identify sequences less than 300 amino acids in length. These short, secreted mature peptides were then sorted by cysteine content to identify sequences containing at least four cysteine residues. Differentially expressed genes predicted to be short, secreted, and cysteine rich were noted by Noar and Daub[46]. Since *M. fijiensis* homologs of the *C. fulvum* effectors Avr4, Ecp2, and Ecp6 have previously been described[102,104], the expression of the corresponding genes was also reported. JGI gene IDs for *MfAvr4*, *MfEcp2* and *Ecp6* are Mycfi1.estExt_fgenesh1_pg.C_60009, gw1.3.823.1, and estExt_fgenesh1_kg.C_70127, respectively.

Phylogenetic Tree of Fusicoccadiene Synthases

Once homologs were identified of the *M. fijiensis* fusicoccadiene synthase using blastp, the *M. fijiensis* sequence and its top 49 homologs[46] were aligned using MUSCLE v3.8.31[201] in Mesquite v3.04[202]. ModelGenerator v0.85[203] was used to select the best substitution model using the Aikaike and and Bayesian Information Criteria. RaxmlGUI v1.3.1[205] was used to create a maximum likelihood tree using the JTT+I+G substitution model[291] with slow bootstrap, the autoMRE function, and no outgroup.

Identification of Gene Clusters

To identify clusters of differentially expressed genes and to analyze how differentially expressed genes are distributed across the *M. fijiensis* genome scaffolds, scaffold number and position on the scaffold was downloaded from JGI, and matched to each gene ID. To identify clusters of differentially expressed genes, loci were identified that contain at least three adjacent, similarly differentially expressed genes (Figures 7, 9, S5, S6)[46].

Distribution of Differentially Expressed Genes Across Genome Scaffolds

To analyze the distribution of differentially expressed genes across scaffolds, differentially expressed genes were sorted according to scaffold number, and the number of differentially expressed genes on a scaffold was compared to the total number of genes on the scaffold (Figure 10). For proteins encoded by genes on scaffolds 15 and 21, blastp searches were done using NCBI's non-redundant protein database (4/26/2016) to identify conserved domains and the top 10 homologs by bitscore of each sequence of interest (maximum E-value=$1x10^{-5}$) (Figure S7)[46].

To determine whether scaffolds 13 and 20 are present in isolate 14H1-11A, three genes were chosen on scaffold 13 and scaffold 20 for PCR assays. Genomic DNA was extracted from isolate 14H1-11A and CIRAD86 using DNeasy Plant Mini Kit (Qiagen), according to manufacturer's instructions. PCR amplifications were done using OneTaq 2X Master Mix with Standard Buffer (NEB), according to manufacturer's instructions with 35 cycles, with primer sequences, annealing temperatures, extension times, and expected

product sizes indicated by Noar and Daub[46]. To obtain the sequence of the gene on scaffold 20 which amplified for both isolates (accession XM_007934411.1), the high fidelity DNA polymerase iProof (Bio-Rad) was used to amplify this gene from both isolates, according to manufacturer's instructions with an annealing temperature of 60°C, an extension time of 20 seconds, and 35 cycles. The band was gel purified from both isolates using a QIAquick gel extraction kit (Qiagen), and Sanger sequencing was done by Eton Bioscience using the forward primer[46]. A blastn search was done using NCBI's nucleotide collection database for each sequence against the *M. fijiensis* genome to verify amplification of the correct gene (Table 4). A blastx search was also done with the sequence from isolate 14H1-11A to determine how changes in nucleotide sequence in this isolate affect the sequence of the predicted protein. The program Chimera v1.10.2[292] was used to display the sequence alignment and the degree of conservation for each residue in the alignment.

List of Abbreviations

PDB: Potato Dextrose Broth; NCBI: National Center for Biotechnology Information; CAZy: Carbohydrate-Active Enzymes; GO: Gene Ontology; log2FC: $\log_2$ fold change; DUF3328: Domain of Unknown Function 3328; CFEM: Common in Fungal Extracellular Membrane; HsbA: Hydrophobic Surface Binding Protein A; PR-1: Pathogenesis-Related Protein 1; MFS: Major Facilitator Superfamily; FDR: False Discovery Rate; NRPS: non-ribosomal peptide synthase; ABC: ATP-Binding Cassette; PBSA: polybutylene succinate-co-adipate; BAP: 6-benzylaminopurine; JGI: Joint Genome Institute

Declarations

Acknowledgements

We thank Gert Kema, Wageningen University and Research Centre, The Netherlands for providing the *M. fijiensis* isolate CIRAD86, and Jean Ristaino, North Carolina State University for isolate 14H1-11A. We also thank Elizabeth Thomas from North Carolina State University for consultation with qPCR assays.

Funding

Funding for this study was provided by Dole Food Company (to MED), a National Science Foundation Graduate Research Fellowship (to RDN), and a Molecular Biotechnology training grant traineeship to RDN from the National Institutes of Health. The funding bodies played no role in the design of the study, the collection, analysis, or interpretation of data, or in the writing of the manuscript.

Availability of Data and Material

The transcriptome dataset supporting the conclusions of this article is available in the NCBI database, with the Sequence Read Archive accession SRP075820, at http://www.ncbi.nlm.nih.gov/sra/SRP075820. All other datasets supporting the conclusions of this article are included within the article and its additional files.

Authors' Contributions

RDN conceived the study. RDN and MED planned the experiments. RDN did the laboratory work. RDN and MED analyzed the data. RDN and MED wrote the manuscript. Both authors read and approved the final manuscript.

Competing Interests

The authors declare that they have no competing interests.

Consent for Publication

Not applicable

Ethics Approval and Consent to Participate

Not applicable.

Open Access

Additional Files

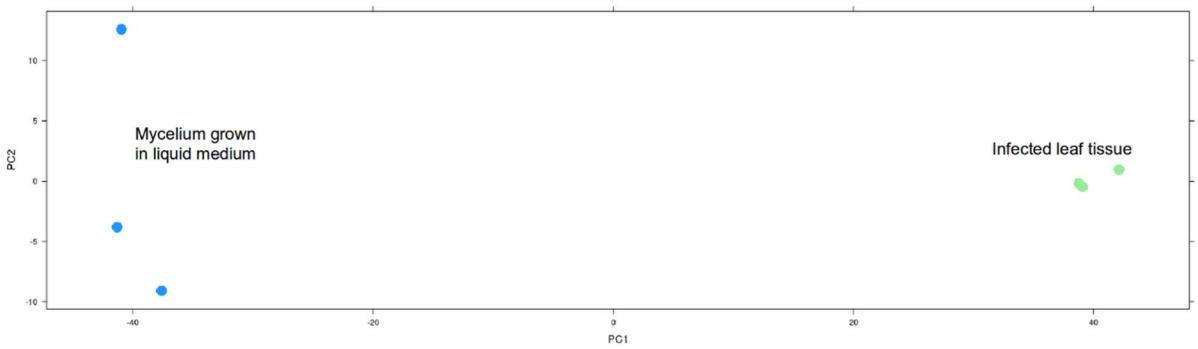


Figure S1. Two-dimensional principal component analysis of infected leaf samples versus mycelium grown in liquid culture.
Infected leaf samples are shown as green dots, and samples of mycelium grown in PDB medium are shown as blue dots.

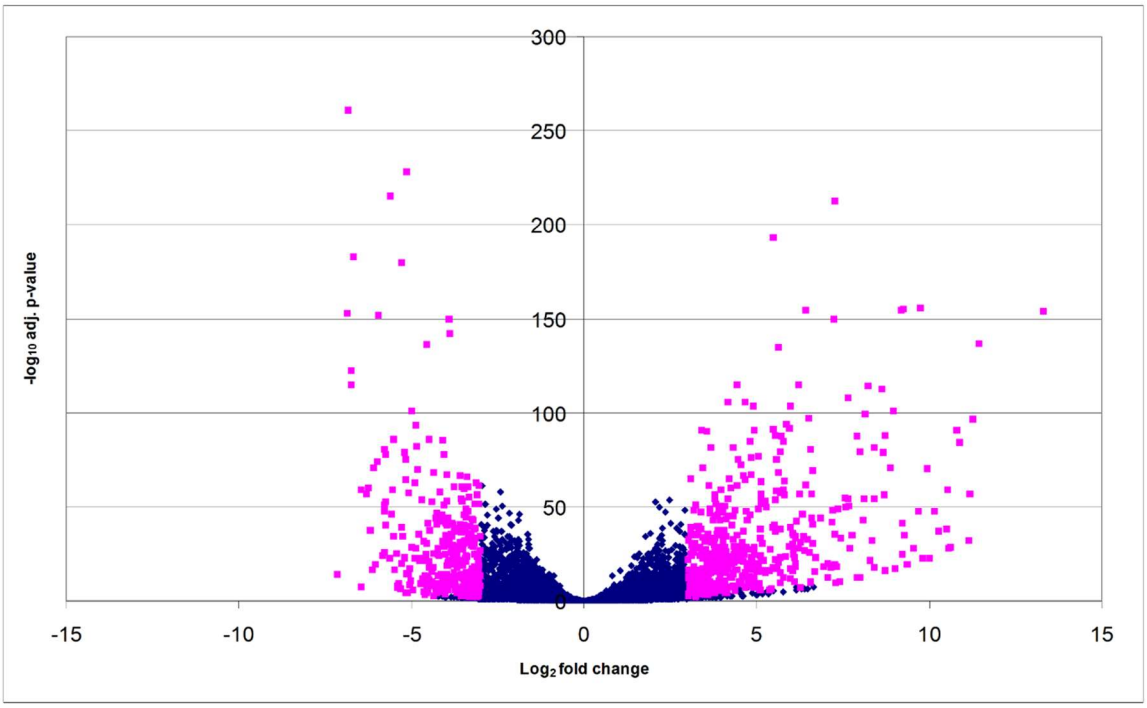


Figure S2. Volcano plot showing the distribution of differentially expressed sequences. The horizontal axis shows the $\log_2$ fold change for expression in infected leaf tissue versus the fungus grown in liquid medium, and the vertical axis shows the -log10 (p-value). Each transcript is represented by a dot, which is colored pink for each of the differentially expressed transcripts listed by Noar and Daub[46], and colored blue for transcripts that are not differentially expressed.

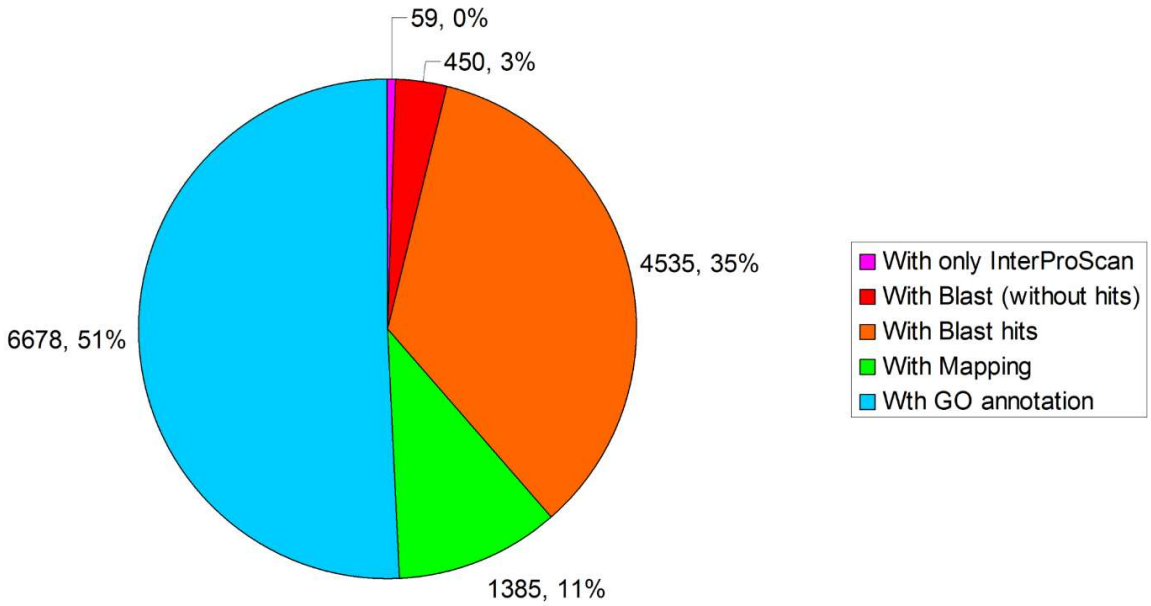


Figure S3. Annotation results of all predicted genes in the *M. fijiensis* genome, using Blast2GO.

Red = Sequences for which blast analysis was done, but no hits were found. Orange = Sequences for which blast analysis was done and hits were found, but no GO terms were mapped to the hits. Green = Sequences for which blast analysis was done and GO terms were successfully mapped to the hits. Blue = Query sequences to which GO terms were successfully annotated. Pink = Sequences for which only the InterProScan results were obtained.

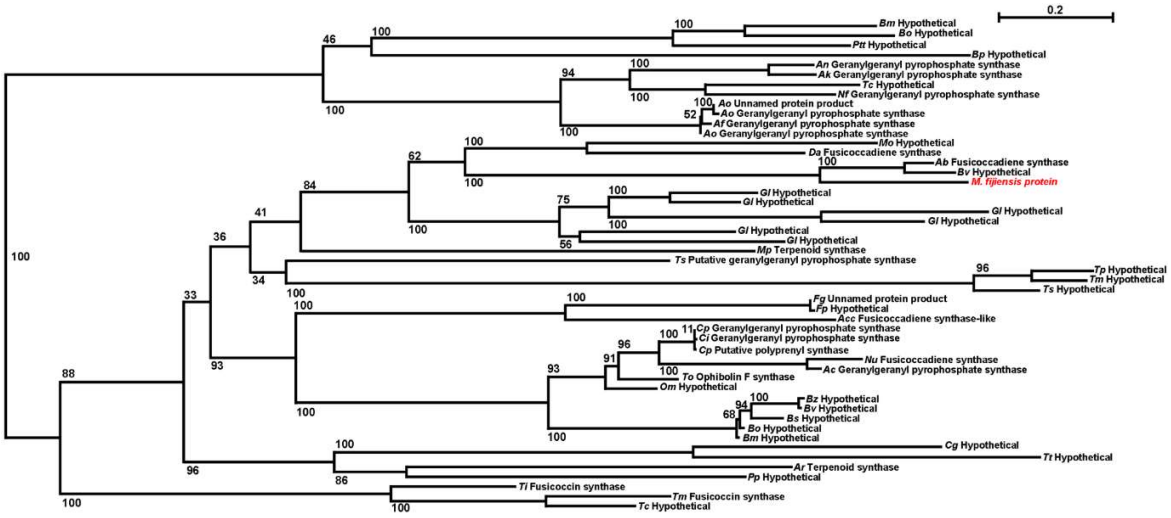


Figure S4. Phylogenetic tree of fusicoccadiene synthase protein sequences.
A maximum likelihood tree was created of the *M. fijiensis* fusicoccadiene synthase sequence and its top 50 hits using blastp with the non-redundant protein sequence database on NCBI. Bootstrap values are indicated on the tree, and the scale bar of branch lengths indicate substitutions per site. A description of each blast hit is shown, along with an abbreviation for species. Ab = *Alternaria brassicicola*; Ac = *Acremonium chrysogenum*; Af = *Aspergillus flavus*; Ak = *Aspergillus kawachii*; An = *Aspergillus niger*; Ao = *Aspergillus oryzae*; Ar = *Aspergillus ruber*; Bm = *Bipolaris maydis*; Bo = *Bipolaris oryzae*; Bp = *Baudoinia panamericana*; Bs = *Bipolaris sorokiniana*; Bv = *Bipolaris victoriae*; Bz = *Bipolaris zeicola*; Cg = *Chaetomium globosum*; Ci = *Coccidioides immitis*; Cp = *Coccidioides posadasii*; Da = *Diaporthe amygdali*; Fg = *Fusarium graminearum*; Fp = *Fusarium pseudograminearum*; Gl = *Gymnopus luxurians*; Mo = *Magnaporthe oryzae*; Mp = *Macrophomina phaseolina*; Nf = *Neosartorya fischeri*; Nu = *Neosartorya udagawae*; Om = *Oidiodendron maius*; Pp = *Pseudogymnoascus pannorum*; Ptt = *Pyrenophora teres* f. *teres*; Tc = *Talaromyces cellulolyticus*; Ti = *Talaromyces islandicus*; Tm = *Talaromyces marneffei*; To = *Tolypocladium ophioglossoides*; Ts = *Talaromyces stipitatus*; Tt = *Thielavia terrestris*.

Figure S5. Gene clusters up-regulated in infected leaf tissue compared to liquid medium.
The description of each gene as determined by blastp of the corresponding protein is shown along with its $\log_2$ fold change of expression in infected leaf tissue versus expression in liquid medium. Black bars are proportional to the $\log_2$ fold change. Gene expression differences that are significant at $p < 0.01$ are shown with two asterisks above the corresponding bar, and those significant at $p < 0.05$ are shown with a single asterisk. NE = no expression detected. Genes in the putative gene cluster are indicated by blue boxes, and genes flanking the cluster are indicated by gray boxes. A) Gene cluster on scaffold 7; B) Gene cluster on scaffold 21; C) Gene cluster on scaffold 6; D) Gene cluster on scaffold 10; E) Gene cluster on scaffold 1; F) Gene cluster on scaffold 6.

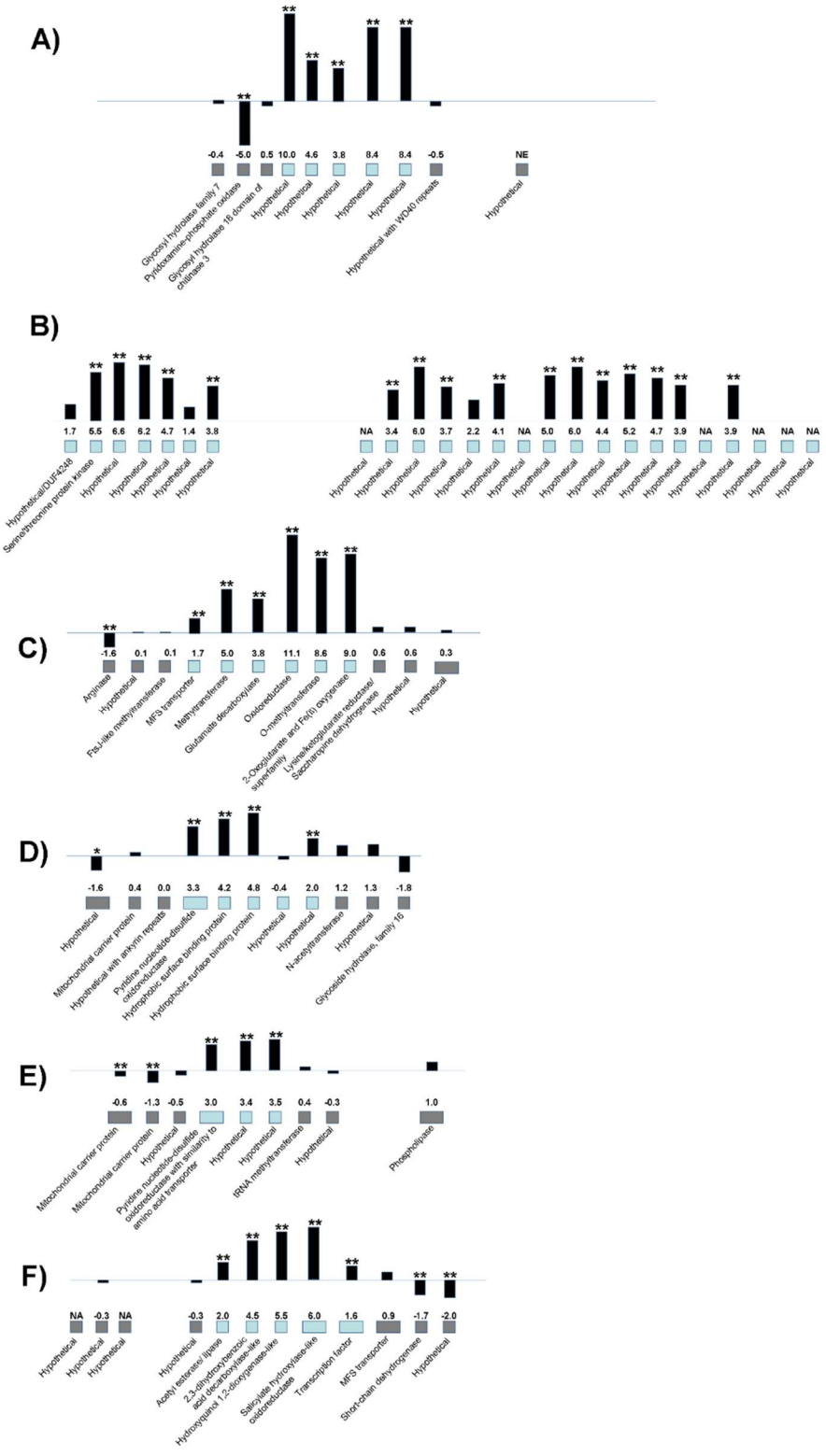
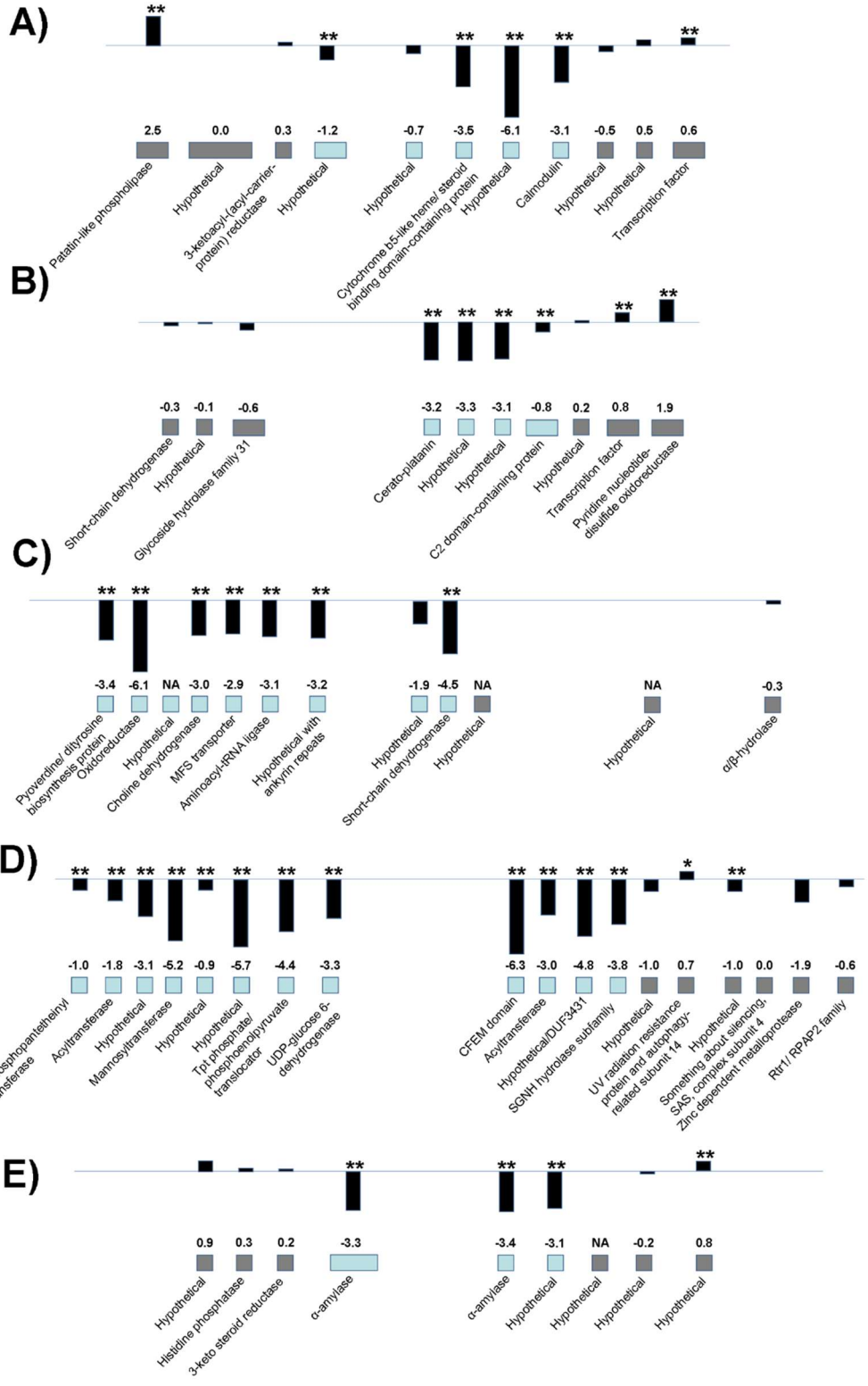

A)
-0.4 -5.0 0.5 10.0 4.6 3.8 8.4 8.4 -0.5 NE
Glycosyl hydrolase family 7
Pyridoxamine-phosphate oxidase
Glycosyl hydrolase 18 domain of chitinase 3
Hypothetical
Hypothetical with WD40 repeats
B)
1.7 5.5 6.6 6.2 4.7 1.4 3.8 NA 3.4 6.0 3.7 2.2 4.1 NA 5.0 6.0 4.4 5.2 4.7 3.9 NA 3.9 NA NA NA
Hypothetical/DUF4348
Serine/threonine protein kinase
Hypothetical
C)
-1.6 0.1 0.1 1.7 5.0 3.8 11.1 8.6 9.0 0.6 0.6 0.3
Arginase
Hypothetical
Fsu-like methyltransferase
MFS transporter
Methyltransferase
Glutamate decarboxylase
Oxidoreductase
O-methyltransferase
2-Oxoglutarate and Fe(II) oxygenase superfamily
Lysine/ketoglutarate reductase/ Saccharopine dehydrogenase
D)
-1.6 0.4 0.0 3.3 4.2 4.8 -0.4 2.0 1.2 1.3 -1.8
Hypothetical
Mitochondrial carrier protein
Hypothetical with ankyrin repeats
Pyridine nucleotide-disulfide oxidoreductase
Hydrophobic surface binding protein
N-acetyltransferase
Glycoside hydrolase, family 16
E)
-0.6 -1.3 -0.5 3.0 3.4 3.5 0.4 -0.3 1.0
Mitochondrial carrier protein
Pyridine nucleotide-disulfide oxidoreductase with similarity to amino acid transporter
tRNA methyltransferase
Phospholipase
F)
NA -0.3 NA -0.3 2.0 4.5 5.5 6.0 1.6 0.9 -1.7 -2.0
Hypothetical
Acetyl esterase/ lipase
2,3-dihydroxybenzoic acid decarboxylase-like
Hydroxyquinol 1,2-dioxygenase-like
Salicylate hydroxylase-like oxidoreductase
Transcription factor
MFS transporter
Short-chain dehydrogenase

Figure S6. Gene clusters down-regulated in infected leaf tissue compared to liquid medium.
The description of each gene as determined by blastp of the corresponding protein is shown along with its $\log_2$ fold change value of expression in infected leaf tissue versus expression in liquid medium. Black bars are proportional to the $\log_2$ fold change value. Gene expression differences that are significant at $p < 0.01$ are shown with two asterisks above the corresponding bar, and those significant at $p < 0.05$ are shown with a single asterisk. NE = no expression detected. Genes in the putative gene cluster are indicated by blue boxes, and genes flanking the cluster are indicated by gray boxes. A) Gene cluster on scaffold 4; B) Gene cluster on scaffold 7; C) Gene cluster on scaffold 8; D) Gene cluster on scaffold 3; E) Gene cluster on scaffold 2


A)
**
**
**
**
**
**
**
2.5
0.0
0.3
-1.2
-0.7
-3.5
-6.1
-3.1
-0.5
0.5
0.6
Patatin-like phospholipase
Hypothetical
3-ketoacyl-(acyl-carrier-protein) reductase
Hypothetical
Hypothetical
Cytochrome b5-like heme/ steroid binding domain-containing protein
Hypothetical
Calmodulin
Hypothetical
Hypothetical
Transcription factor
B)
**
**
**
**
**
**
-0.3
-0.1
-0.6
-3.2
-3.3
-3.1
-0.8
0.2
0.8
1.9
Short-chain dehydrogenase
Hypothetical
Glycoside hydrolase family 31
Cerato-platanin
Hypothetical
Hypothetical
C2 domain-containing protein
Hypothetical
Transcription factor
Pyridine nucleotide-disulfide oxidoreductase
C)
**
**
**
**
**
**
**
-3.4
-6.1
NA
-3.0
-2.9
-3.1
-3.2
-1.9
-4.5
NA
NA
-0.3
Pyoverdine/ dityrosine biosynthesis protein Oxidoreductase
Hypothetical
Choline dehydrogenase
MFS transporter
Aminoacyl-tRNA ligase
Hypothetical with ankyrin repeats
Hypothetical
Short-chain dehydrogenase
Hypothetical
Hypothetical
α/β-hydrolase
D)
**
**
**
**
**
**
**
**
**
**
**
**
*
**
-1.0
-1.8
-3.1
-5.2
-0.9
-5.7
-4.4
-3.3
-6.3
-3.0
-4.8
-3.8
-1.0
0.7
-1.0
0.0
-1.9
-0.6
4'-phosphopantetheinyl transferase
Acyltransferase
Hypothetical
Mannosyltransferase
Hypothetical
Hypothetical
Tpt phosphate/ phosphoenolpyruvate translocator
UDP-glucose 6-dehydrogenase
CFEM domain
Acyltransferase
Hypothetical/DUF3431
SGNH hydrolase subfamily
Hypothetical
UV radiation resistance protein and autophagy-related subunit 14
Hypothetical
Something about silencing, SAS, complex subunit 4
Zinc dependent metalloprotease
Rtr1/ RPAP2 family
E)
**
**
**
**
0.9
0.3
0.2
-3.3
-3.4
-3.1
NA
-0.2
0.8
Hypothetical
Histidine phosphatase
3-keto steroid reductase
α-amylase
α-amylase
Hypothetical
Hypothetical
Hypothetical
Hypothetical

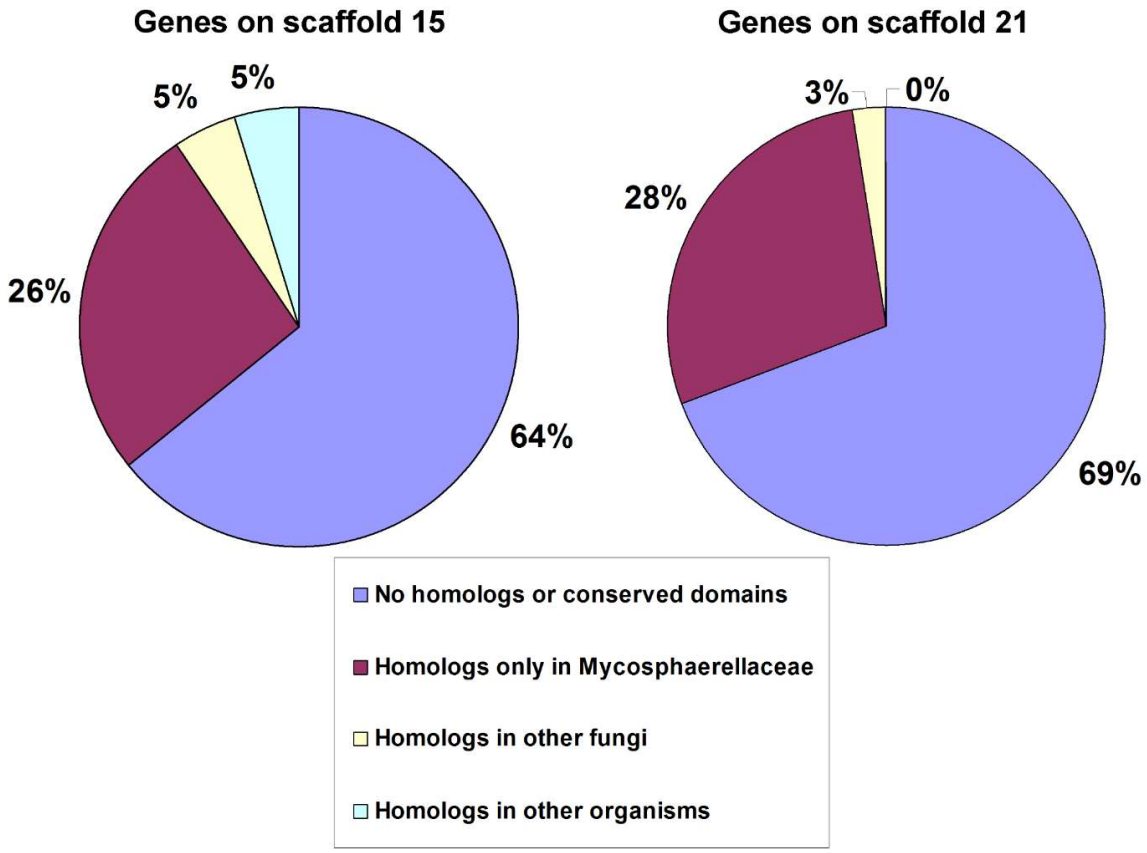


Figure S7. Homologs of genes on scaffolds 15 and 21 of the *M. fijiensis* genome. For each gene on scaffolds 15 and 21, a blastp search was done using the NCBI non-redundant protein sequence database to identify homologs and conserved domains. The charts indicate the percent of genes on the scaffolds with: no homologs or conserved domains identified (dark blue), homologs present in species within Mycosphaerellaceae (purple), homologs present in other fungi but not in non-fungal organisms (yellow), and homologs present in non-fungal organisms (light blue).

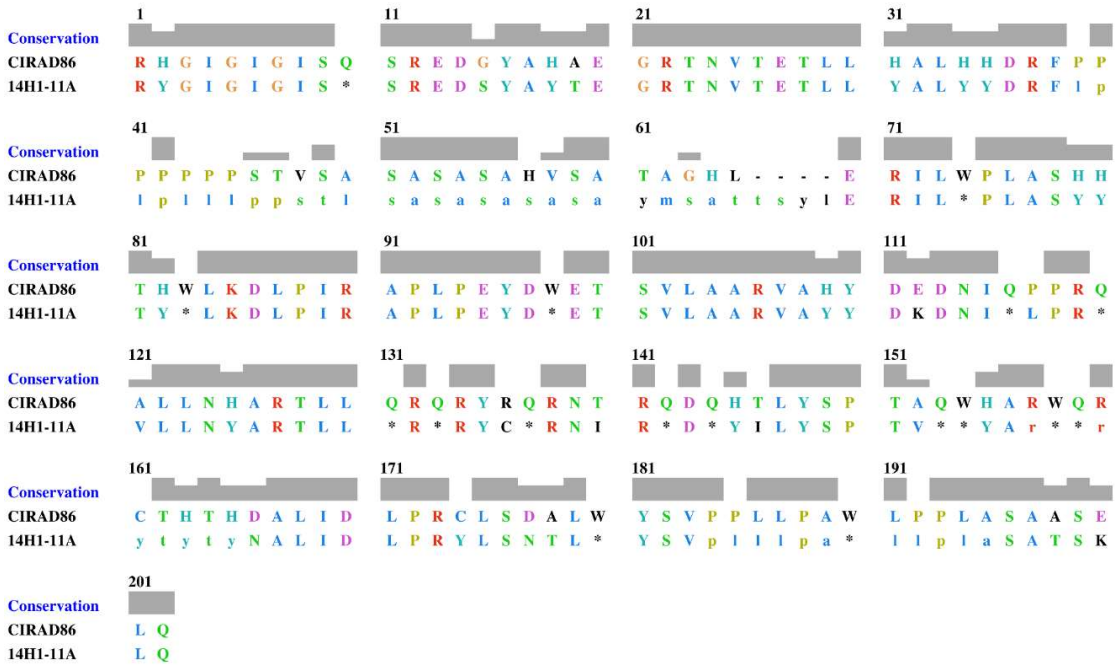


Figure S8. Translated nucleotide sequence alignment of Scaffold 20 hypothetical gene in CIRAD86 and 14H1-11A isolates.
Chimera 1.10.2 was used to display the degree of conservation for each residue within the alignment of translated nucleotide sequences from the scaffold 20 hypothetical gene from isolates CIRAD86 and 14H1-11A. Amino acid residues corresponding to translated nucleotide sequences are color coded according to the ClustalX[293] color scheme. Asterisks indicate stop codons in the 14H1-11A sequence. Clustal histogram bars are used to indicate sequence conservation, so that larger histogram bars correspond to amino acid residues with more similar physio-chemical properties.

CHAPTER 4

A Polyketide Synthase Gene Cluster Associated with the Sexual Reproductive Cycle of the Banana Pathogen, *Mycosphaerella fijiensis*

Roslyn D. Noar, Deyu Xie, Morgan E. Carter, Dongming Ma, and Margaret E. Daub

Noar, R. D., Xie, D., Carter, M. E., Ma, D., and Daub, M. E. (2016). A Polyketide Synthase Gene Cluster Associated with the Sexual Reproductive Cycle of the Banana Pathogen, *Mycosphaerella fijiensis*. *Molecules*, special issue "Polyketides," submitted.

Abstract

Disease spread of *Mycosphaerella fijiensis*, causal agent of the black Sigatoka disease of banana, depends on ascospores produced through the sexual reproductive cycle. We used phylogenetic analysis to identify *M. fijiensis* homologs (PKS8-4 and Hybrid8-3) to the PKS4 polyketide synthases (PKS) from *Neurospora crassa* and *Sordaria macrospora* involved in sexual reproduction. These sequences also clustered with lovastatin, compactin, and betaenone-producing PKS enzymes. Transcriptome analysis showed that both *Hybrid8-3* and *PKS8-4* have higher expression in infected leaf tissue compared to in culture. Domain analysis showed that PKS8-4 is more similar than Hybrid8-3 to PKS4. *PKS8-4-GFP* transcriptional fusion mutants showed expression of GFP in developing pseudothecia, consistent with a role in sexual reproduction. A disruption mutant of *pks8-4* retained normal

pathogenicity on banana and produced pseudothecia when paired with a wild-type *M. fijiensis* isolate of the opposite mating type. GC-MS profiling identified significant changes in non-polar metabolites between the wild type and *pks8-4* mutant including changes in saturated fatty acid methyl esters and alkene and alkane derivatives. Overall, our study shows that a homolog of *PKS4* is expressed in developing mating structures in *M. fijiensis*. Future studies are needed to determine the precise role of this biosynthetic pathway on sexual reproduction.

Introduction

Black Sigatoka, caused by the Dothideomycete fungus *Mycosphaerella fijiensis*, is considered the most economically damaging disease of banana[5]. Fungicide sprays to control this disease account for up to 30% of the total production cost[5]. If not treated with fungicides, yield losses range from 20-80% depending on climate and growing conditions[5,294]. Because black Sigatoka is a major limiting factor to banana production, a better understanding of the molecular basis of the plant-fungal interaction is needed so that new control strategies can be developed.

Many fungi, including close relatives of *M. fijiensis*, produce polyketide secondary metabolites with various roles in the fungal life cycle[57,295]. Some polyketides promote pathogenicity in plants and animals[101,227,296,297]. Other polyketides have antimicrobial properties and are thought to be involved in competition with other microbes in the environment[161,295,298]. In many cases, polyketide production is tightly linked with and may play important roles in cell differentiation and developmental processes, such as the

development of sexual and asexual structures[295]. Polyketide production has long been thought to be important for the biology of *M. fijiensis*, and several studies have attempted to identify polyketides and other secondary metabolites produced by this fungus[60,61,228]. These studies have primarily focused on toxic secondary metabolites that could contribute to the leaf necrosis observed in black Sigatoka disease.

The first phytotoxic secondary metabolite identified from *M. fijiensis* was fijiensin (also called vermistatin)[60], which may be produced from a polyketide biosynthetic pathway[71]. Fijiensin was identified from fungal cultures grown in modified M-1-D liquid medium, and was shown to induce necrotic lesions on both resistant and susceptible banana cultivars, but not non-host species[60]. Additional phytotoxic secondary metabolites were later identified from *M. fijiensis* grown in the same medium. These metabolites include 2,4,8-trihydroxytetralone (2,4,8-THT), juglone, and 4-hydroxyscytalone[61], which are melanin shunt metabolites derived from a polyketide biosynthetic pathway[156,216]. Host selectivity on susceptible versus resistant banana cultivars was detected for 2,4,8-THT[61]. Application of the fungicide tricyclazole blocks melanin biosynthesis and leads to an accumulation of melanin shunt pathway metabolites such as 2,4,8-THT[74,299]. Tricyclazole treatment of *M. fijiensis*-infected banana plants resulted in larger necrotic leaf spots[74]; therefore, 2,4,8-THT was initially believed to be an important polyketide pathogenicity factor of *M. fijiensis.*

Melanin itself is an important pathogenicity factor for many fungi. It is crucial for penetration of cells by some fungi that produce appressoria[78,79,183]. It also is thought to aid in pathogenesis by promoting hyphal tip protrusion and protecting against reactive oxygen species[80,300]. Despite early suggestions that melanin and melanin shunt metabolites may be

important for pathogenicity, disruption of the polyketide synthase of this pathway did not affect pathogenicity of *M. fijiensis*[5].

Sequencing of the *M. fijiensis* genome has recently allowed bioinformatics predictions of polyketide synthase (PKS) genes from this fungus[45,54,301]. In previous work we showed that the *M. fijiensis* PKS8-2, PKS10-2, and PKS2-1 sequences have homology with PKS sequences in other fungi that produce fumonisin, solanapyrone, and alternapyrone, respectively[45]. Fumonisin is a pathogenicity factor produced by *Fusarium* spp. that works by perturbing sphingolipid biosynthesis[171,194]. Solanapyrone is a phytotoxic polyketide produced by *Alternaria solani*, but there have been conflicting data on the role of solanapyrone in pathogenesis[190–192]. For alternapyrone, it has been shown that the PKSN protein from *A. solani* produces alternapyrone in a heterologous system[170]. However, it is unknown whether alternapyrone in *A. solani* is modified by other enzymes to generate a different polyketide, and the role of the polyketide from this pathway in fungal biology is unknown[170].

In our previous work, we also used RT-PCR assays and transcriptome analysis to compare expression of polyketide biosynthetic genes in *M. fijiensis* in infected leaf tissue versus growth in culture medium in order to identify PKS genes that may be involved in pathogenicity[45]. The above described *PKS8-2* and *PKS10-2* genes, and their associated clusters, were more highly expressed in infected leaf material, consistent with a role for these polyketides in disease development[45]. By contrast, *PKS2-1* was more strongly expressed in culture as was the PKS (*PKS10-1*) for melanin biosynthesis[45]. Another PKS cluster that lacked any close homologs in the phylogenetic analysis (*PKS7-1*) was highly expressed in leaf tissue relative to culture[45].

Studies to date have primarily focused on the possible role of *M. fijiensis* polyketides in pathogenicity, and have not addressed the role of polyketides in fungal development and morphogenesis, despite the clear association of polyketide production with the stages of the fungal life cycle in other species. LaeA in *A. nidulans* and its orthologs in other fungi regulate the production of many polyketides and other secondary metabolites[302]. LaeA forms a complex with VeA and VelB, which coordinate the production of secondary metabolites with development of sexual and asexual structures[303]. Several examples of polyketide fruiting body or spore pigments have been identified. Many fungi produce melanin in these structures, which acts both as a pigment, and as protection against UV radiation and other stresses[300,304,305]. Other polyketide pigments in fruiting bodies and spores include fusarubins which are produced in fruiting bodies (perithecia) of *Fusarium* spp.[306], the perithecial pigment produced by PKSN from *Nectria haematococca*[307], and the sexual ascospore pigment ascoquinone A from *Aspergillus nidulans*[308].

In addition to polyketides that serve as pigments for sexual structures, a mutant of the gene *PKS4* was found to result in female sterility in *Neurospora crassa*[98]. Deletion of its ortholog in *Sordaria macrospora* also results in sterility, whereas overexpression of *PKS4* results in the development of enlarged, misshapen perithecia[99]. Representative genomes from diverse fungal species, including members of the Sordariomycetes, the Leotiomycetes, the Dothideomycetes, and the Eurotiomycetes, each encoded two homologs of PKS4[99]. In each species, one of the homologs had domains characteristic of a PKS enzyme, and the other had domains characteristic of a hybrid PKS/non-ribosomal peptide synthase (NRPS) enzyme[99]. Most of these homologs have not been characterized, with the exception of the *Fusarium* spp.

homologs, which produce fusarin C (hybrid PKS/NRPS called PKS10, GzFus1) and fusarielins (PKS9, FSL1)[309,310]. Although *PKS9* is expressed in conditions promoting development of perithecial, no perithecial or ascospore development phenotypes were observed for deletion mutants of either *PKS9* or *GzFus1*[309].

The goal of this study was to identify homologs in *M. fijiensis* of *PKS4* in *S. macrospora* and *N. crassa* in order to better understand sexual reproduction in this species. Here we show that the previously identified *PKS8-4* and *Hybrid8-3*[45] are homologs of these genes. We further show that PKS8-4 is expressed in pseudothecia during the sexual reproductive cycle. GC-MS profiling of a *pks8-4* mutant identified distinct differences in non-polar metabolites.

Results

Phylogenetic Analysis of PKS Protein Sequences

In previous work[45], we conducted a phylogenetic analysis of the polyketide synthases in *M. fijiensis* using full-length sequences of polyketide synthases with known products important in plant pathogenicity such as toxins and melanin. The *M. fijiensis* PKS sequences were predicted using Secondary Metabolites Unique Regions Finder (SMURF) analysis[47]. Our phylogenetic analysis did not include sequences for PKS proteins involved in development such as PKS4. Thus we re-ran the phylogenetic analysis to include the PKS4 sequences from *Neurospora crassa* and *Sordaria macrospora* (Accessions XP_011395279.1 and XP_003348600.1, respectively). In addition, we utilized the antiSMASH program[311] to identify possible additional homologous PKS genes, and included the corresponding protein

sequences in this analysis. The resulting tree of PKS protein sequences is shown in Figure 1. The *N. crassa* and *S. macrospora* PKS4 sequences were separated into a clade, with a bootstrap value of 100, with the *M. fijiensis* PKS8-4 and Hybrid8-3 sequences. This clade also includes the lovastatin and compactin nonaketide synthases from *Aspergillus terreus* and *Penicillium citrinum,* respectively, as well as the betaenone-producing PKS protein sequence from *Phoma betae*.

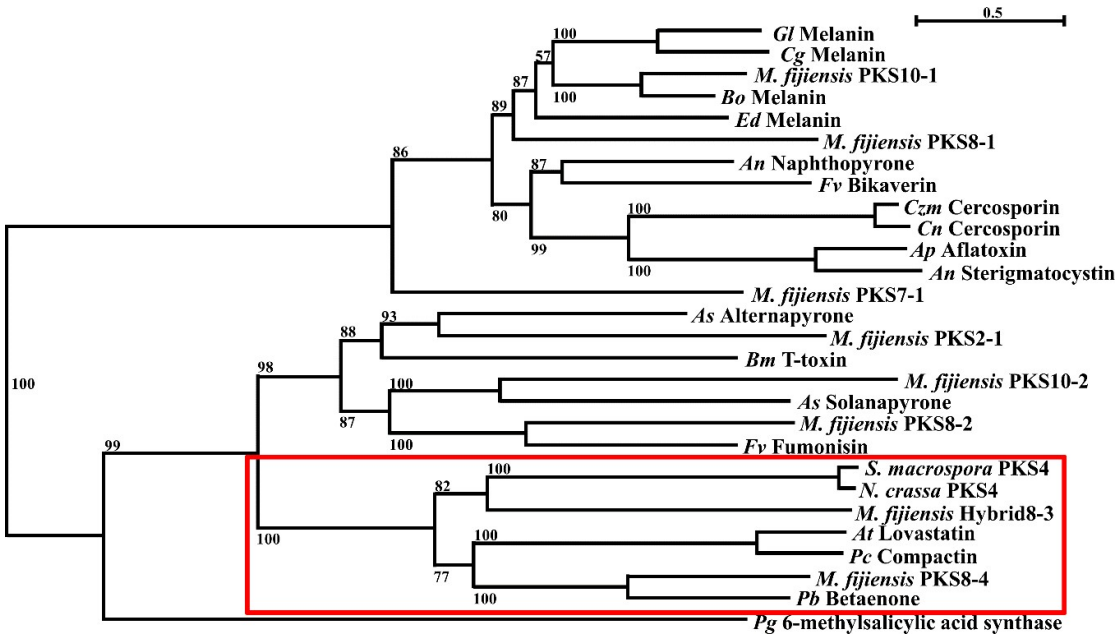


Figure 1. Phylogenetic analysis of *M. fijiensis* PKS protein sequences.
A maximum likelihood tree was created of the *M. fijiensis* PKS protein sequences and well-characterized PKS sequences from other species, including PKS4 from *Sordaria macrospora* and *Neurospora crassa,* involved in sexual reproduction[98,99]. Boxed area shows PKS4 clade. The tree indicates bootstrap values for each relationship, and the scale bar indicates substitutions per site. A description of the polyketide produced by each PKS is indicated, along with an abbreviation of the species name. *An = Aspergillus nidulans*; *Ap = Aspergillus parasiticus*; *Cn = Cercospora nicotianae*; *Czm = Cercospora zeae-maydis*; *Fv = Fusarium verticillioides*; *Cg = Colletotrichum graminicola*; *Gl = Glarea lozoyensis*; *Bo = Bipolaris oryzae*; *Ed = Exophiala dermatitidis*; *Pb = Phome betae*; *Pg = Penicillium griseofulvum*; *Pc = Penicillium citrinum*; *At = Aspergillus terreus*; *As = Alternaria solani*; *Bm = Bipolaris maydis.*

Conserved Domain Analysis of PKS Protein Sequences

To further characterize the functions of the *M. fijiensis* PKS8-4 and Hybrid8-3 enzymes, we conducted a PKS domain analysis using NCBI's Conserved Domain Database[168] and the antiSMASH program[311]. In these analyses we compared conserved domains of the *M. fijiensis* proteins[45] to the PKS proteins that clustered with them in the

phylogenetic analysis: the *S. macrospora* and *N. crassa* PKS4 enzymes, and the PKS enzymes that produce compactin in *P. citrinum*, and lovastatin in *A. terreus*, and betaenone in *P. betae* (Figure 2, Table S1). AntiSMASH did not identify the *Sordaria* PKS4, thus no analysis is shown for this protein (Figure 2B). Although the phylogenetic tree indicated that Hybrid8-3 is more similar than PKS8-4 to the *N. crassa* and *S. macrospora* PKS4 sequences, both analyses showed that the *M. fijiensis* PKS8-4 has a more similar domain structure with the PKS4 enzymes than does Hybrid8-3 (Figure 2). Hybrid8-3 contains several additional domains characteristic of non-ribosomal peptide synthases that the PKS4 sequences lack, and is more similar to the compactin and lovastatin nonaketide PKS domain structure (Figure 2). In addition to the similarity to the PKS4 proteins, both analyses identified the *M. fijiensis* PKS8-4 as having the same domain structure as the *P. betae* betaenone PKS (Bet1), consistent with the phylogenetic analysis. The NCBI Conserved Domain Database identified only one difference in domains between PKS8-4 and the PKS4 sequence from *N. crassa*, identifying a thioester reductase domain in PKS8-4 that PKS4 lacks (Figure 2A). AntiSMASH calls this a terminal domain (Figure 2B). Thioester reductase domains are used by some PKS enzymes to promote dissociation of the polyketide from the PKS[312]. Overall, domain analysis shows a closer association of the *M. fijiensis* PKS8-4 rather than Hybrid8-3 with the PKS4 enzymes involved in sexual reproduction.

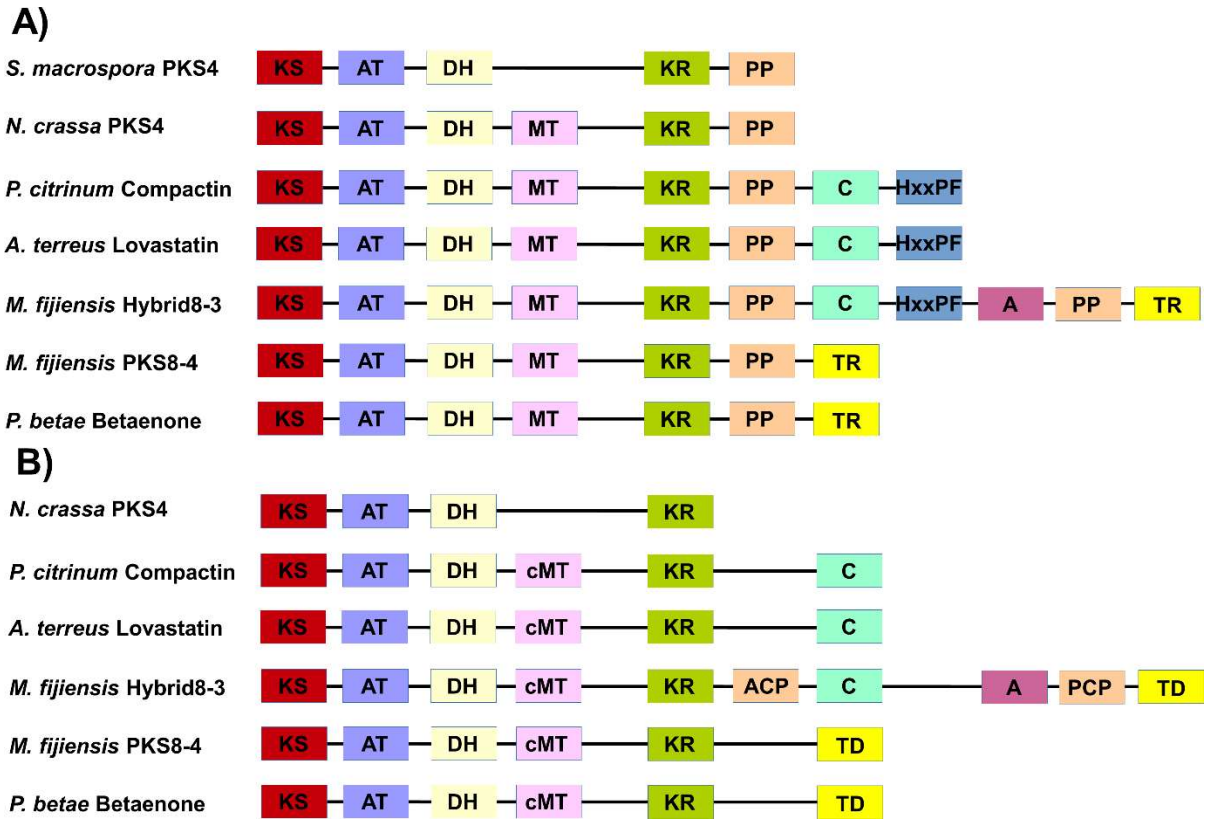


Figure 2. Domains identified from each PKS or hybrid PKS/NRPS enzyme. KS = ketosynthase; AT = acyltransferase; DH = dehydratase; MT/cMT = methyltransferase; KR = ketoreductase; ACP = acyl carrier protein; PCP = peptidyl carrier protein; PP = phosphopantetheine attachment site for ACP or PCP domain; C = condensation; HxxPF = HxxPF domain; A = adenylation; TR = thioester reductase; TD = terminal domain. A) Domains identified using NCBI's Conserved Domain Database (Table S1 and Noar and Daub[45]). B) Domains identified using the antiSMASH program[311].

Comparison of PKS Gene Clusters

The genomes of *M. fijiensis*, *N. crassa*, *S. macrospora*, and *A. terreus* have been sequenced and are annotated with predicted genes and protein sequences. Genes encoding secondary metabolite pathways are often clustered together in fungal genomes[47]. Therefore, genes adjacent to the PKS gene were identified from each annotated genome. Protein

functions were predicted based on results from blastp and from the conserved domain analysis for each protein sequence. Also, the betaenone-producing gene cluster has been characterized from *P. betae*[313], and this gene cluster was included in the comparison (Figure 3, Table S2). This analysis agreed with previous research showing that the lovastatin nonaketide PKS gene, the betaenone PKS gene, and the *M. fijiensis Hybrid8-3* and *PKS8-4* genes are clustered with types of genes common in secondary metabolite gene clusters, such as those encoding cytochrome P450s, enoyl reductases, beta lactamases, transporters, and transcription factors (Figure 3)[45,47,313–315]. Overall, the *PKS8-4* cluster was most similar to the betaenone cluster. In contrast, the *PKS4* genes from *N. crassa* and *S. macrospora* were not clustered with genes common in secondary metabolite clusters, consistent with previously reported work (Figure 3)[99,253].

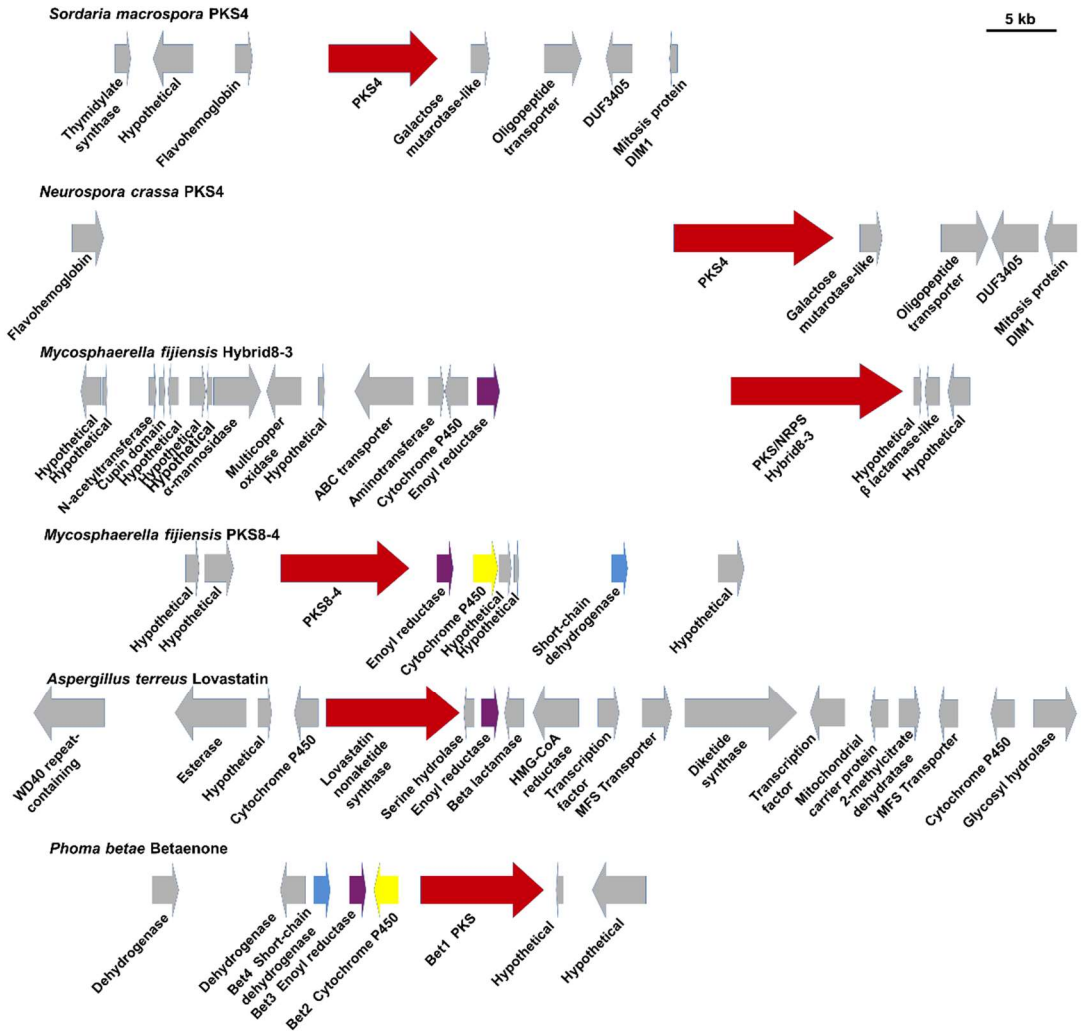


Figure 3. PKS gene clusters.
Each PKS gene is shown along with adjacent genes in the genome. Genes are labeled with putative functions of the corresponding protein, as determined by blastp and conserved domain analysis (Table S2). Gene orientations are indicated by direction of arrows. The PKS or hybrid PKS/NRPS gene is shown with a red arrow, putative orthologous genes are shown with the same color, and other adjacent genes are shown with gray arrows.

Although secondary metabolite genes for a given pathway are typically clustered together in fungal genomes[47], in some cases the secondary metabolite pathway genes are distributed to multiple loci[316]. As shown in Figure 3, there were no obvious secondary

metabolite clusters associated with the *Sordaria* and *Neurospora PKS4* genes. Therefore, to identify possible orthologs of *M. fijiensis Hybrid8-3* and *PKS8-4* cluster genes in the genomes of *N. crassa* and *S. macrospora*, blastp searches were performed of each protein encoded by putative *M. fijiensis* cluster genes against these species. *A. terreus*, *P. citrinum*, and *P. betae* sequences were also included, and results are shown in Table S3. For each protein with conserved domains encoded by the *M. fijiensis* PKS or hybrid gene cluster, a homolog was found from *N. crassa* and *S. macrospora.* However, protein sequence was higher for the other species. The *N. crassa* and *S. macrospora* homologs are spread throughout the genomes. Because of the lack of clustering with the PKS and the relatively low protein sequence similarity, it is unclear whether these sequences are orthologous.

Expression Analysis of Genes in *PKS8-4* and *Hybrid8-3* Gene Clusters

Our previous research indicated by RT-PCR analysis that *PKS8-4* in *M. fijiensis* isolate 10CR1-24 was expressed in more infected leaf tissue samples than culture samples, whereas *Hybrid8-3* was strongly expressed in both conditions[45]. To better characterize expression of these genes and the gene clusters we analyzed transcriptome data from leaves infected with *M. fijiensis* isolate 14H1-11A as compared to culture samples (Figure 4). Our transcriptome analysis confirms that *PKS8-4* has much higher expression in infected leaf tissue compared to culture medium, with a $\log_2$ fold change value (log2FC) of 4.4 (Figure 4A). Furthermore, the nearby enoyl reductase, cytochrome P450, dehydrogenase, and one hypothetical gene also had higher expression in infected leaf tissue compared to culture medium (Figure 4A), consistent with our previous prediction of the cluster composition[45].

The transcriptome analysis also shows that *Hybrid8-3* has higher expression in infected leaf tissue compared to culture medium, but the log2FC was smaller (log2FC=2.3) (Figure 4B). The nearby genes encoding an enoyl reductase, a cytochrome P450, an ABC transporter, and one hypothetical sequence also had higher expression in infected leaf tissue, again supporting our previous hypothesis that these genes are part of the biosynthetic cluster with *Hybrid8-3* (Figure 4B)[45].

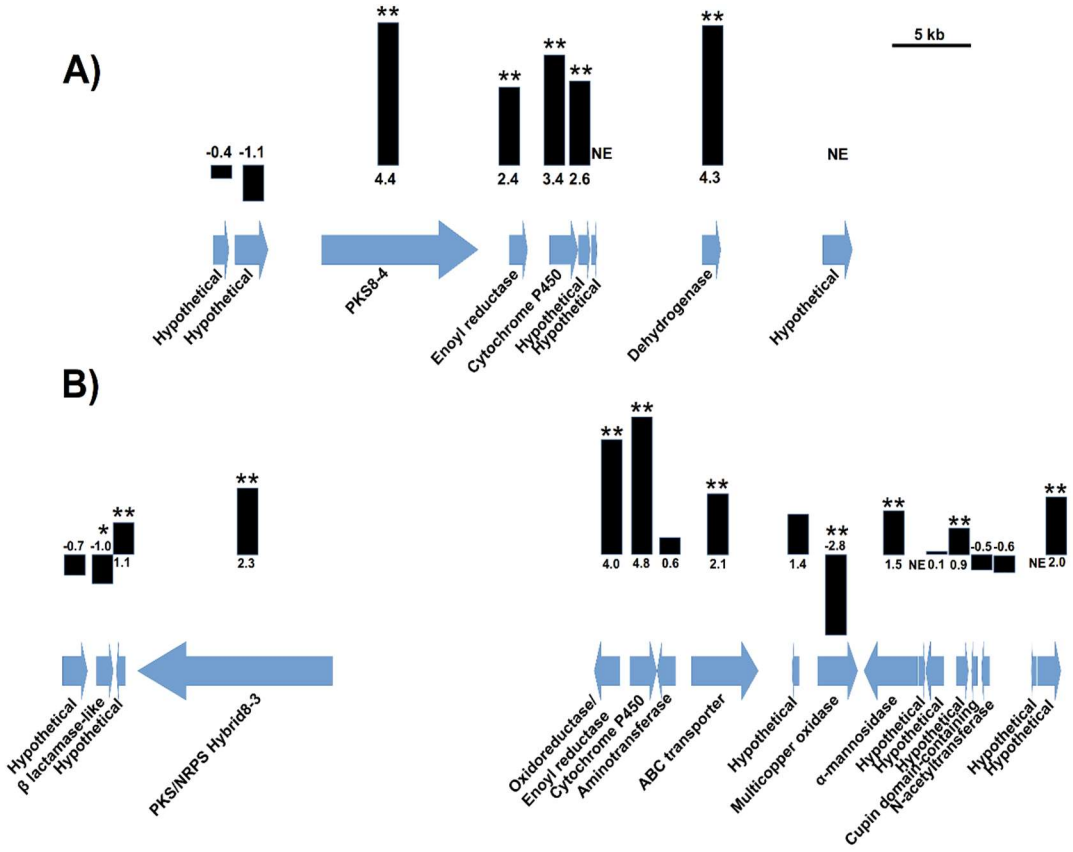


Figure 4. RNA-Seq analysis of expression of PKS clusters.
Data shown are relative expression in infected leaf tissue compared to growth in *in vitro* culture. The PKS or hybrid PKS/NRPS gene is shown with the adjacent genes in the *M. fijiensis* genome. Each gene is labeled with its putative function based on blastp analysis and conserved domains. Arrows indicate gene orientation. Black bars are proportional to the log2FC value. Scale bar indicates 1 kb. Single asterisks indicate significance at $p < 0.05$, whereas double asterisks indicate significance at $p < 0.01$. A) *PKS8-4* gene cluster; B) *Hybrid8-3* gene cluster.

GFP Transcriptional Fusion of PKS8-4 Shows Promoter Activity in Pre-pseudothecia

The greater differential expression of *PKS8-4* in infected leaf tissue relative to culture (Figure 4 and Noar and Daub[45]), along with the PKS domain analysis showing similarity to PKS enzymes involved in sexual reproduction, led us to choose PKS8-4 for further analysis.

We first conducted promoter fusion analysis to localize expression. A GFP transcriptional fusion construct was created (Figure S1) and transformed into *M. fijiensis* isolates 10CR1-24 and CR12. Transformants for each *M. fijiensis* isolate were grown on Potato Dextrose Agar (PDA) and observed using fluorescence microscopy. With each isolate, tiny circular areas of GFP fluorescence were observed, typically in melanized colony extensions, under a dissecting microscope using fluorescence microscopy. These small circular areas of GFP fluorescence were dissected from the colony and observed under confocal microscopy (Figure 5). This structure is consistent with the expected structure of pseudothecia initials[317]. As a control, we used mycelial cultures grown under conditions used for RNA-Seq and did not see fluorescence (data not shown), which is consistent with our data showing that expression of genes in the *PKS8-4* gene cluster was much higher in infected leaf tissue compared to in culture medium.

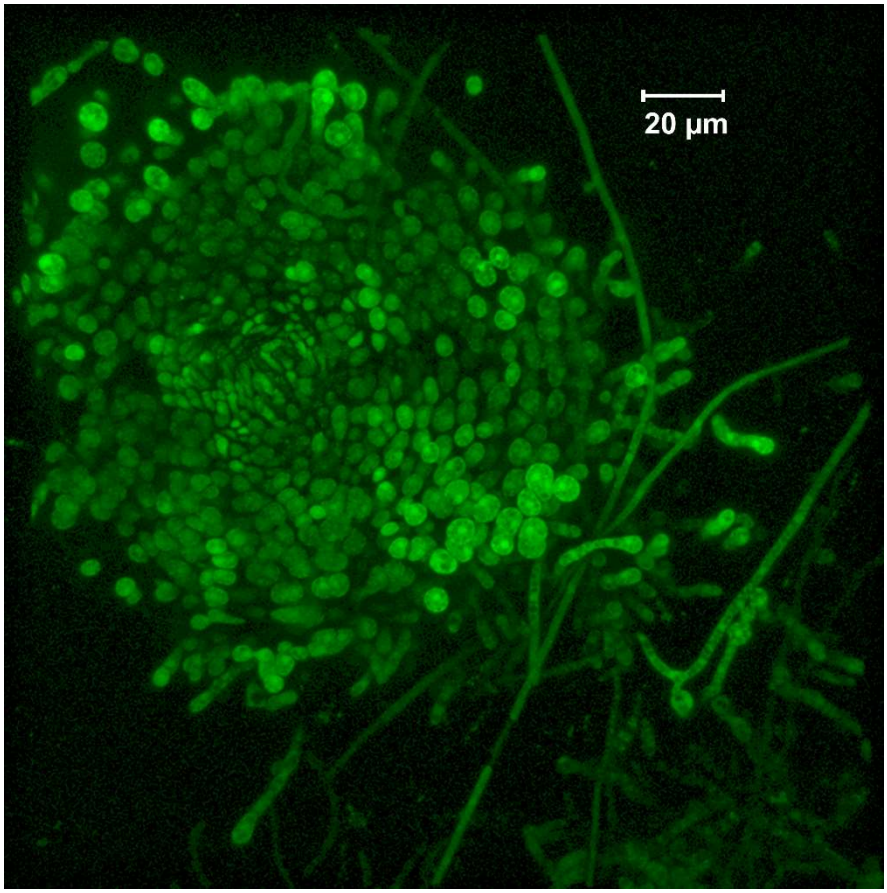


Figure 5. *PKS8-4* promoter activity in reproductive structures.
Small circular area of GFP fluorescence as observed in *PKS8-4-GFP* transcriptional fusion in *M. fijiensis* isolate 10CR1-24 when grown on PDA medium.

To better observe GFP fluorescence during the maturation of pseudothecia, we first determined the mating types of many of our *M. fijiensis* isolates by amplifying sequences from the *mat1-1* and *mat1-2* idiomorphs (Table 1)[318]. We paired *PKS8-4-GFP* transformants of isolates 10CR1-24 and CR12 (mating types 1 and 2, respectively) with wild-type isolates of the opposite mating type. After 3 weeks, we observed abundant tiny circular areas of GFP fluorescence on the mating plates (Figure 6). These areas were then imaged under confocal

microscopy (Figure 7), which showed GFP expression in the developing reproductive structures.

Table 1. Mating types of *M. fijiensis* isolates determined through amplification of mating type idiomorphs, *mat1-1* and *mat1-2*.

Mating type 1	Mating type 2
10CR1-24	CR12
CIRAD86	14H1-11A
Asia8837	CAM275
P587	10CR1-30
10CR1-31	10CR1-32
14H1-14F	10CR1-33
14H1-33A	10CR1-34
14H1-8A	10CR1-36
14H1-37A	14H1-13C
14H1-14E	14H1-9A
14H1-10C	14H1-1C
14H116C	14H10-B
	14H22-B

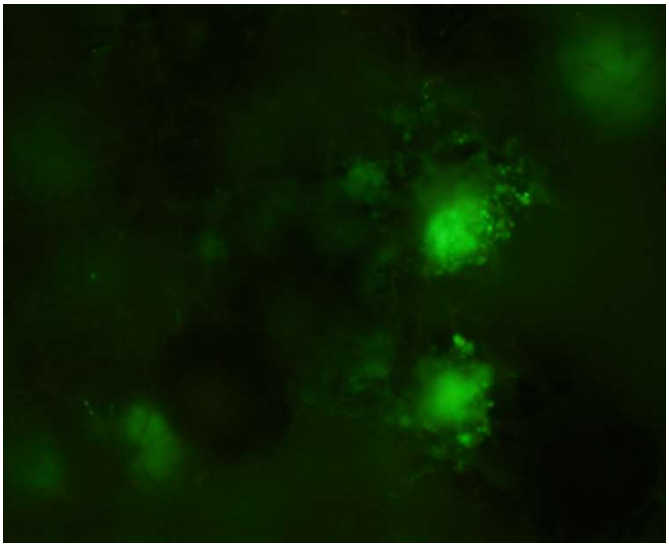


Figure 6. PKS8-4-GFP expression in mating plates viewed using fluorescence microscopy.
Small circular areas of GFP fluorescence as observed with *PKS8-4-GFP* transcriptional fusion in *M. fijiensis* isolate 10CR1-24 (mating type 1) paired with wild-type isolate CR12 (mating type 2), when grown on conditions previously shown to induce mating[319]. Images were taken using a dissecting microscope.

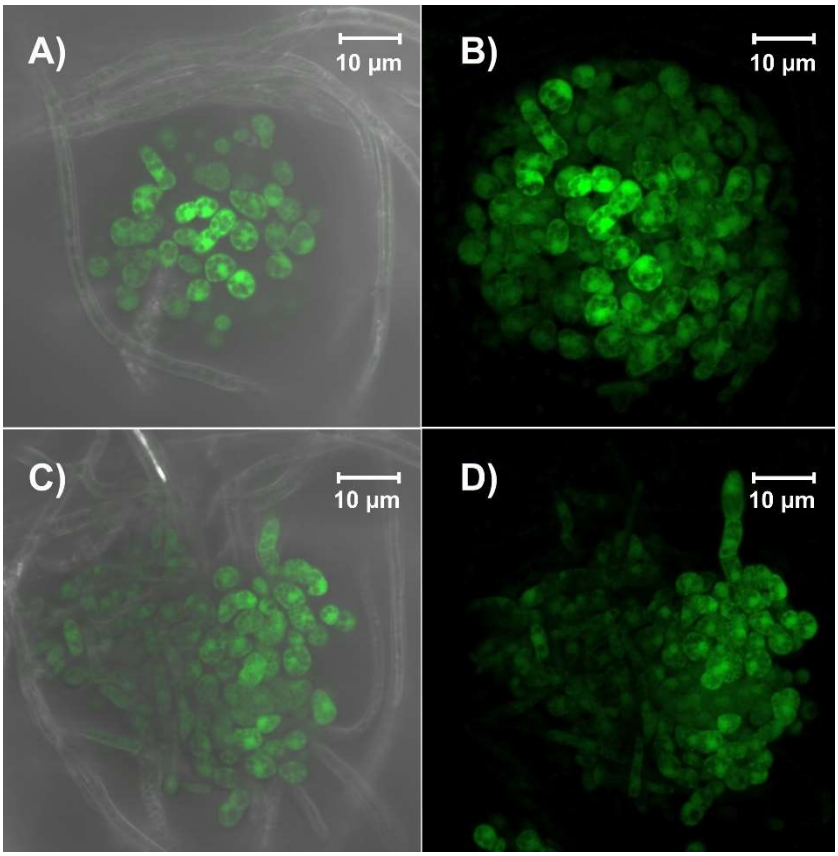


Figure 7. Confocal images of developing pseudothecia.
Small circular area of GFP fluorescence as observed in *PKS8-4-GFP* transcriptional fusion in *M. fijiensis* isolate 10CR1-24 mated with wild-type isolate CR12, when grown under conditions previously shown to induce mating[319]. Images were taken using a Zeiss LSM 710 confocal microscope. A) GFP fluorescence shown as single optical slice, with transmitted light; B) Z-stack projection of GFP fluorescence only, of same structure observed in Figure 7A; C) GFP fluorescence of another example of the structure, shown as a single optical slice, with transmitted light; D) Z-stack projection of GFP fluorescence only, of same structure observed in Figure 7C.

To determine whether the *PKS8-4* promoter is still active later into the maturation of pseudothecia, mating plates of the *PKS8-4-GFP* transcriptional fusion in isolate 10CR1-24 mated with wild-type CR12 were produced and allowed to develop for 7 weeks until

melanized pseudothecia developed. GFP fluorescence was still observed in the mature, melanized pseudothecia (Figure 8).

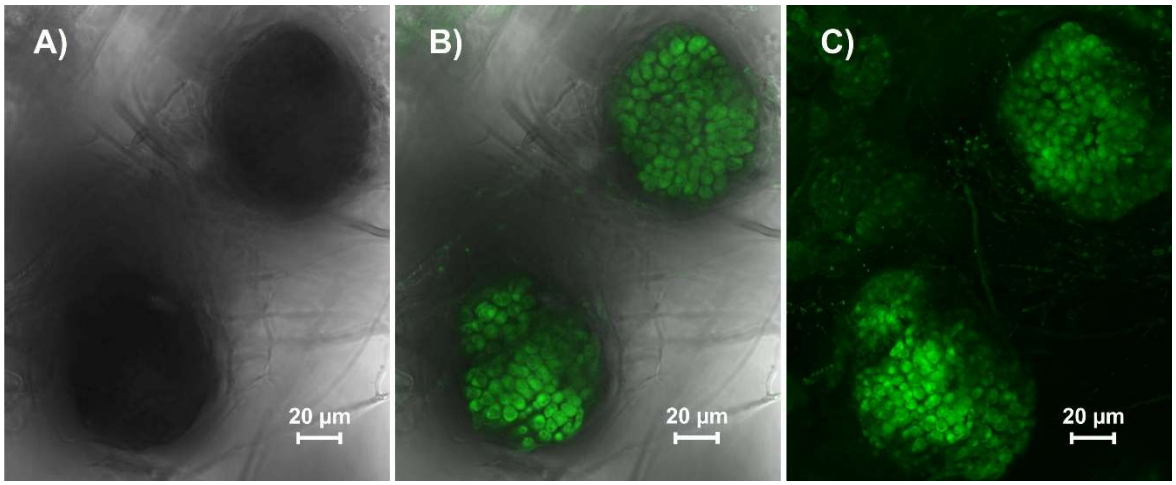


Figure 8. Expression of PKS8-4-GFP in mature pseudothecia.
GFP fluorescence as observed in melanized pseudothecia produced from *PKS8-4-GFP* transcriptional fusion in isolate 10CR1-24 mated with wild-type isolate CR12. Images were taken using a Zeiss LSM 710 confocal microscope. A) Melanized pseudothecia as observed with transmitted light only, with a single optical slice; B) Single optical slice as observed with transmitted light and GFP fluorescence; C) Z-stack of GFP fluorescence only.

pks8-4 Disruption Mutant Has Normal Morphology, Conidiation, and Pathogenicity

To further characterize the function of PKS8-4, we developed an *Agrobacterium tumefaciens*-mediated transformation procedure and used it to create a *pks8-4* disruption mutant. A disruption construct was created (Figure S2) and transformed into *M. fijiensis*. Sixty-five transformants were obtained and screened to identify disruptants. Of these, one disruptant was identified (Figure 9).

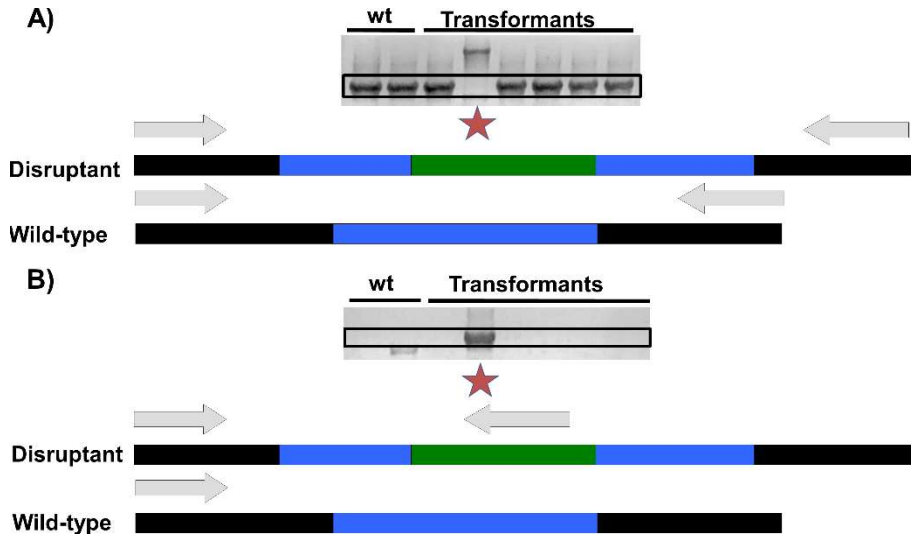


Figure 9. Identification of *pks8-4* disruptants by PCR.
Wild-type and transformants are indicated on the gel images. Black boxes indicate expected product sizes for wild-type (A) versus *pks8-4* disruptant (B). A red star is used to indicate the *pks8-4* disruptant. Arrows indicate the primer binding sites for each reaction for both the wild-type and the *pks8-4* disruptant. The green segment indicates the hygromycin resistance cassette, the blue segments indicate the part of the *PKS8-4* sequence that was used to create the disruption construct, and the black segments indicate the neighboring region in the genome that was not amplified to create the disruption construct. A) Primers were used that span the region where the disruption construct should integrate by homologous recombination. Transformants with a wild-type copy of *PKS8-4* should yield a much smaller PCR product than transformants with a hygromycin resistance cassette inserted into this gene; B) Primers were used so that one targets the hygromycin resistance cassette, and the other targets a region of *PKS8-4* that is distal to the sequence used to create the disruption construct.

Characterization of the *pks8-4* mutant did not identify any phenotypic differences. No differences were observed in colony growth rate or color and appearance of the disruptant versus the wild-type on PDA (Figure 10A). We also tested the ability of the disruptant to produce asexual conidiospores compared to wild-type. The mutant produced conidia, and there were no differences in the numbers of conidia produced, or the size and appearance of the conidia (Figure 10B). We then tested for changes in pathogenicity. Conidia of the *pks8-4*

disruptant were inoculated onto tissue culture plants; non-disrupted transformants were used as controls. After 5 weeks, plants inoculated with the mutant developed characteristic necrotic lesions of black Sigatoka disease (Figure 10C), and there were no differences in symptoms or timing, indicating that the *pks8-4* disruptant is still pathogenic.

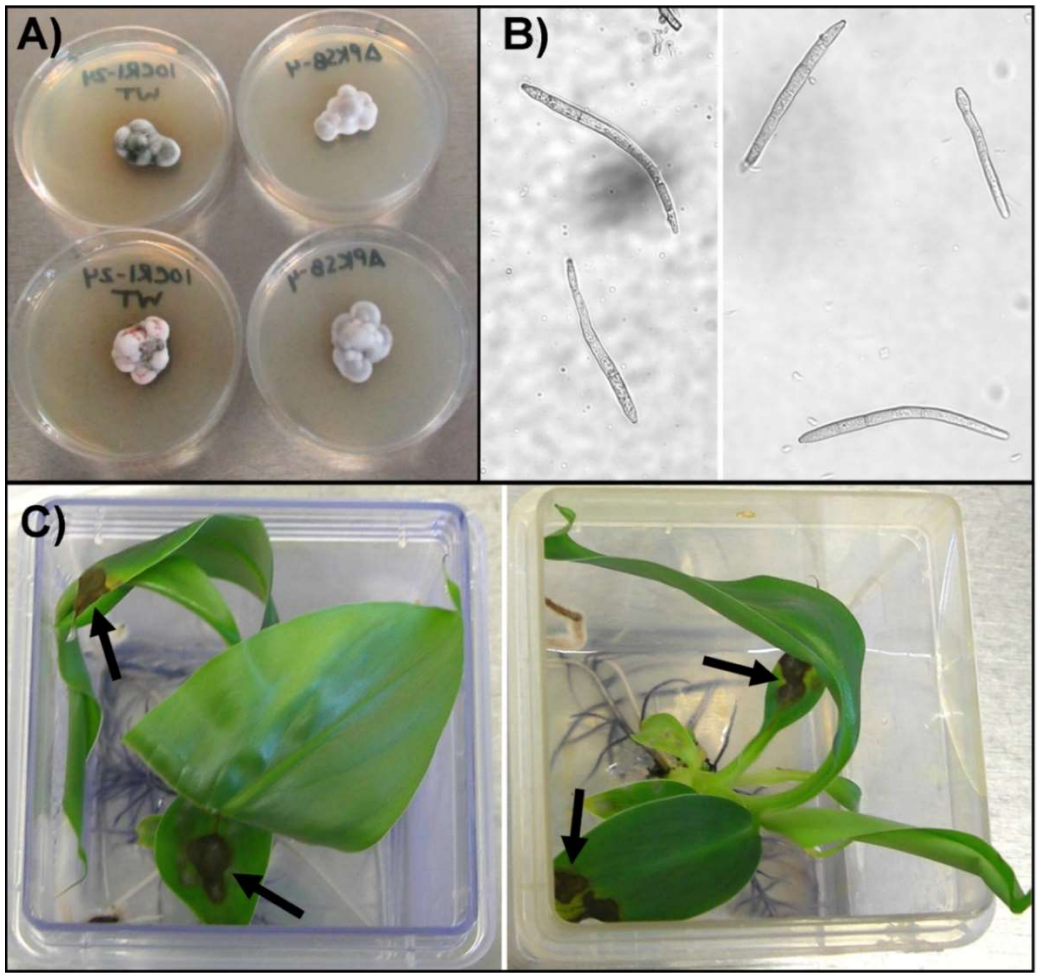


Figure 10. Characterization of the *pks8-4* disruptant.
A) 3 week old *M. fijiensis* colonies of (left) 10CR1-24 wild-type versus (right) *pks8-4* disruptant; B) *M. fijiensis* conidia of (left) 10CR1-24 wild-type versus (right) *pks8-4* disruptant; C) Inoculated plants with arrows showing typical necrotic lesions characteristic of *M. fijiensis* infection from (left) 10CR1-24 with ectopic insertion control versus (right) *pks8-4* disruptant.

Finally, we investigated the ability of the *pks8-4* disruptant to produce their sexual fruiting bodies (pseudothecia) when paired with an isolate of the opposite mating type. Conidia of wild-type isolate 10CR1-24 or *pks8-4* disruptant (mating type 1) were combined in equal numbers with conidia from the compatible mating type 2 isolate 14H1-11A and used to inoculate potted banana plants under greenhouse conditions. At 9 weeks post-inoculation, symptomatic leaves were harvested and scored for the presence or absence of pseudothecia. For the mating of isolate 14H1-11A with 10CR1-24 wild-type and *pks8-4* mutant, pseudothecia were observed on 78% and 79% of the leaves, respectively. Although no differences were observed in the number or appearance of pseudothecia between the wild-type and *pks8-4* mutant, in a mating, both isolates may serve as the female or male parent. Thus the resulting pseudothecia may have been formed from the wild type 14H1-11A strain.

Annotation of Untargeted Metabolites

In initial analysis, we used high performance liquid chromatography-photodiode array/electrospray ionization-mass spectrometry (HPLC-PDA/ESI-MS) for non-targeted profiling of polyketides and other polar secondary metabolites. No dramatic differences were observed in metabolite profiles between *pks8-4* mutant and wild-type samples (data not shown). GC-MS based profiling was then conducted. Our analysis detected more than 100 non-polar metabolite peaks in wild-type samples (Figure S4). Peak deconvolution annotated 38 metabolites with more than 80% mass spectrum identity to metabolites in the NIST 11 library. Based on their structures, these metabolites were categorized into esters (fatty acid esters and other esters), alkane and alkene derivatives, and other metabolites (Table S4). The

peak sizes of many metabolites were reduced in the *pks8-4* mutant compared to the wild-type (Figures S4 and 11A).

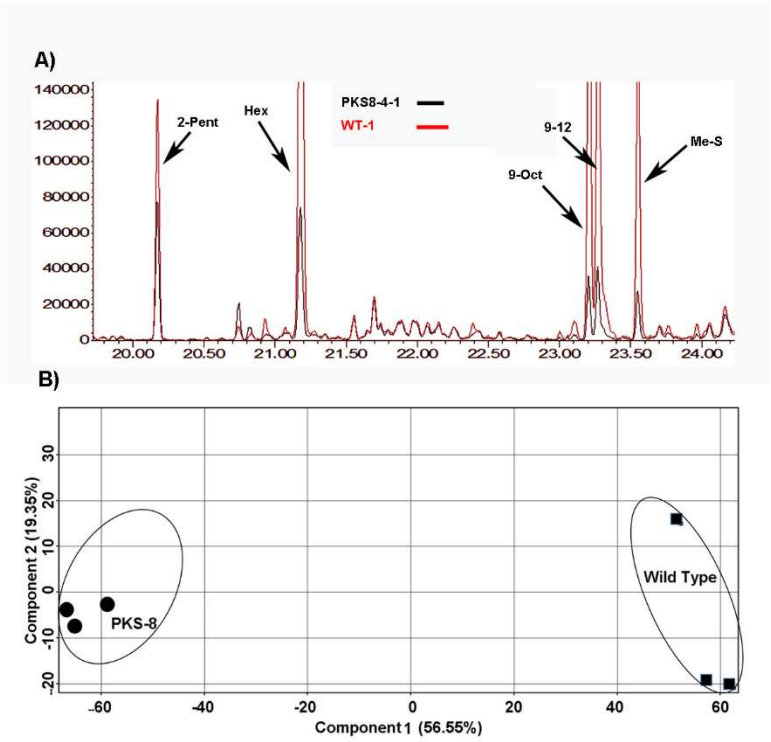


Figure 11. GC-MS based profiling and principal component analysis (PCA).
Hexane extracts of the *pks8-4* mutant and wild-type (WT) control samples were analyzed using GC-MS. Metabolites were annotated using their mass spectra finger printing matched to a standard library. A) A total ion chromatograph comparing metabolite profiles between *pks8-4* and WT extracts from the retention time 20.2 min to 24.4 min. B) A PCA plot showing metabolic differentiation between the *pks8-4* mutant and WT samples. Abbreviations, 2-Pent: 2-pentadecanone, 6, 10, 14-trimethyl; 9-Oct: 9-octadecanoid acid, methyl ester; 9,12-Oct: 9,12-octadecadienoic acid, methyl ester, (E, E)-; Hex: hexadecanoic acid, methyl ester; Me-S: methyl stearate.

PCA and Hierarchical Analysis

Thirty-five metabolites were used for principal component analysis (PCA) and hierarchical analysis. Log_2 values of these metabolites were used as a data matrix for PCA in the Mass Profiler Professional (MPP) software. An ordination plot was developed to show that these metabolite profiles in the *pks8-4* mutant samples were separated from those in wild-type samples (Figure 11B). This result indicates that metabolic activities controlling these non-polar metabolites were altered in the *pks8-4* mutant.

Hierarchical analysis generated a heatmap together with a clustering tree (Figure 12A). Fold changes of metabolite levels in the samples are indicated by different colors. Although variations existed in different biological samples, fold changes showed reduction of nine metabolites in the *pks8-4* mutant samples (Figures 12A and B). Three are saturated fatty acid methyl esters: methyl stearate (stearic acid methyl ester, octadecanoic acid methyl ester), 9-octadecenoic acid methyl ester, and hexadecanoic acid methyl ester. Six other metabolites include three alkene or alkane derivatives, two other esters, and one phenol compound (phenol, 2,4-bis(1,1-dimethylethyl)-) (Figure 12B). Only one annotated metabolite (pentanoic acid, 5-hydroxy-, 2,4-di-t-butylphenyl ester) was increased in abundance in the *pks8-4* mutant compared to the wild-type. This compound results from condensation of pentanoic acid and phenol, 2,4-bis(1,1-dimethylethyl).

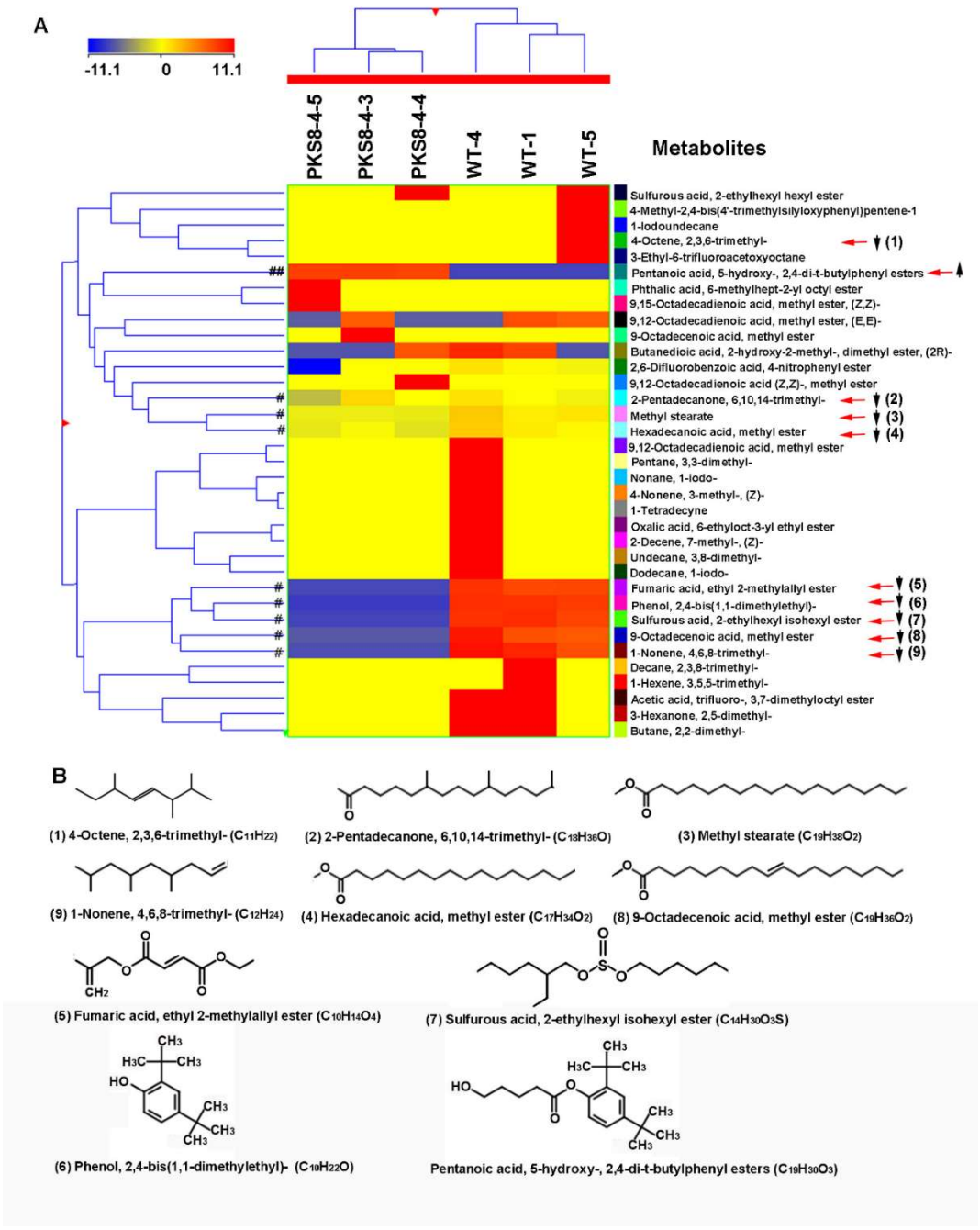


Figure 12. A heatmap and clustering analysis visualizing metabolic differentiation of non-polar metabolites between *pks8-4* and wild-type (WT) control samples.
Thirty-five metabolites annotated by GC-MS analysis were used to generate heatmaps and clustering using the MPP software. A) A heatmap and clustering image show patterns of 35 metabolites. B) Structures of nine metabolites that are reduced in the *pks8-4* mutant (numbered 1-9), and one increased metabolite (not numbered).

Discussion

In *N. crassa* and *S. macrospora*, it was previously reported that PKS4 is involved in development of perithecia[98,99]. In our study, we identified the *M. fijiensis* homologs PKS8-4 and Hybrid8-3, consistent with the previous report that many fungal species have two homologs of PKS4, encoding one PKS and one hybrid PKS/NRPS enzyme[99]. Our analysis showed that *M. fijiensis* PKS8-4 and Hybrid8-3 form a clade with lovastatin and compactin-producing nonaketide synthases, the betaenone-producing PKS (Bet1) from *P. betae*, and PKS4 from *N. crassa* and *S. macrospora* (Figure 1). PKS8-4 had a very similar domain structure and gene cluster compared to Bet1 (Figure 2), suggesting that PKS8-4 may produce a product with similar structure to betaenone. The role of betaenone in fungal biology and pathogenicity has not been fully characterized. However, betaenone is phytotoxic[320], and a betaenone derivative was shown to act as a protein kinase inhibitor[321]. Lovastatin and compactin have been studied extensively for their uses in medicine[322,323]. Less is known about the role these polyketides normally play in fungal biology, however some studies are consistent with a role in sexual reproduction. In *Aspergillus spp.*, overexpression of LaeA results both in increased lovastatin production and increased numbers of fruiting bodies (cleistothecia) as compared to the wild-type, whereas deletion of LaeA completely blocks lovastatin production and results in production of cleistothecia that are only one fifth of the normal size[302,324]. In *Eurotium repens*, the teleomorph (sexual) stage produces compactin, while the anamorph (asexual) stage does not[325]. Therefore, it is plausible that *M. fijiensis* PKS8-4 and Hybrid8-3, the lovastatin and compactin producing PKS enzymes, and the PKS4 enzymes represent a clade that play a role in fungal sexual reproduction. Based on the

phytotoxicity of betaenone, more research is needed to determine if the product of PKS8-4 may be toxic and possibly play a protective role for the sexual structures.

We showed that PKS8-4 has more similar domains to PKS4 compared to Hybrid8-3, and that *PKS8-4* is more highly expressed than *Hybrid8*-3 in infected leaf tissue compared to liquid medium. Since *M. fijiensis* pseudothecia develop on infected banana leaf tissue[17], we concluded that PKS8-4 is the more likely enzyme to be involved in sexual reproduction in this fungus. We generated a *pks8-4* mutant and showed that it is still pathogenic and is able to produce pseudothecia when paired with an isolate of a compatible mating type. In *N. crassa*, disruption of *PKS4* was shown to prevent development of perithecia because of female sterility. *M. fijiensis* is an obligate out-crosser, and in mating, either parent may serve as the female parent. If *M. fijiensis* PKS8-4 also affects female fertility, then it may be that the non-disrupted parent strain served as the female parent for all the pseudothecia observed in our experiment. Further research will be required to investigate this possibility.

Overall, our analysis showed that a homolog of *PKS4* is expressed in developing pseudothecia in *M. fijiensis*. Although we were unable to conclusively show that *PKS8-4* is important for development of pseudothecia like its ortholog in *N. crassa* and *S. macrospora*, we did provide data in support of this hypothesis. This is an exciting possibility because in field conditions, wind dispersal of ascospores is considered the most important means of spread of *M. fijiensis*, as ascospores have been documented to spread several kilometers via wind[5,10]. Therefore, identification of a gene important for pseudothecia and ascospore development would provide a target for new disease control methods. One possibility for targeting such a gene is host-induced gene silencing. In this method, the host plant is

transformed with a silencing construct that targets a fungal gene. This method of gene targeting and disease control is effective for a wide range of fungal pathogens, including biotrophs such as *Blumeria graminis*[124], hemibiotrophs such as *Fusarium spp*[126,326], and necrotrophs such as *Sclerotinia sclerotiorum*[127]. Another new and relevant technology is to spray dsRNAs targeting this gene on the leaves of infected plants. This new technology is being investigated as a control method for insect pests and viral and fungal pathogens[136–138], and it has been shown that dsRNAs on leaf surfaces can maintain biological activity for over 4 weeks[136]. Finally, chemical targeting of this pathway may also be possible. For example, the fungicide tricyclazole has been shown to shut down melanin biosynthesis, which is another polyketide biosynthetic pathway[327]. Characterization of the PKS8-4 biosynthetic pathway could help facilitate research into chemical inhibition of one or more steps in this pathway.

In addition to our study of the expression of *PKS8-4* and its effect on sexual reproduction, we also conducted non-targeted metabolomics to characterize the effect of the *pks8-4* disruption on metabolites produced. Non-targeted (or untargeted) metabolomics is a powerful, unbiased approach to understand metabolomes in different organisms or cells[328–334]. We first used HPLC-PDA/ESI-MS for profiling of polyketides and other polar secondary metabolites. However, no clear differences were observed in metabolite profiles between the *pks8-4* mutant and wild-type samples (data not shown). As a functional interaction between FAS and PKS enzymes has been identified in fungi[335], we then utilized a GC-MS based non-targeted profiling approach to comparatively analyze non-polar metabolites between the wild-type and the *pks8-4* mutant. This analysis revealed major differences in the non-polar

metabolomes of the wild-type compared to the *pks8-4* mutant. We were able to annotate structures for 35 of the metabolites. Of these, three are methyl esters of saturated fatty acids and their levels were highly reduced in the mutant (Figure 12B). This result suggests a possible interplay between the activity of PKS8-4 with fatty acid formation. Well characterized, fungal PKS and fatty acid synthase (FAS) enzymes share a similar architecture of iterative domains that catalyze a series of reactions to load acetyl-CoA and then add malonyl-CoA to elongate a growing polyketide chain to produce diverse final metabolites, although the final products of PKS and FAS enzymes are different[336–338]. These changes are consistent with there being a functional interaction between FAS and PKS enzymes, as was reported in the biosynthesis of aflatoxin B-1[335]. In addition to fatty acid esters, the levels of a few alkane and alkene esters were significantly reduced in *pks8-4* mutants (Figure 12B). As documented, alkanes and alkenes are derived from fatty acids or the fatty acid pathway via acetyl-CoA and acyl-ACP[339,340]. The reduction of these metabolites in the *pks8-4* mutant also supports the association of PKS8-4 with the biosynthesis of fatty acids. Further research will be necessary to characterize the product of the PKS8-4 biosynthetic cluster.

Materials and Methods

Phylogenetic Analysis of PKS Protein Sequences

Full-length PKS protein sequences from *M. fijiensis* and well-characterized sequences from other species were aligned with the *S. macrospora* PKS4 (accession XP_003348600.1), the *N. crassa* PKS4 (accession XP_011395279.1), and the *P. betae* Bet1 (accession

BAQ25466.1) sequences, and a maximum likelihood phylogenetic tree was generated as described previously, using RaxmlGUI v1.3.1 (Figure 1)[45].

Comparison of PKS Conserved Domains and Gene Clusters

For the protein sequences for *M. fijiensis* Hybrid8-3 and PKS8-4, *S. macrospora* and *N. crassa* PKS4, *A. terreus* LovB, *P. citrinum* MlcA, and *P. betae* Bet1, NCBI's Conserved Domain Database[168] was used to identify conserved domains (Figure 2, Table S1). To compare PKS gene clusters, annotated genomes for *M. fijiensis* (NCBI Genome ID 10962), *N. crassa* OR74A (NCBI Genome ID 19), *S. macrospora* (NCBI Genome ID 2242), and *A. terreus* (NCBI Genome ID 53) were used to identify genes adjacent to the PKS. The previously characterized gene cluster from *P. betae* for betaenone[313] was also compared. The antiSMASH 3.0 program[311] was also used to predict PKS gene clusters and domains from PKS enzymes for *M. fijiensis*, *N. crassa*, *S. macrospora*, and *A. terreus*. Gene cluster and PKS domain information was downloaded from the 'Minimum Information about a Biosynthetic Gene cluster' repository[341] for compactin and betaenone. To predict the functions of the corresponding protein sequences, blastp searches were performed using NCBI's non-redundant protein sequences database, and conserved domains were identified using NCBI's conserved domain database (Figure 3, Table S2)[168]. To identify homologs of the proteins encoded by the *PKS8-4* and *Hybrid8-3* gene clusters, blastp searches were done of the corresponding protein sequences against the *N. crassa*, *S. macrospora*, *A. terreus*, *P. citrinum*, and *P. betae* non-redundant protein sequences (Table S3).

Transcriptome Analysis

Samples and transcriptome analysis of *M. fijiensis* isolate 14H1-11A in infected banana leaf tissue versus Potato Dextrose Broth (PDB) medium have been described previously[45,46]. Briefly, *M. fijiensis* conidia were obtained[46], diluted to a concentration of $5.2x10^4$/mL in 0.5% sterile Tween 20, and 25 mL of this suspension was used to inoculate each potted banana plant, which were kept at 25 °C under an 18h light/6h dark photoperiod under cool white light. To maintain high humidity conditions, plants were covered in clear plastic bags until one week post-inoculation, then symptomatic leaves were harvested at 6 weeks post-inoculation. Flasks containing 50 mL of PDB medium were also inoculated with a total of $1.3x10^4$ conidia, and were incubated at 25 °C for 1 week in the dark. RNA was isolated and sequenced using an Illumina HiSeq machine as described previously[45,46]. Transcriptome data are available via SRP075820 through NCBI. Sequences were mapped to both the banana and *M. fijiensis* genomes and differentially expressed genes were identified as previously described[45,46]. For each gene in the putative *PKS8-4* and *Hybrid8-3* gene clusters, $\log_2$ fold change values and p-values from this analysis were reported in Figure 4.

Generation of *Agrobacterium*-Compatible Transformation Vectors

To create the *PKS8-4* promoter – *GFP* transcriptional fusion construct, promoterless *GFP* with a *trpC* terminator was amplified from the vector pRG2[342] using the primers 5'-ACGGTAACTAGTGCTTGAGCAGACATCACC-3' and 5'-TTAATTAAGATTAAGTTGGGTAACGCCA-3'. The PCR product was digested with HindIII and SpeI, and inserted into pEarleyGate 100[343] using the compatible HindIII and

XbaI sites. The Hph selectable marker was amplified from plasmid pCB1636[344] using the primers 5'-CGACTGAAGCTTTCGACGTTAACTGGTTCCC-3' and 5'-GCATATAAGCTTCGTTAACTGATATTGAAGGAGCA-3' that add HindIII restriction sites. A HindIII digest of the PCR product was inserted into the modified pEarleyGate 100 vector. Finally, a region of approximately 1.5 kb upstream of the *PKS8-4* start codon was amplified by PCR using the primers 5'-GCATAGGAATTCAGCAGTCTATATACTAGAGGCT-3' and 5'-TCACGAGAATTCCATGGGGGCGTCCTGGCTGC-3' that add EcoRI sites. An EcoRI digest of this product was used to insert the promoter and create the final vector, containing pEarleyGate 100 modified to contain *PKS8-4* promoter-*GFP-trpC* terminator with a selectable marker for hygromycin (Figure S1).

To create the *pks8-4* disruption mutant, PCR was used to amplify approximately 2.5 kb close to the 5' end of the *PKS8-4* coding sequence, with primers to add attB sites for cloning into the vector pDONR221 (Invitrogen) using Gateway technology. The Hph selectable marker was amplified from the vector pCB1636[344], with primers to add restrictions sites for HindIII (primer sequences indicated previously). HindIII was used to cut the Entry clone in the middle of the *PKS8-4* sequence and insert the Hph PCR product. The *Agrobacterium*-compatible vector pEarleyGate 100[343] was digested with SacI and XhoI, treated with Klenow enzyme to create blunt ends, and ligated back together to remove the *Bar* gene, a selectable marker for plant transformation. Gateway LR reactions were used to transfer the disruption construct into the modified pEarleyGate vector (Figure S2).

To create the constitutive GFP construct, plasmid pRG2[342] and pEarleyGate 100 previously modified to contain promoterless *GFP* and the hygromycin resistance cassette were both digested with EcoRI and ligated together to insert the constitutive glyceraldehyde 3-phosphate dehydrogenase (*gpd*) promoter into the modified pEarleyGate 100 (Figure S6).

Generating Mutants of *M. fijiensis*

M. fijiensis was transformed using *A. tumefaciens* strain EHA105, based on the protocol by Utermark and Karlovsky[345]. Briefly, EHA105 with the appropriate plasmid is grown in liquid Lysogeny Broth (LB) medium with 50 µg/mL each kanamycin and rifampicin to an OD_{600} of 0.5 to 0.9, washed in Induction Medium (IM) and resuspended in IM supplemented with 200 µM acetosyringone. Cells are then grown to an OD_{600} of 0.3, mixed with *M. fijiensis* conidia or mycelial fragments, and spread onto cellophane covering solid IM plates. Plates are incubated for one week at room temperature, and then cellophane was transferred to PDA with 125 mg/L hygromycin and 0.56 g/L ticarcillin, with additional PDA with hygromycin and ticarcillin poured on top of the cellophane to select for transformants and to kill the *A. tumefaciens*. After about 3 weeks, colonies appeared and were transferred to new plates for further analysis.

Identification of Mating Types from *M. fijiensis* Isolates

To determine the mating types of *M. fijiensis* isolates, PCR assays were used to amplify sequences from *mat1-1* and *mat1-2* genes[318]. Primers 2820Mt1-F 5'-CGACCGCTCAACTCCTGGATGG-3' and 3313Mt1-R 5'-

GTCGAGGCTTGGGGTGAAGAGG-3' were used to amplify a 493 bp product from the *mat1-1* gene, as described previously[318]. For the *mat1-2* gene, a 763 bp product was amplified using primers 1Mt2-F 5'-GATGGCTACTCAGGTCACTGC-3' and 762Mt2-R 5'-ATGGCTTGCGTGGCTGGTA-3'. PCR assays were done using the OneTaq enzyme (NEB) using manufacturer's instructions.

PKS8-4 Promoter Activity Characterization via GFP

PKS8-4-GFP transcriptional fusion mutants were grown on PDA medium for initial observations. To test whether the *PKS8-4* promoter is active in sexual structures, the protocol by Etebu et al[319] was used to mate the transcriptional fusion mutant in the 10CR1-24 background with wild-type CR12, or vice versa. To test whether the *PKS8-4* promoter is active in culture, mycelium from a transcriptional fusion mutant in the 10CR1-24 background was macerated and grown in a rotary shaker in 50 mL PDB flasks for 11 days at 28 °C in the dark. GFP was observed under fluorescence microscopy using a dissecting microscope and a Zeiss LSM 710 confocal microscope.

Characterization of *pks8-4* Disruptant

Two PCR reactions were used to confirm *pks8-4* disruption. PCR reactions were done on genomic DNA with primer sites flanking the region used to create the disruption construct using primers 5'-ATGCTCGTCTTCGCTAGTGG-3' and 5'-CGTGATGTATGCCTTGATGT-3'. Disruption by insertion of a hygromycin resistance cassette results in a larger product (Figure 9A). Additionally, primers 5'-

GGCAAAGGAATAGAGTAGAT-3' and 5'-CGTGATGTATGCCTTGATGT-3' were used to amplify from the hygromycin resistance cassette into the flanking genomic sequence (Figure 9B); this reaction produces a product only if the *PKS8-4* gene has been disrupted by a hygromycin resistance cassette.

The *pks8-4* disruptant and the 10CR1-24 wild-type were both grown on PDA in the dark at 25 °C for 3 weeks to initially compare the colony appearance (Figure 10A). Ability of the wild-type 10CR1-24 and *pks8-4* disruptant to produce conidia was assessed by generating conidia as described by Peraza-Echeverría et al[207] (Figure 10B). Grand Nain banana tissue culture plants were maintained as previously described on modified Murashige and Skoog medium on an 18h light/6h dark photoperiod with cool white fluorescent light at 25 °C[45]. Medium was amended with 6-benzylaminopurine for bud proliferation, or was unamended for rooting, as previously described[45]. Pathogenicity of the *pks8-4* disruptant was assessed by inoculating $5x10^4$ conidia/mL of the 10CR1-24 wild-type or the *pks8-4* disruptant in 10 µL droplets on the leaves of tissue culture banana plants and observing lesion development after 5 weeks.

The ability of the *pks8-4* disruptant to produce pseudothecia was assessed. Tissue cultured banana plants were transferred to potting mix and acclimated to greenhouse conditions for two weeks, with a 12h light/12h dark photoperiod, at 26 °C during the day, and 22 °C during the night. Conidia of 14H1-11A and 10CR1-24 or the *pks8-4* disruptant were harvested and combined 1:1, such that each isolate was at a final concentration of $2.5x10^5$ conidia/mL. Plants were inoculated by atomizing 5 mL of this inoculum onto 25 plants for each mating pair. After inoculation, plants were covered with clear plastic bags for

one week to maintain high humidity conditions. Symptomatic leaves were harvested from the inoculation at 9 weeks post-inoculation, and were scored visually for presence or absence of pseudothecia.

Preparation of Samples for Chemical Analysis

To compare the metabolic profiles of the 10CR1-24 wild-type and the *pks8-4* mutant, 10CR1-24 or *pks8-4* mutant were first maintained on PDA medium. Banana leaf pieces of approximately 1 in^2 were autoclaved and placed on top of 1% water agar, as described for induction of mating by Etebu et al[319]. To create the inoculum for each plate, 300 mg mycelium was excised from each plate and macerated in 6 mL water, using a mortar and pestle. Five 20 µL droplets of inoculum were pipetted onto the banana leaf piece on each plate. Plates were incubated at 25 °C for 3 weeks, then pseudothecia initials were scraped from each banana leaf piece into a tube using a scalpel. Samples were pooled such that each of five biological replicates represents pseudothecia initials from ten plates.

Extraction of Non-Polar Metabolites

Dried samples were ground into fine powder at room temperature, and 200 µL hexane (HPLC grade, EMD, NJ, USA) was used to extract 10 mg powdered sample in a 1.5 mL tube for 10 min at room temperature. Samples were sonicated for 10 min in a water bath, then tubes were centrifuged at 12,000 rpm for 10 min. The resulting supernatant was pipetted to a new 1.5 mL tube. The remaining pellet was suspended in another 200 µL hexane for a second extraction using the previously described steps. The two hexane extractions for each

sample were pooled together to obtain 400 µL of extract for GC-MS analysis. Two hundred µL were then pipetted into a 250-µL insert in a 2-mL vial for GC-MS analysis described below.

Untargeted Gas Chromatograph-Mass Spectrometry Analysis

Metabolite analysis was conducted using a gas chromatograph 6890 coupled with 5975C MSD (Agilent Technologies, USA). A HP-5 MS 5% phenyl methyl siloxane column (30 m × 0.25 mm × 0.25 µm) was used to separate metabolites. A splitless mode was used to inject samples. The inlet and detector temperatures were set at 250°C. The oven temperature program was initially set at 40°C for 1 min, then ramped to 280°C with a consistent rate of 8°C/min and then held for 5 min. Pure helium was used as the carrier gas, with a flow rate of 1 mL/min. A positive electron impact ion source (70 eV) was used to ionize metabolites. Mass fragments were scanned in the range of 40-800 (*m/z*).

As we described previously[332,346], metabolite peaks detected by GC-MS were deconvoluted using the NIST 11 library and the Agilent MassHunter Mass Profile (MHMP) and Mass Profiler Professional (MPP) software. In brief, untargeted and unknown peaks were deconvoluted and annotated to metabolites using both ChemiStation and the NIST 11 standard library. Metabolite mass data files from ChemiStation were translated into MassHunter data files using Agilent MassHunter GC/MS Translator software (version B.05.02) and then were deconvoluted using the MassHunter Qualitative Analysis software (Version B.06.00). Based on values of more than 80% mass spectrum identity to a standard compound in the library, untargeted peaks were annotated to metabolites. All annotated

metabolite peak data were then exported to "cef" files for principal component analysis (PCA) and construction of heatmap and clustering analysis.

Metabolite cef files were imported to MPP software for PCA, heatmap construction, and hierarchical clustering analysis. Before these analyses, all data were aligned, normalized (to $\log_2$ value), and baselined to the median level of all samples for each experiment. The fold change for each metabolite between the *pks8-4* mutant and the 10CR1-24 wild-type samples was based on $\log_2$ value in each biological sample. Log_2 values of metabolites were used for PCA, heatmap and clustering analysis to generate plots to visualize differentiation between samples.

Acknowledgments

We thank Miguel Muñoz (Dole Food Company, San José, Costa Rica) for providing tissue-cultured banana plants. This work was funded by a gift from Dole Food Company to MED and by graduate fellowships from National Institutes of Health and the National Science Foundation to RDN.

Author Contributions

R.D.N., M.E.D. and D.X. conceived and designed the experiments; R.D.N., M.E.C., D.X., and D.M. performed the experiments; R.D.N., D.X., D.M., M.E.C., and M.E.D. analyzed the data; M.E.D. and D.X. contributed reagents/materials/analysis tools; R.D.N., D.X., and M.E.D. wrote the paper.

Conflicts of Interest

The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Supplementary Materials

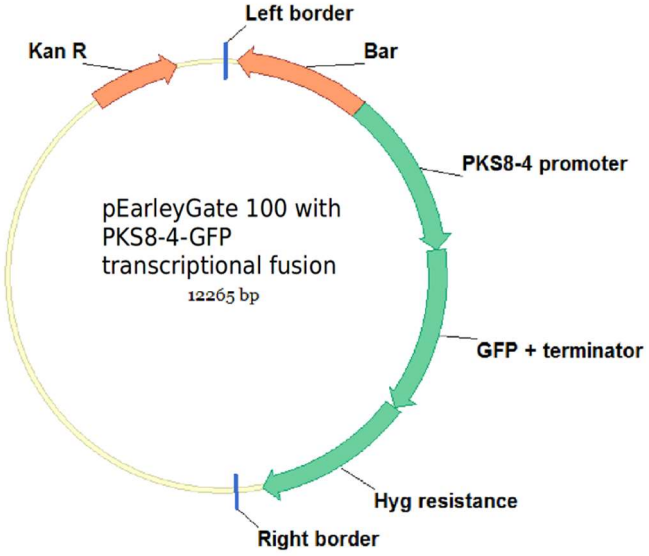


Figure S1. *PKS8-4-GFP* transcriptional fusion construct.
A vector was created to analyze the promoter activity of *PKS8-4* by fusing the promoter to a sequence that encodes GFP, followed by a *trpC* terminator. A hygromycin resistance cassette was used as a selectable marker. This construct was inserted into a modified pEarleyGate 100 vector backbone.

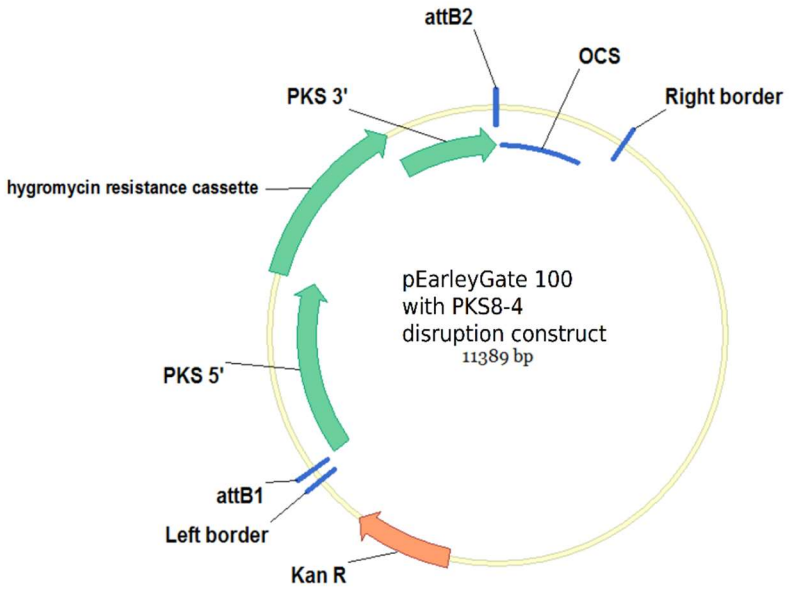


Figure S2. *PKS8-4* disruption construct.
A vector was created containing the *PKS8-4* sequence interrupted by a hygromycin resistance cassette, in a modified pEarleyGate 100 vector backbone.

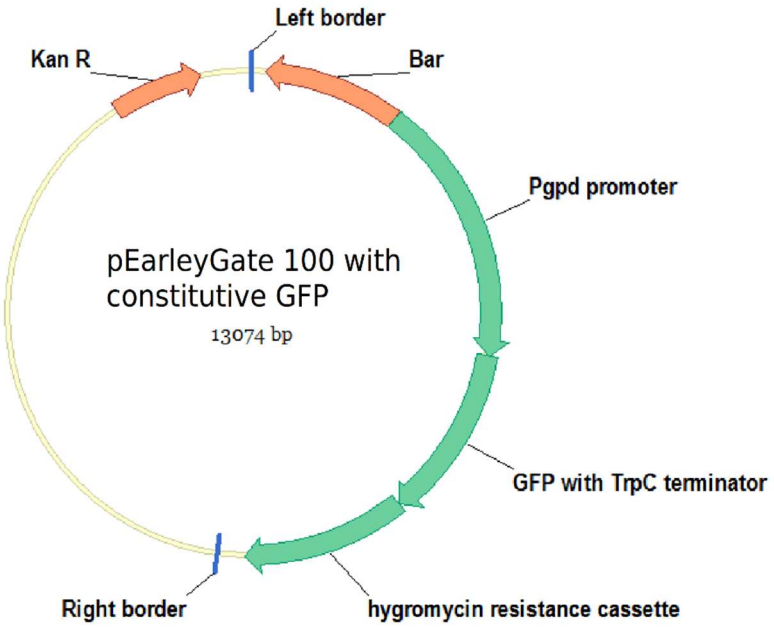


Figure S3. Constitutive GFP construct.
A vector was created to constitutively express GFP under the control of a *gpd* promoter, with a *trpC* terminator. A hygromycin resistance cassette was used as a selectable marker. This construct was inserted into a modified pEarleyGate 100 vector backbone.

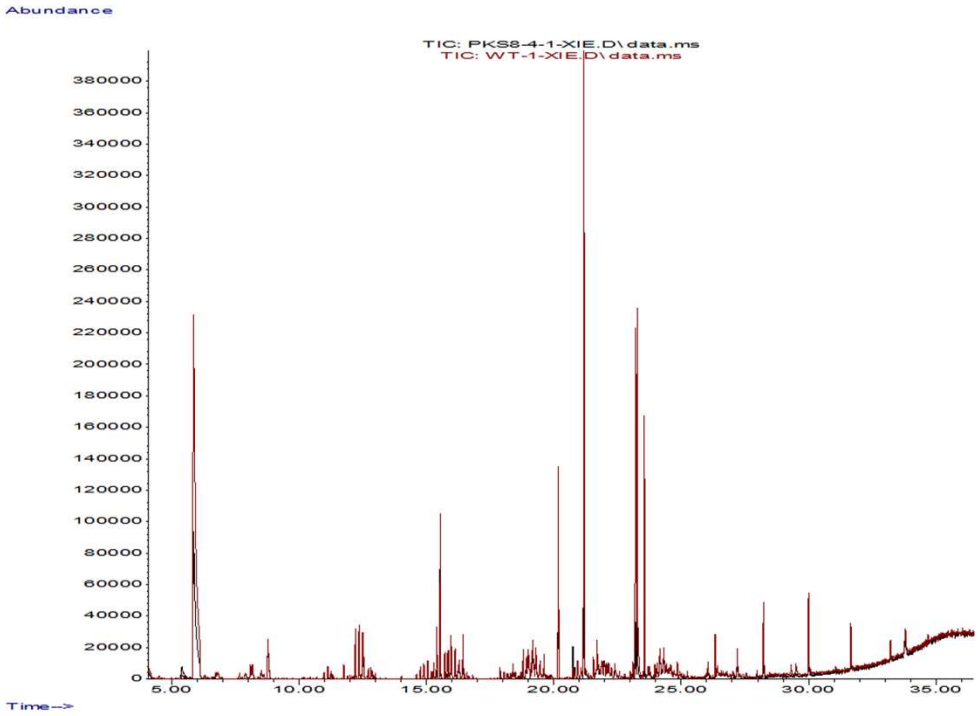


Figure S4. Overview of total ion chromatographs.
An overview of total ion chromatographs shows alterations of non-polar metabolite profiles in *pks8-4* mutant (black color) compared to wild-type control (WT) (red color).

Table S1. Conserved domains in each PKS enzyme.
For each PKS or hybrid PKS/NRPS protein sequence, the table indicates the domains present and their associated E-values, as determined by NCBI's Conserved Domain Database. KS = ketosynthase domain; AT = acyltransferase domain; DH = dehydratase domain; MT = methyltransferase domain; KR = ketoreductase domain; PP1 = first phosphopantetheine attachment site; C = condensation domain; HxxPF = HxxPF domain; A = adenylation domain; PP2 = second phosphopantetheine attachment site; TR = thioester reductase domain. For some of the domains, the Conserved Domain Database also predicts whether active sites, NAD(P) binding sites, S-adenosylmethionine (SAM) binding sites, AMP-binding sites, and Acyl-activating enzyme consensus motifs are present within the sequence. The presence (+) or absence (-) of these features is noted within the table.

	KS	AT	DH	MT	KR	PP1	C	HxxPF	A	PP2	TR
Sordaria macrospora PKS4	7.76E-145 Act +	1.07E-102	3.83E-33	N/A	1.01E-56 Act + NAD(P) +	2.21E-06	N/A	N/A	N/A	N/A	N/A
Neurospora crassa PKS4	2.46E-145 Act +	4.55E-102	1.82E-32	1.86E-03 SAM -	4.60E-54 Act + NAD(P) +	1.02E-06	N/A	N/A	N/A	N/A	N/A
Penicillium citrinum Compactin nonaketide synthase	8.74E-144 Act +	6.39E-92	1.26E-45	7.00E-19 SAM +	6.55E-61 Active + NAD(P) +	1.02E-10	2.71E-34	2.98E-04	N/A	N/A	N/A
Aspergillus terreus Lovastatin nonaketide synthase	1.31E-145 Act +	8.82E-92	5.86E-42	5.05E-20 SAM +	3.16E-57 Act + NAD(P) +	3.51E-09	1.12E-27	1.49E-04	N/A	N/A	N/A
Phoma betae Betaenone synthase	3.13E-138 Active site +	3.43E-91	3.21E-29	1.03E-17 SAM +	8.34E-56 Act + NAD(P) +	0.00766	N/A	N/A	N/A	N/A	7.83E-33

Table S2. Blastp and conserved domain analysis for PKS proteins and proteins encoded by neighboring genes in the genome.

For the PKS or hybrid PKS/NRPS and proteins encoded by neighboring genes in the genome, the tables indicate the description of the sequence in Figure 3, the location and orientation of the gene within the genome, accession numbers for the transcript and corresponding protein, conserved domains identified using NCBI's Conserved Domain Database, and the ten best homologs identified using blastp with NCBI's non-redundant protein sequences database. Blastp results include the species from which the hit was found, the description of the sequence, the bitscore, the E-value, percent identity and similarity, and the accession number of the sequence. A) PKS4 and proteins encoded by neighboring genes in the *S. macrospora genome*, with location in the genome, accession numbers, and conserved domains of each sequence; B) Blastp hits for PKS4 and proteins encoded by neighboring genes in the *S. macrospora* genome; C) PKS4 and proteins encoded by neighboring genes in the *N. crassa* genome, with location in the genome, accession numbers, and conserved domains of each sequence; D) Blastp hits for PKS4 and proteins encoded by neighboring genes in the *N. crassa* genome; E) Lovastatin nonaketide synthase and proteins encoded by neighboring genes in the *A. terreus* genome, with location in the genome, accession numbers, and conserved domains of each sequence; F) Blastp hits for lovastatin nonaketide synthase and proteins encoded by neighboring genes in the *A. terreus* genome; G) Betaenone-producing Bet1 and proteins encoded by neighboring genes in the *P. betae* genome, with location in genome, accession numbers, and conserved domains of each sequence; H) Blastp hits for betaenone-producing Bet1 and proteins encoded by neighboring genes in the *P. betae* genome.

Table S2A

Description	Location	Accessions	Conserved domains
mitosis protein DIM1	NW_003546212.1 290312-290877 +	XM_003348548.1 XP_003348596.1	Mitosis protein DIM1 (E-value=3.35e-86); U5 snRNP protein, DIM1 family (E-value=8.00e-86); Dim1 family (E-value=2.59e-71); Thiol-disulfide isomerase or thioredoxin (E-value=1.42e-04)
DUF3405	NW_003546212.1 293487-295387 +	XM_003348549.1 XP_003348597.1	Protein of unknown function (DUF3405) (E-value=0e+00)
OPT oligopeptide transporter	NW_003546212.1 297079-299753 -	XM_003348550.1 XP_003348598.1	small oligopeptide transporter, OPT family (E-value=7.53e-174); OPT oligopeptide transporter protein (E-value=1.24e-153)
galactose mutarotase-like	NW_003546212.1 303573-304935 -	XM_003348551.1 XP_003348599.1	galactose mutarotase_like (E-value=7.84e-105); galactose mutarotase (E-value=1.44e-49); aldose 1-epimerase (E-value=1.66e-49); Galactose mutarotase or related enzyme (E-value=2.23e-48); Aldose 1-epimerase (E-value=3.34e-44)
PKS4	NW_003546212.1 307277-315100 -	XM_003348552.1 XP_003348600.1	PKS (E-value=0e+00); Acyl transferase domain (AT) in PKS (E-value=0e+00); Beta-ketoacyl synthase (KS) (E-value=7.76e-145); AT (E-value=1.07e-102); KS, N-terminal (E-value=1.75e-89); Polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=9.62e-68); Ketoreductase domain (KR) (E-value=1.01e-56); KR (E-value=1.40e-51); AT (E-value=1.07e-44); KR domain of fatty acid synthase (E-value=4.75e-40); PKS dehydratase (DH) (E-value=3.83e-33); 3-oxoacyl-acyl carrier protein synthase (E-value=2.31e-26); Malonyl CoA-acyl carrier protein transacylase (E-value=5.27e-25); DH in PKS (E-value=4.23e-19); malonyl CoA-acyl carrier protein transacylase (E-value=2.83e-18); short chain dehydrogenase (E-value=5.31e-14); 3-ketoacyl-(acyl-carrier-protein) reductase (E-value=3.10e-13); NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family (E-value=3.12e-13); 3-oxoacyl-(acyl-carrier-protein) reductase (E-value=1.04e-11); [acyl-carrier protein] S-malonyltransferase (E-value=1.03e-08); Phosphopantetheine attachment site (PP) (E-value=2.21e-06); iterative type I PKS product template domain (E-value=3.28e-06); gluconate 5-dehydrogenase (E-value=3.81e-05); PP (E-value=1.04e-03); thiolase, N-terminal (E-value=2.19e-03)
Flavo-hemoglobin	NW_003546212.1 320468-321727 -	XM_003348553.1 XP_003348601.1	Bifunctional nitric oxide dioxygenase/dihydropteridine reductase 2 (E-value=0e+00); FAD_NAD(P)H binding domain of flavohemoglobin (E-value=7.57e-103); Globin domain of flavohemoglobins (flavoHbs) (E-value=1.07e-78); Hemoglobin-like flavoprotein (E-value=5.92e-71); Ferredoxin-NADP reductase (E-value=3.03e-56); phenylacetate-CoA oxygenase/reductase, PaaK subunit (E-value=1.76e-21); Globin (E-value=2.82e-08); Oxidoreductase FAD-binding domain (E-value=2.73e-06)
Hypothetical	NW_003546212.1 324635-326761 +	XM_003348554.1 XP_003348602.1	No conserved domains
Thymidylate synthase	NW_003546212.1 328301-329434 -	XM_003348555.1 XP_003348603.1	Bifunctional dihydrofolate reductase-thymidylate synthase (E-value=6.65e-165); Thymidylate synthase (E-value=2.29e-157); thymidylate synthase (E-value=1.66e-117); Thymidylate synthase (E-value=2.63e-111); thymidylate synthase (E-value=1.98e-108); Thymidylate synthase and pyrimidine hydroxymethylase (E-value=1.13e-96)

Table S2B

***Sordaria macrospora* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Mitosis protein DIM1	*Madurella mycetomatis*	Mitosis protein dim1	293	1.00E-100	98.00%	99.00%	KXX79361.1
	Podospora anserina S mat+	hypothetical protein	277	2.00E-94	94.00%	96.00%	XP_001904152.1
	Fusarium pseudograminearum CS3096	hypothetical protein FPSE_02797	276	4.00E-94	91.00%	96.00%	XP_009254191.1
	Nectria haematococca mpVI 77-13-4	predicted protein	275	1.00E-93	90.00%	95.00%	XP_003051647.1
	Fusarium oxysporum Fo5176	hypothetical protein	275	2.00E-93	91.00%	96.00%	EGU81843.1
	Microdochium bolleyi	mitosis protein DIM1	274	3.00E-93	87.00%	96.00%	KXJ92162.1
	Eutypa lata UCREL1	putative mitosis protein dim1 protein	274	3.00E-93	88.00%	95.00%	XP_007791937.1
	Fusarium avenaceum	thioredoxin-like protein 4a	273	5.00E-93	89.00%	95.00%	KIL90829.1
	Colletotrichum gloeosporioides Nara gc5	mitosis protein dim1	271	4.00E-92	87.00%	95.00%	XP_007278115.1
	Phaeoacremonium minimum UCRPA7	putative mitosis protein dim1 protein	270	8.00E-92	86.00%	95.00%	XP_007916280.1
DUF3405	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_123696	1143	0.00E+00	100.00%	95.00%	XP_009853079.1
	Neurospora crassa OR74A	hypothetical protein NCU08396	1142	0.00E+00	93.00%	94.00%	XP_963235.1
	Podospora anserina S mat+	hypothetical protein	686	0.00E+00	55.00%	70.00%	XP_001904151.1
	Myceliophthora thermophila ATCC 42464	hypothetical protein MYCTH_2308049	654	0.00E+00	53.00%	69.00%	XP_003664856.1
	Chaetomium thermophilum var. thermophilum DSM 1495	biotin synthase-like protein	638	0.00E+00	52.00%	67.00%	XP_006692443.1
	Chaetomium globosum CBS 148.51	hypothetical protein CHGG_10102	633	0.00E+00	52.00%	69.00%	XP_001228029.1

Table S2B (continued)

***Sordaria macrospora* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% identity**	**% similarity**	**Accession**
DUF3405	*Thielavia terrestris NRRL 8126*	hypothetical protein THITE_2096267	629	0.00E+00	52.00%	68.00%	XP_003654710.1
	Madurella mycetomatis	hypothetical MMYC01_204715	610	0.00E+00	50.00%	63.00%	KXX79358.1
	Purpureocillium lilacinum	biotin synthase-like protein	544	0.00E+00	48.00%	63.00%	OAQ95426.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_06751	543	0.00E+00	47.00%	63.00%	KEY72922.1
OPT oligopeptide transporter	*Sordaria macrospora k-hell*	unnamed protein product	1710	0.00E+00	97.00%	97.00%	CCC12734.1
	Neurospora crassa OR74A	hypothetical protein NCU08397	1553	0.00E+00	90.00%	93.00%	XP_963236.2
	Neurospora tetrasperma FGSC 2508	hypothetical protein NEUTE1DRAFT_66680	1552	0.00E+00	90.00%	94.00%	XP_009853078.1
	Phaeoacremonium minimum UCRPA7	putative oligopeptide transporter 2 protein	1235	0.00E+00	69.00%	80.00%	XP_007919500.1
	Magnaporthiopsis poae ATCC 64411	oligopeptide transporter 2	1160	0.00E+00	66.00%	78.00%	KLU87224.1
	Phaeomoniella chlamydospora	putative opt oligopeptide transporter	1159	0.00E+00	67.00%	80.00%	KKY16838.1
	Pestalotiopsis fici W106-1	hypothetical protein PFICI_08391	1155	0.00E+00	65.00%	78.00%	XP_007835163.1
	Gaeumannomyces graminis var. tritici R3-111a-1	oligopeptide transporter 2	1151	0.00E+00	65.00%	78.00%	XP_009218180.1
	Colletotrichum gloeosporioides Cg-14	OPT oligopeptide transporter	1146	0.00E+00	66.00%	77.00%	EQB46928.1
	Colletotrichum gloeosporioides Nara gc5	OPT oligopeptide transporter	1145	0.00E+00	66.00%	77.00%	XP_007278974.1

Table S2B (continued)

***Sordaria macrospora* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% identity**	**% similarity**	**Accession**
Galactose mutarotase-like	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_102663	768	0.00E+00	94.00%	97.00%	XP_003348599.1
	Neurospora crassa OR74A	Aldose 1-epimerase	752	0.00E+00	92.00%	96.00%	XP_009853077.1
	Chaetomium globosum CBS 148.51	hypothetical protein CHGG_10106	589	0.00E+00	75.00%	82.00%	XP_963237.2
	Madurella mycetomatis	Aldose 1-epimerase	582	0.00E+00	76.00%	85.00%	XP_001228033.1
	Pestalotiopsis fici W106-1	hypothetical protein PFICI_05028	578	0.00E+00	73.00%	82.00%	KXX79359.1
	Myceliophthora thermophila ATCC 42464	aldose epimerase-like protein	571	0.00E+00	73.00%	82.00%	XP_007831800.1
	Podospora anserina S mat+	hypothetical protein	566	0.00E+00	74.00%	84.00%	XP_003664855.1
	Thielavia terrestris NRRL 8126	hypothetical protein THITE_2050221	565	0.00E+00	73.00%	84.00%	XP_001904150.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_04403	545	0.00E+00	72.00%	81.00%	XP_003654713.1
	Stachybotrys chlorohalonata IBT 40285	hypothetical S40285_05676	544	0.00E+00	73.00%	81.00%	KEY72813.1
PKS4	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_85360	4789	0.00E+00	89.00%	94.00%	XP_009853076.1

Table S2B (continued)

***Sordaria macrospora* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% identity**	**% similarity**	**Accession**
PKS4	*Neurospora crassa OR74A*	polyketide synthase 4	4763	0.00E+00	89.00%	94.00%	XP_011395279.1
	Madurella mycetomatis	Nonribosomal peptide synthetase 14	2828	0.00E+00	56.00%	71.00%	KXX76983.1
	Podospora anserina S mat+	Putative polyketide synthase	2728	0.00E+00	54.00%	69.00%	CDP29013.1
	Colletotrichum gloeosporioides Cg-14	hypothetical protein CGLO_17938	2629	0.00E+00	52.00%	69.00%	EQB43400.1
	Colletotrichum tofieldiae	Beta-ketoacyl synthase domain-containing protein	2607	0.00E+00	52.00%	69.00%	KZL74145.1
	Colletotrichum incanum	Beta-ketoacyl synthase domain-containing protein	2597	0.00E+00	52.00%	68.00%	KZL86980.1
	Colletotrichum salicis	Beta-ketoacyl synthase domain-containing protein	2591	0.00E+00	52.00%	69.00%	KXH33580.1
	Colletotrichum graminicola M1.001	Beta-ketoacyl synthase domain-containing protein	2588	0.00E+00	52.00%	68.00%	XP_008097088.1
	Colletotrichum fioriniae PJ7	Beta-ketoacyl synthase domain-containing protein	2583	0.00E+00	52.00%	69.00%	XP_007591014.1
Flavohemoglobin	*Neurospora crassa OR74A*	flavohemoglobin	782	0.00E+00	90.00%	95.00%	XP_957939.2
	Neurospora tetrasperma FGSC 2508	hypothetical protein NEUTE1DRAFT_117694	778	0.00E+00	90.00%	95.00%	XP_009853075.1
	Podospora anserina S mat+	hypothetical protein	573	0.00E+00	68.00%	80.00%	XP_001904160.1
	Madurella mycetomatis	flavohemoglobin	564	0.00E+00	67.00%	80.00%	KXX76984.1
	Chaetomium globosum CBS 148.51	hypothetical CHGG_10094	561	0.00E+00	66.00%	78.00%	XP_001228021.1
	Chaetomium thermophilum var. thermophilum DSM 1495	hypothetical protein CTHT_0019520	536	0.00E+00	65.00%	77.00%	XP_006692438.1
	Myceliophthora thermophila ATCC 42464	Oxidoreductase-like protein	531	0.00E+00	65.00%	76.00%	XP_003664863.1

Table S2B (continued)

***Sordaria macrospora* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% identity**	**% similarity**	**Accession**
Flavohemoglobin	*Thielavia terrestris NRRL 8126*	hypothetical protein THITE_128077	523	0.00E+00	64.00%	76.00%	XP_003654702.1
	Magnaporthe oryzae 70-15	flavohemoglobin	517	7.00E-180	63.00%	76.00%	XP_003718919.1
	Gaeumannomyces graminis var. tritici R3-111a-1	flavohemoglobin	496	1.00E-170	59.00%	73.00%	XP_009218197.1
Hypothetical	*Neurospora crassa OR74A*	hypothetical protein NCU10052	853	0.00E+00	62.00%	73.00%	XP_957940.1
	Neurospora tetrasperma FGSC 2508	hypothetical protein NEUTE1DRAFT_102660	852	0.00E+00	62.00%	73.00%	XP_009853074.1
Thymidylate synthase	*Neurospora crassa OR74A*	thymidylate synthase	605	0.00E+00	86.00%	89.00%	XP_957941.1
	Neurospora tetrasperma FGSC 2508	thymidylate synthase	593	0.00E+00	87.00%	91.00%	XP_009853073.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_01818	548	0.00E+00	75.00%	82.00%	KEY65302.1
	Stachybotrys chlorohalonata IBT 40285	hypothetical protein S40285_01809	537	0.00E+00	75.00%	81.00%	KFA63367.1
	Myceliophthora thermophila ATCC 42464	hypothetical protein MYCTH_2308073	528	0.00E+00	74.00%	80.00%	XP_003664862.1
	Fusarium pseudograminearum CS3096	hypothetical protein FPSE_10728	525	0.00E+00	74.00%	81.00%	XP_009262120.1
	Fusarium poae	hypothetical protein FPOA_00759	523	0.00E+00	72.00%	79.00%	OBS26816.1
	Microdochium bolleyi	bifunctional dihydrofolate reductase-thymidylate synthase	523	0.00E+00	71.00%	80.00%	KXJ95282.1
	Neonectria ditissima	thymidylate synthase	523	0.00E+00	72.00%	80.00%	KPM35620.1
	Nectria haematococca mpVI 77-13-4	predicted protein	521	0.00E+00	73.00%	81.00%	XP_003054223.1

Table S2C

Description in Figure 3	Location in genome	Accessions	Conserved domains
Flavo-hemoglobin	NC_026507.1 1037324-1039598 +	XM_952846.3 XP_957939.2	Bifunctional nitric oxide dioxygenase/dihydropteridine reductase 2 (E-value=0e+00); FAD_NAD(P)H binding domain of flavohemoglobin (E-value=1.26e-104); Globin domain of flavohemoglobins (flavoHbs) (E-value=1.86e-79); Hemoglobin-like flavoprotein (E-value=1.42e-70); Ferredoxin-NADP reductase (E-value=8.46e-58); phenylacetate-CoA oxygenase/reductase, PaaK subunit (E-value=7.81e-23); Globin (E-value=8.78e-08); Oxidoreductase NAD-binding domain (E-value=1.03e-06)
PKS4	NC_026507.1 1080398-1091850 +	XM_011396977.1 XP_011395279.1	PKS (E-value=0e+00); AT in PKS (E-value=0e+00); KS (E-value=2.46e-145); AT domain (E-value=4.55e-102); KS, N-terminal (E-value=8.93e-91); polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=1.21e-70); KR domain (E-value=4.60e-54); KR domain (E-value=3.84e-49); AT domain (E-value=5.09e-44); KR domain of fatty acid synthase (E-value=7.61e-39); DH domain of PKS (E-value=1.82e-32); 3-oxoacyl-ACP synthase (E-value=2.34e-25); Malonyl CoA-ACP transacylase (E-value=3.04e-23); DH domain in PKS (E-value=1.90e-17); Malonyl CoA-ACP transacylase (E-value=1.12e-16); 3-ketoacyl-(ACP) reductase (E-value=2.43e-12); NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family (E-value=4.54e-12); short chain dehydrogenase (E-value=5.44e-12); 3-oxoacyl-(ACP) reductase (E-value=1.42e-11); iterative type I PKS product template domain (E-value=5.82e-07); PP (E-value=1.02e-06); [ACP] S-malonyltransferase (E-value=3.39e-05); gluconate 5-dehydrogenase (E-value=9.39e-05); PP (E-value=1.34e-04); MT domain (E-value=1.86e-03); Thiolase, N-terminal domain (E-value=2.19e-03)
PKS4, variant	NC_026507.1 1080398-1091850 +	XM_011396978.1 XP_011395280.1	PKS (E-value=0e+00); AT in PKS (E-value=0e+00); KS (E-value=2.46e-145); AT in PKS (E-value=4.55e-102); KS, N-terminal (E-value=8.93e-91); Polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=1.21e-70); KR domain (E-value=4.60e-54); KR domain (E-value=3.84e-49); AT domain (E-value=5.09e-44); KR domain of fatty acid synthase (E-value=7.61e-39); DH domain of PKS (E-value=1.82e-32); 3-oxoacyl-ACP synthase (E-value=2.34e-25); Malonyl CoA-ACP transacylase (E-value=3.04e-23); DH domain in PKS (E-value=1.90e-17); Malonyl CoA-ACP transacylase (E-value=1.12e-16); 3-ketoacyl-(ACP) reductase (E-value=2.43e-12); NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family (E-value=4.54e-12); short chain dehydrogenase (E-value=5.44e-12); 3-oxoacyl-(ACP) reductase (E-value=1.42e-11); iterative type I PKS product template domain (E-value=5.82e-07); PP (E-value=1.02e-06); [ACP] S-malonyltransferase (E-value=3.39e-05); gluconate 5-dehydrogenase (E-value=9.39e-05); PP (E-value=1.34e-04); MT domain (E-value=1.86e-03); Thiolase, N-terminal (E-value=2.19e-03)
Galactose mutarotase-like	NC_026507.1 1093660-1095284 +	XM_958144.3 XP_963237.2	Galactose mutarotase-like (E-value=6.25e-105); aldose 1-epimerase (E-value=4.54e-52); galactose mutarotase (E-value=1.77e-51); Galactose mutarotase or related enzyme (E-value=8.40e-49); Aldose 1-epimerase (E-value=2.20e-47)
OPT oligopeptide transporter	NC_026507.1 1093660-1095284 +	XM_958144.3 XP_963236.2	small oligopeptide transporter, OPT family (E-value=0e+00); OPT oligopeptide transporter protein (E-value=5.47e-168)

Table S2C (continued)

Description in Figure 3	Location in genome	Accessions	Conserved domains
DUF3405	NC_026507.1 1103044-1106408 -	XM_958142.2 XP_963235.1	Protein of unknown function (DUF3405) (E-value=0e+00)
mitosis protein DIM1	NC_026507.1 1106760-1109089 -	XM_011396975.1 XP_011395277.1	Mitosis protein DIM1 (E-value=3.35e-86); U5 snRNP protein, DIM1 family (E-value=8.00e-86); Dim1 family (E-value=2.59e-71); Thiol-disulfide isomerase or thioredoxin (E-value=1.42e-04)
mitosis protein DIM1	NC_026507.1 1107894-1109065 -	XM_011396976.1 XP_011395278.1	Mitosis protein DIM1 (E-value=3.35e-86); U5 snRNP protein, DIM1 family (E-value=8.00e-86); Dim1 family (E-value=2.59e-71); Thiol-disulfide isomerase or thioredoxin (E-value=1.42e-04)

Table S2D

***Neurospora crassa* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Flavohemoglobin	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_117694	846	0.00E+00	98.00%	99.00%	XP_009853075.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05696	782	0.00E+00	90.00%	95.00%	XP_003348601.1
	Podospora anserina S mat+	hypothetical protein	558	0.00E+00	68.00%	78.00%	XP_001904160.1
	Madurella mycetomatis	Flavohemoprotein	555	0.00E+00	67.00%	79.00%	KXX76984.1
	Chaetomium globosum CBS 148.51	hypothetical protein CHGG_10094	544	0.00E+00	64.00%	77.00%	XP_001228021.1
	Chaetomium thermophilum var. thermophilum DSM 1495	hypothetical protein CTHT_0019520	530	0.00E+00	65.00%	76.00%	XP_006692438.1
	Myceliophthora thermophila ATCC 42464	Oxidoreductase-like protein	528	0.00E+00	64.00%	76.00%	XP_003664863.1
	Thielavia terrestris NRRL 8126	hypothetical protein THITE_128077	516	6.00E-179	64.00%	76.00%	XP_003654702.1
	Magnaporthe oryzae 70-15	Flavohemoprotein	506	7.00E-175	61.00%	76.00%	XP_003718919.1
	Gaeumannomyces graminis var. tritici R3-111a-1	Flavohemoprotein	491	7.00E-168	59.00%	73.00%	XP_009218197.1
PKS4	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_85360	5195	0.00E+00	97.00%	98.00%	XP_009853076.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05695	4810	0.00E+00	89.00%	94.00%	XP_003348600.1
	Madurella mycetomatis	Nonribosomal peptide synthetase 14	2847	0.00E+00	56.00%	71.00%	KXX76983.1
	Podospora anserina S mat+	Putative polyketide synthase	2744	0.00E+00	54.00%	70.00%	CDP29013.1
	Colletotrichum tofieldiae	Beta-ketoacyl synthase domain-containing protein	2633	0.00E+00	52.00%	69.00%	KZL74145.1
	Colletotrichum incanum	Beta-ketoacyl synthase domain-containing protein	2625	0.00E+00	52.00%	69.00%	KZL86980.1
	Colletotrichum gloeosporioides Cg-14	hypothetical protein CGLO_17938	2622	0.00E+00	52.00%	69.00%	EQB43400.1
	Colletotrichum salicis	Beta-ketoacyl synthase domain-containing protein	2614	0.00E+00	52.00%	69.00%	KXH33580.1
	Colletotrichum nymphaeae SA-01	Beta-ketoacyl synthase domain-containing protein	2607	0.00E+00	52.00%	69.00%	KXH61164.1

Table S2D (continued)

Neurospora crassa PKS4 biosynthetic cluster	Blastp hits						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
PKS4	*Colletotrichum fioriniae PJ7*	Beta-ketoacyl synthase domain-containing protein	2605	0.00E+00	52.00%	69.00%	XP_007591014.1
PKS4, variant	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_85360	5195	0.00E+00	97.00%	98.00%	XP_009853076.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05695	4810	0.00E+00	89.00%	94.00%	XP_003348600.1
	Madurella mycetomatis	Nonribosomal peptide synthetase 14	2847	0.00E+00	56.00%	71.00%	KXX76983.1
	Podospora anserina S mat+	Putative polyketide synthase	2744	0.00E+00	54.00%	70.00%	CDP29013.1
	Colletotrichum tofieldiae	Beta-ketoacyl synthase domain-containing protein	2633	0.00E+00	52.00%	69.00%	KZL74145.1
	Colletotrichum incanum	Beta-ketoacyl synthase domain-containing protein	2625	0.00E+00	52.00%	69.00%	KZL86980.1
	Colletotrichum gloeosporioides Cg-14	hypothetical protein CGLO_17938	2622	0.00E+00	52.00%	69.00%	EQB43400.1
	Colletotrichum salicis	Beta-ketoacyl synthase domain-containing protein	2614	0.00E+00	52.00%	69.00%	KXH33580.1
	Colletotrichum nymphaeae SA-01	Beta-ketoacyl synthase domain-containing protein	2607	0.00E+00	52.00%	69.00%	KXH61164.1
	Colletotrichum fioriniae PJ7	Beta-ketoacyl synthase domain-containing protein	2605	0.00E+00	52.00%	69.00%	XP_007591014.1
Galactose mutarotase-like	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_102663	776	0.00E+00	96.00%	98.00%	XP_009853077.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05694	752	0.00E+00	92.00%	96.00%	XP_003348599.1
	Chaetomium globosum CBS 148.51	hypothetical protein CHGG_10106	595	0.00E+00	75.00%	82.00%	XP_001228033.1
	Madurella mycetomatis	Aldose 1-epimerase	582	0.00E+00	76.00%	86.00%	KXX79359.1
	Myceliophthora thermophila ATCC 42464	aldose epimerase-like protein	574	0.00E+00	74.00%	82.00%	XP_003664855.1
	Thielavia terrestris NRRL 8126	hypothetical protein THITE_2050221	572	0.00E+00	75.00%	84.00%	XP_003654713.1
	Pestalotiopsis fici W106-1	hypothetical protein PFICI_05028	565	0.00E+00	71.00%	82.00%	XP_007831800.1
	Podospora anserina S mat+	hypothetical protein	562	0.00E+00	74.00%	84.00%	XP_001904150.1

Table S2D (continued)

***Neurospora crassa* PKS4 biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Galactose mutarotase-like	*Chaetomium thermophilum var. thermophilum DSM 1495*	hypothetical protein CTHT_0019580	545	0.00E+00	72.00%	81.00%	XP_006692444.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_04403	541	0.00E+00	72.00%	80.00%	KEY72813.1
OPT oligopeptide transporter	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_66680	1757	0.00E+00	98.00%	99.00%	XP_009853078.1
	Sordaria macrospora k-hell	unnamed protein product	1645	0.00E+00	93.00%	96.00%	CCC12734.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05693	1581	0.00E+00	90.00%	93.00%	XP_003348598.1
	Phaeoacremonium minimum UCRPA7	putative oligopeptide transporter 2 protein	1283	0.00E+00	70.00%	83.00%	XP_007919500.1
	Phaeomoniella chlamydospora	putative opt oligopeptide transporter	1218	0.00E+00	67.00%	81.00%	KKY16838.1
	Gaeumannomyces graminis var. tritici R3-111a-1	oligopeptide transporter 2	1206	0.00E+00	67.00%	82.00%	XP_009218180.1
	Colletotrichum gloeosporioides Cg-14	OPT oligopeptide transporter	1206	0.00E+00	67.00%	79.00%	EQB46928.1
	Colletotrichum gloeosporioides Nara gc5	OPT oligopeptide transporter	1204	0.00E+00	67.00%	80.00%	XP_007278974.1
	Pestalotiopsis fici W106-1	hypothetical protein PFICI_08391	1204	0.00E+00	67.00%	81.00%	XP_007835163.1
	Colletotrichum orbiculare MAFF 240422	OPT oligopeptide transporter	1203	0.00E+00	67.00%	79.00%	ENH84976.1
DUF3405	*Neurospora tetrasperma FGSC 2508*	hypothetical protein NEUTE1DRAFT_123696	1220	0.00E+00	99.00%	99.00%	XP_009853079.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_05692	1142	0.00E+00	93.00%	94.00%	XP_003348597.1
	Podospora anserina S mat+	hypothetical protein	706	0.00E+00	58.00%	72.00%	XP_001904151.1
	Myceliophthora thermophila ATCC 42464	hypothetical protein MYCTH_2308049	682	0.00E+00	55.00%	71.00%	XP_003664856.1
	Chaetomium thermophilum var. thermophilum DSM 1495	biotin synthase-like protein	664	0.00E+00	54.00%	69.00%	XP_006692443.1

Table S2D (continued)

***Neurospora crassa* PKS4 biosynthetic cluster**	**Blastp**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
DUF3405	*Chaetomium globosum CBS 148.51*	hypothetical protein CHGG_10102	663	0.00E+00	54.00%	71.00%	XP_001228029.1
	Thielavia terrestris NRRL 8126	hypothetical protein THITE_2096267	659	0.00E+00	54.00%	70.00%	XP_003654710.1
	Madurella mycetomatis	hypothetical protein MMYC01_204715	640	0.00E+00	52.00%	66.00%	KXX79358.1
	Purpureocillium lilacinum	biotin synthase-like protein	559	0.00E+00	53.00%	66.00%	OAQ95426.1
	Purpureocillium lilacinum	biotin synthase-like protein	556	0.00E+00	53.00%	66.00%	OAQ87467.1
Mitosis protein DIM1	*Madurella mycetomatis*	Mitosis protein dim1	293	1.00E-100	98.00%	99.00%	KXX79361.1
	Podospora anserina S mat+	hypothetical protein	277	2.00E-94	94.00%	96.00%	XP_001904152.1
	Fusarium pseudograminearum CS3096	hypothetical protein FPSE_02797	276	4.00E-94	91.00%	96.00%	XP_009254191.1
	Nectria haematococca mpVI 77-13-4	predicted protein	275	1.00E-93	90.00%	95.00%	XP_003051647.1
	Fusarium oxysporum Fo5176	hypothetical protein FOXB_07638	275	2.00E-93	91.00%	96.00%	EGU81843.1
	Microdochium bolleyi	mitosis protein DIM1	274	3.00E-93	87.00%	96.00%	KXJ92162.1
	Eutypa lata UCREL1	putative mitosis protein dim1 protein	274	3.00E-93	88.00%	95.00%	XP_007791937.1
	Fusarium avenaceum	Thioredoxin-like protein 4a	273	5.00E-93	89.00%	95.00%	KIL90829.1
	Colletotrichum gloeosporioides Nara gc5	mitosis protein dim1	271	4.00E-92	87.00%	95.00%	XP_007278115.1
	Phaeoacremonium minimum UCRPA7	putative mitosis protein dim1 protein	270	8.00E-92	86.00%	95.00%	XP_007916280.1
Mitosis protein DIM1	*Madurella mycetomatis*	Mitosis protein dim1	293	1.00E-100	98.00%	99.00%	KXX79361.1
	Podospora anserina S mat+	hypothetical protein	277	2.00E-94	94.00%	96.00%	XP_001904152.1

Table S2D (continued)

***Neurospora crassa* PKS4 biosynthetic cluster**	**Blastp**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Mitosis protein DIM1	*Fusarium pseudograminearum CS3096*	hypothetical protein FPSE_02797	276	4.00E-94	91.00%	96.00%	XP_009254191.1
	Nectria haematococca mpVI 77-13-4	predicted protein	275	1.00E-93	90.00%	95.00%	XP_003051647.1
	Fusarium oxysporum Fo5176	hypothetical protein FOXB_07638	275	2.00E-93	91.00%	96.00%	EGU81843.1
	Microdochium bolleyi	mitosis protein DIM1	274	3.00E-93	87.00%	96.00%	KXJ92162.1
	Eutypa lata UCREL1	putative mitosis protein dim1 protein	274	3.00E-93	88.00%	95.00%	XP_007791937.1
	Fusarium avenaceum	Thioredoxin-like protein 4a	273	5.00E-93	89.00%	95.00%	KIL90829.1
	Colletotrichum gloeosporioides Nara gc5	mitosis protein dim1	271	4.00E-92	87.00%	95.00%	XP_007278115.1
	Phaeoacremonium minimum UCRPA7	putative mitosis protein dim1 protein	270	8.00E-92	86.00%	95.00%	XP_007916280.1

Table S2E

Description in Figure 3	Scaffold	Transcript accession	Conserved domains
WD40 domain-containing	NT_165939.1 17638-22719 -	XM_001209259.1 XP_001209259.1	WD40 domain (E-value=2.74e-77); WD40 repeat (E-value=5.39e-63); N-terminal domain of NWD NACHT-NTPase (E-value=1.91e-65); WD40 domain (E-value=8.06e-64); protein SPA1-RELATED (E-value=9.57e-12); WD40 repeats (E-value=1.15e-09); WD domain, G-beta repeat (E-value=6.79e-09); NACHT domain (E-value=1.69e-08); WD40 repeats (E-value=3.41e-08); WD domain, G-beta repeat (E-value=6.29e-08); protein SPA1-RELATED (E-value=1.19e-06); tol-pal system beta propeller repeat protein TolB (E-value=1.32e-05); AAA+-type ATPase, SpoVK/Ycf46/Vps4 family (E-value=1.44e-04); Nup133 N terminal like (E-value=2.36e-03); Lipoprotein LpqB beta-propeller domain (E-value=6.14e-03)
Esterase	NT_165939.1 27708-32854 -	XM_001209260.1 XP_001209260.1	SEST_like (E-value=5.61e-59); SGNH_hydrolase subfamily (E-value=2.33e-47); GDSL-like lipase/acylhydrolase (E-value=5.92e-22); lysophospholipase L1 or related esterase (E-value=6.85e-08); Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella (E-value=6.08e-07); Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella (E-value=9.23e-05); Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella (E-value=1.76e-03); Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella (E-value=2.57e-03); GDSL-like Lipase/Acylhydrolase family (E-value=3.93e-03); Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella (E-value=7.08e-03)
Hypothetical	NT_165939.1 33617-34603 +	XM_001209261.1 XP_001209261.1	No conserved domains
Cytochrome P450	NT_165939.1 36186-37918 -	XM_001209262.1 XP_001209262.1	Cytochrome P450 (E-value=5.39e-30); Cytochrome P450 (E-value=3.40e-18); flavonoid 3'-monooxygenase (E-value=2.44e-14); carotene beta-ring hydroxylase (E-value=7.38e-10); cytochrome P450, cyclodipeptide synthase-associated (E-value=8.19e-07)

Table S2E (continued)

Description	Location	Accessions	Conserved domains
Nonaketide synthase	NT_165939.1 38401-47965 +	XM_001209263.1 XP_001209263.1	PKS (E-value=0e+00); Acyl transferase domain (AT) in PKS enzymes (E-value=0e+00); Beta-ketoacyl synthase (KS) (E-value=1.49e-145); AT in PKS (E-value=2.64e-92); KS, N-terminal (E-value=2.62e-86); polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=1.24e-65); Ketoreductase (KR) (E-value=2.64e-57); KR (E-value=7.23e-55); KR domain of fatty acid synthase (E-value=6.49e-44); PKS dehydratase (DH) (E-value=1.17e-41); AT domain (E-value=6.56e-40); peptide synthase (E-value=2.30e-33); 3-oxoacyl-[acyl-carrier-protein (ACP)] synthase (E-value=6.34e-29); Malonyl CoA-ACP transacylase (E-value=6.21e-31); Condensation domain (E-value=3.79e-29); Malonyl CoA-ACP transacylase (E-value=1.05e-20); Non-ribosomal peptide synthetase component F (E-value=9.71e-20); Methyltransferase domain (MT) (E-value=2.52e-19); DH in PKS (E-value=5.56e-16); KS C-terminal extension (E-value=3.59e-10); polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=2.15e-09); [ACP] S-malonyltransferase (E-value=3.82e-09); short chain dehydrogenase (E-value=1.71e-08); Ubiquinone/menaquinone biosynthesis C-methylase UbiE (E-value=3.89e-07); S-adenosylmethionine-dependent MT (E-value=5.43e-07); Thiolase, N-terminal (E-value=2.67e-06); ubiquinone/menaquinone biosynthesis MT (E-value=4.31e-05); Phosphopantetheine attachment site (PP) (E-value=1.29e-04); iterative type I PKS product template domain (E-value=3.59e-04); ubiquinone/menaquinone biosynthesis MT (E-value=4.69e-04); KR and fatty acid synthase, complex SDRs (E-value=1.25e-03); short chain dehydrogenase (E-value=2.78e-03); HxxPF-repeated domain (E-value=4.50e-03); KR domain of fatty acid synthase (E-value=5.58e-03)
Serine hydrolase	NT_165939.1 48199-48969 -	XM_001209264.1 XP_001209264.1	Serine hydrolase (FSH1) (E-value=1.52e-32); Alpha/beta superfamily hydrolase (E-value=4.15e-03)
Enoyl reductase	NT_165939.1 49367-50619 +	XM_001209265.1 XP_001209265.1	Enoyl reductase-like (E-value=6.06e-137); zinc-binding alcohol dehydrogenase family protein (E-value=6.21e-25); Enoylreductase (E-value=7.77e-22); alcohol dehydrogenase (E-value=1.52e-19); Zn-dependent alcohol dehydrogenase (E-value=2.96e-15); Alcohol dehydrogenase GroES-like domain (E-value=7.28e-10)
Beta-lactamase	NT_165939.1 50999-52392 -	XM_001209266.1 XP_001209266.1	Beta-lactamase (E-value=7.55e-44); CubicO group peptidase, beta-lactamase class C family (E-value=1.24e-30); putative periplasmic esterase (E-value=2.41e-13)
Hydroxy-methyl-glutaryl-coenzyme A reductase	NT_165939.1 52947-56256 -	XM_001209267.1 XP_001209267.1	Class I hydroxymethylglutaryl-coenzyme A reductase (E-value=0e+00); Hydroxymethylglutaryl-coenzyme A reductase (E-value=0e+00); 3-hydroxy-3-methylglutaryl-coenzyme A reductase (E-value=0e+00); 3-hydroxy-3-methylglutaryl Coenzyme A reductase (E-value=5.15e-178); Hydroxymethylglutaryl-CoA reductase (E-value=1.09e-148); N-terminal domain with HPIH motif (E-value=6.63e-44); Sterol-sensing domain of SREBP cleavage-activation (E-value=1.59e-08)
Transcription factor	NT_165939.1 57546-59057 +	XM_001209268.1 XP_001209268.1	GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain (E-value=3.10e-09); GAL4-like Zn(II)2Cys6 (or C6 zinc) binuclear cluster DNA-binding domain (E-value=2.90e-08); Fungal Zn(2)-Cys(6) binuclear cluster domain (E-value=1.54e-07)

Table 2SE (continued)

Description	Location	Accessions	Conserved domains
Drug resistance transporter, EmrB/QacA subfamily	NT_165939.1 60624-62784 +	XM_001209269.1 XP_001209269.1	Drug resistance transporter, EmrB/QacA subfamily (E-value=8.94e-38); Major Facilitator Superfamily (E-value=6.69e-20); The Major Facilitator Superfamily (E-value=4.15e-19); putative transporter (E-value=2.48e-17); Predicted arabinose efflux permease, MFS family (E-value=6.25e-13); Multidrug resistance protein (E-value=2.24e-10)
Diketide synthase	NT_165939.1 63664-71705 +	XM_001209270.1 XP_001209270.1	AT in PKS (E-value=3.29e-173); PKS (E-value=8.11e-156); Enoylreductase (ER) (E-value=3.30e-121); ER of PKS (E-value=5.66e-114); KS domain (E-value=8.59e-107); KS, N-terminal (E-value=3.00e-70); AT in PKS (E-value=1.74e-68); KR domain (E-value=5.15e-64); KR domain (E-value=6.09e-60); KR domain of fatty acid synthase (E-value=4.04e-56); PKS DH domain (E-value=2.68e-54); putative NAD(P)H quinone oxidoreductase, PIG3 family (E-value=1.75e-45); NADPH:quinone reductase or related Zn-dependent oxidoreductase (E-value=8.99e-45); polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=7.72e-40); DH in PKS (E-value=6.30e-36); 3-oxoacyl-ACP synthase (E-value=1.47e-25); Malonyl CoA-ACP transacylase (E-value=5.38e-22); AT domain (E-value=1.72e-21); alcohol dehydrogenase (E-value=6.05e-21); MT domain (E-value=1.94e-20); Zinc-binding dehydrogenase (E-value=9.87e-16); Malonyl CoA-ACP transacylase (E-value=1.13e-15); S-adenosylmethionine-dependent MT (E-value=6.90e-13); 3-oxoacyl-(ACP) reductase (E-value=1.04e-10); PP (E-value=3.65e-10); 3-ketoacyl-(ACP) reductase (E-value=5.80e-10); PP (E-value=6.15e-08); Ubiquinone/menaquinone biosynthesis C-methylase UbiE (E-value=4.15e-07); short chain dehydrogenase (E-value=2.73e-06); Thiolase, N-terminal (E-value=1.57e-04); SAM-dependent MT (E-value=1.81e-04); [ACP] S-malonyltransferase (E-value=6.97e-04); KR domain of fatty acid synthase (E-value=7.65e-04); KS C-terminal extension (E-value=1.07e-03); MT in PKS (E-value=1.92e-03); ubiquinone/menaquinone biosynthesis methyltransferases (E-value=8.89e-03)
C6 Transcription factor	NT_165939.1 73159-75637 -	XM_001209271.1 XP_001209271.1	Fungal transcription factor regulatory middle homology region (E-value=2.03e-25); GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain (E-value=4.32e-10); GAL4-like Zn(II)2Cys6 binuclear cluster DNA-binding domain (E-value=9.04e-10); Fungal Zn(2)-Cys(6) binuclear cluster domain (E-value=1.02e-09); Fungal specific transcription factor domain (E-value=6.41e-08); Fungal specific transcription factor domain (E-value=4.44e-06)
Mitochondrial carrier protein	NT_165939.1 77394-78692 -	XM_001209272.1 XP_001209272.1	Mitochondrial carrier protein (E-value=5.49e-18); Mitochondrial carrier protein (E-value=2.22e-14); Mitochondrial carrier protein (E-value=2.13e-12); Mitochondrial carrier protein (E-value=1.82e-11)
2-methylcitrate dehydratase PrpD-like	NT_165939.1 79382-80910 +	XM_001209273.1 XP_001209273.1	MmgE/PrpD family (E-value=9.94e-111); 2-methylcitrate dehydratase PrpD (E-value=6.56e-45)
MFS transporter	NT_165939.1 82183-83574 -	XM_001209274.1 XP_001209274.1	Major Facilitator Superfamily (E-value=2.45e-27); Multidrug resistance protein (E-value=3.82e-14); Major Facilitator Superfamily (E-value=1.72e-27); drug resistance transporter, Bcr/CflA subfamily (E-value=2.84e-15); Predicted arabinose efflux permease, MFS family (E-value=4.76e-13); bicyclomycin/multidrug efflux system (E-value=9.61e-12)

Table 2SE (continued)

Description	Location	Accessions	Conserved domains
Cytochrome P450	NT_165939.1 85872-87559 -	XM_001209275.1 XP_001209275.1	Cytochrome P450 (E-value=1.43e-49); Cytochrome P450 (E-value=6.55e-33); Cytochrome P450 (E-value=2.00e-30); Cytochrome P450, cyclodipeptide synthase-associated (E-value=1.13e-12); carotene beta-ring hydroxylase (E-value=1.82e-10)
Glycosyl hydrolase family 115	NT_165939.1 88864-91946 +	XM_001209276.1 XP_001209276.1	Glycosyl hydrolase family 115 (E-value=2.15e-174)

Table S2F

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
WD40 domain-containing	*Pseudogymnoascus sp. VKM F-4515 (FW-2607)*	hypothetical protein V496_02264	1044	0.00E+00	42.00%	57.00%	KFY65896.1
	Chaetomium globosum CBS 148.51	hypothetical protein CHGG_08200	1026	0.00E+00	42.00%	56.00%	XP_001225856.1
	Metarhizium robertsii ARSEF 23	NACHT and WD40 domain protein	969	0.00E+00	41.00%	56.00%	XP_007826000.1
	Metarhizium brunneum ARSEF 3297	NACHT and WD40 domain protein	954	0.00E+00	41.00%	55.00%	XP_014544162.1
	Metarhizium anisopliae	NACHT and WD40 domain protein	954	0.00E+00	40.00%	55.00%	KFG79444.1
	Metarhizium guizhouense ARSEF 977	NACHT and WD40 domain protein	952	0.00E+00	40.00%	55.00%	KID82919.1
	Madurella mycetomatis	Vegetative incompatibility protein HET-E-1	932	0.00E+00	40.00%	54.00%	KXX78699.1
	Pseudogymnoascus sp. VKM F-4517 (FW-2822)	hypothetical protein V498_05551, partial	929	0.00E+00	43.00%	57.00%	KFY91248.1
	Trichoderma gamsii	hypothetical protein TGAM01_10307	905	0.00E+00	38.00%	54.00%	KUE95230.1
	Trichoderma harzianum	hypothetical protein THAR02_10976	875	0.00E+00	38.00%	54.00%	KKO96923.1
Esterase	*Aspergillus terreus*	esterase	2831	0.00E+00	90.00%	91.00%	AAD34550.1
	Aspergillus lentulus	hypothetical protein ALT_9506	1919	0.00E+00	87.00%	92.00%	GAQ12185.1
	Aspergillus fumigatus var. RP-2014	esterase	1714	0.00E+00	87.00%	92.00%	KEY78975.1
	Aspergillus fumigatus A1163	esterase, putative	1709	0.00E+00	88.00%	93.00%	EDP55661.1
	Aspergillus fumigatus Z5	esterase, putative	1706	0.00E+00	88.00%	93.00%	KMK57412.1
	Aspergillus fumigatus Af293	esterase	1705	0.00E+00	88.00%	93.00%	XP_750067.1
	Aspergillus nidulans FGSC A4	hypothetical protein AN9260.2	1487	0.00E+00	78.00%	84.00%	XP_682529.1
	Aspergillus nidulans FGSC A4	hypothetical protein AN6464.2	1097	0.00E+00	62.00%	74.00%	XP_664068.1
	Pseudogymnoascus sp. VKM F-4519 (FW-2642)	hypothetical protein V501_04321	953	0.00E+00	57.00%	69.00%	KFZ12219.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Esterase	*Penicillium freii*	hypothetical protein ACN42_g1883	902	0.00E+00	53.00%	67.00%	KUM65192.1
Hypothetical	*Pseudogymnoascus sp. VKM F-4517 (FW-2822)*	hypothetical protein V498_00485	161	2.00E-44	38.00%	54.00%	KFY99837.1
	Neonectria ditissima	hypothetical protein AK830_g4815	155	8.00E-42	36.00%	56.00%	KPM41720.1
	Pseudogymnoascus sp. VKM F-4520 (FW-2644)	hypothetical protein V502_04103	153	4.00E-41	37.00%	54.00%	KFZ18427.1
	Oidiodendron maius Zn	hypothetical protein OIDMADRAFT_35258	150	3.00E-40	35.00%	50.00%	KIM94224.1
	Colletotrichum orbiculare MAFF 240422	hypothetical protein Cob_02205	143	5.00E-37	37.00%	54.00%	ENH76480.1
	Colletotrichum gloeosporioides Cg-14	hypothetical protein CGLO_04770	142	9.00E-37	35.00%	54.00%	EQB55315.1
	Pseudogymnoascus sp. 05NY08	hypothetical protein VF21_09178	142	1.00E-36	38.00%	50.00%	OBT72111.1
	Oidiodendron maius Zn	hypothetical protein OIDMADRAFT_134785	140	2.00E-36	36.00%	50.00%	KIM94817.1
	Pseudogymnoascus sp. 23342-1-I1	hypothetical protein VE03_09356	140	1.00E-35	34.00%	50.00%	OBT61485.1
	Podospora anserina S mat+	hypothetical protein	141	2.00E-35	37.00%	52.00%	XP_001910653.1
Cytochrome P450	*Monascus pilosus*	Dihydromonacolin L monooxygenase mokC	957	0.00E+00	86.00%	92.00%	Q3S2T8.1
	Penicillium citrinum	DihydroML-246B monooxygenase mlcC	792	0.00E+00	73.00%	86.00%	Q8J0F6.1
	Colletotrichum higginsianum	P450 monooxygenase	773	0.00E+00	70.00%	82.00%	CCF38811.1
	Fusarium langsethiae	cytochrome P450 monooxygenase	422	1.00E-139	42.00%	59.00%	KPA36514.1
	Colletotrichum gloeosporioides Nara gc5	cytochrome P450 monooxygenase	399	1.00E-130	42.00%	61.00%	XP_007275304.1
	Colletotrichum gloeosporioides Cg-14	cytochrome P450	394	5.00E-129	42.00%	62.00%	EQB56061.1
	Aspergillus niger ATCC 1015	hypothetical protein ASPNIDRAFT_182561	389	2.00E-127	40.00%	62.00%	EHA24759.1
	Aspergillus niger CBS 513.88	cytochrome P450 monooxygenase	388	1.00E-126	39.00%	62.00%	XP_001397042.2

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Cytochrome P450	*Oidiodendron maius Zn*	hypothetical protein OIDMADRAFT_117721	387	3.00E-126	38.00%	59.00%	KIN03514.1
	Colletotrichum simmondsii	cytochrome P450	385	3.00E-125	39.00%	59.00%	KXH49511.1
Nonaketide synthase	*Monascus pilosus*	Lovastatin nonaketide synthase mokA	4723	0.00E+00	76.00%	84.00%	Q3S2T9.1
	Colletotrichum higginsianum IMI 349063	Beta-ketoacyl synthase domain-containing protein	3964	0.00E+00	64.00%	77.00%	OBR06526.1
	Penicillium citrinum	Compactin nonaketide synthase mokA	3895	0.00E+00	63.00%	76.00%	Q8J0F7.1
	Colletotrichum fioriniae PJ7	Beta-ketoacyl synthase domain-containing protein	3196	0.00E+00	54.00%	70.00%	XP_007602284.1
	Colletotrichum salicis	Beta-ketoacyl synthase domain-containing protein	3189	0.00E+00	54.00%	69.00%	KXH63779.1
	Colletotrichum simmondsii	Beta-ketoacyl synthase domain-containing protein	3187	0.00E+00	54.00%	69.00%	KXH48923.1
	Colletotrichum incanum	Beta-ketoacyl synthase domain-containing protein	3054	0.00E+00	52.00%	67.00%	KZL81493.1
	Colletotrichum incanum	polyketide synthase	3038	0.00E+00	52.00%	67.00%	KZL80660.1
	Colletotrichum tofieldiae	polyketide synthase	3023	0.00E+00	52.00%	67.00%	KZL71928.1
	Colletotrichum nymphaeae SA-01	Beta-ketoacyl synthase domain-containing protein	3014	0.00E+00	52.00%	67.00%	KXH44578.1
Serine hydrolase	*Monascus pilosus*	Esterase mokD/Monacolin K biosynthesis protein D	357	4.00E-122	68.00%	76.00%	Q3S2U0.1
	Penicillium citrinum	Esterase mlcF/Compactin biosynthesis protein F	292	6.00E-97	57.00%	71.00%	Q8J0F8.1
	Colletotrichum higginsianum	oxidoreductase	261	3.00E-84	51.00%	63.00%	CCF43998.1
	Colletotrichum sublineola	hypothetical protein CSUB01_00033	229	3.00E-72	46.00%	62.00%	KDN72350.1
	Colletotrichum graminicola M1.001	hypothetical protein GLRG_11512	206	3.00E-63	42.00%	60.00%	XP_008100387.1
	Colletotrichum incanum	oxidoreductase	205	7.00E-63	41.00%	59.00%	KZL86154.1
	Colletotrichum nymphaeae SA-01	oxidoreductase	200	5.00E-61	39.00%	59.00%	KXH44575.1
	Colletotrichum fioriniae PJ7	oxidoreductase	200	8.00E-61	39.00%	59.00%	XP_007602287.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Serine hydrolase	*Colletotrichum orbiculare MAFF 240422*	phospholipase carboxylesterase	195	1.00E-57	39.00%	58.00%	ENH82921.1
	Colletotrichum simmondsii	alcohol dehydrogenase GroES-like domain-containing protein	200	4.00E-57	39.00%	59.00%	KXH48925.1
Enoyl reductase	*Monascus pilosus*	Deshydrogenase mokE/Enoyl reductase	626	0.00E+00	84.00%	91.00%	Q3S2U1.1
	Colletotrichum higginsianum IMI 349063	Enoyl reductase	572	0.00E+00	78.00%	86.00%	OBR06532.1
	Colletotrichum higginsianum	Enoyl reductase	570	0.00E+00	77.00%	86.00%	CCF43999.1
	Penicillium citrinum	Deshydrogenase mlcG/Compactin biosynthesis protein G	553	0.00E+00	73.00%	83.00%	Q8J0F9.1
	Colletotrichum orbiculare MAFF 240422	Zinc-binding dehydrogenase family	466	5.00E-162	63.00%	75.00%	ENH82920.1
	Colletotrichum incanum	Enoyl reductase	461	4.00E-160	62.00%	75.00%	KZL80653.1
	Colletotrichum tofieldiae	Enoyl reductase	461	6.00E-160	62.00%	74.00%	KZL71906.1
	Colletotrichum nymphaeae SA-01	alcohol dehydrogenase GroES-like domain-containing protein	459	3.00E-159	61.00%	74.00%	KXH44576.1
	Colletotrichum fioriniae PJ7	alcohol dehydrogenase GroES-like domain-containing protein	456	3.00E-158	61.00%	73.00%	XP_007602286.1
	Pyrenophora teres f. teres 0-1	hypothetical protein PTT_16764	457	4.00E-158	61.00%	75.00%	XP_003304242.1
Beta-lactamase	*Monascus pilosus*	Acyltransferase mokF/Lovastatin hydrolase	659	0.00E+00	88.00%	94.00%	Q3S2U2.1
	Penicillium citrinum	Acyltransferase mlcH/Compactin biosynthesis protein H	652	0.00E+00	75.00%	87.00%	Q8J0G0.1
	Colletotrichum higginsianum	Beta-lactamase	613	0.00E+00	71.00%	85.00%	CCF38810.1
	Colletotrichum higginsianum IMI 349063	Beta-lactamase	601	0.00E+00	61.00%	80.00%	OBR06528.1
	Colletotrichum gloeosporioides Cg-14	Beta-lactamase	475	6.00E-164	55.00%	72.00%	EQB52990.1
	Colletotrichum higginsianum IMI 349063	Beta-lactamase	469	1.00E-161	55.00%	73.00%	OBR11266.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Beta-lactamase	*Colletotrichum orbiculare MAFF 240422*	transesterase	451	2.00E-154	54.00%	70.00%	ENH86454.1
	Colletotrichum gloeosporioides Nara gc5	transesterase	409	4.00E-139	56.00%	72.00%	XP_007280257.1
	Ascochyta rabiei	hypothetical protein ST47_g1757	279	3.00E-87	40.00%	58.00%	KZM27105.1
	Parastagonospora nodorum SN15	hypothetical protein SNOG_09364	274	2.00E-85	38.00%	57.00%	XP_001799659.1
Hydroxy-methylglutaryl-coenzyme A reductase	*Aspergillus rambellii*	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1614	0.00E+00	76.00%	85.00%	KKK24297.1
	Monascus pilosus	3-hydroxy-3-methylglutaryl-coenzyme A reductase mokG	1528	0.00E+00	73.00%	82.00%	Q3S2U3.1
	Aspergillus oryzae RIB40	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1494	0.00E+00	72.00%	81.00%	XP_001826888.2
	Aspergillus flavus AF70	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1492	0.00E+00	72.00%	81.00%	KOC17804.1
	Aspergillus flavus NRRL3357	HMG-CoA reductase, putative	1484	0.00E+00	72.00%	81.00%	XP_002385162.1
	Aspergillus oryzae 3.042	3-hydroxy-3-methylglutaryl-CoA reductase	1481	0.00E+00	72.00%	81.00%	EIT73568.1
	Aspergillus nomius NRRL 13137	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1349	0.00E+00	67.00%	76.00%	XP_015410655.1
	Byssochlamys spectabilis No. 5	HMG-CoA reductase	1072	0.00E+00	52.00%	65.00%	GAD92459.1
	Rasamsonia emersonii CBS 393.64	Hydroxymethylglutaryl-CoA reductase (NADPH)	1071	0.00E+00	52.00%	66.00%	XP_013327653.1
	Aspergillus clavatus NRRL 1	HMG-CoA reductase	1070	0.00E+00	51.00%	67.00%	XP_001272815.1
Transcription factor	*Aspergillus rambellii*	hypothetical protein ARAM_007686	505	2.00E-173	55.00%	64.00%	KKK18794.1
	Monascus pilosus	Transcription factor mokH/Monacolin K biosynthesis protein H	451	2.00E-152	55.00%	62.00%	Q3S2U4.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Transcription factor	*Penicillium citrinum*	Transcription factor mlcR/Compactin biosynthesis protein H	294	5.00E-91	39.00%	51.00%	Q8J0F2.1
	Colletotrichum higginsianum IMI 349063	hypothetical protein CH63R_10651	127	1.00E-28	28.00%	42.00%	OBR06531.1
	Sordaria macrospora k-hell	hypothetical protein SMAC_01201	76.3	2.00E-11	26.00%	40.00%	XP_003350304.1
	Pestalotiopsis fici W106-1	hypothetical protein PFICI_01575	65.9	4.00E-08	23.00%	35.00%	XP_007828347.1
	Fonsecaea pedrosoi CBS 271.37	hypothetical protein Z517_05667	63.9	2.00E-07	54.00%	68.00%	XP_013286448.1
	Cladophialophora immunda	hypothetical protein PV07_02976	64.3	2.00E-07	54.00%	68.00%	XP_016251535.1
	Cladophialophora bantiana CBS 173.52	hypothetical protein Z519_04310	63.5	3.00E-07	54.00%	68.00%	XP_016622394.1
	Fonsecaea erecta	hypothetical protein AYL99_11332	63.2	4.00E-07	54.00%	68.00%	OAP54231.1
Drug resistance transporter, EmrB/QacA subfamily	*Monascus pilosus*	Efflux pump mokI/Monacolin K biosynthesis protein I	834	0.00E+00	80.00%	87.00%	Q3S2U5.1
	Penicillium citrinum	Efflux pump mlcE/Compactin biosynthesis protein E	746	0.00E+00	69.00%	79.00%	Q8J0F3.1
	Xylona heveae TC161	efflux pump	737	0.00E+00	68.00%	79.00%	KZF26031.1
	Hirsutella minnesotensis 3608	hypothetical protein HIM_08274	684	0.00E+00	64.00%	76.00%	KJZ72348.1
	Penicillium expansum	Major facilitator superfamily domain, general substrate transporter	647	0.00E+00	60.00%	75.00%	KGO47551.1
	Penicillium expansum	Major facilitator superfamily domain, general substrate transporter	645	0.00E+00	60.00%	75.00%	XP_016597104.1
	Colletotrichum higginsianum	efflux pump	613	0.00E+00	58.00%	73.00%	CCF43997.1
	Oidiodendron maius Zn	hypothetical protein OIDMADRAFT_99363	596	0.00E+00	57.00%	72.00%	KIN06910.1
	Pseudogymnoascus sp. VKM F-4246	hypothetical protein V492_08489	578	0.00E+00	54.00%	72.00%	KFY05508.1
	Pseudogymnoascus sp. VKM F-4513 (FW-928)	hypothetical protein V494_08341	575	0.00E+00	55.00%	72.00%	KFY29990.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Diketide synthase	*Monascus pilosus*	Lovastatin diketide synthase mokB/Monacolin K biosynthesis protein B	3546	0.00E+00	70.00%	79.00%	Q3S2U6.1
	Penicillium citrinum	Compactin diketide synthase mokB	2915	0.00E+00	58.00%	71.00%	Q8J0F5.1
	Colletotrichum higginsianum IMI 349063	Type I polyketide synthase	2707	0.00E+00	55.00%	69.00%	OBR06529.1
	Aspergillus niger ATCC 1015	hypothetical protein ASPNIDRAFT_188817	2281	0.00E+00	48.00%	63.00%	EHA28244.1
	Aspergillus niger CBS 513.88	polyketide synthase	2274	0.00E+00	48.00%	63.00%	XP_001393508.2
	Aspergillus niger	unnamed protein product	2267	0.00E+00	47.00%	63.00%	CAK40131.1
	Penicillium digitatum Pd1	hypothetical protein PDIP_46770	2042	0.00E+00	42.00%	60.00%	XP_014534966.1
	Oidiodendron maius Zn	hypothetical protein OIDMADRAFT_207730	2029	0.00E+00	43.00%	60.00%	KIM94092.1
	Thielavia terrestris NRRL 8126	hypothetical protein THITE_2088794	1999	0.00E+00	43.00%	59.00%	XP_003653639.1
	Podospora anserina S mat+	hypothetical protein	1982	0.00E+00	42.00%	59.00%	XP_001908964.1
C6 Transcription factor	*Aspergillus kawachii IFO 4308*	C6 transcription factor	434	9.00E-139	37.00%	56.00%	GAA88110.1
	Aspergillus niger	C6 transcription factor	432	5.00E-138	37.00%	56.00%	GAQ45093.1
	Aspergillus niger	unnamed protein product	428	2.00E-136	38.00%	55.00%	CAK45760.1
	Aspergillus niger ATCC 1015	hypothetical protein ASPNIDRAFT_120084	417	3.00E-133	39.00%	56.00%	EHA18675.1
	Aspergillus luchuensis	C6 transcription factor	417	5.00E-132	36.00%	55.00%	GAT22624.1
	Aspergillus clavatus NRRL 1	C6 transcription factor, putative	394	7.00E-123	36.00%	56.00%	XP_001274697.1
	Aspergillus fumigatus Z5	C6 transcription factor	391	1.00E-121	36.00%	53.00%	KMK62905.1
	Aspergillus fumigatus var. RP-2014	transcription factor C6	389	1.00E-120	36.00%	53.00%	KEY79761.1
	Aspergillus udagawae	Phosphate-repressible phosphate permease	385	1.00E-117	37.00%	55.00%	GAO83017.1
	Aspergillus fumigatus A1163	C6 transcription factor, putative	380	2.00E-117	35.00%	52.00%	EDP52162.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Mitochondrial carrier protein	*Rosellinia necatrix*	putative tricarboxylate transporter mitochondrial carrier protein	248	3.00E-78	46.00%	63.00%	GAP93451.1
	Capronia semi-immersa	hypothetical protein PV04_05254	242	8.00E-76	46.00%	64.00%	KIW69373.1
	Sclerotinia borealis F-4157	tricarboxylate transport protein	243	2.00E-75	45.00%	65.00%	ESZ96799.1
	Ceratocystis platani	putative mitochondrial carrier C19G12.05	240	6.00E-75	43.00%	62.00%	KKF97159.1
	Stachybotrys chlorohalonata IBT 40285	hypothetical protein S40285_03300	239	1.00E-74	46.00%	64.00%	KFA70073.1
	Cyphellophora europaea	hypothetical protein HMPREF1541_07556	239	2.00E-74	44.00%	64.00%	XP_008720102.1
	Cladophialophora yegresii CBS 114405	hypothetical protein A1O7_04054	238	2.00E-74	45.00%	63.00%	XP_007756260.1
	Cladophialophora bantiana CBS 173.52	hypothetical protein Z519_11998	238	4.00E-74	44.00%	63.00%	XP_016614031.1
	Cladophialophora psammophila CBS 110553	hypothetical protein A1O5_03039	237	5.00E-74	45.00%	64.00%	XP_007741842.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_09292	238	5.00E-74	45.00%	64.00%	KEY65502.1
2-methylcitrate dehydratase PrpD-like	*Talaromyces cellulolyticus*	Immune-responsive protein	532	0.00E+00	59.00%	74.00%	GAM43161.1
	Aspergillus nomius NRRL 13137	Immune-responsive protein	526	0.00E+00	55.00%	70.00%	XP_015412298.1
	Aspergillus flavus AF70	Immune-responsive protein	523	2.00E-180	55.00%	69.00%	KOC13246.1
	Aspergillus oryzae RIB40	Immune-responsive protein	522	4.00E-180	55.00%	69.00%	XP_001827196.1
	Rosellinia necatrix	putative family protein	510	3.00E-175	53.00%	67.00%	GAP92989.1
	Xylona heveae TC161	2-methylcitrate dehydratase PrpD	504	3.00E-173	53.00%	67.00%	KZF20967.1
	Neofusicoccum parvum UCRNP2	putative cis-aconitic acid decarboxylase protein	501	9.00E-172	54.00%	67.00%	XP_007587253.1
	Aspergillus parasiticus SU-1	MmgE/PrpD family protein	516	3.00E-171	55.00%	69.00%	KJK67520.1
	Aspergillus niger	Cis-aconitic acid decarboxylase	493	4.00E-168	50.00%	66.00%	GAQ43869.1
	Aspergillus kawachii IFO 4308	Cis-aconitic acid decarboxylase	493	8.00E-168	51.00%	67.00%	GAA88205.1
MFS transporter	*Aspergillus oryzae RIB40*	MFS transporter	568	0.00E+00	71.00%	84.00%	XP_001819137.1
	Aspergillus flavus NRRL3357	MFS transporter, putative	568	0.00E+00	71.00%	84.00%	XP_002382138.1

Table S2F (continued)

***Aspergillus terreus* Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
MFS transporter	*Aspergillus flavus AF70*	MFS transporter	568	0.00E+00	71.00%	83.00%	KOC10872.1
	Aspergillus parasiticus SU-1	Major Facilitator Superfamily protein	565	0.00E+00	71.00%	83.00%	KJK63559.1
	Aspergillus luchuensis	MFS transporter	506	6.00E-176	64.00%	77.00%	GAT28158.1
	Aspergillus udagawae	uncharacterized transporter C750.02c/PB2B2.16c	498	2.00E-172	64.00%	78.00%	GAO90936.1
	Penicillium roqueforti FM164	Major facilitator superfamily	489	1.00E-168	65.00%	77.00%	CDM36253.1
	Aspergillus fumigatus Af293	MFS transporter	484	2.00E-167	64.00%	77.00%	XP_749665.1
	Aspergillus niger	hypothetical protein ABL_07290	505	3.00E-167	64.00%	77.00%	GAQ44629.1
	Penicillium freii	hypothetical protein ACN42_g9230	486	4.00E-167	65.00%	76.00%	KUM57941.1
Cytochrome P450	*Aspergillus parasiticus SU-1*	cytochrome P450	893	0.00E+00	78.00%	90.00%	KJK65535.1
	Aspergillus fischeri NRRL 181	cytochrome P450 phenylacetate 2-hydroxylase, putative	882	0.00E+00	78.00%	90.00%	XP_001266359.1
	Aspergillus lentulus	Phenylacetate 2-hydroxylase	878	0.00E+00	79.00%	90.00%	GAQ08779.1
	Aspergillus flavus AF70	putative cytochrome P450 phenylacetate 2-hydroxylase	876	0.00E+00	79.00%	90.00%	KOC07376.1
	Aspergillus fumigatus Af293	cytochrome P450 phenylacetate 2-hydroxylase	875	0.00E+00	77.00%	90.00%	XP_748171.1
	Aspergillus flavus NRRL3357	cytochrome P450 phenylacetate 2-hydroxylase, putative	874	0.00E+00	79.00%	90.00%	XP_002373986.1
	Apsergillus udagawae	Phenylacetate 2-hydroxylase	872	0.00E+00	76.00%	90.00%	GAO84546.1
	Aspergillus niger CBS 513.88	Phenylacetate 2-hydroxylase	868	0.00E+00	76.00%	88.00%	XP_001397416.1
	Aspergillus niger ATCC 1015	hypothetical protein ASPNIDRAFT_49153	868	0.00E+00	76.00%	88.00%	EHA21697.1
	Aspergillus kawachii IFO 4308	cytochrome P450 phenylacetate 2-hydroxylase	868	0.00E+00	76.00%	88.00%	GAA84736.1
Glycosyl hydrolase family 115	*Pestalotiopsis fici W106-1*	hypothetical protein PFICI_00949	914	0.00E+00	49.00%	65.00%	XP_007827721.1
	Pyrenochaeta sp. DS3sAY3a	hypothetical protein IQ07DRAFT_558914	887	0.00E+00	47.00%	66.00%	OAL54666.1

Table S2F (continued)

Aspergillus terreus **Lovastatin biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Glycosyl hydrolase family 115	*Pseudogymnoascus sp. VKM F-4515 (FW-2607)*	hypothetical protein V496_10041	886	0.00E+00	47.00%	65.00%	KFY49400.1
	Stachybotrys chartarum IBT 40293	hypothetical protein S40293_01071	885	0.00E+00	48.00%	63.00%	KFA56478.1
	Stachybotrys chlorohalonata IBT 40285	hypothetical protein S40285_02546	885	0.00E+00	49.00%	65.00%	KFA68067.1
	Pseudogymnoascus sp. VKM F-4517 (FW-2822)	hypothetical protein V498_09471	885	0.00E+00	47.00%	65.00%	KFY76946.1
	Stachybotrys chartarum IBT 7711	hypothetical protein S7711_01638	884	0.00E+00	48.00%	63.00%	KEY75198.1
	Aspergillus nomius NRRL 13137	hypothetical protein ANOM_001506	880	0.00E+00	48.00%	64.00%	XP_015411468.1
	Paraphaeosphaeria sporulosa	hypothetical protein CC84DRAFT_1119522	879	0.00E+00	46.00%	64.00%	OAG05956.1
	Magnaporthe oryzae 70-15	hypothetical protein MGG_03407	872	0.00E+00	47.00%	66.00%	XP_003716533.1

Table S2G

Description	Location	Accessions	Conserved domains
Dehydrogenase	LC011911.1 8443-10353 +	BAQ25461.1	choline dehydrogenase (E-value=1.77e-95); choline dehydrogenase or related flavoprotein (E-value=1.26e-93); choline dehydrogenase (E-value=4.10e-72); GMC oxidoreductase (E-value=7.06e-44); GMC oxidoreductase (E-value=5.34e-32)
Dehydrogenase	LC011911.1 17557-19380 -	BAQ25462.1	FAD binding domain (E-value=4.97e-14); Berberine and berberine like (E-value=4.28e-10); FAD/FMN-containing dehydrogenase (E-value=5.14e-10); cytokinin dehydrogenase (E-value=9.97e-04)
bet4 Short-chain dehydrogenase	LC011911.1 19969-21148 +	BAQ25463.1	retinol dehydrogenase (retinol-DH), light dependent protochlorophyllide (Pchlide) (E-value=5.38e-50); oxidoreductase (E-value=6.34e-30); NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family (E-value=2.67e-21); short chain dehydrogenase (E-value=5.10e-21); short chain dehydrogenase (E-value=7.83e-18); light-dependent protochlorophyllide reductase (E-value=1.38e-11); rhamnulose-1-phosphate aldolase/alcohol dehydrogenase (E-value=2.24e-09); KR domain (E-value=3.65e-06); Short-chain dehydrogenase involved in D-alanine esterification of teichoic acids (E-value=1.63e-05); KR domain (E-value=8.59e-05)
bet3 Enoyl reductase	LC011911.1 21812-23000 +	BAQ25464.1	enoyl_reductase_like (E-value=5.82e-105); alcohol dehydrogenase GroES-like domain (E-value=2.42e-07); NADPH:quinone reductase or related Zn-dependent oxidoreductase (E-value=9.07e-37); enoylreductase (E-value=1.42e-13); alcohol dehydrogenase (E-value=7.32e-13); putative NAD(P)H quinone oxidoreductase, PIG3 family (E-value=1.46e-12)
bet2 Cytochrome P450	LC011911.1 23246-25006 -	BAQ25465.1	cytochrome P450 (E-value=2.10e-62); cytochrome P450 (E-value=1.14e-36); cytochrome P450 (E-value=9.16e-19); carotene beta-ring hydroxylase (E-value=6.76e-16)
bet1 PKS	LC011911.1 25772-34622 +	BAQ25466.1	PKS (E-value=0e+00); AT domain (E-value=3.43e-91); beta-ketoacyl synthase, N-terminal domain (E-value=1.56e-82); KR domain (E-value=8.34e-56); KR domain (E-value=1.77e-51); beta-ketoacyl reductase domain (E-value=2.41e-45); AT domain (E-value=5.44e-35); extended short-chain dehydrogenase, subgroup 1 (E-value=8.02e-35); PKS dehydratase (E-value=3.21e-29); male sterility protein (E-value=1.27e-26); malonyl CoA-ACP transacylase (E-value=2.23e-24); MT domain (E-value=1.03e-17); malonyl CoA-ACP transacylase (E-value=6.01e-17); dehydratase domain (E-value=1.46e-15); ubiquinone/menaquinone biosynthesis C-methylase UbiE (E-value=4.36e-07); S-adenosylmethionine-dependent methyltransferases (E-value=1.13e-06); [acyl-carrier protein] S-malonyltransferase (E-value=6.24e-06); ubiquinone/menaquinone biosynthesis methyltransferases (E-value=3.84e-04); acetyacetyl-CoA reductase (E-value=6.72e-04); PP attachment site (E-value=7.66e-03); AT domain (E-value=0e+00); beta-ketoacyl synthase (E-value=3.13e-138); polyketide-type polyunsaturated fatty acid synthase PfaA (E-value=1.57e-70); thioester reductase domain (E-value=7.83e-33); 3-oxoacyl-ACP synthase (E-value=4.20e-30); thioester reductase domain of alpha aminoadipate reductase Lys2 and NRPSs (E-value=1.86e-27); acetoacetyl-CoA reductase (E-value=6.57e-11); short chain dehydrogenase (E-value=5.60e-10); thiolase, N-terminal domain (E-value=3.04e-06); methyltransferase (E-value=4.77e-03)

Table S2G (continued)

Description	Location	Accessions	Conserved domains
Hypothetical	LC011911.1 35031-35549 -	BAQ25467.1	No conserved domains
Hypothetical	LC011911.1 36547-40380 -	BAQ25468.1	No conserved domains

Table S2H

Phoma betae **Betaenone biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
	Alternaria alternata	alcohol oxidase	994	0.00E+00	79%	87%	XP_018387630.1
	Neofusicoccum parvum UCRNP2	putative aryl-alcohol dehydrogenase protein	845	0.00E+00	67%	79%	XP_007587319.1
	Aspergillus niger CBS 513.88	aryl-alcohol dehydrogenase	696	0.00E+00	55%	73%	XP_001399178.1
	Aspergillus niger ATCC 1015	hypothetical protein ASPNIDRAFT 36617	696	0.00E+00	55%	73%	EHA22573.1
	Aspergillus niger	aryl-alcohol dehydrogenase	690	0.00E+00	56%	72%	GAQ35257.1
	Aspergillus kawachii IFO 4308	aryl-alcohol dehydrogenase	686	0.00E+00	55%	73%	GAA92522.1
	Aspergillus udagawae	alcohol dehydrogenase [acceptor]	677	0.00E+00	56%	71%	GAO81598.1
	Aspergillus fumigatus Z5	aryl-alcohol dehydrogenase	669	0.00E+00	56%	70%	KMK58320.1
	Aspergillus fumigatus var. RP-2014	aryl alcohol dehydrogenase	669	0.00E+00	55%	70%	KEY78899.1
Dehydrogenase	*Aspergillus fumigatus* Af293	aryl-alcohol dehydrogenase	669	0.00E+00	56%	70%	XP_746395.1
	Alternaria alternata	FAD-binding domain-containing protein	940	0.00E+00	78%	86%	XP_018382150.1
	Ascochyta rabiei	flavin adenine dinucleotide binding	939	0.00E+00	78%	87%	KZM21154.1
	Stemphylium lycopersici	isoamyl alcohol oxidase	925	0.00E+00	76%	86%	KNG46186.1
	Setosphaeria turcica Et28A	hypothetical protein SETTUDRAFT 100397	850	0.00E+00	71%	82%	XP_008031658.1
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 20037	841	0.00E+00	69%	83%	XP_003306800.1
	Pyrenophora tritici-repentis Pt-1C-BFP	isoamyl alcohol oxidase	800	0.00E+00	69%	81%	XP_001934279.1
	Glonium stellatum	putative FAD-dependent isoamyl alcohol oxidase	583	0.00E+00	52%	67%	OCL04327.1
Dehydrogenase	*Metarhizium rileyi* RCEF 4871	FAD-binding, type 2	566	0.00E+00	51%	65%	OAA34872.1

Table S2H (continued)

***Phoma betae* Betaenone biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
Dehydrogenase	*Bipolaris maydis* ATCC 48331	hypothetical protein COCC4DRAFT 65599	500	6.00E-173	81%	89%	XP_014074140.1
	Marssonina brunnea f. sp. *'multigermtubi'* MB m1	FAD binding domain protein	499	8.00E-168	47%	61%	XP_007296766.1
bet4 Short-chain dehydrogenase	*Ascochyta rabiei*	oxidoreductase	627	0.00E+00	86%	92%	KZM21152.1
	Setosphaeria turcica Et28A	hypothetical protein SETTUDRAFT 158084	603	0.00E+00	83%	90%	XP_008031659.1
	Bipolaris maydis ATCC 48331	hypothetical protein COCC4DRAFT 151132	603	0.00E+00	84%	91%	XP_014074141.1
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 20036	549	0.00E+00	78%	88%	XP_003306799.1
	Alternaria alternata	NAD(P)-binding protein	478	2.00E-168	87%	91%	XP_018382151.1
	Parastagonospora nodorum SN15	hypothetical protein SNOG 07862	472	2.00E-165	75%	84%	XP_001798189.1
	Pyrenophora tritici-repentis Pt-1C-BFP	retinol dehydrogenase 12	463	5.00E-162	75%	87%	XP_001934280.1
	Colletotrichum gloeosporioides Cg-14	short-chain dehydrogenase	422	3.00E-145	60%	77%	EQB51186.1
	Colletotrichum gloeosporioides Nara gc5	short-chain dehydrogenase, putative	421	2.00E-144	60%	77%	XP_007283081.1
	Diaporthe ampelina	putative short-chain dehydrogenase	417	1.00E-143	62%	78%	KKY31728.1
bet3 Enoyl reductase	*Stemphylium lycopersici*	zinc-binding dehydrogenase family oxidoreductase	641	0.00E+00	86%	91%	KNG46188.1
	Alternaria alternata	zinc-binding dehydrogenase family oxidoreductase	623	0.00E+00	87%	91%	XP_018382153.1
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 20034	613	0.00E+00	82%	88%	XP_003306797.1
	Ascochyta rabiei	oxidoreductase	610	0.00E+00	87%	92%	KZM21155.1
	Pyrenophora tritici-repentis Pt-1C-BFP	zinc-binding dehydrogenase family oxidoreductase	605	0.00E+00	81%	87%	XP_001934282.1

Table S2H (continued)

***Phoma betae* Betaenone biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
bet3 Enoyl reductase	*Setosphaeria turcica* Et28A	hypothetical protein SETTUDRAFT 100917	598	0.00E+00	84%	90%	XP_008031661.1
	Bipolaris maydis ATCC 48331	hypothetical protein COCC4DRAFT 150992	597	0.00E+00	81%	89%	XP_014074143.1
	Parastagonospora nodorum SN15	hypothetical protein SNOG 07864	588	0.00E+00	78%	86%	XP_001798191.1
	Colletotrichum gloeosporioides Nara gc5	zinc-binding dehydrogenase family	475	1.00E-165	63%	76%	XP_007283077.1
	Colletotrichum gloeosporioides Cg-14	hypothetical protein CGLO 09286	471	6.00E-164	63%	76%	EQB51190.1
bet2 Cytochrome P450	*Stemphylium lycopersici*	cytochrome P450	948	0.00E+00	87%	91%	KNG46189.1
	Ascochyta rabiei	heme binding	945	0.00E+00	87%	91%	KZM21156.1
	Alternaria alternata	cytochrome P450	932	0.00E+00	90%	94%	XP_018382154.1
	Bipolaris maydis ATCC 48331	hypothetical protein COCC4DRAFT 65603	926	0.00E+00	87%	91%	XP_014074144.1
	Setosphaeria turcica Et28A	hypothetical protein SETTUDRAFT 158087	916	0.00E+00	85%	92%	XP_008031662.1
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 20033	890	0.00E+00	81%	90%	XP_003306796.1
	Parastagonospora nodorum SN15	hypothetical protein SNOG 07865	871	0.00E+00	83%	91%	XP_001798192.1
	Pyrenophora tritici-repentis Pt-1C-BFP	cytochrome P450 monooxygenase	871	0.00E+00	81%	90%	XP_001934283.1
	Colletotrichum gloeosporioides Cg-14	cytochrome P450	610	0.00E+00	55%	74%	EQB51189.1
	Colletotrichum gloeosporioides Nara gc5	cytochrome P450	608	0.00E+00	55%	74%	XP_007283078.1
bet1 PKS	*Alternaria alternata*	polyketide synthase PksF	5064	0.00E+00	84%	90%	AFN68297.1
	Stemphylium lycopersici	polyketide synthase	5038	0.00E+00	82%	88%	KNG46190.1
	Alternaria alternata	polyketide synthase PksF	5002	0.00E+00	83%	89%	XP_018382155.1
	Setosphaeria turcica Et28A	hypothetical protein SETTUDRAFT 174734	4915	0.00E+00	82%	89%	XP_008031663.1

Table S2H (continued)

Phoma betae **Betaenone biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
	Ascochyta rabiei	transferase	4880	0.00E+00	81%	88%	KZM21157.1
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 20032	4725	0.00E+00	78%	87%	XP_003306795.1
	Bipolaris maydis ATCC 48331	hypothetical protein COCC4DRAFT 65604	4690	0.00E+00	82%	89%	XP_014074145.1
	Parastagonospora nodorum SN15	hypothetical protein SNOG 07866	4396	0.00E+00	73%	83%	XP_001798193.1
	Bipolaris maydis	polyketide synthase	3998	0.00E+00	83%	90%	AAR90270.1
bet1 PKS	*Colletotrichum gloeosporioides* Nara gc5	lovastatin nonaketide synthase	3490	0.00E+00	59%	73%	XP_007283082.1
	Stemphylium lycopersici	hypothetical protein TW65 91146	72.8	5.00E-14	57%	59%	KNG51649.1
	Pyrenochaeta sp. DS3sAY3a	hypothetical protein IQ07DRAFT 591301	54.3	2.00E-07	51%	57%	OAL45490.1
	Alternaria alternata	hypothetical protein CC77DRAFT 1004892	54.3	6.00E-07	57%	59%	XP_018390679.1
	Rosellinia necatrix	predicted protein	49.3	2.00E-04	44%	52%	GAP91188.1
	Bipolaris oryzae ATCC 44560	hypothetical protein COCMIDRAFT 26394	43.5	1.00E-02	49%	52%	XP_007688029.1
	Alloactinosynnema iranicum	cell envelope-related function transcriptional attenuator common domain-containing protein	44.7	1.10E-02	44%	45%	SDC90266.1
	Emmonsia parva UAMH 139	hypothetical protein EMPG 13278	43.9	1.60E-02	46%	59%	KLJ11537.1
	Diaporthe ampelina	hypothetical protein UCDDA912 g07449	43.1	2.00E-02	50%	54%	KKY32555.1
	Grosmannia clavigera kw1407	hypothetical protein CMQ 5314	42.7	2.60E-02	43%	54%	XP_014174534.1
Hypothetical	*Bipolaris zeicola* 26-R-13	hypothetical protein COCCADRAFT 958	41.6	3.40E-02	44%	46%	XP_007707419.1

Table S2H (continued)

Phoma betae **Betaenone biosynthetic cluster**	**Blastp hits**						
Description in Figure 3	**Species**	**Description**	**Max score**	**E-value**	**% Identity**	**% Similarity**	**Accession**
	Pyrenophora teres f. *teres* 0-1	hypothetical protein PTT 16149	407	1.00E-124	39%	54%	XP_003303789.1
	Pyrenophora tritici-repentis Pt-1C-BFP	predicted protein	380	3.00E-114	37%	51%	XP_001931686.1
	Alternaria alternata	hypothetical protein CC77DRAFT 1076552	366	1.00E-109	35%	51%	XP_018390681.1
	Stemphylium lycopersici	hypothetical protein TW65 91144	281	2.00E-79	31%	47%	KNG51647.1
	Paraphaeosphaeria sporulosa	hypothetical protein CC84DRAFT 54969	226	4.00E-58	28%	45%	XP_018042173.1
	Ascochyta rabiei	pectate lyase	128	2.00E-26	30%	49%	KZM28577.1
	Leptosphaeria maculans JN3	predicted protein	94	5.00E-16	29%	51%	XP_003844856.1
	Pyrenochaeta sp. DS3sAY3a	hypothetical protein IQ07DRAFT 683913	72	4.00E-09	33%	53%	OAL45485.1
Hypothetical	*Ascochyta rabiei*	hypothetical protein ST47 g7855	49.7	2.00E-02	31%	43%	KZM20990.1

Table S3. Blastp searches for homologs of *M. fijiensis* PKS8-4 and Hybrid8-3 cluster sequences.
The table indicates the JGI protein ID and description of the *M. fijiensis* protein sequence used as a query, and blastp results from *N. crassa*, *S. macrospora*, *A. terreus*, *P. citrinum*, and *P. betae*, including the description of the best hit from that species, the bitscore (Bit), the E-value, percent sequence identity and similarity, the accession of the hit, and the location and orientation (Dir) of the corresponding gene in the genome. A) Blastp searches for homologs of proteins encoded by the *PKS8-4* gene cluster in the *N. crassa* genome; B) Blastp searches for homologs of proteins encoded by the *PKS8-4* gene cluster in the *S. macrospora* genome; C) Blastp searches for homologs of proteins encoded by the *PKS8-4* gene cluster in the *A. terreus* genome; D) Blastp searches for homologs of proteins encoded by the *PKS8-4* gene cluster in sequences from *P. citrinum*; E) Blastp searches for homologs of proteins encoded by the *PKS8-4* gene cluster in sequences from *P. betae*; F) Blastp searches for homologs of proteins encoded by the *Hybrid8-3* gene cluster in the *N. crassa* genome; G) Blastp searches for homologs of proteins encoded by the *Hybrid8-3* gene cluster in the *S. macrospora* genome; H) Blastp searches for homologs of proteins encoded by the *Hybrid8-3* gene cluster in the *A. terreus* genome; I) Blastp searches for homologs of proteins encoded by the *Hybrid8-3* gene cluster in sequences from *P. citrinum*; J) Blastp searches for homologs of proteins encoded by the *Hybrid8-3* gene cluster in sequences from *P. betae*.

Table S3A

Proteins encoded by *PKS8-4* gene cluster		**Blast of *Neurospora crassa* OR74A**								
JGI protein ID of *M. fijiensis* sequence	**Description of *M. fijiensis* sequence**	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
190533	PKS8-4	polyketide synthase 4	1155	0.00E+00	33.00%	49.00%	XP_011395279.1	NC_026507.1	1080398-1091850	+
37626	Enoyl reductase	hypothetical protein NCU06080	60.8	1.00E-10	28.00%	44.00%	XP_959542.3	NC_026507.1	1875898-1877483	-
190535	Cytochrome P450	trichothecene C-15 hydroxylase	214	1.00E-63	30.00%	50.00%	XP_964385.2	NC_026501.1	7386748-7388890	-
178317	Hypothetical	-	-	-	-	-	-	-	-	-
178318	Hypothetical	-	-	-	-	-	-	-	-	-
37548	Dehydrogenase	retinol dehydrogenase 12	70.5	5.00E-14	31.00%	48.00%	XP_959690.1	NC_026507.1	940644-942441	+

Table S3B

Proteins encoded by *PKS8-4* gene cluster		**Blast of *Sordaria macrospora***								
JGI protein ID of *M. fijiensis* sequence	**Description of *M. fijiensis* sequence**	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
190533	PKS8-4	hypothetical protein SMAC_05695	1153	0.00E+00	33.00%	49.00%	XP_003348600.1	NW_003546212.1	307277-315100	-
37626	Enoyl reductase	hypothetical protein SMAC_08552	52.8	7.00E-08	21.00%	40.00%	XP_003346050.1	NW_003546176.1	92917-94489	-
190535	Cytochrome P450	hypothetical protein SMAC_06042	236	4.00E-72	30.00%	49.00%	XP_003349347.1	NW_003546221.1	79463-81194	+
178317	Hypothetical	-	-	-	-	-	-	-	-	-
178318	Hypothetical	-	-	-	-	-	-	-	-	-
37548	Dehydrogenase	hypothetical protein SMAC_07498	73.9	4.00E-15	32.00%	51.00%	XP_003345510.1	NW_003546165.1	151953-153041	+

Table S3C

Proteins encoded by *PKS8-4* gene cluster		**Blast of *Aspergillus terreus* NIH2624**								
JGI protein ID of *M. fijiensis* sequence	**Description of *M. fijiensis* sequence**	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
190533	PKS8-4	hypothetical protein ATEG_09961	1384	0.00E+00	35.00%	52.00%	XP_001209263.1	NT_165939.1	38401-47965	+
37626	Enoyl reductase	hypothetical protein ATEG_09963	154	9.00E-44	34.00%	47.00%	XP_001209265.1	NT_165939.1	49367-50619	+
190535	Cytochrome P450	conserved hypothetical protein	246	2.00E-76	33.00%	53.00%	XP_001217418.1	NT_165936.1	934152-935674	-
178317	Hypothetical	-	-	-	-	-	-	-	-	N/A
178318	Hypothetical	-	-	-	-	-	-	-	-	N/A
37548	Dehydrogenase	conserved hypothetical protein	120	3.00E-32	31.00%	51.00%	XP_001215480.1	NT_165932.1	172723-173661	+

Table S3D

Proteins encoded by *PKS8-4* gene cluster		**Blast of *Penicillium citrinum***								
JGI protein ID of *M. fijiensis* sequence	**Description of *M. fijiensis* sequence**	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
190533	PKS8-4	Compactin nonaketide synthase mokA	1423	0.00E+00	36.00%	53.00%	Q8J0F7.1	N/A	N/A	N/A
37626	Enoyl reductase	Deshydrogenase mlcG/Compactin biosynthesis protein G	154	3.00E-46	32.00%	48.00%	Q8J0F9.1	N/A	N/A	N/A
190535	Cytochrome P450	-	-	-	-	-	-	N/A	N/A	N/A
178317	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
178318	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
37548	Dehydrogenase	-	-	-	-	-	-	N/A	N/A	N/A

Table S3E

Proteins encoded by *PKS8-4* gene cluster		**Blast of *Phoma betae***								
JGI protein ID of *M. fijiensis* sequence	**Description of *M. fijiensis* sequence**	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
190533	PKS8-4	polyketide synthase (Bet1)	2664	0.00E+00	48%	64%	BAQ25466.1	N/A	N/A	N/A
37626	Enoyl reductase	zinc-binding dehydrogenase family protein (Bet3)	391	9.00E-139	53%	69%	BAQ25464.1	N/A	N/A	N/A
190535	Cytochrome P450	cytochrome P450 (Bet2)	370	2.00E-126	38%	55%	BAQ25465.1	N/A	N/A	N/A
178317	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
178318	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
37548	Dehydrogenase	short-chain dehydrogenase (Bet4)	352	2.00E-124	54%	69%	BAQ25463.1	N/A	N/A	N/A

Table S3F

Proteins encoded by *Hybrid8-3* gene cluster		**Blast of *Neurospora crassa* OR74A**								
JGI protein ID of M. fijiensis sequence	**Description of *M. fijiensis sequence***	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
156764	Hypothetical	hypothetical protein NCU06867	86.7	2.00E-19	34.00%	50.00%	XP_011395249.1	NC_026507.1	576344-579844	+
35041	Beta lactamase like	Metallo-beta-lactamase	278	3.00E-93	46.00%	58.00%	XP_963853.1	NC_026501.1	1201437-1202856	-
123063	Hypothetical	hypothetical protein NCU02682	94.7	3.00E-25	42.00%	55.00%	XP_965455.1	NC_026501.1	3184221-3185353	-
93879	Hybrid8-3	polyketide synthase 4	1292	0.00E+00	35.00%	50.00%	XP_011395279.1	NC_026507.1	1080398-1091850	+
212229	Enoyl reductase	oxidoreductase	85.1	2.00E-18	25.00%	42.00%	XP_959878.1	NC_026507.1	4168312-4170729	-
190367	Cytochrome P450	Benzoate 4-monooxygenase	171	5.00E-50	28.00%	46.00%	XP_961698.3	NC_026505.1	2866219-2867861	+
34861	Aminotransferase	Branched-chain-amino-acid aminotransferase	313	1.00E-104	49.00%	62.00%	XP_961082.2	NC_026505.1	5522933-5525365	+
34522	ABC transporter	multidrug resistance protein MDR	1285	0.00E+00	51.00%	70.00%	XP_959059.2	NC_026503.1	4417159-4421983	+

Table S3G

Proteins encoded by *Hybrid8-3* gene cluster		**Blast of *Sordaria macrospora***								
JGI protein ID of M. fijiensis sequence	**Description of *M. fijiensis sequence***	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
156764	Hypothetical	-	-	-	-	-	-	-	-	-
35041	Beta lactamase like	hypothetical protein SMAC_05009	263	3.00E-87	44.00%	56.00%	XP_003352895.1	NW_003546239.1	359285-360426	-
123063	Hypothetical	hypothetical protein SMAC_04339	96.3	3.00E-25	42.00%	58.00%	XP_003351035.1	NW_003546230.1	503079-503969	-
93879	Hybrid8-3	hypothetical protein SMAC_05695	1283	0.00E+00	35.00%	50.00%	XP_003348600.1	NW_003546212.1	307277-315100	-
212229	Enoyl reductase	hypothetical protein SMAC_08552	79	2.00E-16	25.00%	42.00%	XP_003346050.1	NW_003546176.1	92917-94489	-
190367	Cytochrome P450	hypothetical protein SMAC_01414	262	3.00E-84	33.00%	50.00%	XP_003352580.1	NW_003546238.1	810685-812432	+
34861	Aminotransferase	hypothetical protein SMAC_05595	317	2.00E-105	48.00%	61.00%	XP_003349312.1	NW_003546220.1	403326-404967	+
34522	ABC transporter	hypothetical protein SMAC_02009	1257	0.00E+00	50.00%	68.00%	XP_003348988.1	NW_003546213.1	1440058-1444144	+

Table S3H

Proteins encoded by *Hybrid8-3* gene cluster		**Blast of *Aspergillus terreus* NIH2624**								
JGI protein ID of M. fijiensis sequence	**Description of *M. fijiensis sequence***	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
156764	Hypothetical	predicted protein	106	1.00E-26	36.00%	51.00%	XP_001213044.1	NT_165928.1	859788-860651	-
35041	Beta lactamase like	conserved hypothetical protein	186	4.00E-59	57.00%	68.00%	XP_001212392.1	NT_165927.1	1182881-1183492	-
123063	Hypothetical	Beta-xylosidase	70.5	2.00E-15	40.00%	49.00%	XP_001210179.1	NT_165972.1	279828-281795	-
93879	Hybrid8-3	hypothetical protein ATEG_00325	1837	0.00E+00	33.00%	50.00%	XP_001210411.1	NT_165972.1	919949-931962	-
212229	Enoyl reductase	hypothetical protein ATEG_09963	217	4.00E-68	35.00%	52.00%	XP_001209265.1	NT_165939.1	49367-50619	+
190367	Cytochrome P450	predicted protein	200	1.00E-59	28.00%	48.00%	XP_001217488.1	NT_165936.1	1136810-1138339	-
34861	Aminotransferase	conserved hypothetical protein	318	2.00E-106	47.00%	60.00%	XP_001213656.1	NT_165929.1	647719-649026	+
34522	ABC transporter	multidrug resistance protein 1	1316	0.00E+00	53.00%	70.00%	XP_001218046.1	NT_165937.1	1181451-1185599	-

Table S3I

Proteins encoded by *Hybrid8-3* gene cluster		**Blast of *Penicillium citrinum***								
JGI protein ID of M. fijiensis sequence	**Description of *M. fijiensis sequence***	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
156764	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
35041	Beta lactamase like	-	-	-	-	-	-	N/A	N/A	N/A
123063	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
93879	Hybrid8-3	Compactin nonaketide synthase mokA	1493	0.00E+00	33.00%	50.00%	Q8J0F7.1	N/A	N/A	N/A
212229	Enoyl reductase	Deshydrogenase mlcG/Compactin biosynthesis protein G	211	3.00E-68	35.00%	53.00%	Q8J0F9.1	N/A	N/A	N/A
190367	Cytochrome P450	DihydroML-236B monooxygenase mlcC/Compactin biosynthesis proein C	42	2.00E-06	22.00%	41.00%	Q8J0F6.1	N/A	N/A	N/A
34861	Aminotransferase	-	-	-	-	-	-	N/A	N/A	N/A
34522	ABC transporter	-	-	-	-	-	-	N/A	N/A	N/A

Table S3J

Proteins encoded by *Hybrid8-3* gene cluster		**Blast of *Phoma betae***								
JGI protein ID of M. fijiensis sequence	**Description of *M. fijiensis sequence***	**Description of best hit**	**Bit**	**E-value**	**% identity**	**% similarity**	**Accession**	**Scaffold/ chromosome**	**Position**	**Dir**
156764	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
35041	Beta lactamase like	-	-	-	-	-	-	N/A	N/A	N/A
123063	Hypothetical	-	-	-	-	-	-	N/A	N/A	N/A
93879	Hybrid8-3	Polyketide synthase bet1	1356	0.00E+00	35%	52%	BAQ25466.1	N/A	N/A	N/A
212229	Enoyl reductase	Zinc-binding dehydrogenase family protein bet3	147	5.00E-44	32%	47%	BAQ25464.1	N/A	N/A	N/A
190367	Cytochrome P450	P450 monooxygenase No. 1	126	2.00E-34	26%	43%	BAD29966.1	N/A	N/A	N/A
34861	Aminotransferase	-	-	-	-	-	-	N/A	N/A	N/A
34522	ABC transporter	-	-	-	-	-	-	N/A	N/A	N/A

Table S4. Thirty-eight non-polar metabolites annotated by GC-MS analysis. Metabolites were categorized into esters, alkane and alkene derivatives, and other metabolites.

Categories	Compound's name
Fatty acid esters	Methyl stearate (Octadecanoic acid, methyl ester; stearic acid methyl ester)
	9-Octadecenoic acid, methyl ester, (E)-
	9-Octadecenoic acid, methyl ester
	9,12-Octadecadienoic acid, methyl ester
	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
	9,12-Octadecadienoic acid(E,E)-, methyl ester,
	9,15-Octadecadienoic acid (Z,Z)-, methyl ester,
	Hexadecanoic acid, methyl ester (Palmitic acid, methyl ester)
Other esters	Sulfurous acid, 2-ethylhexyl isohexyl ester
	Sulfurous acid, 2-ethylhexyl hexyl ester
	Oxalic acid, 6-ethyloct-3-yl ethyl ester
	2,6-Difluorobenzoic acid, 4-nitrophenyl ester
	Fumaric acid, ethyl 2-methylallyl ester
	Acetic acid, trifluoro-, 3,7-dimethyloctyl ester
	Butanedioic acid, 2-hydroxy-2-methyl-, dimethyl ester, (2R)-
	Phthalic acid, 6-methylhept-2-yl octyl ester
	Acetic acid, trifluoro-, 2,2-dimethylpropyl ester
	Dichloroacetic acid, 6-ethyl-3-octyl ester
	Pentanoic acid, 5-hydroxy-, 2,4-di-t-butylphenyl ester
Alkane and alkene derivatives	Butane, 2,2-dimethyl-
	Pentane, 3,3-dimethyl-
	Undecane, 3,8-dimethyl-
	1-Hexene, 3,5,5-trimethyl-
	3-Hexanone, 2,5-dimethyl-
	Octane, 2,7-dimethyl-
	4-Octene, 2,3,6-trimethyl-
	Nonane, 1-iodo-
	4-Nonene, 3-methyl-, (Z)-
	1-Nonene, 4,6,8-trimethyl-
	Decane, 2,3,8-trimethyl-
	2-Decene, 7-methyl-, (Z)-
	1-Iodoundecane
	Dodecane, 1-iodo-
	2-Pentadecanone, 6,10,14-trimethyl-
Others	1-Tetradecyne
	3-Ethyl-6-trifluoroacetoxyoctane
	4-Methyl-2,4-bis(4'-trimethylsilyloxyphenyl)pentene-1
	Phenol, 2,4-bis(1,1-dimethylethyl)-

CHAPTER 5

A Novel Polyketide Synthase Gene Cluster in the Plant Pathogenic Fungus *Mycosphaerella fijiensis*

Roslyn D. Noar, Elizabeth Thomas, Deyu Xie, and Margaret E. Daub

Introduction

Black Sigatoka disease, caused by the fungus *Mycosphaerella fijiensis*, is considered the most economically damaging disease of banana and plantain. Fungicide sprays accounting for 25-30% of the total banana production cost are currently the primary means of control for this disease, and without them, 50% yield loss can result[5–7,18]. The disease causes loss of photosynthetic capacity due to necrotic streaking on the banana leaves, as well as premature fruit ripening[5,6].

Both conidia and ascospores of *M. fijiensis* can germinate on the banana leaf, growing epiphytically before entering the leaf via the stomata[5]. The hyphae encircle the substomatal cavity and grow intercellularly in a biotrophic phase, before switching to a necrotrophic lifestyle and causing the characteristic necrotic leaf streaks[5]. In related fungi, secretion of toxic polyketides facilitates necrotrophic growth. For example, the closely related *Cercospora* spp. and *Ramularia collo-cygni* produce the perylenequinone cercosporin, and the anthraquinone rubellin, respectively, through polyketide biosynthetic pathways. These are both light-activated toxins that produce reactive oxygen species that kill the host tissue[56,57,152].

It has long been suspected that *M. fijiensis* and other related Sigatoka banana pathogens produce a polyketide toxin with a similar mode of action, because Sigatoka disease symptoms are less severe in partial shade[5]. For example, an early method of managing symptoms of yellow Sigatoka was to grow plants in the shade of coffee trees, which increased marketable banana yields by 50%[27]. The incubation period is the same for shaded and unshaded banana plants affected by black Sigatoka, but banana plants in shaded conditions have fewer necrotic leaves[28]. Because of these suggestions that a light-activated toxin may be important for black Sigatoka symptom development, several studies have been conducted to identify toxic metabolites from *M. fijiensis*. The first such metabolite to be identified from *M. fijiensis* was fijiensin (also called vermistatin), which is toxic to resistant and non-resistant banana varieties, but not non-host species[60,61]. Further studies identified phytotoxic metabolites: juglone, 4-hydroxyscytalone, 2,4,8-trihydroxytetralone (2,4,8-THT), 2-carboxy-3-hydroxycinnamic acid, and isoochracinic acid[61]. Juglone, 4-hydroxyscytalone, and 2,4,8-THT are shunt pathway metabolites of melanin biosynthesis[72,73]. 2,4,8-THT was believed to be important for pathogenicity because of host selectivity observed when comparing its phytotoxic effects on resistant and susceptible banana cultivars[61]. Furthermore, application of the fungicide tricyclazole blocks melanin biosynthesis, and causes an accumulation of 2,4,8-THT[74]. Application of tricyclazole to *M. fijiensis*-infected banana plants caused large necrotic leaf spots[74]. Juglone was also suspected as an important pathogenicity factor, because it has light-dependent toxicity and inhibits electron transport in banana chloroplasts[76]. It was thought that this could explain the light-dependence observed for black Sigatoka symptoms. Melanin itself plays assorted roles in pathogenicity in other

fungi, including penetration of cells by fungi that produce appressoria and protection against reactive oxygen species that may be produced by the host plant's oxidative burst defense[78,79,82]. Despite the early indication that melanin and its shunt pathway metabolites could play an important role in black Sigatoka disease symptoms, more recently it has been reported that disruption of the melanin-producing polyketide synthase has no effect on disease symptoms in *M. fijiensis*[5].

Genome sequencing of *M. fijiensis* and its close relatives *M. musicola* and *M. eumusae* has enabled bioinformatics analyses of polyketide biosynthetic genes in these species[45,54,55,301]. Polyketides are synthesized by large, multi-domain polyketide synthase (PKS) enzymes[161]. The polyketide is then modified by other enzymes that are typically encoded together in gene clusters with the polyketide synthase gene in fungal genomes[47]. We previously used bioinformatics to predict seven polyketide synthase genes and one hybrid PKS/NRPS in the *M. fijiensis* genome[45]. Based on an analysis of the domains in each PKS enzyme, PKS7-1, PKS8-1, and PKS10-1 (the melanin-producing PKS enzyme) were predicted to be non-reducing enzymes, and PKS2-1, PKS8-2, Hybrid8-3, PKS8-4, and PKS10-2 were predicted to be reducing enzymes. Since perylenequinones and anthraquinones such as cercosporin, elsinochrome and endocrocin consist largely of aromatic rings, they are produced by non-reducing PKS enzymes[57,347,348]. Therefore, PKS7-1 and PKS8-1 are enzymes of particular interest for the hypothesis that *M. fijiensis* may produce a compound similar to cercosporin and rubellin.

Cercosporin and rubellin are both red-fluorescent compounds that are soluble in dimethyl sulfoxide (DMSO) and acetone[58,100,101,349–352]. It has also been shown that

cercosporin is kept in a reduced, less active, green-fluorescent state inside the hyphae, and then it is oxidized to its active, red-fluorescent form once it is transported out of the hyphae[101,352]. In this study, we demonstrate the production of a red fluorescent compound by *M. fijiensis.* Our goals were to determine if *M. fijiensis* has similar fluorescence properties as its close relative *C. nicotianae*, to characterize the solubility of the fluorescent compound in acetone and DMSO, and to use high performance liquid chromatograph (HPLC) and tandem mass spectrometry to identify molecules extracted in these solvents. We were interested to determine whether fluorescence of *M. fijiensis* is correlated to gene expression of polyketide biosynthetic genes, especially from the gene clusters for *PKS7-1* and *PKS8-1*. If so, then we wanted to determine if over-expression of a transcription factor from the gene cluster would result in higher expression of other genes in the cluster, and if such over-expression would result in phenotypes such as increased production of the fluorescent compound or different disease symptoms when banana plants are inoculated with over-expressers compared to wild-type or vector-only controls. Finally, we wanted to characterize the expression pattern of polyketide biosynthetic genes of interest throughout the course of symptom development and compare expression at those time points to expression in germinating conidia.

Results

Red Fluorescence Observed from *M. fijiensis* Cultures and Extraction of Fluorescing Compound

Many polyketides are fluorescent molecules, including aflatoxin, the perylenequinones cercosporin, elsinochrome, and altertoxins, and the anthraquinones rubellin

and emodin[100,101,353–356]. To investigate the possibility that *M. fijiensis* produces fluorescent polyketides, we screened colonies of *M. fijiensis* isolates under UV fluorescence microscopy. On modified Potato Dextrose Agar (PDA) medium, *M. fijiensis* colonies were typically observed to have yellow autofluorescence, but isolates 10CR1-24 and 10CR1-36 were occasionally observed to fluoresce red (Figure 1). Red fluorescence was typically observed in approximately 2 week old colonies, and daily monitoring of these colonies revealed that red fluorescence was transient, typically lasting for about two days (data not shown).

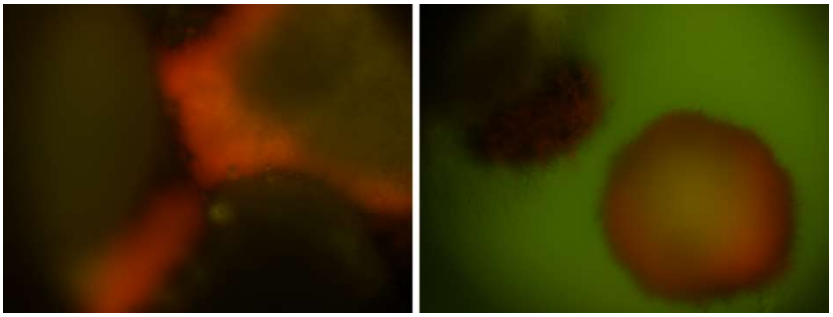


Figure 1. Red fluorescence in *M. fijiensis* colonies.
M. fijiensis isolate 10CR1-24 was grown on modified PDA medium for 2 weeks. Colonies are shown with red-fluorescing edges under UV with a dissecting microscope.

In the closely related fungus *Cercospora nicotianae*, the polyketide perylenequinone cercosporin is produced and kept in a reduced, less active state while inside the hyphae[101]. This reduced form of cercosporin fluoresces green, whereas it fluoresces red once it is transported from the cell in its active state[101]. In another closely related fungus *Pseudocercosporella capsellae*, the hyphae of cercosporin-producing strains fluoresce green from reduced cercosporin, whereas strains that do not produce cercosporin do not have

fluorescing hyphae[357]. To characterize the red fluorescent compound from *M. fijiensis*, confocal microscopy was used to characterize the fluorescence emission spectrum of red-fluorescing hyphae. Hyphae fluorescing yellow under the dissecting microscope were also observed at higher magnification using the confocal microscope, which revealed that some regions of the hyphae were bright green, and others were faint but slightly reddish (Figure 2A). Fluorescence emission spectra were determined from the areas with green fluorescence (Figure 2B). We compared these fluorescence emission spectra from *M. fijiensis* to *C. nicotianae* producing cercosporin, both in green fluorescent hyphae with reduced cercosporin, and in areas with red, active cercosporin (Figure 2C-D, Table 1). This analysis revealed that areas of red fluorescence from both fungi had a very similar fluorescence emission peak of about 630 nm (Table 1). Green fluorescence from both fungi was similar, but not identical, with a fluorescence emission peak from *M. fijiensis* at about 530 nm, whereas the fluorescence emission peak from *C. nicotianae* was about 545 nm (Table 1).

Figure 2. Fluorescence emission spectra from *M. fijiensis* compared to *C. nicotianae* hyphae.
The figure shows confocal microscope images from *M. fijiensis* and *C. nicotianae* hyphae, and the fluorescence emission spectra from several green or red-fluorescing areas of hyphae. A) Confocal microscope image from *M. fijiensis* hyphae from a colony that appeared to fluoresce yellow under a dissecting microscope; B) Fluorescence emission spectra from small green areas of *M. fijiensis* hyphae as seen in Figure 2A; C) Confocal image of red-fluorescing hyphae from *M. fijiensis*; D) Fluorescence emission spectra from small red areas of *M. fijiensis* hyphae as seen in Figure 2C; E) Confocal image of *C. nicotianae* hyphae, showing areas of green fluorescence from reduced cercosporin, and other areas with red fluorescence from oxidized cercosporin; F) Fluorescence emission spectra from small green areas of *C. nicotianae* hyphae as seen in Figure 2E; G) Fluorescence emission spectra from small red areas of *C. nicotianae* hyphae as seen in Figure 2E.

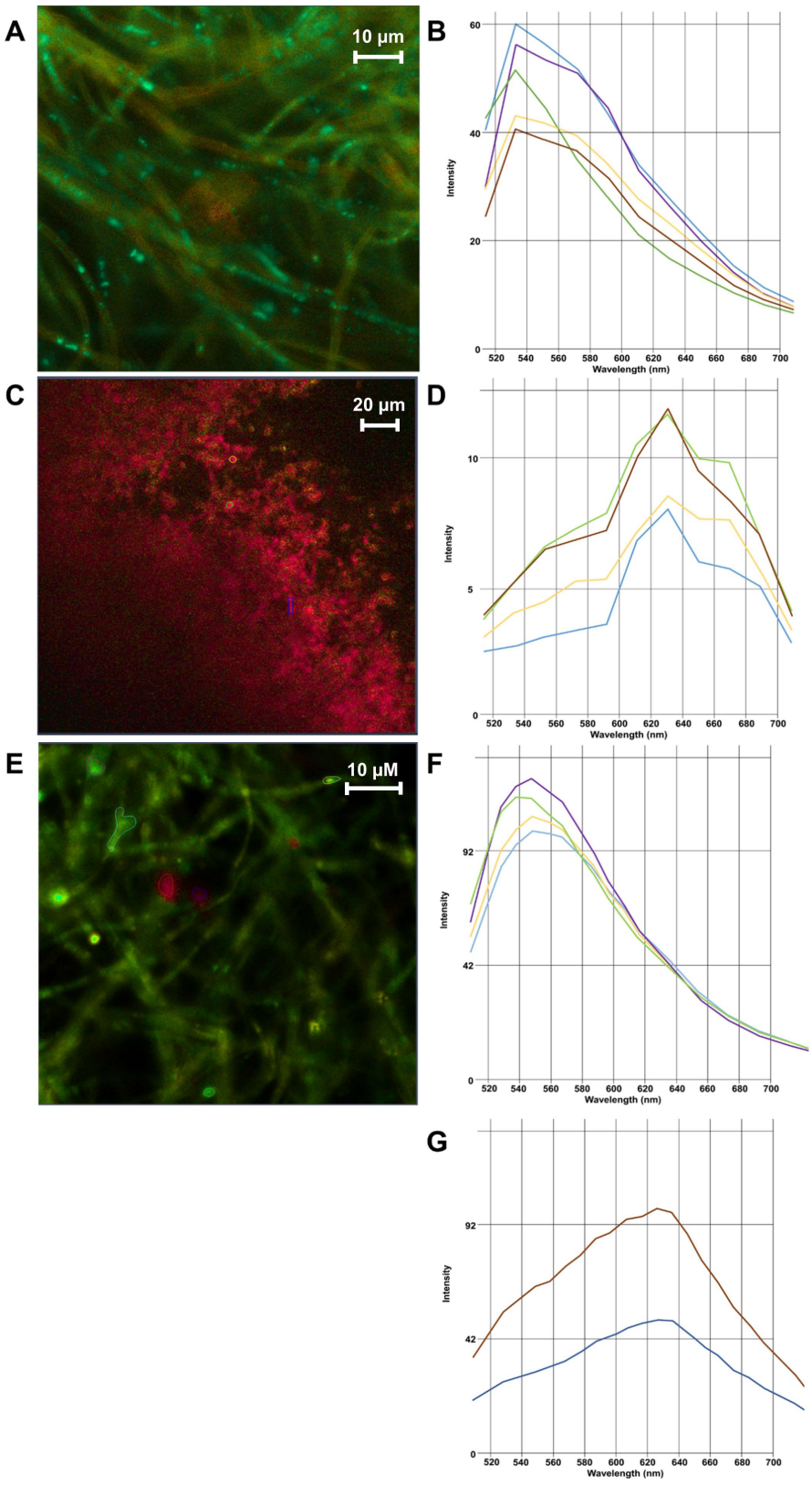

A
10 µm
B
60
40
20
0
Intensity
520 540 560 580 600 620 640 660 680 700
Wavelength (nm)
C
20 µm
D
10
5
0
Intensity
520 540 560 580 600 620 640 660 680 700
Wavelength (nm)
E
10 µM
F
92
42
0
Intensity
520 540 560 580 600 620 640 660 680 700
Wavelength (nm)
G
92
42
0
Intensity
520 540 560 580 600 620 640 660 680 700
Wavelength (nm)

Table 1. Approximate fluorescence emission peaks from *M. fijiensis* compared to *C. nicotianae* hyphae.
Table shows a summary of the approximate fluorescence emission peaks from *M. fijiensis* and *C. nicotianae*, from areas with green or red hyphae (Figure 2).

	Green fluorescence	Red fluorescence
C. nicotianae hyphae producing cercosporin	550 nm	630 nm
M. fijiensis hyphae	530 nm	630 nm

Since the polyketide perylenequinone cercosporin is soluble in acetone and DMSO [350,351], red-fluorescing colony fragments from *M. fijiensis* isolate 10CR1-36 were incubated in these solvents to determine if the red fluorescent compound from *M. fijiensis* is also soluble in these solvents. This resulted in diffusion of the red fluorescence into the solvents, demonstrating that the *M. fijiensis* compound is also soluble in acetone and DMSO. Although the resulting solutions had little visible color, they emitted strong red fluorescence when viewed under UV light. The resulting solutions were analyzed under spectrophotometry, which showed an absorption peak at about 400 nm in both solvents (Figure 3). This absorption peak was in the range of those observed for other anthraquinones and perylenequinones[358–361].

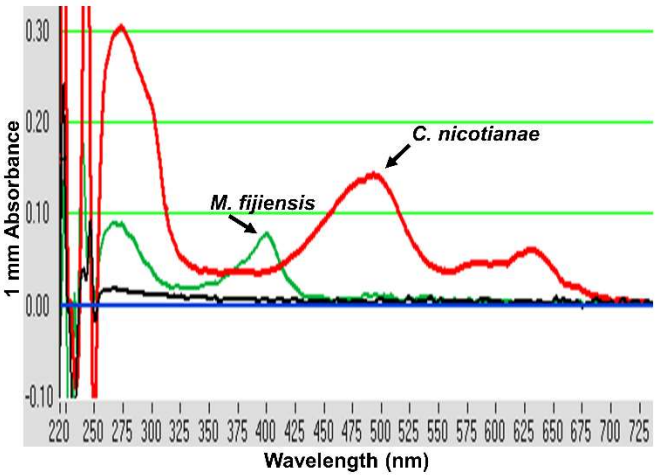


Figure 3. Absorption spectrum of red fluorescent compound.
DMSO was used to extract the red fluorescent compound from *M. fijiensis* isolate 10CR1-36, and cercosporin from *C. nicotianae*, and absorption spectra of both compounds were generated. Arrows indicating absorption spectra of compounds from each species are shown.

Chemical Analysis of Extracts

Liquid chromatography and tandem mass spectrometry were used for further analysis of the extracts with red fluorescence, and the data were searched with an in-house library to obtain potential chemical formulas (Table S1). Since the red fluorescent compound was soluble in both acetone and DMSO, extracts with both solvents were tested. As controls, cercosporin was extracted from *C. nicotianae* mycelium using the same solvents. Also, cercosporin crystals were obtained as described by Daub[362], and these crystals were dissolved in acetone or DMSO as further controls. Compounds identified from each solution are listed in Table S1. From these data, a list of compounds was generated that were extracted from red-fluorescing *M. fijiensis* colonies in both solvents, but were absent from the *C. nicotianae* extractions or from the cercosporin crystals dissolved in either solvent (Table 2). In total, 15

compounds met these criteria. Among these compounds were pulverochromenol, a chromone with similarity to monodictyphenone and related compounds, an idebenone-like metabolite which is a quinone, and 1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene (Table 2). Although this analysis is unable to determine with certainty which compound is the product of the PKS8-1 biosynthetic pathway, we were able to identify several compounds that have not previously been reported from *M. fijiensis* that may be derived from a polyketide biosynthetic pathway.

Table 2. Compounds extracted from *M. fijiensis* but not *C. nicotianae*.
Table provides a list of compounds with their chemical formulas that were identified from both the acetone and DMSO extracts from the red *M. fijiensis* mycelium, but were not identified from *C. nicotianae* extracts.

Compound	**Chemical formula**
1,2-Dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene	$C_{10}H_{10}O_3$
Idebenone Metabolite (QS-4)	$C_{13}H_{16}O_6$
Pulverochromenol	$C_{20}H_{22}O_4$
Rufloxacin	$C_{17}H_{18}FN_3O_3S$
Lamprolobine	$C_{15}H_{24}N_2O_2$
Zidovudine	$C_{10}H_{13}N_5O_4$
4(15)-Hirsutene	$C_{15}H_{24}$
10,13-Octadecadiynoic acid	$C_{18}H_{28}O_2$
9-HOTE (9-hydroxy-10,12,15(E,Z,Z) octadecatrienoate)	$C_{18}H_{30}O_3$
Patchoula-2,4-diene	$C_{15}H_{22}$
9,10-epoxy-12-octadecenoic acid	$C_{18}H_{32}O_3$
(S)-lamenallenic acid	$C_{18}H_{30}O_2$
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$
Deoxynupharidine	$C_{15}H_{23}NO$

Expression Analysis of Colonies with Red Fluorescence

RT-PCR assays were used to determine if the colonies with and without red fluorescence have differences in gene expression of PKS genes (Figure 4). This analysis did not identify differences in PKS gene expression between colonies with and without red fluorescence, but *PKS2-1*, *PKS8-1*, *PKS8-2*, and *Hybrid8-3* were all expressed in the samples tested (Figure 4). RT-PCR assays were also used to test expression of other genes in these putative biosynthetic clusters (Figure 4). This analysis showed higher expression in the red-fluorescing colonies of the MFS transporter gene in the *PKS8-1* cluster (Figure 4). These results led us to further investigate the *PKS8-1* cluster.

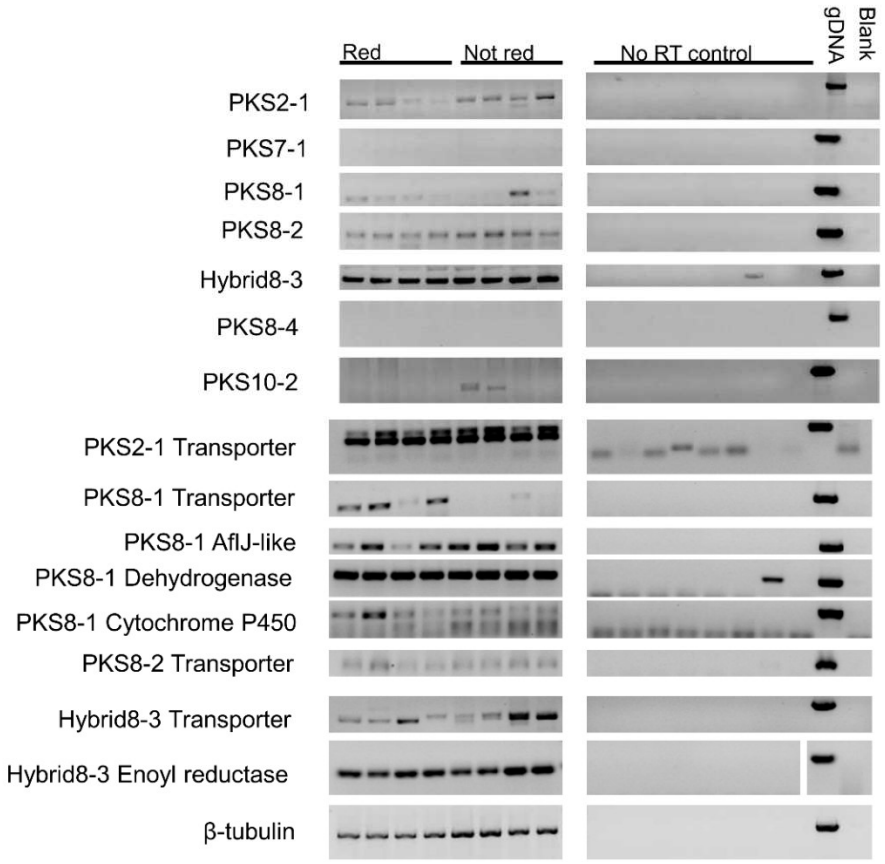


Figure 4. Expression of polyketide biosynthetic genes in red-fluorescing colonies. For genes encoding the PKS or hybrid enzymes as well as other genes in the putative biosynthetic clusters, RT-PCR assays were conducted from *M. fijiensis* colonies grown in modified PDA medium. Four colonies were tested that did or did not show red fluorescence. As controls, the same PCR assay was performed on genomic DNA, no template, or on RNA from each colony that had not been reverse transcribed.

RNA-Seq Analysis of *PKS8-1* Gene Cluster

Previously, we analyzed the expression pattern of *PKS8-1* in isolate 10CR1-24, in infected leaf tissue and in culture medium, using RT-PCR assays, and showed that *PKS8-1* is expressed under both conditions[45]. To learn more about the expression pattern of *PKS8-1* and the adjacent genes in its putative biosynthetic cluster, we used our RNA-Seq dataset[45,46] to

analyze the expression pattern of these genes in the infected leaf tissue versus in Potato Dextrose Broth (PDB) culture medium in isolate 14H1-11A (Figure 5). Tissue samples were harvested from symptomatic leaves at 6 weeks post-inoculation at the initiation of the necrotrophic state. This analysis agreed with previous results in isolate 10CR1-24 using RT-PCR assays[45]. *PKS8-1* was expressed under both conditions, though it had slightly lower expression (log2FC=-1.4) in the infected leaf tissue compared to in PDB (Figure 5). Nearby genes encoding a cytochrome P450, two aldo/keto reductases, a dehydrogenase, and a hypothetical protein also had slightly lower expression in infected leaf tissue (Figure 5). Other nearby genes had negative log2FC values, but these values were not significantly different (genes encoding AflJ-like and AflR-like proteins, beta-lactamase-like proteins, an MFS transporter, another transcription factor, and two hypothetical genes) (Figure 5). Because of the small expression changes, it was not possible to clarify the boundaries of the *PKS8-1* biosynthetic cluster using these data.

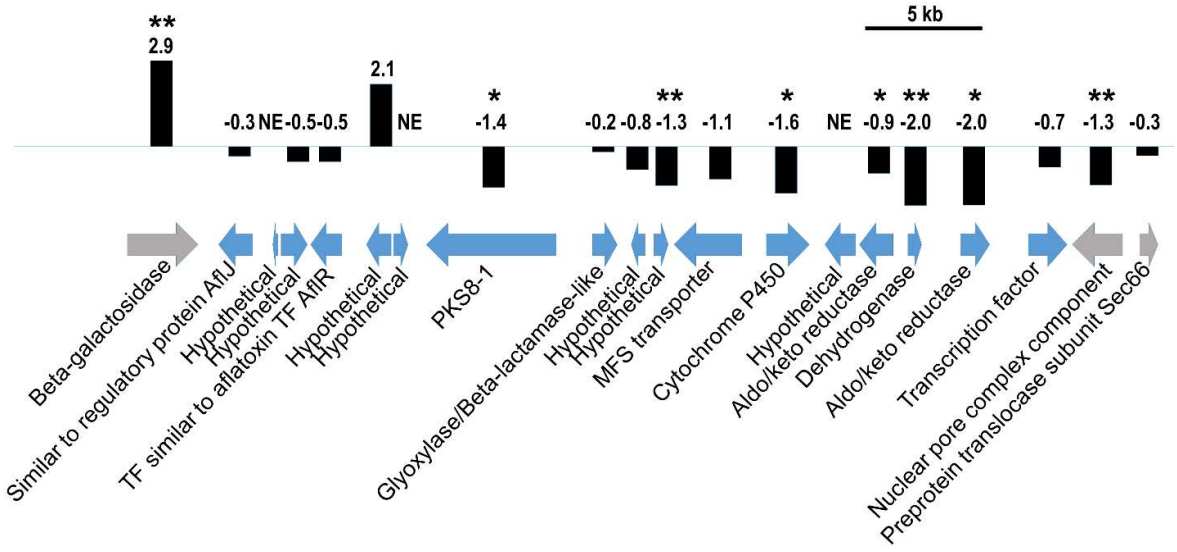


Figure 5. RNA-Seq analysis of *PKS8-1* cluster genes.
PKS8-1 and adjacent genes in the *M. fijiensis* genome are represented by arrows indicating their direction, with descriptions of putative protein functions indicated below the arrows. Blue arrows = genes predicted by Noar and Daub[45] to be part of the *PKS8-1* gene cluster; Gray arrows = genes predicted by Noar and Daub[45] to flank the *PKS8-1* gene cluster. Scale bar indicates 5 kb. Black bars are proportional to log2FC values of expression in leaf tissue vs. culture medium. Single asterisks represent significant expression differences at $p<0.05$, and double asterisks represent significant expression differences at $p<0.01$.

Bioinformatics Analysis of *PKS8-1* and Gene Clusters in Different Species

The genomes of many Dothideomycete species have been sequenced in recent years, which has created opportunities for comparing polyketide biosynthetic gene clusters across different genomes. Since many of these genomes have been sequenced but not yet annotated, we previously conducted tblastn searches of *M. fijiensis* polyketide synthase protein sequences from 103 Dothideomycete genome sequences from NCBI and JGI to identify their ten closest homologs by bitscore[45]. Using these search parameters, the best homologs

included hits from other species within the *Mycosphaerellaceae*, such as *Pseudocercospora pini-densiflorae*, *Septoria populicola*, *Septoria musiva*, and *Cercospora canescens*[45].

To determine if the other species' *PKS8-1* homologs are flanked by similar genes, tblastn searches were done for each predicted protein within the *M. fijiensis* gene cluster, using the same set of Dothideomycete genomes as previously[45], but not limiting the search to the top ten hits. This analysis confirmed that the *PKS8-1* gene cluster is conserved across many of the Dothideomycete species (Table S2)[45]. Although not previously identified due to the cutoff by bitscore of the top 10 PKS8-1 homologs, the genomes of the closely related species *Mycosphaerella musicola* and *Mycosphaerella eumusae* also have orthologous gene clusters (Figure 6) containing a nearly identical combination of genes adjacent to the polyketide synthase.

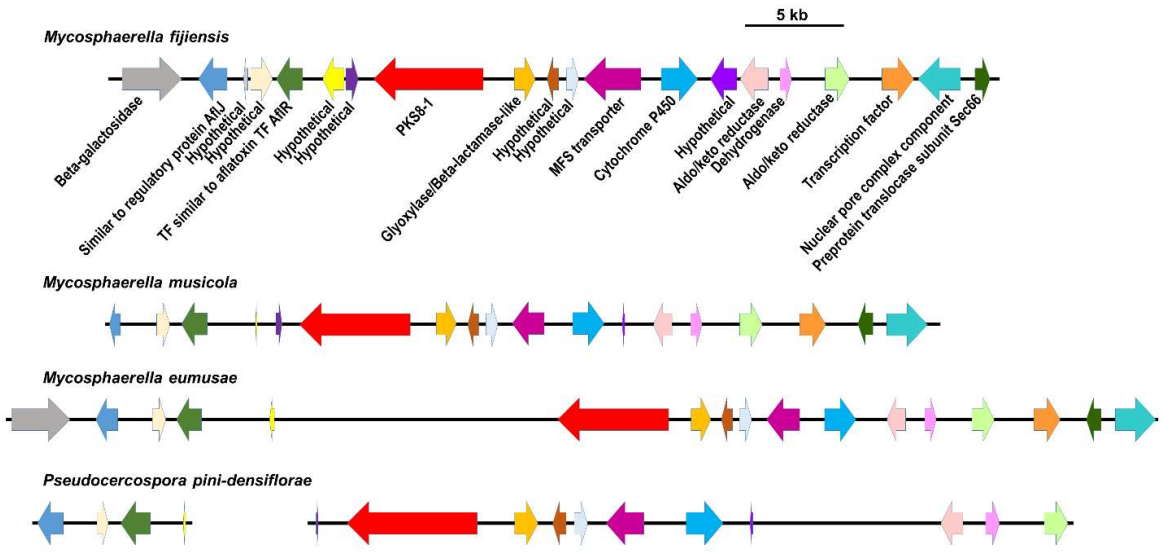


Figure 6. Comparison of *PKS8-1* gene clusters across different species.
For *M. fijiensis*, *M. musicola*, *M. eumusae*, and *P. pini-densiflorae*, *PKS8-1* and adjacent genes in the genome or their putative orthologs are shown by arrows indicating their direction. Putative orthologous genes are shown with the same color of arrow. A description of the putative function of each corresponding protein is shown under the *M. fijiensis* arrows. Scale bar indicates 5 kb.

Comparison to Monodictyphenone-Producing Genes

To predict the type of polyketide that may be produced by PKS8-1, we previously created a phylogenetic tree with PKS protein sequences from *M. fijiensis* and well-characterized PKS sequences from other species[45]. Our phylogenetics analysis showed that the PKS8-1 sequence was part of a clade of non-reducing PKS enzymes, but this analysis did not identify close, well-characterized homologs to PKS8-1[45]. A separate analysis was recently done by Chang et al[301], who reported that there are three non-reducing PKS enzymes encoded by the *M. fijiensis* genome, which was consistent with our prediction that PKS7-1, PKS8-1, and PKS10-1 are non-reducing PKS enzymes in this fungus[45,301]. Based on the

ketosynthase and acyltransferase domains of these PKS enzymes, one non-reducing PKS from *M. fijiensis* was similar to those producing 1,8-dihydroxynaphthalene (DHN)-melanin and elsinochrome, and another had similarity to PKS enzymes producing endocrocin and monodictyphenone[301]. The third non-reducing PKS enzyme did not have similarity to other PKS sequences[301]. To determine whether these predictions would be the same using the entire PKS sequence for analysis, we added sequences for the elsinochrome-producing EfPKS1 from *Elsinoë fawcettii* and the endocrocin and monodictyphenone-producing MdpG from *Aspergillus nidulans* to PKS sequences previously analyzed from *M. fijiensis* and other fungi[45] and created a new maximum likelihood phylogenetic tree (Figure 7). This analysis confirmed that whether the entire PKS sequence or only the KS and AT domains are used to create a phylogenetic tree, PKS8-1 and its homologs in other *Mycosphaerella* species form a clade with a strong bootstrap value with the monodictyphenone-producing MdpG PKS from *A. nidulans* (Figure 7).

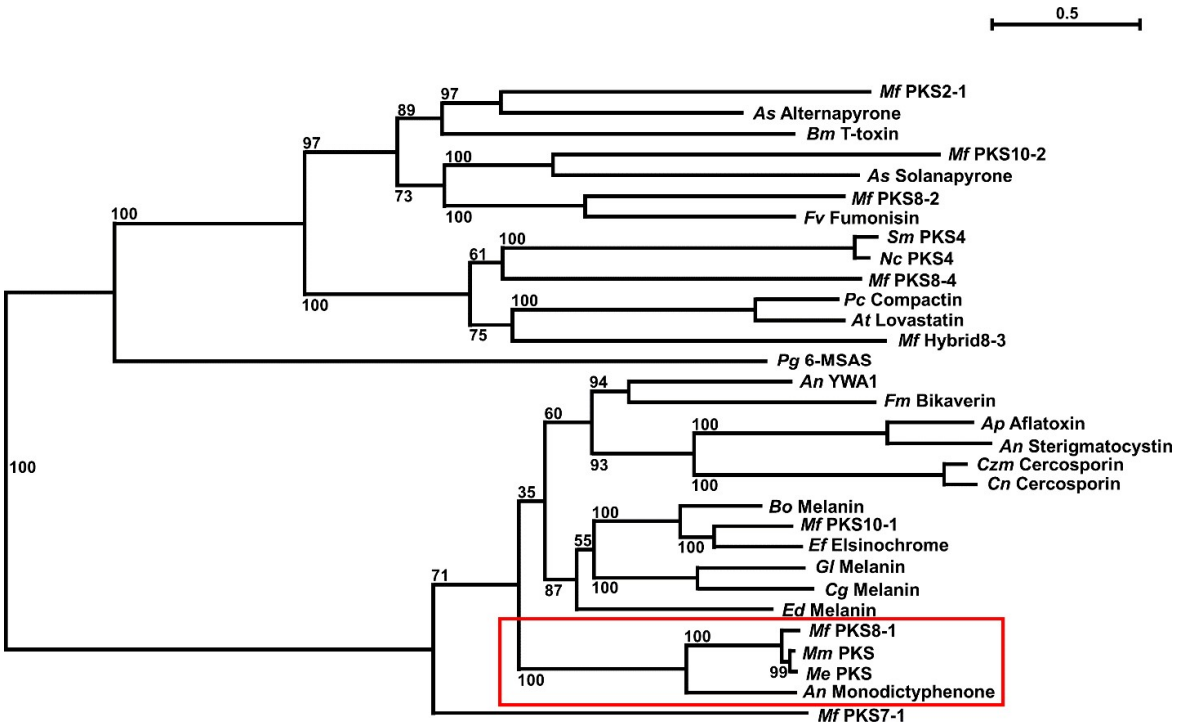


Figure 7. Phylogenetics analysis of PKS8-1.
RAxML was used to create a maximum likelihood phylogenetic tree of PKS protein sequences, using the PKS sequences described previously[45] and others including the monodictyphenone-producing MdpG from *A. nidulans* and the PKS8-1 homolog sequences identified from *M. musicola* and *M. eumusae*. The clade containing PKS8-1, its *M. musicola* and *M. eumusae* homologs, and the monodictyphenone-producing PKS from *A. nidulans* is shown with a red box. Bootstrap values are indicated on the tree, and the scale bar indicates substitutions per site. *As = Alternaria solani*; *Bm = Bipolaris maydis*; *Fv = Fusarium verticillioides*; *Sm = Sordaria macrospora*; *Nc = Neurospora crassa*; *Pc = Penicillium citrinum*; *At = Aspergillus terreus*; *Pg = Penicillium griseofulvum*; *An = Aspergillus nidulans*; *Fm = Fusarium moniliforme*; *Ap = Aspergillus parasiticus*; *Czm = Cercospora zeae-maydis*; *Cn = Cercospora nicotianae*; *Bo = Bipolaris oryzae*; *Ef = Elsinoë fawcettii*; *Gl = Glarea lozoyensis*; *Cg = Colletotrichum graminicola*; *Ed = Exophiala dermatitidis*; *Mm = Mycosphaerella musicola*; *Me = Mycosphaerella eumusae*.

Since the phylogenetic analysis in Figure 7 showed that PKS8-1 and the monodictyphenone-producing MdpG form a clade, the gene clusters for each PKS were compared. Blast searches were used to identify putative orthologs in the *M. fijiensis* genome

for each gene from the monodictyphenone-producing gene cluster from *A. nidulans*[363], and to determine whether these putative orthologs flank the *PKS8-1* gene in the *M. fijiensis* genome (Figure 8, Table S3). This analysis revealed that the two clusters share genes encoding a PKS, a beta-lactamase-like enzyme, a transcription factor with similarity to AflR, a regulatory protein similar to AflJ, and a hypothetical protein (Figure 8). However, they differ in several other enzymes encoded by the respective clusters. The putative *PKS8-1* cluster from *M. fijiensis* contains genes encoding an MFS transporter, a cytochrome P450, two aldo/keto reductases, a dehydrogenase, a transcription factor, and six hypothetical genes that the *A. nidulans* cluster lacks (Figure 8). Conversely, the *A. nidulans* cluster encodes a short-chain dehydrogenase, a glutathione S-transferase, an acyl-CoA synthase, a monooxygenase, a reductase, a dehydrogenase, and an AflY-like hypothetical protein that the *M. fijiensis* cluster lacks (Figure 8)[363]. These results suggest that the two clusters encode similar, but not identical polyketide products.

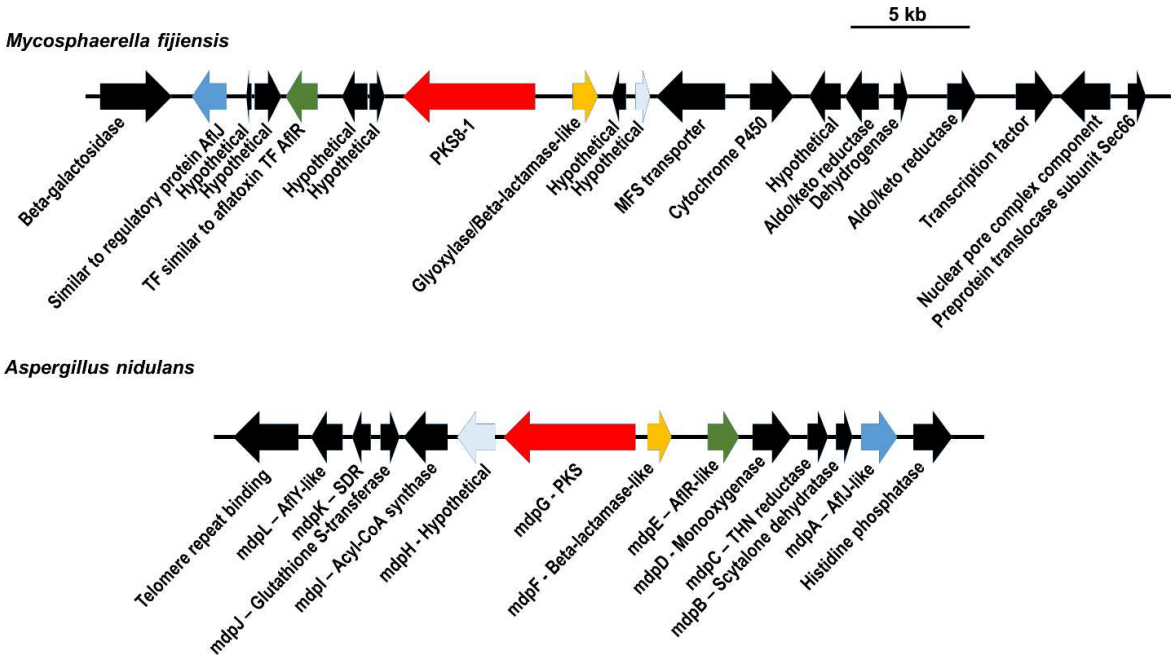


Figure 8. Comparison of *M. fijiensis PKS8-1* cluster and monodictyphenone-producing cluster from *A. nidulans*.
For the *PKS8-1* cluster and the monodictyphenone-producing cluster from *A. nidulans*, each gene is shown in the genome with an arrow indicating its direction. A description of the putative function of each gene is shown below the arrow. Putative orthologs are shown with the same color arrow, and other genes are shown with black arrows.

Defining the *PKS8-1* Gene Cluster Using the *aflR*-Like Transcription Factor Gene Over-expresser

To further characterize the *PKS8-1* cluster, we conducted experiments to over-express the cluster by constitutive expression of a transcription factor gene. There are two genes encoding transcription factors that are located near the *PKS8-1* gene in the *M. fijiensis* genome (Figure 5). One transcription factor is predicted to be AflR-like and it contains a GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain, and the other transcription factor contains a fungal-specific transcription factor domain and a fungal transcription factor

regulatory middle homology region. In the aflatoxin biosynthetic cluster in *Aspergillus flavus*, over-expression of *aflR* results in increased transcription of the biosynthetic genes and leads to increased aflatoxin production[364]. We therefore chose the *aflR*-like transcription factor gene for further analysis. We transformed *M. fijiensis* isolate 10CR1-24 with over-expression construct for the *aflR*-like transcription factor gene. Transformants of the transcription factor over-expression construct were grown in Potato Dextrose Broth (PDB) flasks and tested by RT-PCR assays for expression of *PKS8-1* compared to wild-type and to transformants with the hygromycin resistance cassette only (hyg-only) (Figure 9). This analysis revealed that while the wild-type and the hyg-only controls did not express *PKS8-1*, several of the transcription factor over-expressers do express *PKS8-1* in these conditions (Figure 9). Therefore, a similar experiment was set up to use qPCR assays to test which genes are under the control of the *aflR*-like transcription factor. Based on the expression of *PKS8-1* in Figure 9, we chose the transcription factor over-expressers #4 and 8 for expression analysis, along with the hyg-only controls #1 and 2, and the wild-type control. Our analysis revealed that in culture conditions, these transformants with the *aflR*-like over-expression construct produced about 10-fold more transcript of the *aflR*-like transcription factor than the wild-type (Figure 10). These over-expression mutants also had increased expression of genes encoding PKS8-1 (about 4.3-fold), the beta-lactamase-like protein (about 3.5-fold), the MFS transporter (about 4.9-fold), the cytochrome P450 (about 5-fold), and one of the aldo/keto reductases (about 3.8-fold) (Figure 10). In contrast, the AflJ-like, dehydrogenase, and another aldo/keto reductase were not differentially expressed (Figure 10).

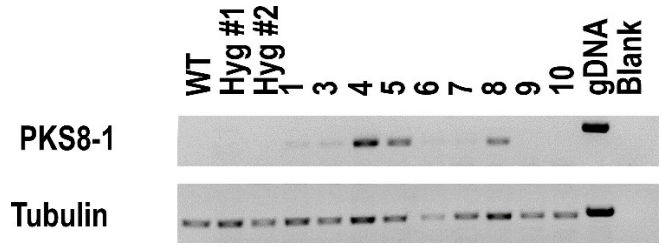


Figure 9. RT-PCR assays of *PKS8-1* expression in transcription factor over-expressers. For the wild-type, two transformants of hyg-only control (Hyg #1 and 2), and nine over-expressers of the *aflR*-like transcription factor gene in the *PKS8-1* gene cluster, RT-PCR assays were used to test expression of the β-tubulin and *PKS8-1* genes. *M. fijiensis* isolate 10CR1-24 genomic DNA and no added template were used as positive and negative controls, respectively. *PKS8-1* primers were designed to span an intron such that products from genomic DNA would be slightly larger than products from cDNA.

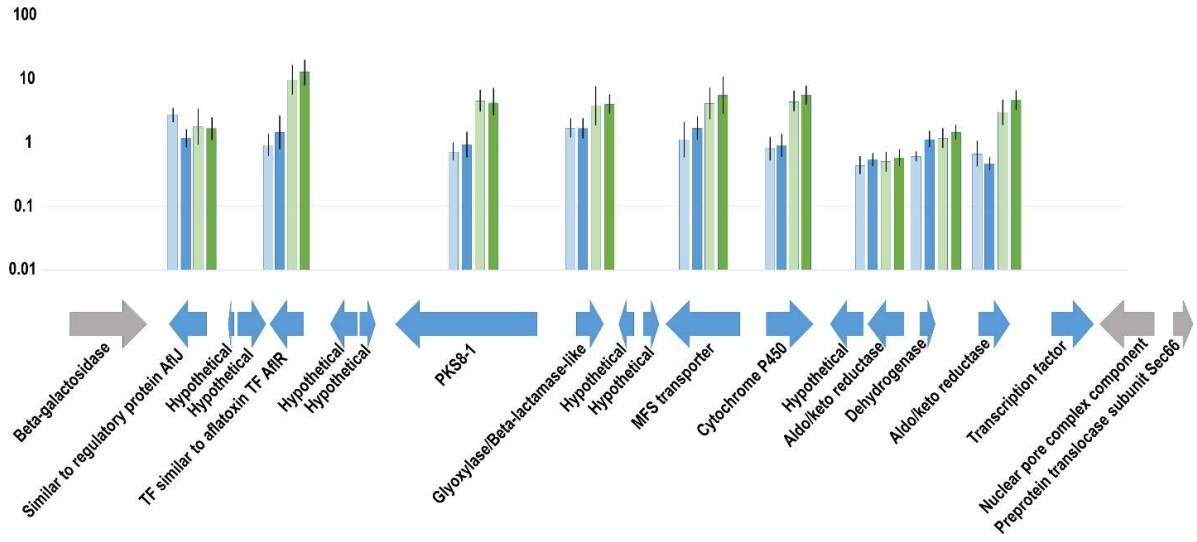


Figure 10. RT-qPCR assays of *PKS8-1* cluster genes in over-expressers compared to wild-type.
The wild-type, two transformants of hyg-only control (Hyg #1 and 2), and the transcription factor over-expresser transformants #4 and 8 were grown in PDB flasks and RT-qPCR assays were performed on these samples for genes predicted previously[45] to be part of the *PKS8-1* biosynthetic cluster. Arrows represent directionality of each gene in the genome, and a description of the putative function of each gene as determined previously[45] is shown below the arrow. Above each gene is a log fold change graph of the expression of each transformant relative to wild-type. Light blue=Hyg #1 transformant; Dark blue=Hyg#2 transformant; Light green=transcription factor over-expresser transformant #4; Dark green=transcription factor over-expresser transformant #8. Error bars indicate standard error from six biological replicates.

Over-expression of the *aflR*-Like Transcription Factor and MFS Transporter Genes Does Not Result in More Red Fluorescence

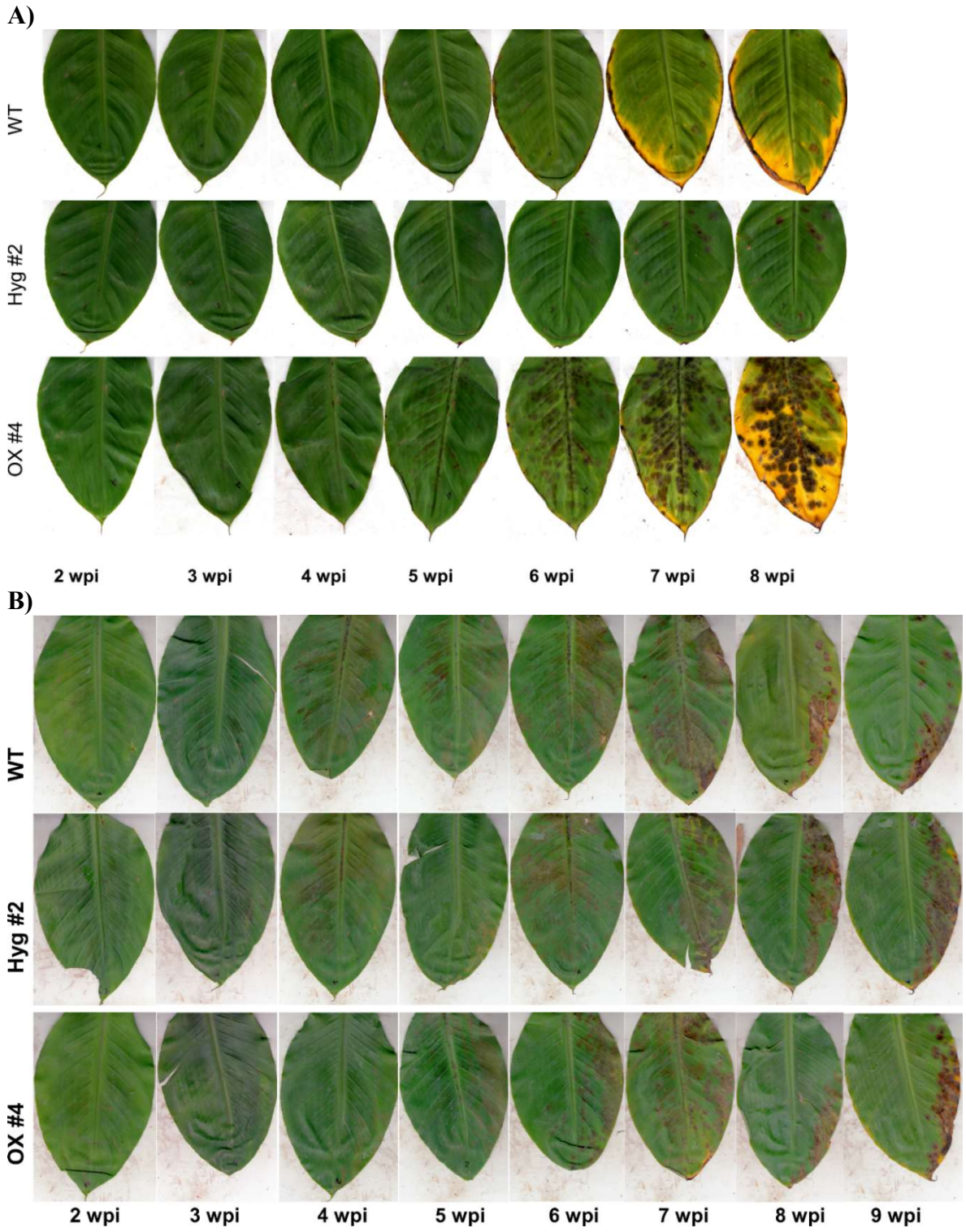
Our qPCR results (Figure 10) showed that constitutive expression of the *aflR*-like transcription factor gene increased expression of the PKS8-1 PKS and MFS transporter genes as well as other genes predicted to be in the cluster. We thus assayed transformants to determine if transformants produced red fluorescence, indicative of a connection between this cluster and the production of the red fluorescing compound. Transformants were screened under a dissecting microscope for red fluorescence, but we did not observe any increase in red fluorescence in these transformants (data not shown). Since we showed in Figure 4 that the MFS transporter in the cluster was the most differentially expressed between cultures producing or not producing red fluorescence, we also tested transformants constitutively expressing the MFS transporter gene. As with the *aflR*-like gene, there was no increase in the incidence of production of red fluorescence (data not shown).

Pathogenicity of *aflR*-Like Transcription Factor Over-expressers

Since many polyketides are phytotoxic, we hypothesized that the *aflR*-like transcription factor over-expressers may cause more severe disease symptoms when inoculated onto a host plant. Therefore, we conducted inoculation experiments to assay for any changes in pathogenicity. Young banana plants were inoculated with conidia, and inoculated leaves were scanned weekly to monitor disease progress over time. In the first two experiments, we tested wild-type, Hyg #1, Hyg #2, and transcription factor over-expressers #4 and #8; the third experiment included only wild-type, Hyg #2 and over-expresser #4.

Results were highly variable between experiments. In the first two experiments, plants inoculated with the transcription factor over-expressers had the highest amount of disease (Figure 11A). However, in experiment 3, the wild type and Hyg #2 isolates were as pathogenic as was the over-expressor #4 (Figure 11B).

Figure 11. Disease from *aflR*-like transcription factor over-expresser.
A) Images from first inoculation experiment. Figure shows representative leaves infected with wild-type, Hyg #2, and the transcription factor over-expresser #4 (OX #4), from 2 weeks post-inoculation to 8 weeks post-inoculation. Results for the second inoculation were very similar, showing greater disease severity with the over-expresser as compared to wild type and the hyg-only transformant (data not shown). B) Images from the third inoculation experiment. For the wild-type, Hyg #2, and the transcription factor over-expresser #4 (OX #4), images of a representative leaf are shown at each week, from 2 weeks post-inoculation to 9 weeks post-inoculation. Unlike experiments one and two, disease development by wild type and the hyg-only transformant was similar to that caused by the over-expresser.


A)
WT
Hyg #2
OX #4
2 wpi
3 wpi
4 wpi
5 wpi
6 wpi
7 wpi
8 wpi
B)
WT
Hyg #2
OX #4
2 wpi
3 wpi
4 wpi
5 wpi
6 wpi
7 wpi
8 wpi
9 wpi

Expression of *PKS8-1* Cluster Genes in Time Course Inoculation Experiment

If the product of the *PKS8-1* cluster plays a role in infection or progression of the disease, we hypothesized that *PKS8-1* and other genes in the biosynthetic cluster should have higher expression at time points when this polyketide plays its role in pathogenicity. Therefore, we set up inoculation experiment 3 with the purpose of assaying gene expression. Plants were inoculated with wild-type *M. fijiensis* isolate 10CR1-24, the transcription factor over-expresser #4, and the Hyg #2 control. Leaves from infected banana plants were harvested weekly for expression analysis using RT-qPCR. As shown in Figure 11, the results of this experiment differed from the previous two inoculation experiments in that the wild type and Hyg #2 control resulted in symptoms that were more severe than seen in the previous two inoculation experiments and were as severe as those caused by the over-expresser. Since disease results were inconsistent with the previous two experiments, the identities of fungal genotypes were tested using PCR assays (Figure S1), and were verified to be correct.

In spite of the unexpected disease results, we used RT-qPCR assays to define expression of the wild-type *M. fijiensis* throughout the infection period compared to expression in germinating conidia. The analysis showed that in the wild-type, *PKS8-1* and cytochrome P450 gene expression from 3-8 weeks post-inoculation was about 10-fold higher than in conidia, whereas expression was lower at 9 weeks post-inoculation (Figure 12). Expression of genes encoding the AflJ-like, beta-lactamase-like, and one of the aldo/keto reductase enzymes was also higher throughout the infection than it was in conidia (Figure 12). In contrast, no increase in expression was observed for genes encoding the AflR-like

transcription factor, the MFS transporter, one aldo/keto reductase, and one dehydrogenase throughout the infection as compared to conidia (Figure 12). These results indicate that for wild-type, many, but not all, of the genes in the putative *PKS8-1* gene cluster have higher expression throughout infection than in germinating conidia.

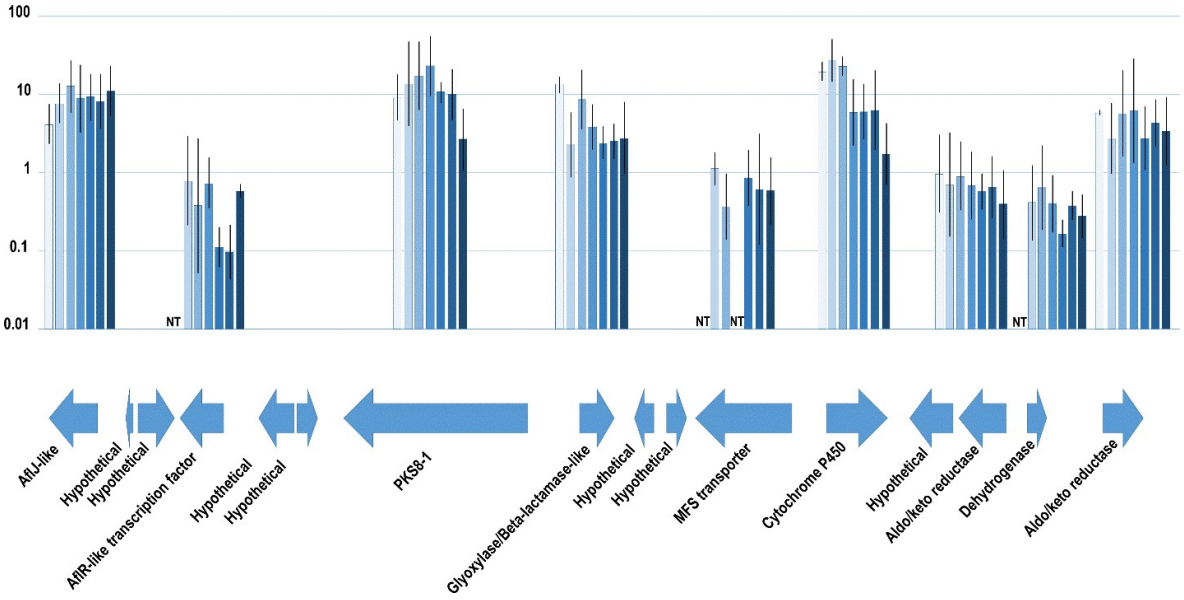


Figure 12. Expression of genes in the putative *PKS8-1* cluster in wild-type *M. fijiensis* after infection compared to germinating conidia.
PKS8-1 and its neighboring genes in the *M. fijiensis* genome are shown as arrows indicating orientation in the genome, with descriptions indicating putative function. For genes in the putative *PKS8-1* gene cluster, RT-qPCR assays were used to analyze expression of wild-type isolate 10CR1-24 from 3 to 9 weeks post-inoculation, compared to expression in germinating conidia. Results for each gene are shown as a bar graph in log scale, with error bars indicating standard error from three biological replicates. The bar in the lightest shade of blue represents results at 3 weeks post-inoculation, and bars are shown in the order of the number of weeks, up to 9 weeks post-inoculation. NT = not tested.

We then used RT-qPCR assays to compare the expression of genes encoding the AflR-like transcription factor, PKS8-1, and the cytochrome P450 in the *aflR*-like

transcription factor over-expresser and the hyg-only control, relative to expression in the wild-type. Expression was assayed in germinating conidia and in leaf tissue at one time point at 6 weeks post-inoculation (Figure 13). In germinating conidia, expression of the *aflR*-like transcription factor gene was higher in the over-expresser compared to the hyg-only control or the wild-type, but at 6 weeks post-inoculation, there was no difference in expression (Figure 13). In both germinating conidia and at 6 weeks post-inoculation, there was no difference in expression of *PKS8-1* in the over-expresser compared to the hyg-only control or the wild-type (Figure 13). Similarly, in germinating conidia, there was no difference in expression of the cytochrome P450 in the over-expresser compared to the hyg-only control or the wild-type (Figure 13). Thus the expression results are consistent with the strain differences (conidia of the *aflR* over-expresser having higher expression of *aflR*) as well as with the lack of difference in both gene expression and disease severity in this experiment. Further research is necessary to define the reasons for the inconsistent disease severity results and to determine if gene expression varies with differences in disease severity. It is possible that environmental conditions or other changes during the inoculation experiment may have induced *PKS8-1* gene expression in wild type, resulting in the lack of difference with the over-expresser. Further research is needed.

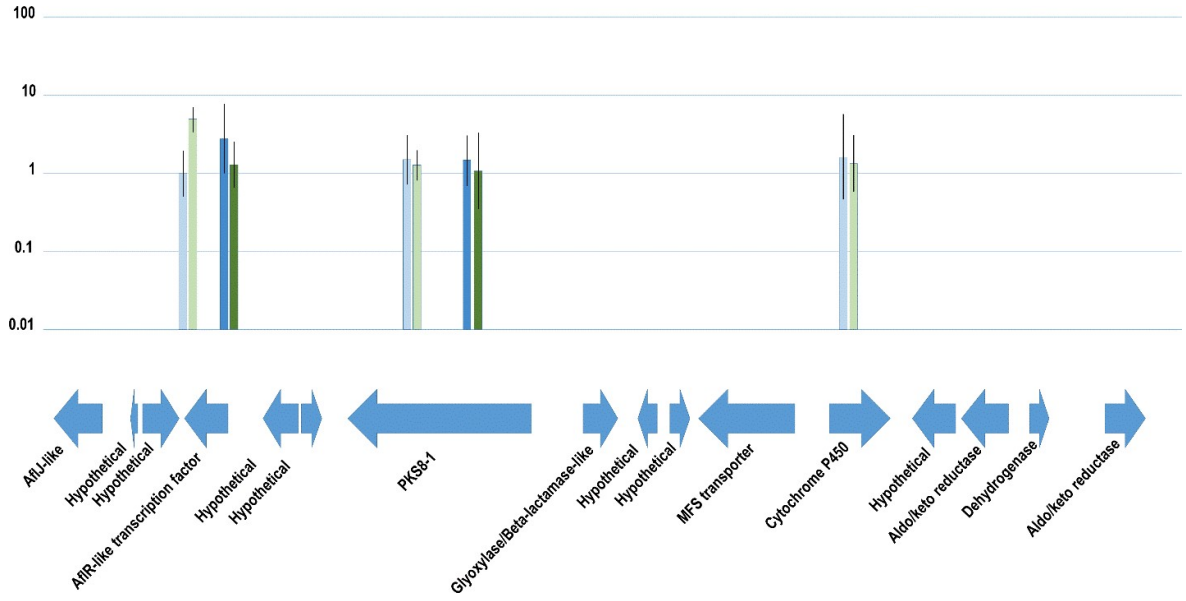


Figure 13. Expression of genes from the putative *PKS8-1* gene cluster in the *aflR*-like transcription factor over-expresser compared to wild-type.
PKS8-1 and its neighboring genes in the *M. fijiensis* genome are shown as arrows indicating orientation in the genome, with descriptions indicating putative function. For the genes encoding the AflR-like transcription factor and PKS8-1, gene expression was tested in germinating conidia and at 6 weeks post-inoculation in the transcription factor over-expresser #4 and the Hyg #2 control, compared to wild-type. For the gene encoding the cytochrome P450, gene expression was tested in germinating conidia. Results for each gene are shown as a bar graph in log scale with values relative to the wild-type, with error bars indicating standard error from three biological replicates. Light blue = Expression in the Hyg #2 control in germinating conidia; Light green = Expression in the *aflR*-like over-expresser #4 in germinating conidia; Dark blue = Expression in the Hyg #2 control at 6 weeks post-inoculation; Dark green = Expression in the *aflR*-like over-expresser #4 at 6 weeks post-inoculation.

Discussion

Overall, our study supported but could not prove the long-standing hypothesis that *M. fijiensis* may produce a similar compound to cercosporin, which is light-activated and may be reduced or oxidized, which modulates its activity and changes its fluorescence emission spectrum[101]. We showed that, like *Cercospora* spp. hyphae producing cercosporin, *M.*

fijiensis hyphae can also fluoresce red or green, and the fluorescence emission spectra are very similar to those observed for cercosporin. We identified fourteen compounds from red-fluorescing *M. fijiensis* colonies that were not extracted from *C. nicotianae* grown in the same conditions (Table 2). This analysis represents the first report of these fourteen compounds from *M. fijiensis* cultures. Among them are: the sesquiterpenoids 4(15)-hirsutene[365] and patchoula-2,4-diene[366]; 1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene, which has a similar structure to fungal melanin; idebenone metabolite QS-4, a benzoquinone with protective effects against lipid peroxidation and ROS-induced damage[367–369]; and the pyranochromone pulverochromenol, which inhibits mammalian cell mitochondrial respiration and has antibacterial activity, and also has a similar structure to the anthraquinones endocrocin and monodictyphenone[370–372] (Table 2). We showed that production of the red-fluorescent compound correlates to expression of the MFS transporter gene from the *PKS8-1* gene cluster, and that this gene cluster is likely to produce a polyketide similar to monodictyphenone. Since we identified pulverochromenol, a compound similar in structure to monodictyphenone, from the extracts of the red compound (Table 2), it is possible that this is the compound produced by the PKS8-1 biosynthetic cluster.

Our expression assays with the *aflR*-like overexpresser suggest that the AflR-like transcription factor regulates the expression of some, but not all genes in the *PKS8-1* gene cluster. We were unable to observe any increase in red fluorescence in over-expression mutants of the *aflR*-like transcription factor or of the MFS transporter. Further research will be required to determine whether the lack of increased red fluorescence is because we were unable to achieve higher expression of all of the PKS8-1 biosynthetic genes.

From three separate inoculation experiments, disease symptoms from the *aflR*-like transcription factor over-expressers were consistently severe (Figures 11 and 12). Due to variability of disease symptoms from the wild-type and the hyg-control, we are unable to determine with certainty whether disease from the *aflR*-like over-expressers is more severe than the controls: in the first two inoculation experiments, the *aflR*-like over-expressers had more lesions than the controls (Figure 11A), but no difference was observed in the third inoculation experiment (Figure 11B). Further research will be needed to determine if there is a difference in gene expression in experiments for which there is a difference in symptoms. We also showed that some genes in the *PKS8-1* gene cluster have higher expression in infected leaf tissue than they do in germinating conidia, and for genes encoding PKS8-1 and the cytochrome P450, expression is highest at about 3 weeks post-inoculation, after which it tapers off. Since necrotic lesions began to appear in our experiments at about 3-4 weeks post-inoculation (Figure 11), perhaps the polyketide produced by the *PKS8-1* gene cluster could be involved in the switch in lifestyle from biotrophy to necrotrophy. Overall from our expression assays, we can conclude that while our *aflR*-like transcription factor overexpression construct is sufficient in culture to cause higher expression of the *aflR*-like gene and some other genes in the gene cluster (Figure 10), there must be other, uncharacterized regulatory factors for this gene cluster since higher expression of the genes in the cluster was not consistent in the third inoculation experiment (Figure 13).

In addition to the data we have generated concerning the red-fluorescent compound and the PKS8-1 biosynthetic cluster, our study represents the first report of over-expression of an *M. fijiensis* gene, since we achieved approximately a 10-fold increase in the expression

of the *aflR*-like transcription factor in culture conditions under the control of the *trpC* promoter (Figure 10). To date, limited work has been done on over-expression of genes in the *Mycosphaerellaceae* species. *In Mycosphaerella graminicola*, the ribosomal protein 27 promoter from *Magnaporthe oryzae* has been used for over-expression of the *CREA*, *NOXa*, and *AREA* genes and has achieved about 3 to 4-fold increases in gene expression[373]. This was using the Agrobacterium-mediated transformation-compatible vector pBHT2, for which restriction enzyme-based cloning methods are necessary[373]. Both pBHT2 and our modified pEarleyGate 100 destination vector have thus been proven effective for Agrobacterium-mediated transformation of over-expression constructs in *Mycosphaerella* species. However, since the vector created for our study is Gateway cloning-compatible (Figure S2D), this vector will provide an important tool for using Agrobacterium-mediated transformation for gene over-expression in a way that is not dependent on the restriction sites present in a gene of interest.

Conclusions

This is the first study to observe red and green fluorescence from *M. fijiensis* hyphae, to show that these have similar but not identical fluorescence emission peaks to oxidized and reduced cercosporin, respectively, and to show that the red fluorescent compound has similar solubility as cercosporin in acetone and DMSO. We showed that the red fluorescence correlates to increased expression of the MFS transporter gene from the *PKS8-1* gene cluster, and we used bioinformatics analyses to predict that the *PKS8-1* gene cluster should produce a similar compound to the anthraquinone monodictyphenone. From acetone and DMSO

extracts from the red-fluorescent *M. fijiensis* cultures, this study identified fourteen compounds that had not previously been reported from this fungus, including: 1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene, the sesquiterpenoids 4(15)-hirsutene and patchoula-2,4-diene, the benzoquinone idebenone metabolite QS-4, and the pyranochromone pulverochromenol, which has a similar structure to the anthraquinone monodictyphenone. We over-expressed the *aflR*-like transcription factor from the *PKS8-1* gene cluster, and showed that this resulted in higher expression of some, but not all genes in the putative *PKS8-1* gene cluster. Inoculation experiments with the *aflR*-like over-expresser suggest, but were unable to conclusively prove that the over-expressers may be more pathogenic than the wild-type *M. fijiensis*. Together, these results provide an exciting new avenue of research for the long-standing hypothesis that *M. fijiensis* may produce a similar metabolite as its close relatives *C. nicotianae* and *R. collo-cygni*, which produce quinones that cause phytotoxicity in the light.

Methods

Microscopy of *M. fijiensis* Colonies with Red Fluorescence

M. fijiensis isolates and *C. nicotianae* were grown on PDA with 0.132% DMSO added. Colonies were screened every few days under UV fluorescence microscopy with a dissecting microscope for approximately 4 weeks. *M. fijiensis* hyphae fluorescing or not fluorescing red were then imaged using a Zeiss LSM 710 confocal microscope with a 488 nm argon laser reflecting off a 488 nm main beam splitter. The same field of view at different emission wavelengths was imaged between 513 and 709 nm with a 20 nm step size to

generate a λ-stack. ZEN software was used to determine the fluorescence emission spectra at several different locations within the hyphae.

Acetone or DMSO was added to mycelia from *M. fijiensis* colonies that were fluorescing red, as well as mycelia from *C. nicotianae* colonies that were producing cercosporin. The mycelia were observed under a dissecting microscope with UV fluorescence microscopy to show that the red fluorescent compound diffuses into the solvent. Absorption spectra of the resulting extracts were generated using a NanoDrop™ ND-1000 spectrophotometer.

HPLC-Q-TOF-MS/MS Analysis

DMSO and acetone extracts from red-fluorescing *M. fijiensis* isolate 10CR1-36 colonies or cercosporin-producing *C. nicotianae* colonies were analyzed using a high resolution Agilent 6520 LC-Q-TOF-MS/MS instrument. As a further control, cercosporin crystals were obtained using the methods described previously by Daub[362]. The crystals were dissolved in acetone or DMSO, and the resulting solutions were analyzed using the same methods. Tandem mass analysis was carried out to obtain accurate fragment of mass formulas.

To predict which compound in Table S1 may be responsible for the red fluorescence observed in the colonies, a list of compounds was generated that were identified from both the DMSO and the acetone extractions from *M. fijiensis*, but that were not identified from either of the *C. nicotianae* extractions, or from the purified cercosporin crystals that were dissolved in acetone or DMSO (Table 2).

Expression Analysis from Colonies with Red Fluorescence

Colony edges with or without red fluorescence from 2 week old colonies of *M. fijiensis* isolate 10CR1-24 were excised and flash frozen in liquid N_2. RNA was isolated using a Spectrum Plant Total RNA kit (Sigma), DNase treated using RNase-free DNase (Qiagen), and reverse transcribed using iScript Select cDNA synthesis kit (Bio-Rad), according to manufacturer's instructions. A control reaction was set up the same way, without reverse transcriptase enzyme added to check for genomic DNA contamination. Semi-quantitative RT-PCR assays were performed using OneTaq DNA polymerase (NEB) according to manufacturer's instructions, with primer sequences, annealing temperatures, extension times, cycles and expected product sizes indicated in Table S4.

RNA-Seq Analysis of *PKS8-1* Gene Cluster

Transcriptome sequencing of *M. fijiensis* grown in Potato Dextrose Broth (PDB) culture medium versus infected banana leaf tissue was previously described[45]. 'Grand Nain' banana tissue culture plants were grown in modified Murashige and Skoog medium, and then were transferred to greenhouse conditions in potting mix. Once plants grew to about 20 cm in height, they were transferred to an incubator with a 18 h light/6 h dark photoperiod with cool-white fluorescent lights at 25 °C; these plants and Potato Dextrose Broth (PDB) flasks were inoculated with conidia of *M. fijiensis* isolate 14H1-11A. Plants were inoculated by spraying with 20 mL of 5.2 x 10^4/mL conidia in 0.5% Tween 20, and then were covered with clear plastic bags for 1 week to maintain high humidity conditions. Symptomatic banana tissue at 6 weeks post-inoculation was flash frozen in liquid N_2 for RNA extraction and

transcriptome sequencing. Flasks were inoculated with 10 µL of 1.3×10^6/mL conidia and incubated in the dark in a rotary shaker at 25–30 °C for one week. Mycelium was harvested by filtering through Miracloth and flash freezing in liquid N_2. Three biological replicates from each condition were analyzed.

The Spectrum Plant Total RNA kit (Sigma) was used for RNA extraction, and samples were DNase treated using DNase I (Roche). The Genomic Sciences Laboratory (North Carolina State University) conducted total RNA sequencing. An Agilent Bioanalyzer and gel electrophoresis were used to confirm RNA quality, and strand-specific libraries were generated using an NEBNext Ultra Directional library prep kit (NEB). An Illumina HiSeq 2500 instrument was used to generate 125-base single-end reads.

RNA-Seq read quality was verified using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and CutAdapt v1.7 was used to trim Illumina TruSeq adapter sequences and low-quality bases, with a minimum sequence length of 36 and a quality cutoff of 20[209]. Reads were mapped to both the *M. fijiensis*[54,160] and banana genomes[34,210] using the program Tophat v2.0.9[211]. HTSeq v0.6.0 was used to determine gene expression levels[212], and DESeq2 v1.4.5 was used to identify differentially expressed genes[213]. RT-qPCR assays were used with several genes to validate results from the RNA-Seq analysis[46]. For *PKS8-1* and adjacent genes in the putative biosynthetic cluster, $\log_2$ fold change values were summarized in Figure 5.

Comparing *M. fijiensis* PKS8-1 Biosynthetic Genes with Putative Orthologs

Tblastn searches were done using BLAST+[206] for each predicted protein within the *M. fijiensis* gene cluster, using the same set of 103 Dothideomycete genomes as was used for a previous analysis[45] (Table S2). This information was used to determine if *PKS8-1* homologs were flanked by similar genes in other species.

A phylogenetic tree was generated using PKS protein sequences described previously[45], PKS4 sequences from *Sordaria macrospora* and *Neurospora crassa* (NCBI accessions XP_003348600.1 and XP_011395279.1, respectively), PKS8-1 homologs identified in Table S2 from *Mycosphaerella musicola* and *Mycosphaerella eumusae* (NCBI accessions KXT14292.1 and KXS96182.1, respectively), and three PKS sequences used to generate a phylogenetic tree by Chang et al[301]: the elsinochrome-producing EfPKS1 from *Elsinoë fawcettii* (NCBI accession ABU63483.1), the endocrocin and monodictyphenone-producing MdpG from *Aspergillus nidulans* (NCBI accession CBF90097.1), and the naphthopyrone YWA1-producing PKS from *Aspergillus nidulans* (NCBI accession CAA46695.2). PKS protein sequences were aligned using Mesquite v3.04[202] with MUSCLE v3.8.31[201]. ModelGenerator v0.85[203] was used to identify LG+I+G+F[204] as the best evolutionary model, using both the Akaike and Bayesian Information Criteria. RaxmlGUI v1.3.1[205] was used to generate a maximum likelihood phylogenetic tree with the LG+I+G+F evolutionary model, slow bootstrap, no outgroup, and the autoMRE function.

To compare the *M. fijiensis PKS8-1* gene cluster with the *MdpG* gene cluster from *A. nidulans*, gene annotations from the *A. nidulans* genome were downloaded from NCBI (Genome ID 22572). Blastp searches for protein sequences encoded by each gene in the

monodictyphenone-producing gene cluster from *A. nidulans*[363] were performed against *M. fijiensis* isolate CIRAD86 sequences using the NCBI non-redundant protein sequences database. This information was used to determine whether the putative orthologs in *M. fijiensis* flank the *PKS8-1* gene in the *M. fijiensis* genome (Figure 8, Table S3).

Creating *aflR*-like Transcription Factor and MFS Transporter Gene Over-expressers

The fungal *trpC* terminator was amplified from pTROYA[374] using primers to add XbaI and HindIII sites. The XbaI and HindIII sites were used to replace the OCS terminator in pEarleyGate 100[343]. A hygromycin resistance cassette was amplified from plasmid pCB1636[344] with primers to add HindIII restriction sites to both ends of the PCR product. The HindIII site in the modified pEarleyGate 100 vector was then used to insert the hygromycin resistance cassette as a fungal selectable marker. The *trpC* promoter was amplified from pTROYA[374] with primers to add SacI and XhoI sites. The SacI and XhoI sites in the modified pEarleyGate 100 were used to replace the 35S:Bar plant selectable marker with a fungal *trpC* promoter, to create the Gateway destination vector. The *aflR*-like transcription factor and the MFS transporter genes were amplified from isolate 96CAM275 and 10CR1-36, respectively, and inserted into pDONR221 via Gateway BP clonase (Life Technologies) reaction to create Entry clones containing the transcription factor or MFS transporter sequences. Finally, the transcription factor and MFS transporter sequences were moved into the destination vector using Gateway LR clonase (Life Technologies) to create the expression vectors. To create a hygromycin cassette-only control, modified pEarleyGate 100 with promoterless GFP and a hygromycin resistance cassette (Chapter 4) was digested

with EcoRI and XbaI to remove the GFP sequence. The vector was treated with Klenow enzyme to generate blunt ends, and was ligated back together to generate a modified pEarleyGate 100 with a hygromycin resistance cassette in the T-DNA (Figure S2).

The *aflR*-like transcription factor gene with the constitutive *trpC* promoter was transformed into *M. fijiensis* isolate 10CR1-24 as previously described (Chapter 4). Transformants were recovered on PDA with 125 mg/L hygromycin.

Preparation of *M. fijiensis* Samples Grown in PDB

M. fijiensis isolates were maintained on PDA with or without 125 mg/L hygromycin, as appropriate. To test gene expression in liquid culture, 50 mg samples of mycelium were macerated and suspended in 125-mL flasks with 50 mL PDB medium. Flasks were grown for 8 days at 28 °C at 200 rpm in the dark. In the initial experiment for semi-quantitative RT-PCR assays, one flask each of wild-type, hyg-only control transformants #1 and 2, and transcription factor over-expressers #1,3,4,5,6,7,8,9, and 10 were tested. In the second experiment to compare gene expression by RT-qPCR, six flasks each of wild-type, hyg-only control transformants #1 and 2, and transcription factor over-expressers #4 and 8 were tested. Mycelium was harvested by filtering through Miracloth and flash freezing in liquid nitrogen.

Banana Inoculations

Banana tissue culture plants of the Grand Nain cultivar were kindly provided by Miguel Muñoz, Dole Food Company, and were maintained on modified Murashige and Skoog medium as previously described, on an 18h light/6h dark photoperiod with cool-white

fluorescent lights at 25-30 °C[45]. Rooted in-vitro cultured banana plants were transferred to potting mix and grown in the North Carolina State University Phytotron facilities with greenhouse conditions on a 12h light/12h dark photoperiod at 26 °C and 22 °C in the day and night, respectively. *M. fijiensis* conidia were produced and harvested as previously described[45], and 5 mL of $2x10^4$ conidia/mL were atomized onto young banana plants (about 20 cm in height). In the first two inoculation experiments, eight to ten plants were inoculated per fungal genotype (wild-type, Hyg #1, Hyg #2, transcription factor over-expresser #4 and #8). In the third experiment, at least 24 plants were inoculated per fungal genotype (wild-type, Hyg #2, transcription factor over-expresser #4). For one week post-inoculation, each plant was covered in a clear plastic bag to maintain high humidity conditions. Plants were harvested from the third inoculation experiment for expression analysis by flash freezing symptomatic leaf tissue in liquid nitrogen. For inoculation experiments 1 and 2, each inoculated leaf was scanned weekly using a CanoScan LiDE 220 scanner. For inoculation experiment 3, each inoculated leaf was scanned just before harvesting for RNA extraction.

Expression Assays of Genes in the *PKS8-1* Gene Cluster

RNA was isolated using a Spectrum Plant Total RNA kit (Sigma), DNase treated using RNase-free DNase (Qiagen), and reverse transcribed using iScript Select cDNA synthesis kit (Bio-Rad), according to manufacturer's instructions. Semi-quantitative RT-PCR assays were performed using OneTaq DNA Polymerase (NEB) according to manufacturer's instructions, with primer sequences, annealing temperatures, extension times, cycles, and expected product sizes indicated in Table S4. iQ SYBR Green Supermix was used for RT-

qPCR reactions, with an initial denaturation of 95°C for 2 minutes, followed by 50 cycles of 95°C for 10 seconds, 57°C for 30 seconds, 72°C for 30 seconds with a plate read, and then 78°C for 30 seconds with a plate read. Melt curves were used to verify amplification of a single product for each reaction. Primer sequences are indicated in Table S5. Each gene of interest was normalized against two *M. fijiensis* reference genes having the same efficiency as the gene of interest, and fold-change was calculated using the $2^{-\Delta\Delta C_T}$ method[290].

Verification of Fungal Genotypes from Third Inoculation Experiment

To verify that plants were inoculated with conidia of the correct genotype in the third inoculation experiment, symptomatic leaf tissue from three plants inoculated with each genotype was harvested by flash freezing in liquid nitrogen. Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen), according to manufacturer's instructions. One PCR assay was done to amplify part of the hygromycin resistance cassette, and another PCR assay was done to amplify from the *trpC* promoter to the *aflR*-like transcription factor gene (Figure S1). PCR assays were conducted using OneTaq DNA polymerase (NEB) according to manufacturer's instructions. Primer sequences, annealing temperatures, extension times, and number of cycles are indicated in Table S4.

Acknowledgements

We thank Dr. Miguel Muñoz (Dole Food Company) for providing tissue-cultured banana plants, Dr. Sonia Herrero (Syngenta Biotechnology) for assistance with gene cloning, and Morgan Carter (Cornell University) for assistance screening for red fluorescence. We also

thank the Phytotron staff at North Carolina State University for their assistance with caring for the plants.

Supplemental Figures and Tables

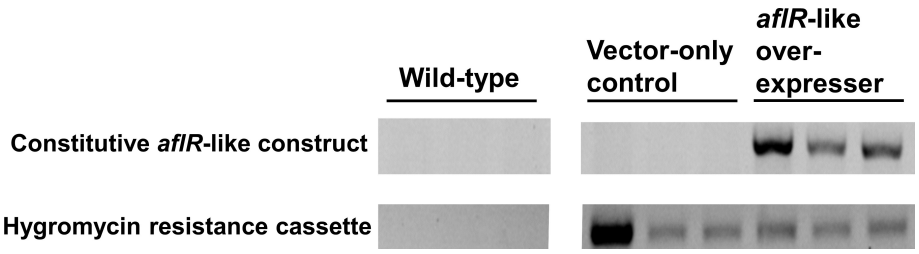


Figure S1. Verification of genotypes from the third inoculation experiment.
To verify that the banana plants from the third inoculation experiment were inoculated with *M. fijiensis* of the correct genotype, leaves with lesions were harvested from three plants that were inoculated with each fungal genotype. Genomic DNA was extracted from the leaf tissue, and for each sample, PCR assays were done to amplify a region from the hygromycin resistance cassette, or a product spanning from the constitutive *trpC* promoter to the *aflR*-like transcription factor sequence.

Figure S2. Cloning plasmids for transformation of *M. fijiensis*.
The OCS terminator from vector pEarleyGate 100[343] (A) was removed using the XbaI and HindIII sites, and was replaced with a PCR product amplified from pTROYA[374] of the fungal *trpC* terminator, to create (B). A hygromycin resistance cassette was amplified from the plasmid pCB1636[344], and was inserted into the modified pEarleyGate vector (B) using the HindIII site, to create (C). A fungal *trpC* promoter was amplified from pTROYA[374] and was used to replace the 35S:Bar cassette in (C), using the SacI and XhoI sites, to create the destination vector (D). The *aflR*-like transcription factor gene was amplified from *M. fijiensis* and was moved into the plasmid pDONR221 via a Gateway BP reaction to create the entry vector (E), and a Gateway LR reaction was used to generate the expression vector (F). The same method was used to create the entry vector (G) and expression vector (H) for the MFS transporter. Finally, the modified pEarleyGate 100 with a promoterless GFP sequence and a hygromycin resistance cassette (Chapter 4) (I) was digested with EcoRI and XbaI, treated with Klenow enzyme, and ligated back together to create the hygromycin resistance cassette-only control vector (J).

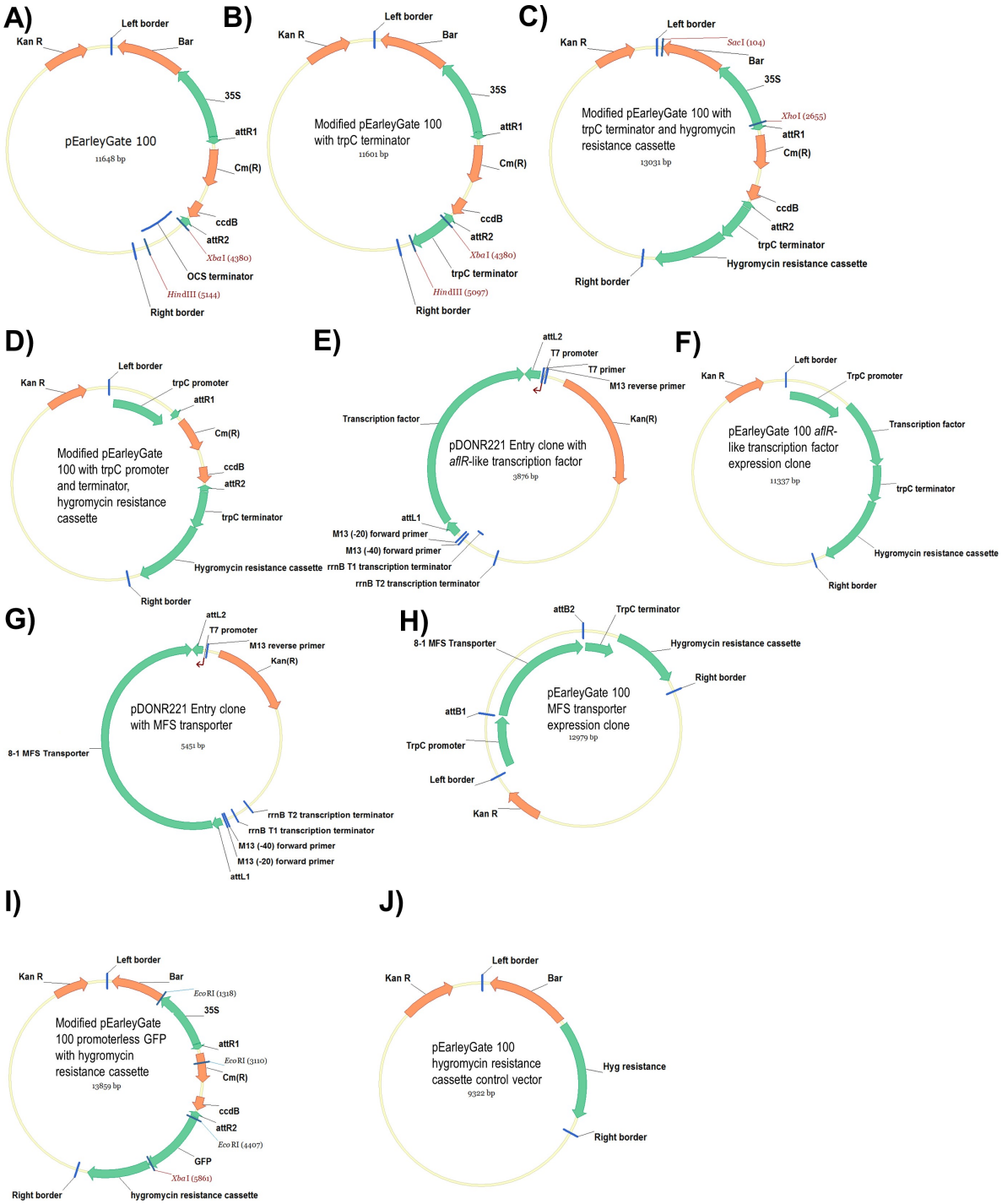

A)
Left border
Kan R
Bar
35S
pEarleyGate 100
11648 bp
attR1
Cm(R)
ccdB
attR2
XbaI (4380)
OCS terminator
HindIII (5144)
Right border
B)
Left border
Kan R
Bar
35S
Modified pEarleyGate 100
with trpC terminator
11601 bp
attR1
Cm(R)
ccdB
attR2
XbaI (4380)
trpC terminator
HindIII (5097)
Right border
C)
Left border
Kan R
SacI (104)
Bar
35S
Modified pEarleyGate 100 with
trpC terminator and hygromycin
resistance cassette
13031 bp
XhoI (2655)
attR1
Cm(R)
ccdB
attR2
trpC terminator
Hygromycin resistance cassette
Right border
D)
Left border
Kan R
trpC promoter
attR1
Cm(R)
Modified pEarleyGate
100 with trpC promoter
and terminator,
hygromycin resistance
cassette
ccdB
attR2
trpC terminator
Hygromycin resistance cassette
Right border
E)
attL2
T7 promoter
T7 primer
M13 reverse primer
Transcription factor
pDONR221 Entry clone with
aflR-like transcription factor
3876 bp
Kan(R)
attL1
M13 (-20) forward primer
M13 (-40) forward primer
rrnB T1 transcription terminator
rrnB T2 transcription terminator
F)
Left border
Kan R
TrpC promoter
Transcription factor
pEarleyGate 100 aflR-
like transcription factor
expression clone
11337 bp
trpC terminator
Hygromycin resistance cassette
Right border
G)
attL2
T7 promoter
M13 reverse primer
Kan(R)
pDONR221 Entry clone
with MFS transporter
5451 bp
8-1 MFS Transporter
rrnB T2 transcription terminator
rrnB T1 transcription terminator
M13 (-40) forward primer
M13 (-20) forward primer
attL1
H)
attB2
TrpC terminator
8-1 MFS Transporter
Hygromycin resistance cassette
Right border
pEarleyGate 100
MFS transporter
expression clone
12979 bp
attB1
TrpC promoter
Left border
Kan R
I)
Left border
Kan R
Bar
EcoRI (1318)
35S
Modified pEarleyGate
100 promoterless GFP
with hygromycin
resistance cassette
13859 bp
attR1
EcoRI (3110)
Cm(R)
ccdB
attR2
EcoRI (4407)
GFP
XbaI (5861)
Right border
hygromycin resistance cassette
J)
Left border
Kan R
Bar
pEarleyGate 100
hygromycin resistance
cassette control vector
9322 bp
Hyg resistance
Right border

Table S1. Chemicals identified from *M. fijiensis* and *C. nicotianae* extracts with DMSO or acetone.

Red-fluorescing colony edges from *M. fijiensis* isolate 10CR1-36 were excised and the red fluorescent compound was extracted from the tissue using DMSO or acetone. As a control, DMSO or acetone was used to extract cercosporin from *C. nicotianae* colonies. Cercosporin crystals were also purified using the methods described by Daub[362], and the crystals were dissolved in DMSO or acetone. HPLC and tandom mass spectrometry analysis was performed to identify compounds from each of the resulting solutions. For each compound identified, the table indicates the name of the compound, its chemical formula, and its molecular mass. A) Purified cercosporin crystal dissolved in DMSO; B) Purified cercosporin crystal dissolved in acetone; C) Crude extract from cercosporin-producing *C. nicotianae* colony in DMSO; D) Crude extract from cercosporin-producing *C. nicotianae* colony in acetone; E) Crude extract from red-fluorescing *M. fijiensis* colony in DMSO; F) Crude extract from red-fluorescing *M. fijiensis* colony in acetone.

Table S1A

Name	Formula	Mass
(22E)-3alpha,12alpha-Dihydroxy-5beta-chol-22-en-24-oic Acid	$C_{24}H_{38}O_4$	390.277
(22E)-3alpha,12alpha-Dihydroxy-5beta-chol-22-en-24-oic Acid	$C_{24}H_{38}O_4$	390.277
(22E)-3alpha,12alpha-Dihydroxy-5beta-chol-22-en-24-oic Acid	$C_{24}H_{38}O_4$	390.277
Methoxyflurane	$C_3H_4Cl_2F_2O$	163.962
12-amino-dodecanoic acid	$C_{12}H_{25}NO_2$	215.189
Heptaminol	$C_8H_{19}NO$	145.147
Mercaptoethanol	C_2H_6OS	78.0136
4,5-Dihydroxyphthalate	$C_8H_6O_6$	198.016
Fosamine	$C_3H_8NO_4P$	153.017
Thiometon	$C_6H_{15}O_2PS_3$	245.997
5-(3-Buten-1-ynyl)-2,2'-bithiophene	$C_{12}H_8S_2$	216.006
8-Hydroxy-3,4,6-trichlorodibenzofuran	$C_{12}H_5Cl_3O_2$	285.935
Fosfosal	$C_7H_7O_6P$	217.998
L-Cysteine	$C_3H_7NO_2S$	121.02
5-Ureido-4-imidazole carboxylate	$C_5H_6N_4O_3$	170.044
2-nonene-4,6,8-triynal	C_9H_4O	128.026
alpha-Methyl-m-tyramine	$C_9H_{13}NO$	151.1
Pirimicarb	$C_{11}H_{18}N_4O_2$	238.142
5-hydroxy caproaldehyde	$C_6H_{12}O_2$	116.084
Diethyl phthalate	$C_{12}H_{14}O_4$	222.089
Trietazine	$C_9H_{16}ClN_5$	229.111
R-4-benzyl-3-isobutyryloxazolidin-2-one	$C_{14}H_{17}NO_3$	247.121
Tenuazonic acid	$C_{10}H_{15}NO_3$	197.105
Leu Lys Ala	$C_{15}H_{30}N_4O_4$	330.226
Ser Gly Arg	$C_{11}H_{22}N_6O_5$	318.166
decanamide	$C_{10}H_{21}NO$	171.162
Anapheline	$C_{13}H_{24}N_2O$	224.189
n-Pentadecylamine	$C_{15}H_{33}N$	227.261
Deacetylspironolactone	$C_{22}H_{30}O_3S$	374.193
Leu Lys Ala	$C_{15}H_{30}N_4O_4$	330.226
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
Ethephon	$C_2H_6ClO_3P$	143.975
3-Methoxy-D-homoestra-1,3,5(10),8-tetraen-17abeta-ol	$C_{20}H_{26}O_2$	298.194
Cercosporin	$C_{29}H_{26}O_{10}$	534.153
Aurasperone D	$C_{31}H_{24}O_{10}$	556.134
N-Acetyl-D-mannosaminolactone	$C_8H_{13}NO_6$	219.074
Ethephon	$C_2H_6ClO_3P$	143.975
ent-Corey PG-Lactone Diol	$C_{15}H_{24}O_4$	268.168
Diclomezine	$C_{11}H_8Cl_2N_2O$	254.001
Specionin	$C_{20}H_{26}O_8$	394.164
Chrysanthetriol	$C_{15}H_{26}O_3$	254.189
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.141
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.141
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.127
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
Trp Gly Ile	$C_{19}H_{26}N_4O_4$	374.196
Syzygiol	$C_{18}H_{18}O_5$	314.116

Table S1B

Name	Formula	Mass
Fosfosal	$C_7H_7O_6P$	217.997
Tioconazole	$C_{16}H_{13}Cl_3N_2OS$	385.982
Benzaldehyde	C_7H_6O	106.042
alpha-Methyl-m-tyramine	$C_9H_{13}NO$	151.1
5-hydroxy caproaldehyde	$C_6H_{12}O_2$	116.084
5-hydroxy caproaldehyde	$C_6H_{12}O_2$	116.084
R-4-benzyl-3-isobutyryloxazolidin-2-one	$C_{14}H_{17}NO_3$	247.121
Ramentaceone	$C_{11}H_8O_3$	188.046
tert-Butylbicyclophosphorothionate	$C_8H_{15}O_3PS$	222.047
Dihydroartemisinin	$C_{15}H_{24}O_5$	284.163
Granisetron metabolite 3	$C_{18}H_{24}N_4O_2$	328.19
N,N-Diethylphenylacetamide	$C_{12}H_{17}NO$	191.131
Anapheline	$C_{13}H_{24}N_2O$	224.189
Ramentaceone	$C_{11}H_8O_3$	188.046
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
Ethephon	$C_2H_6ClO_3P$	143.975
Cercosporin	$C_{29}H_{26}O_{10}$	534.154
Aurasperone D	$C_{31}H_{24}O_{10}$	556.135
Diclomezine	$C_{11}H_8Cl_2N_2O$	254.001
ent-Corey PG-Lactone Diol	$C_{15}H_{24}O_4$	268.167
Dalbinol O-glucoside	$C2_9H_{32}O_{13}$	588.187
Specionin	$C_{20}H_{26}O_8$	394.164
5-Methoxy-2,2-dimethyl-8-phenyl-2H,6H-benzo[1,2-b:5,4-b']dipyran-6-one	$C_{21}H_{18}O_4$	334.121
1",2"-Dihydro-O-methylcyclopomiferin	$C_{26}H_{28}O_6$	436.188
2,2-Dimethyl-3,4-bis(4-methoxyphenyl)-2H-1-benzopyran-7-ol acetate	$C_{27}H_{26}O_5$	430.179
Isobergaptene	$C_{12}H_8O_4$	216.041
Arg Pro Thr	$C_{15}H_{28}N_6O_5$	372.214
Deoxynupharidine	$C_{15}H_{23}NO$	233.178
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.127
Benzaldehyde	C_7H_6O	106.042
Fluticasone 17-carboxylic acid	$C_{21}H_{26}F_2O_5$	396.174
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
1-Methoxypyrene	$C_{17}H_{12}O$	232.09
8Z-Dodecenyl acetate	$C_{14}H_{26}O_2$	226.193
Syzygiol	$C_{18}H_{18}O_5$	314.116
Syzygiol	$C_{18}H_{18}O_5$	314.116
Syzygiol	$C_{18}H_{18}O_5$	314.116

Table S1C

Name	Formula	Mass
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.1
Ile Lys His	$C_{18}H_{32}N_6O_4$	396.248
Pro Asp Asn	$C_{13}H_{20}N_4O_7$	344.133
diamino-pimelic acid	$C_7H_{14}N_2O_4$	190.094
Ac-Tyr-OEt	$C_{13}H_{17}NO_4$	251.116
Carmofur	$C_{11}H_{16}FN_3O_3$	257.119
MID74868:Heptanoic acid, 2-amino-, (1)-; Heptanoic acid, 2-amino-, DL-; (1)-2-Aminoheptanoic acid; a	$C_7H_{15}NO_2$	145.11
D-Pantetheine 4'-phosphate	$C_{11}H_{23}N_2O_7PS$	358.098
Kojibiose	$C_{12}H_{22}O_{11}$	342.117
D-Aspartic acid	$C_4H_7NO_4$	133.038
2-Aminobut-2-enoate	$C_4H_7NO_2$	101.047
2'-Deoxymugineic acid	$C_{12}H_{20}N_2O_7$	304.127
Isoglutamate	$C_5H_9NO_4$	147.054
Ile Lys His	$C_{18}H_{32}N_6O_4$	396.248
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.08
PA(13:0/0:0)	$C_{16}H_{33}O_7P$	368.195
N-Methylethanolamine phosphate	$C_3H_{10}NO_4P$	155.034
2-Deoxy-2-dimethylamino-alpha-D-Glucose	$C_8H_{17}NO_5$	207.11
AMT	$C_5H_{10}N_2S$	130.056
1Î±,25-dihydroxy-24a-homo-22-thiavitamin D3 / 1Î±,25-dihydroxy-24a-homo-22-thiacholecalciferol	$C_{27}H_{44}O_3S$	448.302
Idebenone Metabolite (Benzenebutanoic acid, 2,5-dihydroxy-3,4-dimethoxy-6-methyl-)	$C_{13}H_{18}O_6$	270.11
Dulcitol	$C_6H_{14}O_6$	182.079
Quercetagetin 3,5,6,7,3'-pentamethyl ether	$C_{20}H_{20}O_8$	388.116
Heptabarbital	$C_{13}H_{18}N_2O_3$	250.132
Tribenuron methyl	$C_{15}H_{17}N_5O_6S$	395.09
2(5H)-Furanone	$C_4H_4O_2$	84.021
3-hydroxy-3-methyl-2-oxo-pentanoic acid	$C_6H_{10}O_4$	146.058
Dihydrophloroglucinol	$C_6H_8O_3$	128.047
9S,10S,11R-trihydroxy-12Z-octadecenoic acid	$C_{18}H_{34}O_5$	330.239
Hydroquinone	$C_6H_6O_2$	110.036
2-Deoxy-D-galactose	$C_6H_{12}O_5$	164.069
Nonanoic acid, 3-amino-, (R)-	$C_9H_{19}NO_2$	173.14
Î²-vinyl acrylic acid	$C_5H_6O_2$	98.0364
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
3-keto valeric acid	$C_5H_8O_3$	116.048
4-Nitrotoluene	$C_7H_7NO_2$	137.048
Arabinosylhypoxanthine	$C_{10}H_{12}N_4O_5$	268.08
Xylitol	$C_5H_{12}O_5$	152.07
Cyclohexane-1,3-dione	$C_6H_8O_2$	112.053
Nonanoic acid, 3-amino-, (R)-	$C_9H_{19}NO_2$	173.142
Ludovicin A	$C_{15}H_{20}O_4$	264.135
Toxoflavine	$C_7H_7N_5O_2$	193.061
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
8-iso Prostaglandin F1Î²	$C_{20}H_{36}O_5$	356.255
HA-1077	$C_{14}H_{17}N_3O_2S$	291.103

Table S1C (continued)

Name	Formula	Mass
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
2-Methyl-3-oxoadipate	$C_7H_{10}O_5$	174.053
N-Acetyl-leucyl-leucine	$C_{14}H_{26}N_2O_4$	286.189
6-methyltetrahydropterin	$C_7H_{11}N_5O$	181.097
carisoprodol	$C_{12}H_{24}N_2O_4$	260.173
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.099
16b-Hydroxyestrone	$C_{18}H_{22}O_3$	286.157
7,8-Didemethyl-8-hydroxy-5-deazariboflavin	$C_{16}H_{17}N_3O_7$	363.108
5,7,3',5'-Tetrahydroxy-3,6,8,4'-tetramethoxyflavone 3'-glucoside	$C_{25}H_{28}O_{15}$	568.142
Spinosin 6'''-(E)-p-coumarate	$C_{37}H_{38}O_{17}$	754.206
Isonicotinylglycine	$C_8H_8N_2O_3$	180.054
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
Bis (2-hydroxypropyl) amine	$C_6H_{15}NO_2$	133.11
Purine	$C_5H_4N_4$	120.044
3-Deoxyguanosine	$C_{10}H_{13}N_5O_4$	267.096
L-Glutamic acid n-butyl ester	$C_9H_{17}NO_4$	203.116
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
Î³-Glutamyl-Î²-aminopropiononitrile	$C_8H_{13}N_3O_3$	199.096
Fosfosal	$C_7H_7O_6P$	217.998
V-PYRRO/NO	$C_6H_{11}N_3O_2$	157.086
N2,N5-Dibenzoyl-L-ornithine	$C_{19}H_{20}N_2O_4$	340.141
Chlormephos	$C_5H_{12}ClO_2PS_2$	233.973
Isoglutamate	$C_5H_9NO_4$	147.054
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
Î³-Glutamyl-Î²-aminopropiononitrile	$C_8H_{13}N_3O_3$	199.096
Bromobenzene	C_6H_5Br	155.957
5-Amino-4-imidazole carboxylate	$C_4H_5N_3O_2$	127.037
4,4'-Stilbenedicarboxamidine	$C_{16}H_{16}N_4$	264.137
12-amino-dodecanoic acid	$C_{12}H_{25}NO_2$	215.188
Heptaminol	$C_8H_{19}NO$	145.147
Mercaptoethanol	C_2H_6OS	78.0136
Fosamine	$C_3H_8NO_4P$	153.017
Nitrobenzene	$C_6H_5NO_2$	123.033
Dehydro-L-(+)-ascorbic acid dimer	$C_{12}H_{12}O_{12}$	348.032
3-Oxo-3-phenylpropanoate	$C_9H_8O_3$	164.047
Zalcitabine	$C_9H_{13}N_3O_3$	211.096
(+)-trans-alpha-Irone	$C_{14}H_{22}O$	206.166
L-Cysteine	$C_3H_7NO_2S$	121.02
5-Ureido-4-imidazole carboxylate	$C_5H_6N_4O_3$	170.044
Zalcitabine	$C_9H_{13}N_3O_3$	211.096
N,N-Diethylglycine	$C_6H_{13}NO_2$	131.095
Val Phe Ile	$C_{20}H_{31}N_3O_4$	377.231
5-(3-Pyridyl)-2-hydroxytetrahydrofuran	$C_9H_{11}NO_2$	165.079
Val Pro Arg	$C_{16}H_{30}N_6O_4$	370.233
alpha-Methyl-m-tyramine	$C_9H_{13}NO$	151.1
Pirimicarb	$C_{11}H_{18}N_4O_2$	238.142
N(alpha)-t-Butoxycarbonyl-L-leucine	$C_{11}H_{21}NO_4$	231.146
Diethyl phthalate	$C_{12}H_{14}O_4$	222.089

Table S1C (continued)

Name	Formula	Mass
trans-9(S),10(S)-Dihydrodiolphenanthrene	$C_{14}H_{12}O_2$	212.084
Leu Lys Ala	$C_{15}H_{30}N_4O_4$	330.225
Ser Gly Arg	$C_{11}H_{22}N_6O_5$	318.166
decanamide	$C_{10}H_{21}NO$	171.162
Anapheline	$C_{13}H_{24}N_2O$	224.189
n-Pentadecylamine	$C_{15}H_{33}N$	227.261
Leu Lys Ala	$C_{15}H_{30}N_4O_4$	330.225
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
3-Methoxy-D-homoestra-1,3,5(10),8-tetraen-17abeta-ol	$C_{20}H_{26}O_2$	298.194
Cercosporin	$C_{29}H_{26}O_{10}$	534.153
Aurasperone D	$C_{31}H_{24}O_{10}$	556.135
N-Acetyl-D-mannosaminolactone	$C_8H_{13}NO_6$	219.074
Laurenobiolide	$C_{17}H_{22}O_4$	290.152
Diclomezine	$C_{11}H_8Cl_2N_2O$	254.001
Aliarin	$C_{22}H_{24}O_8$	416.146
5-Methoxy-2,2-dimethyl-8-phenyl-2H,6H-benzo[1,2-b:5,4-b']dipyran-6-one	$C_{21}H_{18}O_4$	334.121
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.127
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
8Z-Dodecenyl acetate	$C_{14}H_{26}O_2$	226.194

Table S1D

Name	Formula	Mass
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.1
Ac-Tyr-OEt	$C_{13}H_{17}NO_4$	251.116
Idebenone Metabolite (Benzenebutanoic acid, 2,5-dihydroxy-3,4-dimethoxy-6-methyl-)	$C_{13}H_{18}O_6$	270.11
Quercetagetin 3,5,6,7,3'-pentamethyl ether	$C_{20}H_{20}O_8$	388.115
Chrysoeriol 7-(3"-Z-p-coumaroylglucoside)	$C_{31}H_{28}O_{13}$	608.153
Tribenuron methyl	$C_{15}H_{17}N_5O_6S$	395.088
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
3-hydroxy-3-methyl-2-oxo-pentanoic acid	$C_6H_{10}O_4$	146.058
2-Deoxy-D-galactose	$C_6H_{12}O_5$	164.069
Dulcitol	$C_6H_{14}O_6$	182.079
Dihydrophloroglucinol	$C_6H_8O_3$	128.047
Î²-vinyl acrylic acid	$C_5H_6O_2$	98.0365
O-Carbamoyl-deacetylcephalosporin C	$C_{15}H_{20}N_4O_8S$	416.098
Toxoflavine	$C_7H_7N_5O_2$	193.061
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
9S,10S,11R-trihydroxy-12Z-octadecenoic acid	$C_{18}H_{34}O_5$	330.239
3-keto valeric acid	$C_5H_8O_3$	116.048
Xylitol	$C_5H_{12}O_5$	152.069
2-Oxosuberate	$C_8H_{12}O_5$	188.069
a-N-Acetylglucosamine	$C_8H_{15}NO_6$	221.09
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
N-Acetyl-leucyl-leucine	$C_{14}H_{26}N_2O_4$	286.189
5,7,3',5'-Tetrahydroxy-3,6,8,4'-tetramethoxyflavone 3'-glucoside	$C_{25}H_{28}O_{15}$	568.142
7,8-Didemethyl-8-hydroxy-5-deazariboflavin	$C_{16}H_{17}N_3O_7$	363.106
16b-Hydroxyestrone	$C_{18}H_{22}O_3$	286.156
Spinosin 6'''-(E)-p-coumarate	$C_{37}H_{38}O_{17}$	754.206
Binapacryl	$C_{15}H_{18}N_2O_6$	322.118
pamoate	$C_{23}H_{16}O_6$	388.094
5-Hydroxy-2-oxo-4-ureido-2,5-dihydro-1H-imidazole-5-carboxylate	$C_5H_6N_4O_5$	202.032
Bis (2-hydroxypropyl) amine	$C_6H_{15}NO_2$	133.11
a-N-Acetylglucosamine	$C_8H_{15}NO_6$	221.09
Dulcitol	$C_6H_{14}O_6$	182.079
DL-2-amino-octanoic acid	$C_8H_{17}NO_2$	159.126
L-Glutamic acid n-butyl ester	$C_9H_{17}NO_4$	203.115
Chlormephos	$C_5H_{12}ClO_2PS_2$	233.973
Fosfosal	$C_7H_7O_6P$	217.997
Bromobenzene	C_6H_5Br	155.958
Nitrobenzene	$C_6H_5NO_2$	123.032
Barbital	$C_8H_{12}N_2O_3$	184.086
N(alpha)-t-Butoxycarbonyl-L-leucine	$C_{11}H_{21}NO_4$	231.146
Tepoxalin	$C_{20}H_{20}ClN_3O_3$	385.122
4-keto-n-caproic acid	$C_6H_{10}O_3$	130.063
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
Lauryl hydrogen sulfate	$C_{12}H_{26}O_4S$	266.156
Cercosporin	$C_{29}H_{26}O_{10}$	534.154
Aurasperone D	$C_{31}H_{24}O_{10}$	556.135

Table S1D (continued)

Name	Formula	Mass
6-(Pentylthio)purine	$C_{10}H_{14}N_4S$	222.094
Aliarin	$C_{22}H_{24}O_8$	416.146
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.141
5-Methoxy-2,2-dimethyl-8-phenyl-2H,6H-benzo[1,2-b:5,4-b']dipyran-6-one	$C_{21}H_{18}O_4$	334.121
Eplerenone	$C_{24}H_{30}O_6$	414.205
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.141
11R-HpETE	$C_{20}H_{32}O_4$	336.229
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.127
11R-HpETE	$C_{20}H_{32}O_4$	336.229
8Z-Dodecenyl acetate	$C_{14}H_{26}O_2$	226.194

Table S1E

Name	Formula	Mass
Fructoselysine	$C_{12}H_{24}N_2O_7$	308.158
N4-Acetylaminobutanal	$C_6H_{11}NO_2$	129.078
D-Proline	$C_5H_9NO_2$	115.063
6-(alpha-D-Glucosaminyl)-1D-myo-inositol	$C_{12}H_{23}NO_{10}$	341.133
Amino acid(Arg-)	$C_6H_{14}N_4O_2$	174.112
3-(Pyrazol-1-yl)-L-alanine	$C_6H_9N_3O_2$	155.07
Flavone	$C_{15}H_{10}O_2$	222.069
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.099
Methyl acrylate	$C_4H_6O_2$	86.0368
DMPO	$C_6H_{11}NO$	113.084
sn-glycero-3-Phosphoethanolamine	$C_5H_{14}NO_6P$	215.056
3-Methyl-quinolin-2,8-diol	$C_{10}H_9NO_2$	175.063
D-Asparagine	$C_4H_8N_2O_3$	132.054
D-Serine	$C_3H_7NO_3$	105.043
N-Methyl-N,4-dinitrosoaniline	$C_7H_7N_3O_2$	165.054
Biotin sulfone	$C_{10}H_{16}N_2O_5S$	276.077
4-Oxoproline	$C_5H_7NO_3$	129.042
4-Carboxy-4-hydroxy-2-oxoadipate	$C_7H_8O_8$	220.022
D-Homoserine	$C_4H_9NO_3$	119.058
MID74868:Heptanoic acid, 2-amino-, (1)-; Heptanoic acid, 2-amino-, DL-; (1)-2-Aminoheptanoic acid; a	$C_7H_{15}NO_2$	145.11
diamino-pimelic acid	$C_7H_{14}N_2O_4$	190.095
Ac-Tyr-OEt	$C_{13}H_{17}NO_4$	251.116
6-bromo-5E,9Z-hexacosadienoic acid	$C_{26}H_{47}BrO_2$	470.276
Pseudopelletierine	$C_9H_{15}NO$	153.115
Adenine	$C_5H_5N_5$	135.054
Kojibiose	$C_{12}H_{22}O_{11}$	342.117
Methyl acrylate	$C_4H_6O_2$	86.0358
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
2,4,6,8,10-dodecapentaenal	$C_{12}H_{14}O$	174.104
Quercetagetin 3,5,6,7,3'-pentamethyl ether	$C_{20}H_{20}O_8$	388.116
Idebenone Metabolite (Benzenebutanoic acid, 2,5-dihydroxy-3,4-dimethoxy-6-methyl-)	$C_{13}H_{18}O_6$	270.11
Gonyautoxin 6	$C_{10}H_{17}N_7O_8S$	395.088
5-Acetamidovalerate	$C_7H_{13}NO_3$	159.089
Dulcitol	$C_6H_{14}O_6$	182.079
Lomustine	$C_9H_{16}ClN_3O_2$	233.092
2-Deoxy-D-galactose	$C_6H_{12}O_5$	164.068
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
1,2-Dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene	$C_{10}H_{10}O_3$	178.062
Idebenone Metabolite (QS-4)	$C_{13}H_{16}O_6$	268.095
D-Proline	$C_5H_9NO_2$	115.063
altretamine	$C_9H_{18}N_6$	210.159
8-iso Prostaglandin F1Î²	$C_{20}H_{36}O_5$	356.255
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
6-Acetamido-3-aminohexanoate	$C_8H_{16}N_2O_3$	188.116
N-Acetyl-leucyl-leucine	$C_{14}H_{26}N_2O_4$	286.189
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095

Table S1E (continued)

Name	Formula	Mass
Halofenozide	$C_{18}H_{19}ClN_2O_2$	330.115
3-O-Acetylhamayne	$C_{18}H_{19}NO_5$	329.126
Pulverochromenol	$C_{20}H_{22}O_4$	326.151
Marchantin A	$C_{28}H_{24}O_5$	440.162
N-Acetyl-D-mannosamine 6-phosphate	$C_8H_{16}NO_9P$	301.052
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.1
Rufloxacin	$C_{17}H_{18}FN_3O_3S$	363.105
Lamprolobine	$C_{15}H_{24}N_2O_2$	264.183
5,7,3',5'-Tetrahydroxy-3,6,8,4'-tetramethoxyflavone 3'-glucoside	$C_{25}H_{28}O_{15}$	568.142
Spinosin 6'''-(E)-p-coumarate	$C_{37}H_{38}O_{17}$	754.206
Hydroxyflutamide	$C_{11}H_{11}F_3N_2O_4$	292.066
AG-82	$C_{10}H_6N_2O_3$	202.037
alhpa-tocopheronolactone	$C_{16}H_{22}O_4$	278.152
Bis (2-hydroxypropyl) amine	$C_6H_{15}NO_2$	133.11
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
Pulverochromenol	$C_{20}H_{22}O_4$	326.151
Asebotin	$C_{22}H_{26}O_{10}$	450.153
Zidovudine	$C_{10}H_{13}N_5O_4$	267.097
L-Glutamic acid n-butyl ester	$C_9H_{17}NO_4$	203.115
Dulcitol	$C_6H_{14}O_6$	182.079
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
3-Deoxyguanosine	$C_{10}H_{13}N_5O_4$	267.097
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.094
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
Isoglutamate	$C_5H_9NO_4$	147.054
4-Oxoproline	$C_5H_7NO_3$	129.042
Ismine	$C_{15}H_{15}NO_3$	257.105
Purine	$C_5H_4N_4$	120.044
12-amino-dodecanoic acid	$C_{12}H_{25}NO_2$	215.188
Heptaminol	$C_8H_{19}NO$	145.146
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
2-hydroxychlorpropamide	$C_{10}H_{13}ClN_2O_4S$	292.03
Benzofuran	C_8H_6O	118.042
3-Oxo-3-phenylpropanoate	$C_9H_8O_3$	164.047
Purine	$C_5H_4N_4$	120.044
4-Fluorocatechol	$C_6H_5FO_2$	128.027
L-Cysteine	$C_3H_7NO_2S$	121.02
5-Ureido-4-imidazole carboxylate	$C_5H_6N_4O_3$	170.044
alpha-Methyl-m-tyramine	$C_9H_{13}NO$	151.099
Pro Arg	$C_{11}H_{21}N_5O_3$	271.165
N,N-Diethylglycine	$C_6H_{13}NO_2$	131.095
5-(3-Pyridyl)-2-hydroxytetrahydrofuran	$C_9H_{11}NO_2$	165.079
Gln Glu	$C_{10}H_{17}N_3O_6$	275.112
2-nonene-4,6,8-triynal	C_9H_4O	128.027
Diethyl phthalate	$C_{12}H_{14}O_4$	222.089
Aloperine	$C_{15}H_{24}N_2$	232.194
Aloperine	$C_{15}H_{24}N_2$	232.194
6Z-Nonenoic acid	$C_9H_{16}O_2$	156.116

Table S1E (continued)

Name	Formula	Mass
Apraclonidine	$C_9H_{10}Cl_2N_4$	244.029
Dihydroartemisinin	$C_{15}H_{24}O_5$	284.163
Dihydroartemisinin	$C_{15}H_{24}O_5$	284.163
Granisetron metabolite 3	$C_{18}H_{24}N_4O_2$	328.189
Arg Met	$C_{11}H_{23}N_5O_3S$	305.153
4(15)-Hirsutene	$C_{15}H_{24}$	204.188
decanamide	$C_{10}H_{21}NO$	171.162
Anapheline	$C_{13}H_{24}N_2O$	224.188
Ramentaceone	$C_{11}H_8O_3$	188.046
n-Pentadecylamine	$C_{15}H_{33}N$	227.261
10,13-Octadecadiynoic acid	$C_{18}H_{28}O_2$	276.209
9-HOTE	$C_{18}H_{30}O_3$	294.22
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
3-Methoxy-D-homoestra-1,3,5(10),8-tetraen-17abeta-ol	$C_{20}H_{26}O_2$	298.194
Patchoula-2,4-diene	$C_{15}H_{22}$	202.173
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
SQ 26180	$C_6H_{10}N_2O_6S$	238.027
Aliarin	$C_{22}H_{24}O_8$	416.146
9,10-epoxy-12-octadecenoic acid	$C_{18}H_{32}O_3$	296.235
(S)-lamenallenic acid	$C_{18}H_{30}O_2$	278.225
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.142
Ile Glu Thr	$C_{15}H_{27}N_3O_7$	361.184
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.142
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.1
Deoxynupharidine	$C_{15}H_{23}NO$	233.177
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.126
Fluticasone 17-carboxylic acid	$C_{21}H_{26}F_2O_5$	396.174
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
8Z-Dodecenyl acetate	$C_{14}H_{26}O_2$	226.194

Table S1F

Name	Formula	Mass
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.099
N-Nitrosodi-n-propylamine	$C_6H_{14}N_2O$	130.111
Carmofur	$C_{11}H_{16}FN_3O_3$	257.117
N-Formyl-4-amino-5-aminomethyl-2-methylpyrimidine	$C_7H_{10}N_4O$	166.085
3-hydroxy-3-methyl-2-oxo-pentanoic acid	$C_6H_{10}O_4$	146.058
Adenine	$C_5H_5N_5$	135.054
Idebenone Metabolite (Benzenebutanoic acid, 2,5-dihydroxy-3,4-dimethoxy-6-methyl-)	$C_{13}H_{18}O_6$	270.11
Dulcitol	$C_6H_{14}O_6$	182.079
Dihydrophloroglucinol	$C_6H_8O_3$	128.047
2-Deoxy-D-galactose	$C_6H_{12}O_5$	164.068
Î²-vinyl acrylic acid	$C_5H_6O_2$	98.0366
1,2-Dihydroxy-3,4-epoxy-1,2,3,4-tetrahydronaphthalene	$C_{10}H_{10}O_3$	178.063
p-Hydroxytriamterene	$C_{12}H_{11}N_7O$	269.102
Idebenone Metabolite (QS-4)	$C_{13}H_{16}O_6$	268.095
2-Amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	117.079
3-hydroxy-3-methyl-2-oxo-pentanoic acid	$C_6H_{10}O_4$	146.058
Toxoflavine	$C_7H_7N_5O_2$	193.06
O-Carbamoyl-deacetylcephalosporin C	$C_{15}H_{20}N_4O_8S$	416.097
carisoprodol	$C_{12}H_{24}N_2O_4$	260.174
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
1-Aminocyclohexanecarboxylic acid	$C_7H_{13}NO_2$	143.095
5-Acetylamino-6-formylamino-3-methyluracil	$C_8H_{10}N_4O_4$	226.069
1,2-Bis(4-nitrophenyl)ethane	$C_{14}H_{12}N_2O_4$	272.08
Rufloxacin	$C_{17}H_{18}FN_3O_3S$	363.105
2-Amino-3-methyl-1-butanol	$C_5H_{13}NO$	103.099
16b-Hydroxyestrone	$C_{18}H_{22}O_3$	286.156
5,7,3',5'-Tetrahydroxy-3,6,8,4'-tetramethoxyflavone 3'-glucoside	$C_{25}H_{28}O_{15}$	568.142
Pulverochromenol	$C_{20}H_{22}O_4$	326.15
5-Hydroxy-2-oxo-4-ureido-2,5-dihydro-1H-imidazole-5-carboxylate	$C_5H_6N_4O_5$	202.032
Spinosin 6'''-(E)-p-coumarate	$C_{37}H_{38}O_{17}$	754.207
Zidovudine	$C_{10}H_{13}N_5O_4$	267.097
Chlormephos	$C_5H_{12}ClO_2PS_2$	233.973
Fosfosal	$C_7H_7O_6P$	217.997
Bromobenzene	C_6H_5Br	155.957
N,N-Diethylglycine	$C_6H_{13}NO_2$	131.095
DHAP(8:0)	$C_{11}H_{21}O_7P$	296.103
Pilocarpidine	$C_{10}H_{14}N_2O_2$	194.106
5-(1-Hydroxyethyl)-4-methylthiazole	C_6H_9NOS	143.041
Capryloylglycine	$C_{10}H_{19}NO_3$	201.135
5-(3-Pyridyl)-2-hydroxytetrahydrofuran	$C_9H_{11}NO_2$	165.078
Australine	$C_8H_{15}NO_4$	189.101
Cystamine	$C_4H_{12}N_2S_2$	152.046
4(15)-Hirsutene	$C_{15}H_{24}$	204.188
4(15)-Hirsutene	$C_{15}H_{24}$	204.188
2-Aminoacridone	$C_{13}H_{10}N_2O$	210.08
9-HOTE	$C_{18}H_{30}O_3$	294.22
10,13-Octadecadiynoic acid	$C_{18}H_{28}O_2$	276.209

Table S1F (continued)

Name	Formula	Mass
3-tert-Butyl-5-methylcatechol	$C_{11}H_{16}O_2$	180.114
Decaketide tricyclic intermediate	$C_{20}H_{16}O_8$	384.085
6-(Pentylthio)purine	$C_{10}H_{14}N_4S$	222.094
Patchoula-2,4-diene	$C_{15}H_{22}$	202.172
5,10-Epoxy-muurolane	$C_{15}H_{24}O$	220.182
Campherene-2,13-diol	$C_{15}H_{26}O_2$	238.193
Aliarin	$C_{22}H_{24}O_8$	416.146
4,5-(methanoxyethano)isolongifol-4-ene	$C_{18}H_{28}O$	260.213
9,10-epoxy-12-octadecenoic acid	$C_{18}H_{32}O_3$	296.235
(S)-lamenallenic acid	$C_{18}H_{30}O_2$	278.225
3-Hydroxylidocaine glucuronide	$C_{20}H_{30}N_2O_8$	426.199
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.14
Eplerenone	$C_{24}H_{30}O_6$	414.205
1",2"-Dihydro-O-methylcyclopomiferin	$C_{26}H_{28}O_6$	436.188
epi-Tulipinolide diepoxide	$C_{17}H_{22}O_6$	322.141
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
Proanthocyanidin A2	$C_{30}H_{24}O_{12}$	576.127
1a,1b-dihomo-PGD2	$C_{22}H_{36}O_5$	380.257
15-oxo-hexadecanoic acid	$C_{16}H_{30}O_3$	270.219
11beta-hydroxyprogesterone	$C_{21}H_{30}O_3$	330.221
18-hydroxy-9S,10R-epoxy-stearic acid	$C_{18}H_{34}O_4$	314.246
1a,1b-dihomo-PGD2	$C_{22}H_{36}O_5$	380.257
15-oxo-hexadecanoic acid	$C_{16}H_{30}O_3$	270.219

Table S2. Homologs of putative PKS8-1 biosynthetic cluster genes in selected Dothideomycetes.

For each gene in the putative PKS8-1 biosynthetic cluster from *M. fijiensis*, tblastn searches were done using BLAST+ with the 103 Dothideomycete genomes listed previously[45]. Hits from *Mycosphaerella eumusae*, *Mycosphaerella musicola*, and *Pseudocercospora pini-densiflorae* are shown in the table. The table indicates the species and genome scaffold from which the hit was found, the percent protein similarity (% Sim), the bitscore (Bit), and the E-value. Hits for the A) Beta-galactosidase (accession XP_007929638.1); B) AflJ-like protein (accession XP_007929877.1); C) Hypothetical protein (accession XP_007929879.1); D) AflR-like transcription factor (accession XP_007929880.1); E) Hypothetical protein (accession XP_007929881.1); F) Hypothetical protein (accession XP_007929882.1); G) PKS8-1 (accession XP_007929626.1); H) Glyoxylase/beta-lactamase-like protein (accession XP_007929883.1); I) Hypothetical protein (accession XP_007929629.1); J) Hypothetical protein (accession XP_007929611.1); K) MFS transporter (accession XP_007929884.1); L) Cytochrome P450 (accession XP_007929885.1); M) Hypothetical protein (accession XP_007929886.1); N) Aldo/keto reductase (accession XP_007929887.1); O) Dehydrogenase/ 3-ketoacyl-(acyl-carrier-protein) reductase (accession XP_007929584.1); P) Aldo/keto reductase (accession XP_007929888.1); Q) Transcription factor (accession XP_007929889.1); R) Nuclear pore complex component (accession XP_007929890.1); S) Preprotein translocase subunit Sec66 (accession XP_007929891.1). Hits for the hypothetical protein (accession XP_007929878.1) were not found.

Table S2A

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000045.1	43.26	113	5E-23
Mycosphaerella eumusae	LFZN01000116.1	50.84	351	4E-131
Mycosphaerella eumusae	LFZN01000116.1	56.62	145	4E-131
Mycosphaerella eumusae	LFZN01000184.1	90.26	1715	0.00E+00
Mycosphaerella eumusae	LFZN01000204.1	51.99	252	3E-66
Mycosphaerella eumusae	LFZN01000204.1	57.14	73.6	1E-24
Mycosphaerella eumusae	LFZN01000204.1	38.26	66.6	1E-24
Mycosphaerella musicola	LFZO01000162.1	49.09	298	6E-97
Mycosphaerella musicola	LFZO01000162.1	49.77	84	6E-97
Mycosphaerella musicola	LFZO01000184.1	42.93	117	2E-24
Mycosphaerella musicola	LFZO01000295.1	49.53	239	2E-62
Mycosphaerella musicola	LFZO01000295.1	59.55	74.7	8E-27
Mycosphaerella musicola	LFZO01000295.1	39.39	70.1	8E-27
Mycosphaerella musicola	LFZO01000295.1	83.33	22.7	8E-27
Pseudocercospora pini-densiflorae	AWYD01001454.1	49.39	233	2E-60
Pseudocercospora pini-densiflorae	AWYD01001454.1	61.36	73.9	4E-23
Pseudocercospora pini-densiflorae	AWYD01001454.1	40.96	62.8	1E-07
Pseudocercospora pini-densiflorae	AWYD01001454.1	43.06	61.2	4E-23
Pseudocercospora pini-densiflorae	KI633681.1	43.13	112	1E-22
Pseudocercospora pini-densiflorae	KI634088.1	50.59	353	2E-99
Pseudocercospora pini-densiflorae	KI634088.1	50.78	140	3E-31

Table S2B

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000003.1	39.26	52	0.00006
Mycosphaerella eumusae	LFZN01000003.1	42.54	45.8	0.005
Mycosphaerella eumusae	LFZN01000004.1	59.57	63.9	1.00E-08
Mycosphaerella eumusae	LFZN01000004.1	54.41	40.8	0.19
Mycosphaerella eumusae	LFZN01000025.1	42.17	48.5	0.0007
Mycosphaerella eumusae	LFZN01000030.1	43.07	79	1.00E-13
Mycosphaerella eumusae	LFZN01000184.1	89.7	641	00.00E+00
Mycosphaerella musicola	LFZO01000086.1	96.13	337	1.00E-100
Mycosphaerella musicola	LFZO01000086.1	90.43	331	9.00E-99
Mycosphaerella musicola	LFZO01000091.1	44.07	65.1	4.00E-09
Mycosphaerella musicola	LFZO01000227.1	43.85	107	1.00E-22
Mycosphaerella musicola	LFZO01000244.1	41.99	65.1	4.00E-09
Mycosphaerella musicola	LFZO01000266.1	48.86	36.2	0.1
Mycosphaerella musicola	LFZO01000266.1	51.16	24.3	0.1
Mycosphaerella musicola	LFZO01000392.1	45.45	49.3	0.0004
Mycosphaerella musicola	LFZO01000565.1	52.59	63.5	1.00E-08
Pseudocercospora pini-densiflorae	AWYD01000810.1	91.33	656	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001286.1	45.68	68.6	3.00E-10
Pseudocercospora pini-densiflorae	AWYD01001656.1	41.09	62	3.00E-08
Pseudocercospora pini-densiflorae	AWYD01003569.1	62.89	80.9	2.00E-14
Pseudocercospora pini-densiflorae	AWYD01003569.1	45.08	70.5	7.00E-11
Pseudocercospora pini-densiflorae	AWYD01003636.1	42.86	60.5	1.00E-07
Pseudocercospora pini-densiflorae	AWYD01003684.1	48.38	146	2.00E-35
Pseudocercospora pini-densiflorae	AWYD01003684.1	46.19	52	0.00005
Pseudocercospora pini-densiflorae	KI633506.1	40.37	38.9	0.73
Pseudocercospora pini-densiflorae	KI633679.1	42.35	70.9	2.00E-18
Pseudocercospora pini-densiflorae	KI633679.1	47.79	47.4	2.00E-18
Pseudocercospora pini-densiflorae	KI634054.1	42.81	82	1.00E-14
Pseudocercospora pini-densiflorae	KI634105.1	43.73	64.3	8.00E-09
Pseudocercospora pini-densiflorae	KI634198.1	37.75	46.6	0.003
Pseudocercospora pini-densiflorae	KI634489.1	38.54	67.8	5.00E-10
Pseudocercospora pini-densiflorae	KI634612.1	45.87	35.4	7.2

Table S2C

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000184.1	63.32	174	1.00E-45
Mycosphaerella musicola	LFZO01000086.1	65.77	184	4.00E-49
Pseudocercospora pini-densiflorae	AWYD01000810.1	79.88	196	4.00E-53

Table S2D

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000004.1	69.77	52.4	0.00006
Mycosphaerella eumusae	LFZN01000004.1	65.91	48.9	0.0007
Mycosphaerella eumusae	LFZN01000004.1	51.16	37.7	2.4
Mycosphaerella eumusae	LFZN01000052.1	74	58.2	0.000001
Mycosphaerella eumusae	LFZN01000082.1	37.61	59.7	3.00E-07
Mycosphaerella eumusae	LFZN01000184.1	88.1	692	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	46.94	37.4	2.5
Mycosphaerella musicola	LFZO01000086.1	86.49	682	0.00E+00
Mycosphaerella musicola	LFZO01000088.1	42.34	54.7	0.00001
Mycosphaerella musicola	LFZO01000088.1	58.49	42	0.1
Mycosphaerella musicola	LFZO01000088.1	45.9	38.5	1.2
Mycosphaerella musicola	LFZO01000088.1	51.61	37.7	2.3
Mycosphaerella musicola	LFZO01000227.1	43.43	49.7	0.0004
Mycosphaerella musicola	LFZO01000303.1	39.56	60.1	2.00E-07
Mycosphaerella musicola	LFZO01000392.1	38.51	62.8	4.00E-08
Mycosphaerella musicola	LFZO01000569.1	62.5	58.9	6.00E-07
Pseudocercospora pini-densiflorae	AWYD01000015.1	54.55	49.7	0.0005
Pseudocercospora pini-densiflorae	AWYD01000810.1	89.4	707	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001586.1	63.46	45.4	0.008
Pseudocercospora pini-densiflorae	AWYD01003569.1	67.44	51.2	0.0001
Pseudocercospora pini-densiflorae	AWYD01003684.1	72.92	58.5	8.00E-07
Pseudocercospora pini-densiflorae	AWYD01003684.1	47.71	39.7	0.54
Pseudocercospora pini-densiflorae	AWYD01003889.1	74	58.5	8.00E-07
Pseudocercospora pini-densiflorae	KI633680.1	57.14	50.4	0.0003

Table S2E

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000184.1	68.97	32.7	5.4
Mycosphaerella eumusae	LFZN01000184.1	64	20	5.4
Mycosphaerella musicola	LFZO01000086.1	91.3	35.4	2.3
Pseudocercospora pini-densiflorae	AWYD01000810.1	75	42	0.02

Table S2F

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella musicola	LFZO01000086.1	48.67	39.7	0.074
Pseudocercospora pini-densiflorae	AWYD01001655.1	85.71	39.7	0.09

Table S2G

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000003.1	55.07	635	0.00E+00
Mycosphaerella eumusae	LFZN01000003.1	51.49	384	2.00E-105
Mycosphaerella eumusae	LFZN01000003.1	46.73	329	2.00E-88
Mycosphaerella eumusae	LFZN01000003.1	68.97	227	0.00E+00
Mycosphaerella eumusae	LFZN01000003.1	69.39	52.8	0.0003
Mycosphaerella eumusae	LFZN01000024.1	45.41	330	1.00E-88
Mycosphaerella eumusae	LFZN01000081.1	50.07	348	6.00E-100
Mycosphaerella eumusae	LFZN01000081.1	40.51	45.4	0.05
Mycosphaerella eumusae	LFZN01000081.1	44.32	45.1	6.00E-100
Mycosphaerella eumusae	LFZN01000115.1	59.59	1276	0.00E+00
Mycosphaerella eumusae	LFZN01000115.1	47.44	125	6.00E-44
Mycosphaerella eumusae	LFZN01000115.1	56.52	80.5	6.00E-44
Mycosphaerella eumusae	LFZN01000115.1	41.33	52.8	1.00E-13
Mycosphaerella eumusae	LFZN01000115.1	39.35	51.2	1.00E-13
Mycosphaerella eumusae	LFZN01000115.1	43.31	42.4	0.41
Mycosphaerella eumusae	LFZN01000184.1	94.73	1513	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	95.95	876	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	85.65	673	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	91.09	192	0.00E+00
Mycosphaerella eumusae	LFZN01000241.1	46.84	337	6.00E-91
Mycosphaerella eumusae	LFZN01000252.1	49.85	681	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	97.47	1446	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	87.13	956	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	85.24	667	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	92.08	193	0.00E+00
Mycosphaerella musicola	LFZO01000158.1	52.06	453	6.00E-127
Mycosphaerella musicola	LFZO01000225.1	49.42	670	0.00E+00
Mycosphaerella musicola	LFZO01000331.1	59.8	1276	0.00E+00
Mycosphaerella musicola	LFZO01000331.1	43.31	42	0.61
Mycosphaerella musicola	LFZO01000580.1	50.62	341	1.00E-96
Mycosphaerella musicola	LFZO01000580.1	43.18	40.8	1.00E-96
Pseudocercospora pini-densiflorae	AWYD01001655.1	92.44	2572	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	85.31	670	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01003363.1	52.66	449	7.00E-126
Pseudocercospora pini-densiflorae	AWYD01003684.1	55.18	629	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01003684.1	53	238	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01003684.1	69.39	53.1	0.0002
Pseudocercospora pini-densiflorae	KI633445.1	61.28	1244	0.00E+00
Pseudocercospora pini-densiflorae	KI633445.1	43.31	42.4	0.44
Pseudocercospora pini-densiflorae	KI633506.1	50.22	700	0.00E+00
Pseudocercospora pini-densiflorae	KI633506.1	46.77	319	9.00E-90
Pseudocercospora pini-densiflorae	KI633506.1	77.42	40	9.00E-90

Table S2H

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000020.1	56.1	47.8	0.0001
Mycosphaerella eumusae	LFZN01000020.1	84.62	22.7	0.0001
Mycosphaerella eumusae	LFZN01000092.1	61.34	216	2.00E-59
Mycosphaerella eumusae	LFZN01000184.1	94.59	494	2.00E-165
Mycosphaerella eumusae	LFZN01000184.1	92.19	114	2.00E-165
Mycosphaerella musicola	LFZO01000006.1	54.95	45.8	0.004
Mycosphaerella musicola	LFZO01000037.1	61.57	226	8.00E-63
Mycosphaerella musicola	LFZO01000086.1	95.37	494	1.00E-165
Mycosphaerella musicola	LFZO01000086.1	95	115	1.00E-165
Pseudocercospora pini-densiflorae	AWYD01000147.1	60.22	226	5.00E-63
Pseudocercospora pini-densiflorae	AWYD01000902.1	54.88	46.6	0.002
Pseudocercospora pini-densiflorae	AWYD01001655.1	94.59	494	3.00E-166
Pseudocercospora pini-densiflorae	AWYD01001655.1	92.19	116	3.00E-166

Table S2I

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000101.1	53.91	68.6	3.00E-11
Mycosphaerella eumusae	LFZN01000184.1	80.1	291	5.00E-88
Mycosphaerella musicola	LFZO01000001.1	48.19	39.3	0.084
Mycosphaerella musicola	LFZO01000051.1	46.31	36.6	0.83
Mycosphaerella musicola	LFZO01000086.1	80.95	291	9.00E-88
Pseudocercospora pini-densiflorae	AWYD01000033.1	47.27	32.7	0.007
Pseudocercospora pini-densiflorae	AWYD01000033.1	44.44	29.6	0.007
Pseudocercospora pini-densiflorae	AWYD01001543.1	47.46	40.4	0.042
Pseudocercospora pini-densiflorae	AWYD01001655.1	88.89	252	3.00E-76
Pseudocercospora pini-densiflorae	AWYD01001655.1	93.55	56.6	3.00E-76
Pseudocercospora pini-densiflorae	AWYD01004156.1	48.19	37.7	0.34
Pseudocercospora pini-densiflorae	KI633672.1	55.48	97.4	9.00E-21

Table S2J

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000184.1	89.39	117	2.00E-62
Mycosphaerella eumusae	LFZN01000184.1	86.57	112	2.00E-62
Mycosphaerella eumusae	LFZN01000184.1	100	53.9	2.00E-62
Mycosphaerella musicola	LFZO01000086.1	93.85	130	1.00E-67
Mycosphaerella musicola	LFZO01000086.1	89.23	118	1.00E-67
Mycosphaerella musicola	LFZO01000086.1	100	52.4	1.00E-67
Mycosphaerella musicola	LFZO01000771.1	47.89	65.9	1.00E-10
Pseudocercospora pini-densiflorae	AWYD01001655.1	77.19	169	6.00E-70
Pseudocercospora pini-densiflorae	AWYD01001655.1	89.23	117	6.00E-70
Pseudocercospora pini-densiflorae	AWYD01001985.1	41.94	51.6	0.000006
Pseudocercospora pini-densiflorae	AWYD01002155.1	47.89	68.6	2.00E-11

Table S2K

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000003.1	54.5	235	2.00E-80
Mycosphaerella eumusae	LFZN01000003.1	58.47	90.1	2.00E-80
Mycosphaerella eumusae	LFZN01000043.1	62.55	348	1.00E-101
Mycosphaerella eumusae	LFZN01000043.1	52.5	36.6	6
Mycosphaerella eumusae	LFZN01000091.1	65.58	408	3.00E-122
Mycosphaerella eumusae	LFZN01000184.1	90.76	543	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	91.28	300	0.00E+00
Mycosphaerella musicola	LFZO01000061.1	65.28	362	2.00E-106
Mycosphaerella musicola	LFZO01000061.1	48.72	44.7	0.025
Mycosphaerella musicola	LFZO01000086.1	96.71	547	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	90.77	298	0.00E+00
Mycosphaerella musicola	LFZO01000130.1	55.21	209	3.00E-101
Mycosphaerella musicola	LFZO01000130.1	77.65	186	3.00E-101
Pseudocercospora pini-densiflorae	AWYD01000488.1	55.13	163	4.00E-62
Pseudocercospora pini-densiflorae	AWYD01000488.1	55.8	102	4.00E-62
Pseudocercospora pini-densiflorae	AWYD01001185.1	61.98	187	2.00E-67
Pseudocercospora pini-densiflorae	AWYD01001185.1	59.26	95.5	2.00E-67
Pseudocercospora pini-densiflorae	AWYD01001185.1	64	52	2.00E-11
Pseudocercospora pini-densiflorae	AWYD01001185.1	53.85	43.1	2.00E-11
Pseudocercospora pini-densiflorae	AWYD01001655.1	96.41	547	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	91.28	298	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001761.1	61.66	139	1.00E-63
Pseudocercospora pini-densiflorae	AWYD01001761.1	52.01	131	1.00E-63
Pseudocercospora pini-densiflorae	AWYD01001985.1	55.26	179	3.00E-59
Pseudocercospora pini-densiflorae	AWYD01001985.1	44.1	75.5	3.00E-59
Pseudocercospora pini-densiflorae	KI633948.1	65.94	361	4.00E-106
Pseudocercospora pini-densiflorae	KI633948.1	47.14	36.6	6.8
Pseudocercospora pini-densiflorae	KI634005.1	65.05	396	3.00E-118

Table S2L

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000007.1	40.44	48.5	0.001
Mycosphaerella eumusae	LFZN01000083.1	41.73	52	0.00009
Mycosphaerella eumusae	LFZN01000162.1	46.85	58.2	0.000001
Mycosphaerella eumusae	LFZN01000184.1	96.83	727	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	97.18	147	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	95	79	0.00E+00
Mycosphaerella eumusae	LFZN01000282.1	38.64	73.9	3.00E-22
Mycosphaerella eumusae	LFZN01000282.1	70	57	3.00E-22
Mycosphaerella musicola	LFZO01000035.1	48.1	41.6	0.15
Mycosphaerella musicola	LFZO01000036.1	40.28	42.7	0.066
Mycosphaerella musicola	LFZO01000065.1	40.57	42.7	0.068
Mycosphaerella musicola	LFZO01000086.1	93.63	870	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	97.5	83.2	2.00E-14
Mycosphaerella musicola	LFZO01000215.1	51.76	88.6	4.00E-16
Mycosphaerella musicola	LFZO01000215.1	53.85	38.1	2.1
Mycosphaerella musicola	LFZO01000220.1	46.88	47.4	0.002
Mycosphaerella musicola	LFZO01000541.1	40.71	47.8	0.002
Mycosphaerella musicola	LFZO01000688.1	43.31	49.3	0.0007
Pseudocercospora pini-densiflorae	AWYD01000106.1	37.65	42	0.11
Pseudocercospora pini-densiflorae	AWYD01000490.1	44.09	49.7	0.0006
Pseudocercospora pini-densiflorae	AWYD01000858.1	43.33	45.4	0.01
Pseudocercospora pini-densiflorae	AWYD01001468.1	54.22	100	9.00E-48
Pseudocercospora pini-densiflorae	AWYD01001468.1	47.71	97.8	9.00E-48
Pseudocercospora pini-densiflorae	AWYD01001468.1	52.11	38.9	9.00E-48
Pseudocercospora pini-densiflorae	AWYD01001655.1	95.2	738	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	98.61	153	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	97.5	82	5.00E-14
Pseudocercospora pini-densiflorae	AWYD01001753.1	41.04	41.6	0.16
Pseudocercospora pini-densiflorae	AWYD01002593.1	40.95	43.9	0.034
Pseudocercospora pini-densiflorae	AWYD01002593.1	43.37	39.3	0.74
Pseudocercospora pini-densiflorae	AWYD01002647.1	38.96	45.4	0.01
Pseudocercospora pini-densiflorae	AWYD01002715.1	44.95	49.3	0.0007
Pseudocercospora pini-densiflorae	KI633797.1	40.71	45.8	0.009
Pseudocercospora pini-densiflorae	KI634057.1	36.92	45.4	0.01
Pseudocercospora pini-densiflorae	KI634165.1	42	42	0.12
Pseudocercospora pini-densiflorae	KI634606.1	55.56	139	3.00E-53
Pseudocercospora pini-densiflorae	KI634606.1	50.74	95.1	3.00E-53
Pseudocercospora pini-densiflorae	KI634606.1	38.37	47.8	0.002

Table S2M

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella musicola	LFZO01000086.1	80	59.7	5.00E-09
Pseudocercospora pini-densiflorae	AWYD01001655.1	81.4	66.2	3.00E-11

Table S2N

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000002.1	56.94	180	2.00E-47
Mycosphaerella eumusae	LFZN01000002.1	47.08	142	1.00E-34
Mycosphaerella eumusae	LFZN01000006.1	58.28	233	8.00E-66
Mycosphaerella eumusae	LFZN01000169.1	60.61	251	5.00E-72
Mycosphaerella eumusae	LFZN01000184.1	93.06	590	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	44.12	69.3	8.00E-11
Mycosphaerella eumusae	LFZN01000275.1	48.54	104	4.00E-31
Mycosphaerella eumusae	LFZN01000275.1	65.38	55.5	4.00E-31
Mycosphaerella musicola	LFZO01000012.1	45.25	73.6	3.00E-12
Mycosphaerella musicola	LFZO01000013.1	58.6	231	4.00E-65
Mycosphaerella musicola	LFZO01000086.1	92.72	593	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	45.88	68.2	2.00E-10
Mycosphaerella musicola	LFZO01000100.1	62.03	253	1.00E-72
Mycosphaerella musicola	LFZO01000127.1	51.96	110	2.00E-34
Mycosphaerella musicola	LFZO01000127.1	62.96	60.1	2.00E-34
Mycosphaerella musicola	LFZO01000174.1	44.95	117	2.00E-26
Mycosphaerella musicola	LFZO01000201.1	67.2	200	7.00E-63
Mycosphaerella musicola	LFZO01000201.1	75.86	66.2	7.00E-63
Mycosphaerella musicola	LFZO01000339.1	52.29	99.4	2.00E-20
Mycosphaerella musicola	LFZO01000339.1	54.87	73.9	2.00E-12
Pseudocercospora pini-densiflorae	AWYD01000740.1	50.25	112	3.00E-35
Pseudocercospora pini-densiflorae	AWYD01000740.1	64.2	60.8	3.00E-35
Pseudocercospora pini-densiflorae	AWYD01000770.1	63.85	267	2.00E-77
Pseudocercospora pini-densiflorae	AWYD01000893.1	58.44	244	7.00E-70
Pseudocercospora pini-densiflorae	AWYD01001276.1	71.43	223	1.00E-85
Pseudocercospora pini-densiflorae	AWYD01001276.1	79.57	118	1.00E-85
Pseudocercospora pini-densiflorae	AWYD01001424.1	57.37	233	5.00E-66
Pseudocercospora pini-densiflorae	AWYD01001655.1	97.67	436	8.00E-166
Pseudocercospora pini-densiflorae	AWYD01001655.1	98.84	173	8.00E-166
Pseudocercospora pini-densiflorae	AWYD01001655.1	44.71	65.5	1.00E-09
Pseudocercospora pini-densiflorae	AWYD01004150.1	53.95	164	3.00E-42
Pseudocercospora pini-densiflorae	KI633479.1	58	108	1.00E-38
Pseudocercospora pini-densiflorae	KI633479.1	46.3	77.4	1.00E-38
Pseudocercospora pini-densiflorae	KI633624.1	55.56	183	9.00E-49

Table S2O

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000009.1	60.59	110	7.00E-25
Mycosphaerella eumusae	LFZN01000013.1	50	75.5	2.00E-13
Mycosphaerella eumusae	LFZN01000050.1	47.21	74.7	5.00E-13
Mycosphaerella eumusae	LFZN01000050.1	44.33	35.4	2.7
Mycosphaerella eumusae	LFZN01000050.1	39.64	34.7	4.3
Mycosphaerella eumusae	LFZN01000052.1	52.41	75.9	2.00E-13
Mycosphaerella eumusae	LFZN01000052.1	42.79	55.5	6.00E-07
Mycosphaerella eumusae	LFZN01000154.1	50.73	83.2	7.00E-16
Mycosphaerella eumusae	LFZN01000175.1	53.89	81.3	4.00E-15
Mycosphaerella eumusae	LFZN01000184.1	94.9	330	4.00E-101
Mycosphaerella eumusae	LFZN01000241.1	55.72	96.3	3.00E-20
Mycosphaerella eumusae	LFZN01000253.1	66.02	100	1.00E-28
Mycosphaerella eumusae	LFZN01000253.1	76	38.1	1.00E-28
Mycosphaerella eumusae	LFZN01000253.1	85.71	32	1.00E-28
Mycosphaerella musicola	LFZO01000031.1	55.17	91.7	1.00E-18
Mycosphaerella musicola	LFZO01000086.1	94.9	328	2.00E-100
Mycosphaerella musicola	LFZO01000099.1	49.74	75.5	3.00E-13
Mycosphaerella musicola	LFZO01000148.1	50.97	84.7	3.00E-16
Mycosphaerella musicola	LFZO01000148.1	61.45	50.1	0.00004
Mycosphaerella musicola	LFZO01000320.1	50	82.8	9.00E-16
Mycosphaerella musicola	LFZO01000569.1	52.94	75.9	2.00E-13
Mycosphaerella musicola	LFZO01000569.1	48.25	45.8	0.0009
Pseudocercospora pini-densiflorae	AWYD01000477.1	50	75.9	2.00E-13
Pseudocercospora pini-densiflorae	AWYD01000477.1	55.38	69.3	3.00E-11
Pseudocercospora pini-densiflorae	AWYD01000739.1	50	73.9	8.00E-13
Pseudocercospora pini-densiflorae	AWYD01000902.1	62.31	111	3.00E-25
Pseudocercospora pini-densiflorae	AWYD01001655.1	96.94	341	7.00E-105
Pseudocercospora pini-densiflorae	AWYD01001767.1	51.71	82.8	9.00E-16
Pseudocercospora pini-densiflorae	AWYD01001767.1	56.99	43.9	0.004
Pseudocercospora pini-densiflorae	AWYD01002258.1	54.37	48.5	1.00E-12
Pseudocercospora pini-densiflorae	AWYD01002258.1	65.43	48.1	1.00E-12
Pseudocercospora pini-densiflorae	AWYD01003889.1	52.94	74.7	4.00E-13
Pseudocercospora pini-densiflorae	AWYD01003889.1	43.26	58.5	6.00E-08
Pseudocercospora pini-densiflorae	KI633671.1	54.5	81.3	3.00E-15
Pseudocercospora pini-densiflorae	KI633786.1	68.9	186	6.00E-51
Pseudocercospora pini-densiflorae	KI634513.1	52.85	79	2.00E-14
Pseudocercospora pini-densiflorae	KI634544.1	69.81	78.6	6.00E-30
Pseudocercospora pini-densiflorae	KI634544.1	70.93	76.3	6.00E-30
Pseudocercospora pini-densiflorae	KI634544.1	48.73	67.8	8.00E-11
Pseudocercospora pini-densiflorae	KI634544.1	39.84	41.6	0.023
Pseudocercospora pini-densiflorae	KI634544.1	48.45	39.3	0.15

Table S2P

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000014.1	70.73	171	2.00E-73
Mycosphaerella eumusae	LFZN01000014.1	58.48	130	2.00E-73
Mycosphaerella eumusae	LFZN01000036.1	81.25	286	3.00E-126
Mycosphaerella eumusae	LFZN01000036.1	70.45	192	3.00E-126
Mycosphaerella eumusae	LFZN01000091.1	45.68	99.4	3.00E-20
Mycosphaerella eumusae	LFZN01000107.1	66.07	315	6.00E-93
Mycosphaerella eumusae	LFZN01000141.1	49.36	108	4.00E-23
Mycosphaerella eumusae	LFZN01000184.1	97.38	682	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	42.31	59.7	2.00E-07
Mycosphaerella eumusae	LFZN01000184.1	100	46.2	0.003
Mycosphaerella musicola	LFZO01000022.1	65.87	305	2.00E-89
Mycosphaerella musicola	LFZO01000052.1	69.51	169	4.00E-73
Mycosphaerella musicola	LFZO01000052.1	59.06	131	4.00E-73
Mycosphaerella musicola	LFZO01000086.1	97.38	683	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	41.76	59.3	2.00E-07
Mycosphaerella musicola	LFZO01000086.1	100	46.2	0.003
Mycosphaerella musicola	LFZO01000209.1	43.79	102	5.00E-21
Mycosphaerella musicola	LFZO01000434.1	81.25	285	6.00E-83
Mycosphaerella musicola	LFZO01000434.1	70.86	196	3.00E-52
Pseudocercospora pini-densiflorae	AWYD01000646.1	54.31	77	2.00E-18
Pseudocercospora pini-densiflorae	AWYD01000646.1	53.57	40.4	2.00E-18
Pseudocercospora pini-densiflorae	AWYD01001655.1	97.63	674	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	42.86	59.7	2.00E-07
Pseudocercospora pini-densiflorae	AWYD01001655.1	100	46.6	0.003
Pseudocercospora pini-densiflorae	AWYD01002309.1	43.79	104	1.00E-21
Pseudocercospora pini-densiflorae	AWYD01004107.1	48.36	93.6	2.00E-18
Pseudocercospora pini-densiflorae	KI633613.1	43.45	101	6.00E-21
Pseudocercospora pini-densiflorae	KI633906.1	49.36	108	3.00E-23
Pseudocercospora pini-densiflorae	KI634333.1	80.73	284	2.00E-82
Pseudocercospora pini-densiflorae	KI634333.1	67.35	191	2.00E-50
Pseudocercospora pini-densiflorae	KI634401.1	57.25	288	6.00E-84
Pseudocercospora pini-densiflorae	KI634401.1	34.09	36.6	3.1

Table S2Q

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000006.1	52.28	186	7.00E-48
Mycosphaerella eumusae	LFZN01000011.1	49.63	63.5	2.00E-08
Mycosphaerella eumusae	LFZN01000068.1	57.97	96.7	2.00E-34
Mycosphaerella eumusae	LFZN01000068.1	47.06	75.5	2.00E-34
Mycosphaerella eumusae	LFZN01000184.1	85.75	706	0.00E+00
Mycosphaerella musicola	LFZO01000036.1	47.76	63.5	2.00E-08
Mycosphaerella musicola	LFZO01000039.1	48.42	189	3.00E-49
Mycosphaerella musicola	LFZO01000086.1	84.94	690	0.00E+00
Mycosphaerella musicola	LFZO01000133.1	57.25	90.5	1.00E-33
Mycosphaerella musicola	LFZO01000133.1	45.05	79.3	1.00E-33
Pseudocercospora pini-densiflorae	AWYD01000134.1	57.25	95.9	2.00E-35
Pseudocercospora pini-densiflorae	AWYD01000134.1	43.73	79.3	2.00E-35
Pseudocercospora pini-densiflorae	AWYD01001655.1	85.04	704	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01002647.1	47.76	63.9	2.00E-08
Pseudocercospora pini-densiflorae	AWYD01002787.1	44.82	136	1.00E-31
Pseudocercospora pini-densiflorae	KI633441.1	50.51	189	4.00E-49
Pseudocercospora pini-densiflorae	KI633846.1	60.81	62.8	3.00E-08

Table S2R

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000184.1	93.94	1173	0.00E+00
Mycosphaerella eumusae	LFZN01000184.1	77.27	57.4	3.00E-06
Mycosphaerella musicola	LFZO01000086.1	96.59	1109	0.00E+00
Mycosphaerella musicola	LFZO01000086.1	91.18	62	1.00E-07
Mycosphaerella musicola	LFZO01000086.1	65.85	45.1	0.02
Pseudocercospora pini-densiflorae	AWYD01001655.1	95.99	1133	0.00E+00
Pseudocercospora pini-densiflorae	AWYD01001655.1	82.05	65.5	1.00E-08
Pseudocercospora pini-densiflorae	AWYD01001655.1	77.27	57.4	4.00E-06

Table S2S

Subject species	Subject scaffold	% Sim	Bit	E-value
Mycosphaerella eumusae	LFZN01000184.1	89.04	226	3.00E-64
Mycosphaerella eumusae	LFZN01000184.1	89	143	1.00E-35
Mycosphaerella musicola	LFZO01000086.1	92.86	225	3.00E-95
Mycosphaerella musicola	LFZO01000086.1	88.12	148	3.00E-95
Pseudocercospora pini-densiflorae	AWYD01001655.1	90.34	229	2.00E-95
Pseudocercospora pini-densiflorae	AWYD01001655.1	88.12	144	2.00E-95

Table S3. Homologs in *M. fijiensis* of protein sequences encoded by the monodictyphenone-producing PKS gene cluster in *A. nidulans*.
For each protein sequence encoded by the monodictyphenone-producing PKS gene cluster in *A. nidulans*, a blastp search was done against sequences from *M. fijiensis* isolate CIRAD86, using NCBI's non-redundant protein sequences database. The Table indicates the NCBI accession for the *A. nidulans* protein sequence and a description of its putative function, as well as a description of the blastp hits identified from *M. fijiensis*. For each hit identified, the Table indicates the NCBI accession of the hit, the bitscore (Bit), the E-value, and the percent protein similarity (% Sim). Sequences present in the putative *M. fijiensis PKS8-1* gene cluster are highlighted yellow, and sequences present in the putative *M. fijiensis* melanin gene cluster are highlighted blue.

Protein encoded by *A. nidulans*		**Description of blastp hits from *M. fijiensis* sequences**			
Accession	**Description**	**Bit**	**E-value**	**% Sim**	**Accession**
CBF90087.1	Telomere repeat binding factor	242	3.00E-69	71%	XP_007930812.1
CBF90088.1	MdpL - Hypothetical/ DUF4243	307	4.00E-101	55%	XP_007928879.1
CBF90090.1	MdpK - Short-chain dehydrogenase	28.5	1.6	41%	XP_007924882.1
CBF90092.1	MdpJ - Glutathione S-transferase	134	1.00E-37	57%	XP_007922713.1
CBF90094.1	MdpI - Acyl CoA-synthase	310	2.00E-98	51%	XP_007922408.1
CBF90095.1	MdpH - Hypothetical/ DUF1772	207	1.00E-66	76%	XP_007929611.1
CBF90097.1	MdpG - polyketide synthase	2230	0.00E+00	74%	XP_007929626.1
CBF90099.1	MdpF - Beta-lactamase-like	397	9.00E-140	74%	XP_007929883.1
CBF90101.1	MdpE - Zn(II)2Cys6 Transcription factor	178	1.00E-51	45%	XP_007929880.1
		56.6	3.00E-09	77%	XP_007930446.1
CBF90103.1	MdpD - Monooxygenase	216	7.00E-65	47%	XP_007920821.1
		192	8.00E-56	47%	XP_007922492.1
		181	6.00E-52	46%	XP_007923528.1
		157	2.00E-43	65%	XP_007920532.1
CBF90105.1	MdpC - THN reductase-like	348	2.00E-122	80%	XP_007931484.1
CBF90107.1	MdpB - Scytalone dehydratase-like	202	1.00E-66	75%	XP_007925880.1
CBF90109.1	MdpA - Transcriptional regulator	292	1.00E-95	58%	XP_007929877.1
CBF90111.1	Histidine phosphatase	221	9.00E-65	45%	XP_007930823.1

Table S4. Primer sequences and PCR conditions used for semi-quantitative RT-PCR assays and PCR assays for genotyping.

For each PCR assay shown in Figures 4, 9 and S1, the table indicates the gene assayed, the forward and reverse primer sequences, the annealing temperature (Ann), extension time (Ext), number of cycles (Cyc), and product size in bp. For RT-PCR primers whose products span intron(s), the expected product sizes from gDNA versus cDNA are indicated as: Product size from gDNA/Product size from cDNA. * Product size for primers that do not span intron(s).

Gene	Primer sequences	Ann	Ext	Cyc	Size
β-tubulin	cagctcgagcgcatgaacg ggtgcgaaaccgaccatgaag	59	1:00	36	745 *
PKS2-1	tgatggctggcggacgatgc cctggctgatggtcttcaga	57	1:00	43	780/674
PKS7-1	gccgtctatgactattgaca tggctgttgattacgctctt	52	1:00	43	774/660
PKS8-1	caggacgcatcaactacttc ctcggcggagtggttagttc	54	0:45	43	511/362
PKS8-2	caggaagattgacgaaaggc catagtgttggatcatgtcg	52	1:00	43	574/524
Hybrid8-3	ctcgccgaacttgatggaga tacaggcatcggaacgacgagg	59	0:45	43	585/495
PKS8-4	tatgctctcacctgacgg agattcatagctctcgat	47	1:00	43	886/542
PKS10-2	cttcgtgtttcatagggaac catcaacttcaatcggatcg	51	1:00	43	809/615
PKS2-1 transporter	ccctagtcatctggcaggaa caaggagaagacgggcagta	57	0:30	43	379/311
PKS8-1 transporter	gcgttcatcgtcggcagagc gtggcactgttgcggtctttg	60	0:40	43	549/499
PKS8-1 AflJ-like	caatggtccgcatacagccg ggttgccaggataagagtcg	56	0:30	43	478/420
PKS8-1 dehydrogenase	gcaagacagccatcgttactg ggtaggactgctcgtgtgag	57	0:30	43	370 *
PKS8-1 cytochrome P450	caacattccagccttcacacc gtaacctttggcagacgcttg	57	0:30	43	425/372
PKS8-2 transporter	cagcactcagcatcatggtc gagccgagcaggaagagcag	57	0:30	43	295/229
Hybrid8-3 transporter	catgatgctgggagccttc gaccatgagctgctttctcc	57	1:00	43	872/746
Hybrid8-3 enoyl reductase	cgttcttgtacgagtgcattg caagcacagaccacttgcat	55	0:40	43	595/472
trpC promoter to aflR-like	ccactctccagcggcttctc ggcaaaggaatagagtagat	50	2:00	33	1900
Hygromycin resistance cassette	ggatgcctccgctcgaagta cgttaactgatattgaaggagca	53	1:15	33	1126

Table S5. Primer sequences and reference genes used for RT-qPCR assays.
For each gene assayed in RT-qPCR assays from Figures 10, 13 and 14, the table indicates the gene assayed, the forward and reverse primer sequences, two reference genes for each gene of interest (reference gene listed first was used for Figures 10, 13, and 14), and PCR product sizes. For primers whose products span intron(s), the expected product sizes from gDNA versus cDNA are indicated as: Product size from gDNA/Product size from cDNA. * Product size for primers that do not span intron(s).

Gene	Primer sequences	Ref	Size
PKS8-1	ggcactatggaacggagatgtc	PDA1	337/265
	ctcggcggagtggttagttc	Actin	
AflR-like	gcctcaccgccaaacatcca	PDA1	263 *
	gctgagtctgctccttcgagaca	β-tubulin	
AflJ-like	ctcaaaggcagcattcatgctcg	β-tubulin	175 *
	ccagcagtgacaagtctgtgac	TFC1	
MFS transporter	cgaaagacgaggctccagtc	β-tubulin	177 *
	atccagtaacacctgcggt	UBC6	
Beta-lactamase-like	cgagaccaacaagccatccaag	β-tubulin	190 *
	ccatgaaggttccgaggtcttc	TFC1	
Cytochrome P450	gcaccactcgttgaccaactc	PDA1	201 *
	ggtgttgtaacagcagagaccac	UBC6	
Ketoreductase 1	ggagatgcagtcgtgatcgg	Actin	89 *
	cctcctcgaaagcctggac	β-tubulin	
Ketoreductase 2	gctggaggatgtcgtatgtgg	Actin	100 *
	gtcgcaagtgcttcctctatcc	β-tubulin	
Dehydrogenase	ggacagctccaaggcaagac	β-tubulin	102 *
	cgtggttctcgacgagcaatg	UBC6	
Actin	cttgactccggtgacggtgtcactc	N/A	277 *
	cgtcaggaagctcgtaggacttctc	N/A	
PDA1	gagcagtcacgttgacaagagc	N/A	399/287
	cctgtttggtgattgcgtgc	N/A	
β-tubulin	tggcacgtctgatctccagc	N/A	204/148
	gcggaagagctgaccgaatgg	N/A	

REFERENCES

1. Frison, E. A. *et al.* The global *Musa* genomic consortium: a boost for banana improvement. in *Banana improvement: cellular, molecular biology, and induced mutations* 341–349 (Science Publishers, Inc., 2004).
2. *Banana market review 2013-2014*. (Food and Agriculture Organization of the United Nations, 2015).
3. Shepherd, K. Banana breeding - past and present. *Acta Hortic.* 37–44 (1987). doi:10.17660/ActaHortic.1987.196.3
4. Daniells, J. *et al.* Planet of the Cavendish - understanding the domination. *Acta Hortic.* 219–224 (2013). doi:10.17660/ActaHortic.2013.986.23
5. Churchill, A. C. L. *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Mol. Plant Pathol.* **12,** 307–328 (2011).
6. Ploetz, R. Black Sigatoka of banana: The most important disease of a most important fruit. *The Plant Health Instructor* (2001). Available at: http://www.apsnet.org/publications/apsnetfeatures/Pages/blacksigatoka.aspx. (Accessed: 23rd September 2016)
7. Marín, D. H., Romero, R. A., Guzmán, M. & Sutton, T. B. Black Sigatoka: An increasing threat to banana cultivation. *Plant Dis.* **87,** 208–222 (2003).
8. Hermanto, C., Opina, O. S. & Natural, M. P. Assessment of fungicide resistance of a population of *Mycosphaerella* spp. on Señorita banana variety (Sucrier group). *Tree For. Sci. Biotechnol.* **4,** 85–90 (2010).

9. Stover, R. Distribution and probable origin of *Mycosphaerella fijiensis* in Southeast-Asia. *Trop. Agric.* **55,** 65–68 (1978).

10. Amil, A. F., Heaney, S. P., Stanger, C. & Shaw, M. W. Dynamics of QoI sensitivity in *Mycosphaerella fijiensis* in Costa Rica during 2000 to 2003. *Phytopathology* **97,** 1451–1457 (2007).

11. Jacome, L. H., Schuh, W. & Stevenson, R. E. Effect of temperature and relative humidity on germination and germ tube development of *Mycosphaerella fijiensis* var. *difformis*. *Phytopathology* **81,** 1480–1485 (1991).

12. Balint-Kurti, P. & Churchill, A. C. L. in *Banana Improvement: Cellular, Molecular Biology, and Induced Mutations* (eds. Jain, S. M. & Swennen, R.) 147–159 (Science Publishers, Inc., 2004).

13. Balint-Kurti, P. J., May, G. D. & Churchill, A. C. Development of a transformation system for *Mycosphaerella* pathogens of banana: a tool for the study of host/pathogen interactions. *FEMS Microbiol. Lett.* **195,** 9–15 (2001).

14. Beveraggi, A., Mourichon, X. & Sallé, G. Étude comparée des premières étapes de l'infection chez des bananiers sensibles et résistants infectés par le *Cercospora fijiensis* (*Mycosphaerella fijiensis*) agent responsable de la maladie des raies noires. *Can. J. Bot.* **73,** 1328–1337 (1995).

15. Liberato, J. *et al.* Black sigatoka of banana. *PaDIL* (2016). Available at: http://www.padil.gov.au/pests-and-diseases/pest/main/136601. (Accessed: 13th September 2016)

16. Washington, J. R., Cruz, J., Lopez, F. & Fajardo, M. Infection studies of *Mycosphaerella fijiensis* on banana and the control of black Sigatoka with chlorothalonil. *Plant Dis.* **82,** 1185–1190 (1998).
17. Meredith, D., Lawrence, J. & Firman, I. Ascospore release and dispersal in black leaf streak disease of bananas (*Mycosphaerella fijiensis*). *Trans. Br. Mycol. Soc.* **60,** 547–554 (1973).
18. Abadie, C. *et al.* New approaches to select cultivars of banana with durable resistance to *Mycosphaerella* leaf spot diseases. *Acta Hortic.* **828,** 171–178 (2009).
19. *2014 Banana FRAC Meeting Minutes*. (Banana Working Group Fungicide Resistance Action Committee, 2014).
20. Canas-Gutierrez, G. P. *et al.* Analysis of the CYP51 gene and encoded protein in propiconazole-resistant isolates of *Mycosphaerella fijiensis*. *Pest Manag. Sci.* **65,** 892–899 (2009).
21. Craenen, K. *Black Sigatoka Disease of Banana and Plantain: A Reference Manual*. (IITA, 1998).
22. Kablan, L., Lagauche, A., Delvaux, B. & Legreve, A. Silicon reduces black Sigatoka development in banana. *Plant Dis.* **96,** 273–278 (2012).
23. Mobambo, K., Zuofa, K., Gauhl, F., Adeniji, M. & Pasberggauhl, C. Effect of soil fertility on host response to black leaf streak of plantain (*Musa* spp., AAB group) under traditional farming systems in Southeastern Nigeria. *Int. J. Pest Manag.* **40,** 75–80 (1994).

24. Orozco-Santos, M. *et al.* Efecto de la urea aplicada a hojas afectadas por Sigatoka negra para reducir el nivel de inóculo en banano gran enano. in (2013). doi:10.13140/RG.2.1.1882.5760

25. Coyne, D., Turoop, L., Dubois, T., Nsubuga, E. W. & Kahangi, E. Endophyte-enhanced banana tissue culture: technology transfer through public-private partnerships in Kenya and Uganda. *ATDF J.* **3,** 18–24 (2012).

26. Vézina, A. & Dubois, T. Planting material. *ProMusa* Available at: http://www.promusa.org/Planting+material. (Accessed: 13th September 2016)

27. Chandler, J. V., Abruna, F. & Silva, S. Effect of shade trees on yields of five crops in the humid mountain region of Puerto Rico. *J Agr Univ P. R.* **50,** 218–225 (1966).

28. Dold, C., Staver, C., Pocasangre, L. & Heller, J. *Musa* in shaded perennial crops - response to light interception. in *Conference on International Research on Food Security, Natural Resource Management, and Rural Development* (2008).

29. *Evaluating Bananas: A Global Partnership : Results of IMTP Phase II*. (International Network for the Improvement of Banana and Plantain, 2000).

30. Ortiz, R. & Vuylsteke, D. Inheritance of black sigatoka disease resistance in plantain-banana (*Musa* spp.) hybrids. *Theor. Appl. Genet.* **89,** 146–152 (1993).

31. Orjeda, G. *Post-harvest Characteristics of Black Sigatoka Resistant Banana, Cooking Banana and Plantain Hybrids*. **Technical guidelines INIBAP,** (Bioversity International, 1998).

32. Daniells, J. W. in *International Symposium on Recent Advances in Banana Crop Protection for Sustainable Production and Improved Livelihoods* (eds. Jones, D. & VanDenBergh, I.) **828,** 411–416 (Int Soc Horticultural Science, 2009).
33. Manzo-Sanchez, G. *et al.* Construction of a genetic linkage map of the fungal pathogen of banana *Mycosphaerella fijiensis*, causal agent of black leaf streak disease. *Curr. Genet.* **53,** 299–311 (2008).
34. D'Hont, A. *et al.* The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **488,** 213–217 (2012).
35. Raboin, L. M. *et al.* Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Mol. Breed.* **16,** 333–341 (2005).
36. Ríos, R. O. *Plant Breeding in the Omics Era*. (Springer, 2015).
37. Vilarinhos, A. D. *et al.* Construction and characterization of a bacterial artificial chromosome library of banana (*Musa acuminata* Colla). *Theor. Appl. Genet.* **106,** 1102–1106 (2003).
38. Davey, M. W. *et al.* A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genomics* **14,** 683 (2013).
39. Rodriguez, H. A., Rodriguez-Arango, E., Morales, J. G., Kema, G. & Arango, R. E. Defense gene expression associated with biotrophic phase of *Mycosphaerella fijiensis* M. Morelet infection in banana. *Plant Dis.* PDIS-08-15-0950-RE (2016). doi:10.1094/pdis-08-15-0950-re

40. Timm, E. S. *et al.* Identification of differentially-expressed genes in response to *Mycosphaerella fijiensis* in the resistant *Musa* accession 'Calcutta-4' using suppression subtractive hybridization. *PLOS ONE* **11,** e0160083 (2016).

41. Portal, O. *et al.* Analysis of expressed sequence tags derived from a compatible *Mycosphaerella fijiensis*–banana interaction. *Plant Cell Rep.* **30,** 913–928 (2011).

42. Passos, M. A. N. *et al.* Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development. *BMC Genomics* **14,** 78 (2013).

43. Uma, S. *et al.* Identification of *Mycosphaerella eumusae* responsive unique genes/transcripts from a resistant banana cultivar. *Acta Hortic.* **1114,** 111–118 (2016).

44. Cho, Y., Hou, S. & Zhong, S. Analysis of expressed sequence tags from the fungal banana pathogen *Mycosphaerella fijiensis*. *Open Mycol. J.* **2,** 61–73 (2008).

45. Noar, R. D. & Daub, M. E. Bioinformatics prediction of polyketide synthase gene clusters from *Mycosphaerella fijiensis*. *PLOS ONE* **11,** e0158471 (2016).

46. Noar, R. D. & Daub, M. E. Transcriptome sequencing of *Mycosphaerella fijiensis* during association with *Musa acuminata* reveals candidate pathogenicity genes. *BMC Genomics* **17,** 690 (2016).

47. Khaldi, N. *et al.* SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **47,** 736–741 (2010).

48. Covert, S. F. Supernumerary chromosomes in filamentous fungi. *Curr. Genet.* **33,** 311–319 (1998).

49. Ahn, J. H. & Walton, J. D. Chromosomal organization of TOX2, a complex locus controlling host-selective toxin biosynthesis in *Cochliobolus carbonum*. *Plant Cell* **8,** 887–897 (1996).

50. Han, Y. N., Liu, X. G., Benny, U., Kistler, H. C. & VanEtten, H. D. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. *Plant J.* **25,** 305–314 (2001).

51. Johnson, L. J. *et al.* Spontaneous loss of a conditionally dispensable chromosome from the *Alternaria alternata* apple pathotype leads to loss of toxin production and pathogenicity. *Curr. Genet.* **40,** 65–72 (2001).

52. Temporini, E. D. & VanEtten, H. D. An analysis of the phylogenetic distribution of the pea pathogenicity genes of *Nectria haematococca* MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: *Neocosmospora boniensis*. *Curr. Genet.* **46,** 29–36 (2004).

53. Wittenberg, A. H. J. *et al.* Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*. *PLOS ONE* **4,** (2009).

54. Ohm, R. A. *et al.* Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLOS Pathog.* **8,** e1003037 (2012).

55. Goodwin, S. B. in *Genomics of plant-associated fungi: monocot pathogens* (ed. R. A. Dean) (Springer-Verlag Berlin Heidelberg, 2014).

56. Heiser, I., Sachs, E. & Liebermann, B. Photodynamic oxygen activation by rubellin D, a phytotoxin produced by *Ramularia collo-cygni* (Sutton et Waller). *Physiol. Mol. Plant Pathol.* **62,** 29–36 (2003).

57. Choquer, M. *et al.* The *CTB1* gene encoding a fungal polyketide synthase is required for cercosporin biosynthesis and fungal virulence of *Cercospora nicotianae*. *Mol. Plant. Microbe Interact.* **18,** 468–476 (2005).

58. Miethbauer, S. *et al.* Uredinorubellins and caeruleoramularin, photodynamically active anthraquinone derivatives produced by two species of the genus *Ramularia*. *J. Nat. Prod.* **71,** 1371–1375 (2008).

59. Crous, P. W. *et al.* Unravelling *Mycosphaerella*: do you believe in genera? *Persoonia* **23,** 99–118 (2009).

60. Upadhyay, R. K., Strobel, G. A., Coval, S. J. & Clardy, J. Fijiensin, the 1st phytotoxin from *Mycosphaerella fijiensis*, the causative agent of black Sigatoka disease. *Experientia* **46,** 982–984 (1990).

61. Stierle, A. A., Upadhyay, R., Hershenhorn, J., Strobel, G. A. & Molina, G. The phytotoxins of *Mycosphaerella fijiensis*, the causative agent of black Sigatoka disease of bananas and plantains. *Experientia* **47,** 853–859 (1991).

62. Arai, M. *et al.* Funicone-related compounds, potentiators of antifungal miconazole activity, produced by *Talaromyces flavus* FKI-0076. *J. Antibiot. (Tokyo)* **55,** 172–180 (2002).

63. Dethoup, T. *et al.* Merodrimanes and other constituents from *Talaromyces thailandiasis*. *J. Nat. Prod.* **70,** 1200–1202 (2007).

64. Fuska, J., Fuskova, A. & Nemec, P. Vermistatin, an antibiotic with cytotoxic effects, produced from *Penicillium vermiculatum*. *Biologia (Bratisl.)* **34,** 735–739 (1979).

65. Fuska, J., Uhrin, D., Proksa, B., Voticky, Z. & Ruppeldt, J. The structure of vermistatin, a new metabolite from *Penicillium vermiculatum*. *J. Antibiot. (Tokyo)* **39,** 1605–1608 (1986).

66. Murtaza, N., Husain, S. A., Sarfaraz, T. B., Sultana, N. & Faizi, S. Isolation and identification of vermistatin, ergosterol, stearic acid and mannitol, metabolic products of *Penicillium verruculosum*. *Planta Med.* **63,** 191–191 (1997).

67. Komai, S. *et al.* New vermistatin derivatives isolated from *Penicillium simplicissimum*. *Heterocycles* **65,** 2771–2776 (2005).

68. Liu, Z. M. *et al.* Eurothiocin A and B, sulfur-containing benzofurans from a soft coral-derived fungus *Eurotium rubrum* SH-823. *Mar. Drugs* **12,** 3669–3680 (2014).

69. Xia, X. K. *et al.* Structural and biological properties of vermistatin and two new vermistatin derivatives isolated from the marine-mangrove endophytic fungus *Guignardia* sp No. 4382. *Helv. Chim. Acta* **90,** 1925–1931 (2007).

70. Sassa, T., Nukina, M. & Suzuki, Y. Deoxyfunicone, a new y-pyrone metabolite from a resorcylide-producing fungus (*Penicillium* sp.). *Agric. Biol. Chem.* **55,** 2415–2416 (1991).

71. Nicoletti, R., Manzo, E. & Ciavatta, M. L. Occurence and bioactivities of funicone-related compounds. *Int. J. Mol. Sci.* **10,** 1430–1444 (2009).

72. Bell, A. A., Stipanovic, R. D. & Puhalla, J. E. Pentaketide metabolites of *Verticillium dahliae*: Identification of (+)-scytalone as a natural precursor to melanin. *Tetrahedron* **32,** 1353–1356 (1976).

73. Stipanovic, R. D. & Bell, A. A. Pentaketide metabolites of *Verticillium dahliae*. 2. Accumulation of naphthol derivatives by aberrant-melanin mutant *BRM-2*. *Mycologia* **69,** 164–172 (1977).

74. Hoss, R., Helbig, J. & Bochow, H. Function of host and fungal metabolites in resistance response of banana and plantain in the black Sigatoka disease pathosystem (*Musa* spp. - *Mycosphaerella fijiensis*). *J. Phytopathol.* **148,** 387–394 (2000).

75. Babula, P., Adam, V., Havel, L. & Kizek, R. Noteworthy secondary metabolites naphthoquinones – their occurrence, pharmacological properties and analysis. *Curr. Pharm. Anal.* **5,** 47–68 (2009).

76. Busogoro, J. P. *et al.* Experimental evidence for the action of *M. fijiensis* toxins on banana photosynthetic apparatus. *Banana Improv. Cell. Mol. Biol. Induc. Mutat.* 161–170 (2004).

77. Busogoro, J. P. *et al. Analysis of the mechanisms of action of* Mycosphaerella fijiensis *toxins during the development of black leaf streak disease*. (2004).

78. Howard, R. J. & Ferrari, M. A. Role of melanin in appressorium function. *Exp. Mycol.* **13,** 403–418 (1989).

79. Steiner, U. & Oerke, E. C. Localized melanization of appressoria is required for pathogenicity of *Venturia inaequalis*. *Phytopathology* **97,** 1222–1230 (2007).

80. Brush, L. & Money, N. P. Invasive hyphal growth in *Wangiella dermatitidis* is induced by stab inoculation and shows dependence upon melanin biosynthesis. *Fungal Genet. Biol.* **28,** 190–200 (1999).

81. Nosanchuk, J. D. & Casadevall, A. Impact of melanin on microbial virulence and clinical resistance to antimicrobial compounds. *Antimicrob. Agents Chemother.* **50,** 3519–3528 (2006).

82. Wang, Y. L. & Casadevall, A. Susceptibility of melanized and nonmelanized *Cryptococcus neoformans* to nitrogen-derived and oxygen-derived oxidants. *Infect. Immun.* **62,** 3004–3007 (1994).

83. Morel, J.-B. & Dangl, J. L. The hypersensitive response and the induction of cell death in plants. *Publ. Online 20 Novemb. 1997 Doi101038sjcdd4400309* **4,** 671–683 (1997).

84. Freitas, M., Lima, J. L. F. C. & Fernandes, E. Optical probes for detection and quantification of neutrophils' oxidative burst. A review. *Anal. Chim. Acta* **649,** 8–23 (2009).

85. Korytowski, W., Pilas, B., Sarna, T. & Kalyanaraman, B. Photoinduced generation of hydrogen peroxide and hydroxyl radicals in melanins. *Photochem. Photobiol.* **45,** 185–190 (1987).

86. Chiarelli-Neto, O. *et al.* Generation and suppression of singlet oxygen in hair by photosensitization of melanin. *Free Radic. Biol. Med.* **51,** 1195–1202 (2011).

87. Beltran-Garcia, M. J. *et al.* Singlet molecular oxygen generation by light-activated DHN-melanin of the fungal pathogen *Mycosphaerella fijiensis* in black Sigatoka disease of bananas. *PLOS ONE* **9,** e91616 (2014).

88. Cruz-Cruz, C. A., Garcia-Sosa, K., Escalante-Erosa, F. & Pena-Rodriguez, L. M. Production of hydrophilic phytotoxins by *Mycosphaerella fijiensis*. *J. Gen. Plant Pathol.* **75,** 191–195 (2009).

89. Cruz-Cruz, C. A., Garcia-Sosa, K., Escalante-Erosa, F. & Pena-Rodriguez, L. M. Physiological effects of the hydrophilic phytotoxins produced by *Mycosphaerella fijiensis*, the causal agent of black Sigatoka in banana plants. *J. Gen. Plant Pathol.* **77,** 93–100 (2011).

90. Krasnoff, S. B., Gibson, D. M., Belofsky, G. N., Gloer, K. B. & Gloer, J. B. New destruxins from the entomopathogenic fungus *Aschersonia* sp. *J. Nat. Prod.* **59,** 485–489 (1996).

91. Lira, S. P. *et al.* New destruxins from the marine-derived fungus *Beauveria felina*. *J Antibiot* **59,** 553–563 (2006).

92. Wang, B., Kang, Q., Lu, Y., Bai, L. & Wang, C. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. *Proc. Natl. Acad. Sci.* **109,** 1287–1292 (2012).

93. Buchwaldt, L. & Jensen, J. S. HPLC purification of destruxins produced by *Alternaria brassicae* in culture and leaves of *Brassica napus*. *Phytochemistry* **30,** 2311–2316 (1991).

94. Buchwaldt, L. & Green, H. Phytotoxicity of destruxin B and its possible role in the pathogenesis of *Alternaria brassicae*. *Plant Pathol.* **41,** 55–63 (1992).

95. Venkatasubbaiah, P., Tisserat, N. A. & Chilton, W. S. Metabolites of *Ophiosphaerella herpotricha*, a cause of spring dead spot of bermudagrass. *Mycopathologia* **128,** 155–159 (1994).

96. Parada, R. Y., Oka, K., Yamagishi, D., Kodama, M. & Otani, H. Destruxin B produced by *Alternaria brassicae* does not induce accessibility of host plants to fungal invasion. *Physiol. Mol. Plant Pathol.* **71,** 48–54 (2007).

97. de Boer, A. H. & de Vries-van Leeuwen, I. J. Fusicoccanes: diterpenes with surprising biological functions. *Trends Plant Sci.* **17,** 360–368 (2012).

98. Nowrousian, M. A novel polyketide biosynthesis gene cluster is involved in fruiting body morphogenesis in the filamentous fungi *Sordaria macrospora* and *Neurospora crassa*. *Curr. Genet.* **55,** 185–198 (2009).

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

100. McGrann, G. R. D. *et al.* The genome of the emerging barley pathogen *Ramularia collo-cygni*. *BMC Genomics* **17,** 584 (2016).

101. Daub, M. E., Herrero, S. & Chung, K.-R. Photoactivated perylenequinone toxins in fungal pathogenesis of plants. *FEMS Microbiol. Lett.* **252,** 197–206 (2005).

102. Stergiopoulos, I. *et al.* Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 7610–7615 (2010).

103. Sanchez-Vallet, A. *et al.* Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. *Elife* **2,** e00790 (2013).

104. Bolton, M. D. *et al.* The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Mol. Microbiol.* **69,** 119–136 (2008).

105. Stergiopoulos, I. & de Wit, P. J. Fungal effector proteins. *Annu Rev Phytopathol* **47,** 233–63 (2009).

106. Tan, K., Oliver, R. P., Solomon, P. S. & Moffat, C. S. Proteinaceous necrotrophic effectors in fungal virulence. *Funct. Plant Biol.* **37,** 907–912 (2010).

107. van den Burg, H. A. *et al.* Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent *Cf-4*-mediated resistance, but retain their chitin binding ability. *J. Biol. Chem.* **278,** 27340–27346 (2003).

108. Giraldo, M. C. & Valent, B. Filamentous plant pathogen effectors in action. *Nat Rev Micro* **11,** 800–814 (2013).

109. Bowman, S. M. & Free, S. J. The structure and synthesis of the fungal cell wall. *BioEssays* **28,** 799–808 (2006).

110. Kantun-Moreno, N. *et al.* Genome-wide in silico identification of GPI proteins in *Mycosphaerella fijiensis* and transcriptional analysis of two GPI-anchored beta-1,3-glucanosyltransferases. *Mycologia* **105,** 285–296 (2013).

111. Kitagaki, H., Wu, H., Shimoi, H. & Ito, K. Two homologous genes, *DCW1* (*YKL046c*) and *DFG5*, are essential for cell growth and encode glycosylphosphatidylinositol (GPI)-anchored membrane proteins required for cell wall biogenesis in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **46,** 1011–1022 (2002).

112. Caracuel, Z., Martínez-Rocha, A. L., Di Pietro, A., Madrid, M. P. & Roncero, M. I. G. *Fusarium oxysporum gas1* encodes a putative β-1, 3-glucanosyltransferase required for virulence on tomato plants. *Mol. Plant. Microbe Interact.* **18,** 1140–1147 (2005).

113. Klis, F. M., Sosinska, G. J., Groot, P. W. J. de & Brul, S. Covalently linked cell wall proteins of *Candida albicans* and their role in fitness and virulence. *FEMS Yeast Res.* **9,** 1013–1028 (2009).

114. Sagaram, U. S., Shaw, B. D. & Shim, W.-B. *Fusarium verticillioides GAP1*, a gene encoding a putative glycolipid-anchored surface protein, participates in conidiation and cell wall structure but not virulence. *Microbiology* **153,** 2850–2861 (2007).

115. Stergiopoulos, I., Zwiers, L. H. & De Waard, M. A. The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Mol. Plant. Microbe Interact.* **16,** 689–698 (2003).

116. Couoh-Uicab, Y., I.Islas-Flores, et al. Cloning, in silico structural characterization and expression analysis of *MfAtr4*, an ABC transporter from the banana pathogen *Mycosphaerella fijiensis*. *Afr. J. Biotechnol.* **11,** 54–79 (2012).

117. Arinaitwe, G., Remy, S., Strosse, H., Swennen, R. & Sági, L. in *Banana improvement: cellular, molecular biology, and induced mutations.* (eds. Mohan, J. S. & Swennen, R.) 351–357 (Science Publishers, Inc., 2004).

118. Chakrabarti, A., Ganapathi, T. R., Mukherjee, P. K. & Bapat, V. A. MSI-99, a magainin analogue, imparts enhanced disease resistance in transgenic tobacco and banana. *Planta* **216,** 587–596 (2003).

119. Remy, S. Genetic transformation of banana (*Musa* spp.) for disease resistance by expression of antimicrobial proteins. (KUL, 2000).

120. Remy, S., Deconinck, I., Swennen, R. & Sagi, L. Development of a leaf disc assay to assess fungus tolerance in banana. in *Proceedings of the 9th European Congress on Biotechnology* ***/1/5-***/5/5 (1999).

121. Remy, S., Kovacs, G., Swennen, R. & Panis, B. Genetically modified bananas: past, present and future. *Acta Hortic.* **974,** 71–80 (2013).

122. Kovacs, G. *et al.* Expression of a rice chitinase gene in transgenic banana ('Gros Michel', AAA genome group) confers resistance to black leaf streak disease. *Transgenic Res.* **22,** 117–130 (2013).

123. Keller, H. *et al.* Pathogen-induced elicitin production in transgenic tobacco generates a hypersensitive response and nonspecific disease resistance. *Plant Cell* **11,** 223–235 (1999).

124. Nowara, D. *et al.* HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. *Plant Cell Online* **22,** 3130–3141 (2010).

125. Yin, C. T., Jurgenson, J. E. & Hulbert, S. H. Development of a host-induced RNAi system in the wheat stripe rust fungus *Puccinia striiformis* f. sp *tritici*. *Mol. Plant. Microbe Interact.* **24,** 554–561 (2011).

126. Tinoco, M., Dias, B., Dall'Astta, R., Pamphile, J. & Aragao, F. In vivo trans-specific gene silencing in fungal cells by in planta expression of a double-stranded RNA. *BMC Biol.* **8,** 27 (2010).

127. Andrade, C. M., Tinoco, M. L. P., Rieth, A. F., Maia, F. C. O. & Aragão, F. J. L. Host-induced gene silencing in the necrotrophic fungal pathogen *Sclerotinia sclerotiorum*. *Plant Pathol.* **65,** 626–632 (2016).

128. Laurie, J. D., Linning, R. & Bakkeren, G. Hallmarks of RNA silencing are found in the smut fungus *Ustilago hordei* but not in its close relative *Ustilago maydis*. *Curr. Genet.* **53,** 49–58 (2008).

129. Drinnenberg, I. A. *et al.* RNAi in budding yeast. *Science* **326,** 544–550 (2009).

130. D'Souza, C. A. *et al.* Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *Mbio* **2,** e00342-10 (2011).

131. Mumbanza, F. M. *et al.* In vitro antifungal activity of synthetic dsRNA molecules against two pathogens of banana, *Fusarium oxysporum* f. sp. *cubense* and *Mycosphaerella fijiensis*. *Pest Manag. Sci.* **69,** 1155–1162 (2013).

132. Gao, A. G. *et al.* Fungal pathogen protection in potato by expression of a plant defensin peptide. *Nat. Biotechnol.* **18,** 1307–1310 (2000).

133. Seale JW, V. P. Amino acid sequence variant alfalfa antifungal protein and its use in plant disease control. (2010).

134. McDougall, P. *The cost and time involved in the discovery, development and authorisation of a new plant biotechnology derived trait*. (Crop Life International, 2011).

135. Lucht, J. M. Public acceptance of plant biotechnology and GM crops. *Viruses* **7,** 4254–4281 (2015).

136. San Miguel, K. & Scott, J. G. The next generation of insecticides: dsRNA is stable as a foliar-applied insecticide. *Pest Manag. Sci.* **72,** 801–809 (2016).

137. Robinson, K. E., Worrall, E. A. & Mitter, N. Double stranded RNA expression and its topical application for non-transgenic resistance to plant viruses. *J. Plant Biochem. Biotechnol.* **23,** 231–237 (2014).

138. Kogel, K. H. & Koch, A. RNA as fungicide: Towards an RNAi-based non-transgenic control of plant diseases. in *Crossroads in Science* **Biology and Disease Management-Chemical Control,** 286–P (American Phytopathological Society, 2015).

139. McDonald, B. A. & Linde, C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* **40,** 349–79 (2002).

140. Santana, M. F. *et al.* Abundance, distribution and potential impact of transposable elements in the genome of *Mycosphaerella fijiensis*. *BMC Genomics* **13,** 720 (2012).

141. Stukenbrock, E. H. & Croll, D. The evolving fungal genome. *Fungal Biol. Rev.* **28,** 1–12 (2014).

142. Leach, J. E., Cruz, C. M. V., Bai, J. F. & Leung, H. Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annu. Rev. Phytopathol.* **39,** 187–224 (2001).

143. Smith, M. C., Rutter, J., Burt, P. J. A., Ramirez, F. & Gonzalez, E. H. Black Sigatoka disease of banana: spatial and temporal variability in disease development. *Ann. Appl. Biol.* **131,** 63–77 (1997).

144. Bendini, H. D., Moraes, W. D., da Silva, S., Tezuka, E. S. & Cruvinel, P. E. Risk analysis of black Sigatoka occurrence based on polynomial models: A case study. *Trop. Plant Pathol.* **38,** 35–43 (2013).

145. Guillermet, C., Le Guen, R., Foure, E., Cespedes, C. & de Bellaire, L. D. Adaptation of the forecasting system to control Black Leaf Streak Disease of banana in the specific conditions of Dominican Republic. *Fruits* **69,** 261–278 (2014).

146. Arzanlou, M. *et al.* Molecular diagnostics for the Sigatoka disease complex of banana. *Phytopathology* **97,** 1112–1118 (2007).

147. Melby, J. M. & Stave, J. W. ELISA evaluation of banana black Sigatoka disease status. *J. Clin. Ligand Assay* **18,** 166–170 (1995).

148. Henderson, J. *et al.* Black Sigatoka disease: new technologies to strengthen eradication strategies in Australia. *Australas. Plant Pathol.* **35,** 181–193 (2006).

149. Etienne, L., Steden, C. & Suter, J. in *Diagnosis and Identification of Plant Pathogens* (eds. Dehne, H. W. et al.) **11,** 377–383 (1997).

150. Otero, A. J. *et al.* Monoclonal antibody-based TAS-ELISA for quantitative detection of *Mycosphaerella fijiensis* antigens. *J. Phytopathol.* **155,** 713–719 (2007).

151. Canas-Gutierrez, G. P., Patino, L. F., Rodriguez-Arango, E. & Arango, R. Molecular characterization of benomyl-resistant isolates of *Mycosphaerella fijiensis*, collected in Colombia. *J. Phytopathol.* **154,** 403–409 (2006).

152. Heiser, I. *et al.* Fatty acid peroxidation by rubellin B, C and D, phytotoxins produced by *Ramularia collo-cygni* (Sutton et Waller). *Physiol. Mol. Plant Pathol.* **64,** 135–143 (2004).

153. Bradshaw, R. E. *et al.* Fragmentation of an aflatoxin-like gene cluster in a forest pathogen. *New Phytol.* **198,** 525–535 (2013).

154. Kabir, M. S., Ganley, R. J. & Bradshaw, R. E. Dothistromin toxin is a virulence factor in dothistroma needle blight of pines. *Plant Pathol.* **64,** 225–234 (2015).

155. Holler, U., Gloer, J. B. & Wicklow, D. T. Biologically active polyketide metabolites from an undetermined fungicolous hyphomycete resembling *Cladosporium*. *J. Nat. Prod.* **65,** 876–882 (2002).

156. Wheeler, M. H. & Stipanovic, R. D. Melanin biosynthesis and the metabolism of flaviolin and 2-hydroxyjuglone in *Wangiella dermatitidis*. *Arch. Microbiol.* **142,** 234–241 (1985).

157. Bell, A. A., Puhalla, J. E., Tolmsoff, W. J. & Stipanovic, R. D. Use of mutants to establish (+)-scytalone as an intermediate in melanin biosynthesis by *Verticillium dahliae*. *Can. J. Microbiol.* **22,** 787–799 (1976).

158. Saparrat, M. C. N., Fermoselle, G. E., Stenglein, S. A., Aulicino, M. B. & Balatti, P. A. *Pseudocercospora griseola* causing angular leaf spot on *Phaseolus vulgaris* produces 1,8-dihydroxynaphthalene-melanin. *Mycopathologia* **168,** 41–47 (2009).

159. Jacobson, E. S. & Ikeda, R. Effect of melanization upon porosity of the cryptococcal cell wall. *Med. Mycol.* **43,** 327–333 (2005).

160. D.O.E. Joint Genome Institute. *Mycosphaerella fijiensis* v2.0. (2016). Available at: http://genomeportal.jgi-psf.org/Mycfi2/Mycfi2.home.html. (Accessed: 23rd September 2016)

161. Keller, N. P., Turner, G. & Bennett, J. W. Fungal secondary metabolism - From biochemistry to genomics. *Nat. Rev. Microbiol.* **3,** 937–947 (2005).

162. Keatinge-Clay, A. T. & Stroud, R. M. The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* **14,** 737–748 (2006).

163. Cox, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.* **5,** 2010–2026 (2007).

164. Yoder, O. C. in *Molecular Genetics of Host-Specific Toxins in Plant Disease* (eds. Kohmoto, K. & Yoder, O. C.) 3–15 (Springer Netherlands, 1998). doi:10.1007/978-94-011-5218-1_1

165. Keller, N. P. & Hohn, T. M. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genet. Biol.* **21,** 17–29 (1997).

166. Wiemann, P. *et al.* Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. *Mol. Microbiol.* **72,** 931–946 (2009).

167. Yu, J., Bhatnagar, D. & Cleveland, T. E. Completed sequence of aflatoxin pathway gene cluster in *Aspergillus parasiticus*. *Febs Lett.* **564,** 126–130 (2004).

168. Marchler-Bauer, A. *et al.* CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39,** D225-229 (2011).

169. Finking, R., Mofid, M. R. & Marahiel, M. A. Mutational analysis of peptidyl carrier protein and acyl carrier protein synthase unveils residues involved in protein−protein recognition. *Biochemistry (Mosc.)* **43,** 8946–8956 (2004).

170. Fujii, I., Yoshida, N., Shimomaki, S., Oikawa, H. & Ebizuka, Y. An iterative type I polyketide synthase PKSN catalyzes synthesis of the decaketide alternapyrone with regio-specific octa-methylation. *Chem. Biol.* **12,** 1301–1309 (2005).

171. Du, L. *et al.* Biosynthesis of sphinganine-analog mycotoxins. *J. Ind. Microbiol. Biotechnol.* **35,** 455–464 (2008).

172. Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11,** 660–670 (2015).

173. Seo, J. A., Proctor, R. H. & Plattner, R. D. Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in *Fusarium verticillioides*. *Fungal Genet. Biol.* **34,** 155–165 (2001).

174. Blackwell, B. A., Edwards, O. E., Fruchier, A., ApSimon, J. W. & Miller, J. D. NMR structural studies of fumonisin B1 and related compounds from *Fusarium moniliforme*. *Fumonisins Food* **392,** 75–91 (1996).

175. Branham, B. E. & Plattner, R. D. Alanine is a precursor in the biosynthesis of fumonisin B1 by *Fusarium moniliforme*. *Mycopathologia* **124,** 99–104 (1993).

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

177. Caldas, E. D. *et al.* Biosynthetic studies of fumonisin B_1 and AAL toxins. *J. Agric. Food Chem.* **46,** 4734–4743 (1998).

178. Kasahara, K. *et al.* Solanapyrone synthase, a possible Diels-Alderase and iterative type I polyketide synthase encoded in a biosynthetic gene cluster from *Alternaria solani*. *Chembiochem* **11,** 1245–1252 (2010).

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

180. Jenke-Kodama, H., Sandmann, A., Muller, R. & Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **22,** 2027–2039 (2005).

181. Huang, W. J. *et al.* Crystal structure of β-ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *Embo J.* **17,** 1183–1191 (1998).

182. Bachmann, B. O. & Ravel, J. in *Methods in Enzymology* **Volume 458,** 181–217 (Academic Press, 2009).

183. Chumley, F. G. & Valent, B. Genetic analysis of melanin-deficient, nonpathogenic mutants of *Magnaporthe grisea*. *Mol. Plant. Microbe Interact.* **3,** 135–143 (1990).

184. Perpetua, N. S., Kubo, Y., Takano, Y. & Furusawa, I. Cloning and characterization of a melanin biosynthetic THR1 reductase gene essential for appressorial penetration of *Colletotrichum lagenarium*. *Mol. Plant. Microbe Interact.* **9,** 323–329 (1996).

185. Cho, Y. *et al.* Transcription factor Amr1 induces melanin biosynthesis and suppresses virulence in *Alternaria brassicicola*. *Plos Pathog.* **8,** (2012).

186. Wight, W. D., Kim, K.-H., Lawrence, C. B. & Walton, J. D. Biosynthesis and role in virulence of the histone deacetylase inhibitor depudecin from *Alternaria brassicicola*. *Mol. Plant. Microbe Interact.* **22,** 1258–1267 (2009).

187. Chen, H. Q., Lee, M. H., Daub, M. E. & Chung, K. R. Molecular analysis of the cercosporin biosynthetic gene cluster in *Cercospora nicotianae*. *Mol. Microbiol.* **64,** 755–770 (2007).

188. Proctor, R. H., Brown, D. W., Plattner, R. D. & Desjardins, A. E. Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. *Fungal Genet. Biol.* **38,** 237–249 (2003).

189. Zhang, Y., Keller, N. & Tsitsigiannis, D. in *Handbook of Industrial Mycology* 410–411 (CRC Press, 2005).

190. Mizushina, Y. *et al.* A plant phytotoxin, solanapyrone A, is an inhibitor of DNA polymerase beta and lambda. *J. Biol. Chem.* **277,** 630–638 (2002).

191. Kaur, S. Phytotoxicity of solanapyrones produced by the fungus *Ascochyta rabiei* and their possible role in blight of chickpea (*Cicer arietinum*). *Plant Sci.* **109,** 23–29 (1995).

192. Kim, W. *et al.* Functional analyses of the Diels-Alderase gene *sol5* of *Ascochyta rabiei* and *Alternaria solani* indicate that the solanapyrone phytotoxins are not required for pathogenicity. *Mol. Plant. Microbe Interact.* **28,** 482–496 (2015).

193. Chen, M., Cahoon, E. B., Saucedo-Garcia, M., Plasencia, J. & Gavilanes-Ruiz, M. Plant sphingolipids: structure, synthesis and function. *Lipids Photosynth. Essent. Regul. Funct.* **30,** 77–115 (2009).

194. Sanchez-Rangel, D. & Plasencia, J. The role of sphinganine analog mycotoxins on the virulence of plant pathogenic fungi. *Toxin Rev.* **29,** 73–86 (2010).

195. Marasas, W. F. O. *et al.* Fumonisins disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: A potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. *J. Nutr.* **134,** 711–716 (2004).

196. Pata, M. O., Hannun, Y. A. & Ng, C. K. Y. Plant sphingolipids: decoding the enigma of the Sphinx. *New Phytol.* **185,** 611–630 (2010).

197. Glenn, A. E. *et al.* Transformation-mediated complementation of a *FUM* gene cluster deletion in *Fusarium verticillioides* restores both fumonisin production and pathogenicity on maize seedlings. *Mol. Plant. Microbe Interact.* **21,** 87–97 (2008).

198. Gilchrist, D. G. Mycotoxins reveal connections between plants and animals in apoptosis and ceramide signaling. *Cell Death Differ.* **4,** 689–698 (1997).

199. Grigoriev, I. V. *et al.* Fueling the future with fungal genomics. *Mycology* **2,** 192–209 (2011).

200. Grigoriev, I. V. *et al.* MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* **42,** D699–D704 (2014).

201. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

202. Maddison, W. P. and D. R. M. Mesquite: a modular system for evolutionary analysis. Version 3.04. (2015). Available at: http://mesquiteproject.org. (Accessed: 23rd September 2016)

203. Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J. & McLnerney, J. O. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* **6,** 29 (2006).

204. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25,** 1307–1320 (2008).

205. Silvestro, D. & Michalak, I. raxmlGUI: a graphical front-end for RAxML. *Org. Divers. Evol.* **12,** 335–337 (2012).

206. Camacho, C. *et al.* BLAST plus: architecture and applications. *BMC Bioinformatics* **10,** 421 (2009).

207. Peraza-Echeverría, L., Rodríguez-García, C. & Zapata-Salazar, D. A rapid, effective method for profuse in vitro conidial production of *Mycosphaerella fijiensis*. *Australas. Plant Pathol.* **37,** 460–463 (2008).

208. Murashige, T. & Skoog, F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15,** 473–497 (1962).

209. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17,** 10–12 (2011).

210. Droc, G. *et al.* The Banana Genome Hub. *Database* **2013,** (2013).

211. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111 (2009).

212. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

213. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

214. Moriwaki, A. *et al.* Insertional mutagenesis and characterization of a polyketide synthase gene (*PKS1*) required for melanin biosynthesis in *Bipolaris oryzae*. *Fems Microbiol. Lett.* **238,** 1–8 (2004).

215. Ludwig, N. *et al.* Melanin is not required for turgor generation but enhances cell-wall rigidity in appressoria of the corn pathogen *Colletotrichum graminicola*. *Mol. Plant. Microbe Interact.* **27,** 315–327 (2014).

216. Feng, B. *et al.* Molecular cloning and characterization of *WdPKS1*, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in *Wangiella* (*Exophiala*) *dermatitidis*. *Infect. Immun.* **69,** 1781–1794 (2001).

217. Zhang, A. *et al.* Efficient disruption of a polyketide synthase gene (*pks1*) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. *Mol. Genet. Genomics* **268,** 645–655 (2003).

218. Feng, G. H. & Leonard, T. J. Characterization of the polyketide synthase gene (*pksL1*) required for aflatoxin biosynthesis in *Aspergillus parasiticus*. *J. Bacteriol.* **177,** 6246–6254 (1995).

219. Arndt, B. *et al.* Genetic engineering, high resolution mass spectrometry and nuclear magnetic resonance spectroscopy elucidate the bikaverin biosynthetic pathway in *Fusarium fujikuroi*. *Fungal Genet. Biol.* **84,** 26–36 (2015).

220. Abe, Y. *et al.* Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in *Penicillium citrinum*. *Mol. Genet. Genomics* **267,** 636–646 (2002).

221. Hendrickson, L. *et al.* Lovastatin biosynthesis in *Aspergillus terreus*: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem. Biol.* **6,** 429–439 (1999).

222. Galagan, J. E. *et al.* Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438,** 1105–1115 (2005).

223. Baker, S. E. *et al.* 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 (2006).

224. Beck, J., Ripka, S., Siegner, A., Schiltz, E. & Schweizer, E. 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 (1990).

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

226. de Lapeyre de Bellaire, L. *et al.* Is chemical control of *Mycosphaerella* foliar diseases of banana sustainable? *Acta Hortic.* **828,** 161–170 (2009).

227. Stergiopoulos, I., Collemare, J., Mehrabi, R. & De Wit, P. J. G. M. Phytotoxic secondary metabolites and peptides produced by plant pathogenic *Dothideomycete* fungi. *Fems Microbiol. Rev.* **37,** 67–93 (2013).

228. Strobel, G. A., Stierle, A. A., Upadhyay, R., Hershenhorn, J. & Molina, G. C. in *Biotechnology Applications for Banana and Plantain Improvement* (eds. Wills, B. & Huggan, R. D.) 93–103 (International Network for the Improvement of Banana and Plantain, 1993).

229. O'Connell, R. J. *et al.* Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nat Genet* **44,** 1060–1065 (2012).

230. Rudd, J. J. *et al.* Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiol.* **167,** 1158–1185 (2015).

231. Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends Biochem. Sci.* **28,** 118–121 (2003).

232. Ohtaki, S. *et al.* Novel hydrophobic surface binding protein, HsbA, produced by *Aspergillus oryzae*. *Appl. Environ. Microbiol.* **72,** 2407–2413 (2006).

233. Skamnioti, P. & Gurr, S. J. Cutinase and hydrophobin interplay: A herald for pathogenesis? *Plant Signal. Behav.* **3,** 248–250 (2008).

234. Niderman, T. *et al.* Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. *Plant Physiol.* **108,** 17–27 (1995).

235. Van Loon, L. C. & Van Strien, E. A. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.* **55,** 85–97 (1999).

236. Prados-Rosales, R. C. *et al.* A PR-1-like protein of *Fusarium oxysporum* functions in virulence on mammalian hosts. *J. Biol. Chem.* **287,** 21970–21979 (2012).

237. Teixeira, P. J. P. L. *et al.* The fungal pathogen *Moniliophthora perniciosa* has genes similar to plant PR-1 that are highly expressed during its interaction with cacao. *PLOS ONE* **7,** e45929 (2012).

238. Király, L., Barna, B. & Király, Z. Plant resistance to pathogen infection: forms and mechanisms of innate and acquired resistance. *J. Phytopathol.* **155,** 385–396 (2007).

239. Yamamoto, S., Katagiri, M., Maeno, H. & Hayaishi, O. Salicylate hydroxylase, a monooxygenase requiring flavin adenine dinucleotide: I. Purification and general properties. *J. Biol. Chem.* **240,** 3408–3413 (1965).

240. Ambrose, K. V. *et al.* Functional characterization of salicylate hydroxylase from the fungal endophyte *Epichloë festucae*. *Sci. Rep.* **5,** 10939 (2015).

241. Williams, R. J., Eakin, R. E. & Snell, E. E. The relationship of inositol, thiamin, biotin, pantothenic acid and vitamin B6 to the growth of yeasts. *J. Am. Chem. Soc.* **62,** 1204–1207 (1940).

242. Yamamoto, L. A. & Segel, I. H. The inorganic sulfate transport system of *Penicillium chrysogenum*. *Arch. Biochem. Biophys.* **114,** 523–538 (1966).

243. Rao, T. K. & DeBusk, A. G. An inducible acetate transport system in *Neurospora crassa* conidia. *Biochim. Biophys. Acta BBA - Biomembr.* **470,** 475–483 (1977).

244. Chaure, P. T. & Connerton, I. F. Derepression of the glyoxylate cycle in mutants of *Neurospora crassa* accelerated for growth on acetate. *Microbiology* **141,** 1315–1320 (1995).

245. Jennings, D. H. *The physiology of fungal nutrition*. (Cambridge University Press, 1995).

246. Marzluf, G. A. Molecular genetics of sulfur assimilation in filamentous fungi and yeast. *Annu. Rev. Microbiol.* **51,** 73–96 (1997).

247. Casal, M., Paiva, S., Andrade, R. P., Gancedo, C. & Leão, C. The lactate-proton symport of *Saccharomyces cerevisiae* is encoded by *JEN1*. *J. Bacteriol.* **181,** 2620–2623 (1999).

248. Carlsen, M. & Nielsen, J. Influence of carbon source on α-amylase production by *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* **57,** 346–349 (2001).

249. Waters, B. M., Blevins, D. G. & Eide, D. J. Characterization of FRO1, a pea ferric-chelate reductase involved in root iron acquisition. *Plant Physiol.* **129,** 85–94 (2002).

250. Paiva, S., Devaux, F., Barbosa, S., Jacq, C. & Casal, M. Ady2p is essential for the acetate permease activity in the yeast *Saccharomyces cerevisiae*. *Yeast* **21,** 201–210 (2004).

251. Dumay, Q. C., Debut, A. J., Mansour, N. M. & Saier, M. H., Jr. The copper transporter (Ctr) family of Cu+ uptake systems. *J. Mol. Microbiol. Biotechnol.* **11,** 10–19 (2006).

252. Lagaert, S., Pollet, A., Courtin, C. M. & Volckaert, G. β-Xylosidases and α-l-arabinofuranosidases: Accessory enzymes for arabinoxylan degradation. *Biotechnol. Adv.* **32,** 316–332 (2014).

253. Teichert, I., Wolff, G., Kueck, U. & Nowrousian, M. Combining laser microdissection and RNA-seq to chart the transcriptional landscape of fungal development. *BMC Genomics* **13,** 511 (2012).

254. Grognet, P. *et al.* Maintaining two mating types: structure of the mating type locus and its role in heterokaryosis in *Podospora anserina*. *Genetics* **197,** 421–432 (2014).

255. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25,** 25–29 (2000).

256. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21,** 3674–3676 (2005).

257. Conesa, A. & Götz, S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* **2008,** 619832 (2008).

258. Zdobnov, E. M. & Apweiler, R. InterProScan – an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17,** 847–848 (2001).

259. Quevillon, E. *et al.* InterProScan: protein domains identifier. *Nucleic Acids Res.* **33,** W116–W120 (2005).

260. Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42,** D490–D495 (2014).

261. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc.* **2,** 953–971 (2007).

262. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8,** 785–786 (2011).

263. van den Burg, H. A., Harrison, S. J., Joosten, M. H. A. J., Vervoort, J. & de Wit, P. J. G. M. *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Mol. Plant. Microbe Interact.* **19,** 1420–1430 (2006).

264. Jenke-Kodama, H. & Dittmann, E. Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. *Nat. Prod. Rep.* **26,** 874–883 (2009).

265. Minami, A. *et al.* Identification and functional analysis of brassicicene C biosynthetic gene cluster in *Alternaria brassicicola*. *Bioorg. Med. Chem. Lett.* **19,** 870–874 (2009).

266. Pazzagli, L. *et al.* Purification, characterization, and amino acid sequence of cerato-platanin, a new phytotoxic protein from *Ceratocystis fimbriata* f. sp *platani*. *J. Biol. Chem.* **274,** 24959–24964 (1999).

267. Chen, H., Kovalchuk, A., Kerio, S. & Asiegbu, F. O. Distribution and bioinformatic analysis of the cerato-platanin protein family in Dikarya. *Mycologia* **105,** 1479–1488 (2013).

268. Frischmann, A. *et al.* Self-assembly at air/water interfaces and carbohydrate binding properties of the small secreted protein EPL1 from the fungus *Trichoderma atroviride*. *J. Biol. Chem.* **288,** 4278–4287 (2013).

269. Baccelli, I. Cerato-platanin family proteins: one function for multiple biological roles? *Front. Plant Sci.* **5,** 769 (2014).

270. Gaderer, R., Bonazza, K. & Seidl-Seiboth, V. Cerato-platanins: a fungal protein family with intriguing properties and application potential. *Appl. Microbiol. Biotechnol.* **98,** 4795–4803 (2014).

271. Briza, P., Winkler, G., Kalchhauser, H. & Breitenbach, M. Dityrosine is a prominent component of the yeast ascospore wall. A proof of its structure. *J. Biol. Chem.* **261,** 4288–4294 (1986).

272. Briza, P., Ellinger, A., Winkler, G. & Breitenbach, M. Characterization of a DL-dityrosine-containing macromolecule from yeast ascospore walls. *J. Biol. Chem.* **265,** 15118–23 (1990).

273. Prillinger, H. *et al.* Phytopathogenic filamentous (*Ashbya*, *Eremothecium*) and dimorphic fungi (*Holleya*, *Nematospora*) with needle-shaped ascospores as new members within the Saccharomycetaceae. *Yeast* **13,** 945–960 (1997).

274. Smail, E. H., Briza, P., Panagos, A. & Berenfeld, L. *Candida albicans* cell walls contain the fluorescent cross-linking amino acid dityrosine. *Infect. Immun.* **63,** 4078–4083 (1995).

275. Visca, P., Imperi, F. & Lamont, I. L. Pyoverdine siderophores: from biogenesis to biosignificance. *Trends Microbiol.* **15,** 22–30 (2007).

276. Goodwin, S. B. *et al.* Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLOS Genet.* **7,** e1002070 (2011).

277. Tsuge, T. *et al.* Evolution of pathogenicity controlled by small, dispensable chromosomes in *Alternaria alternata* pathogens. *Physiol. Mol. Plant Pathol.* **95,** 27–31 (2016).

278. Walton, J. D., Ahn, J.-H., Akimitsu, K., Pitkin, J. W. & Ransom, R. in *Advances in Molecular Genetics of Plant-Microbe Interactions: Vol. 3 Proceedings of the 7th International Symposium on Molecular Plant-Microbe Interactions, Edinburgh, U.K., June 1994* (eds. Daniels, M. J., Downie, J. A. & Osbourn, A. E.) 231–237 (Springer Netherlands, 1994).

279. Kistler, H. C., Meinhardt, L. W. & Benny, U. Mutants of *Nectria haematococca* created by a site-directed chromosome breakage are greatly reduced in virulence toward pea. *Mol. Plant. Microbe Interact.* **9,** 804–809 (1996).

280. Wasmann, C. C. & VanEtten, H. D. Transformation-mediated chromosome loss and disruption of a gene for pisatin demethylase decrease the virulence of *Nectria haematococca* on pea. *Mol. Plant. Microbe Interact.* **9,** 793–803 (1996).

281. Enkerli, J., Bhatt, G. & Covert, S. F. Maackiain detoxification contributes to the virulence of *Nectria haematococca* MP VI on chickpea. *Mol. Plant. Microbe Interact.* **11,** 317–326 (1998).

282. Molina, G. C. & Krausz, J. P. Toxin production of *Mycosphaerella fijiensis* var. *difformis*. *Phytopathology* **77,** 1747–1747 (1987).

283. El Hadrami, A., Kone, D. & Lepoivre, P. Effect of juglone on active oxygen species and antioxidant enzymes in susceptible and partially resistant banana cultivars to Black Leaf Streak Disease. *Eur. J. Plant Pathol.* **113,** 241–254 (2005).

284. Turner, N. C. & Graniti, A. Fusicoccin - a fungal toxin that opens stomata. *Nature* **223,** 1070–1071 (1969).

285. MacKinnon, S. L., Keifer, P. & Ayer, W. A. Components from the phytotoxic extract of *Alternaria brassicicola*, a black spot pathogen of canola. *Phytochemistry* **51,** 215–221 (1999).

286. Fang, X. *et al.* NMR studies of molecular structure in fruit cuticle polyesters. *Phytochemistry* **57,** 1035–1042 (2001).

287. Snoeijers, S. S., Pérez-García, A., Joosten, M. H. A. J. & De Wit, P. J. G. M. The effect of nitrogen on disease development and gene expression in bacterial and fungal plant pathogens. *Eur. J. Plant Pathol.* **106,** 493–506 (2000).

288. Talbot, N. J., Ebbole, D. J. & Hamer, J. E. Identification and characterization of *MPG1*, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. *Plant Cell Online* **5,** 1575–90 (1993).

289. Skamnioti, P. & Gurr, S. J. *Magnaporthe grisea* cutinase2 mediates appressorium differentiation and host penetration and is required for full virulence. *Plant Cell* **19,** 2674–2689 (2007).

290. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25,** 402–408 (2001).

291. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci. CABIOS* **8,** 275–282 (1992).

292. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

293. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680 (1994).

294. Anonymous. Major diseases of banana and plantain: an update of their spread, impact and response strategies. in (Food and Agriculture Organization of the United Nations, 2009).

295. Demain, A. L. & Fang, A. in *History of Modern Biotechnology I* (ed. Fiechter, P. D. A.) 1–39 (Springer Berlin Heidelberg, 2000). doi:10.1007/3-540-44964-7_1

296. Pusztahelyi, T., Holb, I. J. & Pócsi, I. Secondary metabolites in fungus-plant interactions. *Front. Plant Sci.* **6,** 573 (2015).

297. Langfelder, K. *et al.* Identification of a polyketide synthase gene (*pksP*) of *Aspergillus fumigatus* involved in conidial pigment biosynthesis and virulence. *Med. Microbiol. Immunol. (Berl.)* **187,** 79–89 (1998).

298. Zakeri, B. & Lu, T. K. Synthetic biology of antimicrobial discovery. *ACS Synth. Biol.* **2,** 358–372 (2013).

299. Viviani, F., Vidal-Cros, A. & Gaudry, M. in *Mechanisms of Plant Defense Responses* (eds. Fritig, B. & Legrand, M.) 461–461 (Springer Netherlands, 1993). doi:10.1007/978-94-011-1737-1_141

300. Jacobson, E. S. Pathogenic roles for fungal melanins. *Clin. Microbiol. Rev.* **13,** 708–717 (2000).

301. Chang, T.-C., Salvucci, A., Crous, P. W. & Stergiopoulos, I. Comparative genomics of the Sigatoka disease complex on banana suggests a link between parallel evolutionary changes in *Pseudocercospora fijiensis* and *Pseudocercospora eumusae* and increased virulence on the banana host. *PLoS Genet.* **12,** e1005904 (2016).

302. Bok, J. W. & Keller, N. P. LaeA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryot. Cell* **3,** 527–535 (2004).

303. Bayram, Ö. *et al.* VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. *Science* **320,** 1504–1506 (2008).

304. Bell, A. A. & Wheeler, M. H. Biosynthesis and functions of fungal melanins. *Annu. Rev. Phytopathol.* **24,** 411–451 (1986).

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

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

307. Graziani, S., Vasnier, C. & Daboussi, M.-J. Novel polyketide synthase from *Nectria haematococca*. *Appl. Environ. Microbiol.* **70,** 2984–2988 (2004).

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

309. Gaffoor, I. *et al.* Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). *Eukaryot. Cell* **4,** 1926–1933 (2005).

310. Song, Z., Cox, R. J., Lazarus, C. M. & Simpson, T. J. Fusarin C biosynthesis in *Fusarium moniliforme* and *Fusarium venenatum*. *ChemBioChem* **5,** 1196–1203 (2004).

311. Weber, T. *et al.* antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43,** W237–W243 (2015).

312. Xu, J. *et al.* Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 18730–18735 (2007).

313. Ugai, T. *et al.* Heterologous expression of highly reducing polyketide synthase involved in betaenone biosynthesis. *Chem. Commun.* **51,** 1878–1881 (2015).

314. Kennedy, J. *et al.* Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **284,** 1368–1372 (1999).

315. Chen, Y.-P. *et al.* Cloning and characterization of monacolin K biosynthetic gene cluster from *Monascus pilosus*. *J. Agric. Food Chem.* **56,** 5639–5646 (2008).

316. Kodama, M. *et al.* The translocation-associated *Tox1* locus of *Cochliobolus heterostrophus* is two genetic elements on two different chromosomes. *Genetics* **151,** 585–596 (1999).

317. Alexopoulos, C. & Mims, C. *Introductory Mycology*. (John Wiley & Sons Inc, 1979).

318. Conde-Ferráez, L. *et al.* The development of mating type-specific primers for *Mycosphaerella fijiensis*. *Australas. Plant Pathol.* **39,** 217–225 (2010).

319. Etebu, E., Pasberg-Gauhl, C., Gauhl, F. & Daniel-Kalio, L. A. Preliminary studies of in vitro stimulation of sexual mating among isolates of *Mycosphaerella fijiensis*, causal agent of black sigatoka disease in bananas and plantains. *Phytoparasitica* **31,** 69–75 (2003).

320. Ichihara, A. *et al.* Structures of betaenones A and B, novel phytotoxins from *Phoma betae* Fr. *J. Am. Chem. Soc.* **105,** 2907–2908 (1983).

321. Brauers, G. *et al.* Anthraquinones and betaenone derivatives from the sponge-associated fungus *Microsphaeropsis* species: novel inhibitors of protein kinases. *J. Nat. Prod.* **63,** 739–745 (2000).

322. Mulder, N. J. *et al.* InterPro, progress and status in 2005. *Nucleic Acids Res.* **33,** D201–D205 (2005).

323. Manzoni, M. & Rollini, N. Biosynthesis and biotechnological production of statins by filamentous fungi and application of these cholesterol-lowering drugs. *Appl. Microbiol. Biotechnol.* **58,** 555–564 (2002).

324. Bayram, Ö. S. *et al.* LaeA control of velvet family regulatory proteins for light-dependent development and fungal cell-type specificity. *PLOS Genet* **6,** e1001226 (2010).

325. AbdEl-Mongy, M. Regulation of *Eurotium repens* reproduction and secondary metabolite production. *Can. J. Pure Appl. Sci.* **6,** 1937–1944 (2012).

326. Koch, A. *et al.* Host-induced gene silencing of cytochrome P450 lanosterol C14 alpha-demethylase-encoding genes confers strong resistance to *Fusarium* species. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 19324–19329 (2013).

327. Tokousbalides, M. C. & Sisler, H. D. Site of inhibition by tricyclazole in the melanin biosynthetic pathway of *Verticillium dahliae*. *Pestic. Biochem. Physiol.* **11,** 64–73 (1979).

328. Daskalchuk, T., Ahiahonu, P., Heath, D. & Yamazaki, Y. in *Plant Metabolomics* (eds. Saito, P. D. K., Dixon, P. D. R. A. & Willmitzer, P. D. L.) 311–325 (Springer Berlin Heidelberg, 2006). doi:10.1007/3-540-29782-0_22

329. Gunnaiah, R. & Kushalappa, A. C. Metabolomics deciphers the host resistance mechanisms in wheat cultivar Sumai-3, against trichothecene producing and non-

producing isolates of *Fusarium graminearum*. *Plant Physiol. Biochem. PPB Société Fr. Physiol. Végétale* **83,** 40–50 (2014).

330. Barkal, L. J. *et al.* Microbial metabolomics in open microscale platforms. *Nat. Commun.* **7,** 10610 (2016).
331. Kaling, M. *et al.* UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. *Plant Cell Environ.* **38,** 892–904 (2015).
332. Ma, D.-M., Gandra, S. V. S., Sharma, N. & Xie, D.-Y. Integration of GC-MS based non-targeted metabolic profiling with headspace solid phase microextraction enhances the understanding of volatile differentiation in tobacco leaves from North Carolina, India and Brazil. *Am. J. Plant Sci.* **03,** 1759–1769 (2012).
333. Macel, M., Van Dam, N. M. & Keurentjes, J. J. B. Metabolomics: the chemistry between ecology and genetics. *Mol. Ecol. Resour.* **10,** 583–593 (2010).
334. Herpen, T. W. J. M. van *et al. Nicotiana benthamiana* as a production platform for artemisinin precursors. *PLOS ONE* **5,** e14222 (2010).
335. Watanabe, C. M. H. & Townsend, C. A. 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 (2002).
336. Brobst, S. W. & Townsend, C. A. The potential role of fatty acid initiation in the biosynthesis of the fungal aromatic polyketide aflatoxin B_1. *Can. J. Chem.* **72,** 200–207 (1994).

337. Carreras, C. W., Pieper, R. & Khosla, C. in *Bioorganic Chemistry Deoxysugars, Polyketides and Related Classes: Synthesis, Biosynthesis, Enzymes* (ed. Rohr, D. J.) 85–126 (Springer Berlin Heidelberg, 1997). doi:10.1007/BFb0119235

338. O'Hagan, D. Biosynthesis of fatty acid and polyketide metabolites. *Nat. Prod. Rep.* **12,** 1–32 (1995).

339. Schirmer, A., Rude, M. A., Li, X., Popova, E. & del Cardayre, S. B. Microbial biosynthesis of alkanes. *Science* **329,** 559–562 (2010).

340. Zhang, F., Rodriguez, S. & Keasling, J. D. Metabolic engineering of microbial pathways for advanced biofuels production. *Curr. Opin. Biotechnol.* **22,** 775–783 (2011).

341. Medema, M. H. *et al.* Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **11,** 625–631 (2015).

342. Georgianna, D. R. Functional and comparative genomics of *Aspergillus flavus* to characterize secondary metabolism. (North Carolina State University, 2009).

343. Earley, K. W. *et al.* Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45,** 616–629 (2006).

344. Sweigard, J. A., Chumley, F., Carroll, A., Farrall, L. & Valent, B. A series of vectors for fungal transformation. *Fungal Genet. Newsl.* **44,** 52–53 (1997).

345. Utermark, J. & Karlovsky, P. Genetic transformation of filamentous fungi by *Agrobacterium tumefaciens*. (2008).

346. Ma, D.-M. *et al.* A genome-wide scenario of terpene pathways in self-pollinated *Artemisia annua*. *Mol. Plant* **8,** 1580–1598 (2015).

347. Liao, H.-L. & Chung, K.-R. Genetic dissection defines the roles of elsinochrome phytotoxin for fungal pathogenesis and conidiation of the citrus pathogen *Elsinoë fawcettii*. *Mol. Plant-Microbe Interact. MPMI* **21,** 469–479 (2008).

348. Lim, F. Y. *et al.* Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* **78,** 4117–4125 (2012).

349. Arnone, A., Camarda, L., Nasini, G. & Assante, G. Secondary mould metabolites. Part 15. Structure elucidation of rubellins A and B, two novel anthraquinone metabolites from *Mycosphaerella rubella*. *J. Chem. Soc. [Perkin 1]* 255–260 (1986). doi:10.1039/P19860000255

350. Cercosporin | Abcam. Available at: http://www.abcam.com/cercosporin-ab144218.html. (Accessed: 6th September 2016)

351. Daub, M. E. & Briggs, S. P. Changes in tobacco cell membrane composition and structure caused by cercosporin. *Plant Physiol.* **71,** 763–766 (1983).

352. Daub, M. E., Leisman, G. B., Clark, R. A. & Bowden, E. F. Reductive detoxification as a mechanism of fungal resistance to singlet oxygen-generating photosensitizers. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 9588–9592 (1992).

353. Bakırdere, S. *et al.* Aflatoxin species: their health effects and determination methods in different foodstuffs. *Cent. Eur. J. Chem.* **10,** 675–685 (2012).

354. Li, C. *et al.* Photophysical and photosensitive properties of Elsinochrome A. *Chin. Sci. Bull.* **51,** 1050–1054

355. *Chromatography of Mycotoxins: Techniques and Applications*. **54,** (Elsevier, 1993).

356. Sevilla, P., García-Blanco, F., García-Ramos, J. V. & Sánchez-Cortés, S. Aggregation of antitumoral drug emodin on Ag nanoparticles: SEF, SERS and fluorescence lifetime experiments. *Phys. Chem. Chem. Phys.* **11,** 8342–8348 (2009).

357. Gunasinghe, N., You, M. P., Cawthray, G. R. & Barbetti, M. J. Cercosporin from *Pseudocercosporella capsellae* and its critical role in white leaf spot development. *Plant Dis.* **100,** 1521–1531 (2016).

358. Sollod, C. C., Jenns, A. E. & Daub, M. E. Cell surface redox potential as a mechanism of defense against photosensitizers in fungi. *Appl. Environ. Microbiol.* **58,** 444–449 (1992).

359. Dumas, S., Jardon, P., Lepretre, J. C. & Jeunet, A. Electrochemical study of Hypocrellin B in acetonitrile and aqueous micellar media. Comparison with Hypocrellin A. *New J. Chem.* **28,** 91–97 (2004).

360. Schwarz, C. *et al.* Characterization of a genotoxic impact compound in *Alternaria alternata* infested rice as Altertoxin II. *Arch. Toxicol.* **86,** 1911–1925 (2012).

361. Jacquemin, D., Assfeld, X., Preat, J. & Perpete, E. A. Comparison of theoretical approaches for predicting the UV/Vis spectra of anthraquinones. *Mol. Phys.* **105,** 325–331 (2007).

362. Daub, M. E. Cercosporin, a photosensitizing toxin from *Cercospora* species. *Phytopathology* **72,** 370 (1982).

363. Chiang, Y.-M. *et al.* Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl. Environ. Microbiol.* **76,** 2067–2074 (2010).

364. Flaherty, J. E. & Payne, G. A. Overexpression of *aflR* leads to upregulation of pathway gene transcription and increased aflatoxin production in *Aspergillus flavus*. *Appl. Environ. Microbiol.* **63,** 3995–4000 (1997).

365. Quin, M. B., Flynn, C. M. & Schmidt-Dannert, C. Traversing the fungal terpenome. *Nat. Prod. Rep.* **31,** 1449–1473 (2014).

366. Fraga, B. M. Natural sesquiterpenoids. *Nat. Prod. Rep.* **16,** 21–38 (1999).

367. Suno, M. & Nagaoka, A. Inhibition of lipid peroxidation by a novel compound (CV-2619) in brain mitochondria and mode of action of the inhibition. *Biochem. Biophys. Res. Commun.* **125,** 1046–1052 (1984).

368. Rauchová, H. *et al.* Inhibition of glycerophosphate-dependent H_2O_2 generation in brown fat mitochondria by idebenone. *Biochem. Biophys. Res. Commun.* **339,** 362–366 (2006).

369. Sugiyama, Y., Fujita, T., Matsumoto, M., Okamoto, K. & Imada, I. Effects of idebenone (CV-2619) and its metabolites on respiratory activity and lipid peroxidation in brain mitochondria from rats and dogs. *J. Pharmacobiodyn.* **8,** 1006–1017 (1985).

370. Baxter, H., Harborne, J. B. & Moss, G. P. *Phytochemical Dictionary: A Handbook of Bioactive Compounds from Plants, Second Edition*. (CRC Press, 1998).

371. Gosálvez, M., Cañero, R. G. & Blanco, M. A screening for selective anticancer agents among plant respiratory inhibitors. *Eur. J. Cancer 1965* **12,** 1003–1009 (1976).

372. Gonzalez, A. G., Darias, V., Estevez, E. & Vivas, J. M. Chemotherapeutic study of chromones from Spanish Cneoraceae. *Planta Med.* **47,** 56–58 (1983).

373. Choi, Y.-E., Lee, C. & Goodwin, S. B. Generation of reactive oxygen species via *NOXa* is important for development and pathogenicity of *Mycosphaerella graminicola*. *Mycobiology* **44,** 38–47 (2016).

374. Shafran, H., Miyara, I., Eshed, R., Prusky, D. & Sherman, A. Development of new tools for studying gene function in fungi based on the Gateway system. *Fungal Genet. Biol.* **45,** 1147–1154 (2008).