THE EVOLUTION OF SECONDARY METABOLISM AND DEVELOPMENT REGULATION IN THE FUNGAL GENUS *ASPERGILLUS*

By

Abigail Lind

Thesis

Submitted to the Faculty of the Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

MASTER OF SCIENCE

in

Biomedical Informatics

December, 2014

Nashville, Tennessee

Approved:

Antonis Rokas

Patrick Abbot

John A. Capra

ACKNOWLEDGEMENTS

I would like to thank my advisor, Antonis Rokas, for the enormous amounts of support, guidance, and opportunities that he has offered me throughout my time at Vanderbilt. I would also like to thank my committee members, Tony Capra and Patrick Abbot, for their valuable suggestions and guidance on this work.

I would like to thank my co-authors who have made this work possible. I would especially like to thank Jennifer Wisecaver for her consistently excellent suggestions and her insights on visually representing results. My colleagues in the Rokas lab have also provided useful feedback on this work as well as much-appreciated support.

I am grateful to the DBMI faculty, staff, and students whose support and guidance has bene invaluable to me over the past two years. I am grateful to the National Library of Medicine for their support (T15 007450).

Finally, I am grateful to my family and friends for their support and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

SM.....Secondary metabolism

VeA.....Velvet A

MtfA.....Master Transcription Factor A

LaeA.....Loss of AflR Expression A

VosA.....Viability of Spores A

VelB.....Velvet-like B

PKS.....Polyketide synthase

NRPS.....Non-ribosomal peptide synthase

Anid.....*Aspergillus nidulans*

Afum.....*Aspergillus fumigatus*

Aory.....*Aspergillus oryzae*

Aniger.....*Aspergillus niger*

AspGD.....Aspergillus Genomes Database

Ogroup.....Orthogroup

CHAPTER I

INTRODUCTION

The filamentous fungal genus *Aspergillus* is comprised of more than 250 known species with diverse lifestyles and ecologies, including many relevant to human health and industry [1]. *Aspergillus* species, most notably *Aspergillus fumigatus*, can cause a suite of fungal infections known as aspergillosis that can be fatal in immune-compromised individuals [2]. Aflatoxin, a highly hepatotoxic and carcinogenic toxin found in moldy foods, is produced by *Aspergillus flavus* and other *Aspergillus* species. [3]. However, other species of *Aspergillus* are non-toxic and vital for industrial enzyme production and food preparation; *Aspergillus oryzae* ferments sake, miso, and soy sauce, while the "cell factory" *Aspergillus niger* produces many industrially important enzymes and metabolites including citric acid [4,5]. The diversity of these traits reflects the saprotrophic lifestyle that all of these fungi share, where they compete with other fast evolving microbiota in the soil [6,7].

A major source of biodiversity in *Aspergillus* species and other filamentous fungi is a rapidly-evolving secondary metabolism system. Secondary metabolism, also called accessory metabolism, describes the production of small molecules that are not strictly necessary for cell growth and survival [8]. Secondary metabolites (SMs) have a wealth of ecological roles in fungi including territory establishment, fungivory protection, defense, and virulence [9–11]. SMs are remarkably divergent throughout fungal species, where most characterized SMs are only produced by a small number of species. For example, three closely related *Aspergillus* species, *A. fumigatus*, *A. fischerianus*, and *A. clavatus* share 80% of genes but only 30% of SM genes [6,7].

Despite the high divergence of SMs in fungi, many transcriptional regulators that activate or repress expression of multiple SM pathways are broadly conserved throughout filamentous fungi (A. Brakhage 2013). The conservation of these regulators contrasted with the diversity of SM pathways suggests that these regulators have been rewired in different fungi to regulate different SMs. Regulatory circuit rewiring can be driven through structural changes in the regulator itself that may include changes in transcription factor binding sites, changes in the interacting partners of the regulator, or gain or loss of regulatory sequences on different genes [13].

In this study we investigate the interplay between the rapid evolution of SMs and the conservation of SM regulators through genome- and transcriptome-based analyses. First, we quantify the level of SM divergence with respect to both gene cluster structure and gene content in *Aspergillus nidulans, A. fumigatus, A. oryzae,* and *A. niger*. We then examined the roles of two master SM regulators, *veA* and *mtfA*, in *A. fumigatus* and *A. nidulans* through transcriptome-wide gene expression comparison of deletion mutants of each regulator with wild-type strains to build an understanding of each regulator's genome-wide transcriptional control.

CHAPTER II

EXAMINING THE EVOLUTION OF THE REGULATORY CIRCUIT CONTROLLING SECONDARY METABOLISM AND DEVELOPMENT IN THE FUNGAL GENUS *ASPERGILLUS*

Authors

Abigail L. Lind, Jennifer H. Wisecaver, Timothy D. Smith, Xuehuan Feng, Ana M. Calvo, and Antonis Rokas

Introduction

Filamentous fungi produce diverse repertoires of small molecules known as secondary metabolites (SMs) [14]. SMs include widely used pharmaceuticals such the antibiotic penicillin [15], the cholesterol-reducing drug lovastatin [16], and the immunosuppressant cyclosporin [17], as well as potent mycotoxins such as aflatoxin [18] and fumonisin [19,20]. SMs play key ecological roles in territory establishment and defense, communication, and virulence [9,10,21–23].

The genes involved in fungal SM pathways are often physically linked in the genome, forming contiguous SM gene clusters [24]. These gene clusters are typically characterized by a backbone gene, such as those encoding nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs), hybrid NRPS-PKS enzymes, and prenyltransferases (PTRs), whose protein products are responsible for synthesizing the proto-SM. Additional genetic components of SM gene clusters include genes for one or more tailoring enzymes that chemically modify SM

precursors, transporter genes responsible for exporting the final product, and transcription factors that drive expression of additional genes in the cluster. For example, the gene cluster responsible for the synthesis of the mycotoxin gliotoxin in the opportunistic human pathogen *Aspergillus fumigatus* contains 13 genes including a non-ribosomal peptide synthase (*gliP*), multiple tailoring enzymes (*gliI*, *gliJ*, *gliC*, *gliM*, *gliG*, *gliN*, *gliF*), a transporter gene (*gliA*), a transcription factor (*gliZ*), and a gliotoxin oxidase gene that protects the fungus from the harmful effects of gliotoxin (*gliT*) [11,25].

Surveys of fungal genomes show that for any given fungus there are many more gene clusters than known SMs, suggesting that the currently characterized SMs might be only a small fraction of the SMs that a fungus can produce [7,26]. For example, 33 of the 37 putative SM gene clusters in *A. fumigatus* have no characterized products, despite evidence from metabolomics surveys suggesting that the fungus produces many SMs [27–29]. This is likely due to the fact that characterizing the function of SM gene clusters and matching them to specific SMs is a non-trivial task; most SM gene clusters contain many genes and are often only activated under specific ecological conditions such as the availability of different nutrients or the presence of other species [12].

Filamentous fungi exhibit a huge amount of SM biochemical diversity. Individual SMs are often known to be produced by only one or a handful of species, and the SM chemotypic profiles of closely related fungi are typically non-overlapping [8,14]. For example, the meroterpenoid fumagillin, originally isolated from *Aspergillus fumigatus*, has only been detected in *Aspergillus fumigatus* and some isolates of *Penicillium raistrickii* [30,31]. The gene cluster required for its production appears to be conserved in the A. fumigatus close relative, *Aspergillus fischerianus*, though only intermediate compounds have been detected from cultures of this and

other closely related species [27,32,33]. In some genera, including *Aspergillus* [34], the extent of fungal SM distribution is so taxonomically narrow that SM chemotypic profiles have been used as unequivocal species-level identifiers.

As might be expected given their key roles in fungal ecology, SM production – and as a consequence SM gene cluster transcriptional activity – is tightly controlled by a complex network of master SM regulators triggered by a wide variety of environmental cues such as temperature, light, pH, and nutrient availability [12]. Among the master SM regulators identified to date are members of the fungal-specific Velvet protein family, which regulate SM production in response to dark conditions in the model filamentous fungus *Aspergillus nidulans* [35–38]. The founding member of the Velvet family, VeA, stimulates production of diverse types of SMs in various fungal genomes under dark conditions, and has been shown to regulate gliotoxin, fumagillin, fumitgremorgin G, and fumigaclavine C gene cluster expression and metabolite production in *A. fumigatus* [39]. Recently, a VeA-dependent regulator of secondary metabolism, MtfA, was identified in *A. nidulans*, which–unlike VeA–is localized in the nucleus regardless of light conditions [40]. MtfA regulates terrequinone, sterigmatocystin, and penicillin in *A. nidulans*; in *A. fumigatus*, MtfA is necessary for normal protease activity, and virulence assays using the moth *Galleria mellonella* suggest it plays a role in pathogenicity [41].

In addition to regulating SM, both of these regulators have been linked to the regulation of asexual and sexual development. Timing of SM production with developmental changes is well established in filamentous fungi, and the presence/absence of certain SMs has been linked with developmental changes [42–44]. It has been suggested that regulators that coordinate SM and development allow filamentous fungi to have more complex lifestyle and diverse natural

products than their unicellular yeast relatives, which lack *veA* as well as backbone synthesis genes necessary for SM production [44–46].

Remarkably, both *veA* and *mtfA* appear to be broadly conserved in filamentous fungi with non-overlapping SM profiles [40,47]. We used four well-studied organisms from the fungal genus *Aspergillus*, a highly diverse genus and producer of some of the most iconic SMs, including gliotoxin and penicillin, to investigate the evolutionary variability in the distribution of SM gene clusters and its interaction with these two broadly conserved global transcriptional regulators that differ in their response to light, *veA* and *mtfA*. Our evolutionary analyses show that although both the SM gene clusters as well as their gene content are poorly conserved between *Aspergillus* species, explaining the narrow taxonomic distribution and distinctiveness of their SM profiles, the effects of the global transcriptional regulators on SM production in response to environmental cues are largely conserved across these same species. In contrast, examination of the role of *veA* and *mtfA* in development, a process that involves genes that are highly conserved between the two species and whose regulation is intimately linked to SM regulation, yields a very different pattern; whereas the role of *veA* is conserved, *mtfA* regulates development in the homothallic *A. nidulans* but not in the heterothallic *A. fumigatus*.

Methods

Genome sequences and orthogroup definitions for A. nidulans, A. fumigatus, A. oryzae, *and* A. niger

All genome sequences and annotations for *A. nidulans* FGSC A4 s10-m02-r03, *A. fumigatus* AF293 s03-m04-r11, *A. oryzae* RIB40 s01-m08-r21 and *A. niger* CBS 513.88 s01-m06-r10 were taken from the Aspergillus Genomes Database (AspGD) [48]. Orthogroups for

these four genomes were taken from AspGD's orthology assignments for 16 *Aspergillus* species, which were generated using a Jaccard clustering approach [49]. AspGD orthogroups contain groups of genes that are thought to have descended from the *Aspergillus* common ancestor; genes from the same species that are part of a given orthogroup are defined as in-paralogs that have duplicated at some later point after the species diverged from the *Aspergillus* common ancestor. Species-specific genes, which were absent from AspGD orthogroups, were organized into species-specific orthogroups using the MCL algorithm in combination with all-versus-all protein BLAST search [50]. Proteins with BLAST hits with 60% query and subject coverage, an e-value of less than $1e^{-5}$, and a percent identity of greater than 60% were subsequently clustered in MCL with an inflation parameter of 2 and were considered species-specific orthogroups. Proteins that did not pass the BLAST cutoffs were considered single-gene, species-specific orthogroups.

Gene category definitions

Genes involved in secondary metabolism were taken from a previous study that expertly annotated secondary metabolic gene clusters in the four species under study [28]. Manually curated gene cluster boundaries were used when available. Primary metabolism genes were annotated using a previously described enzyme classification pipeline which utilizes KEGG Enzyme Commission annotations [51]. Genes involved in development were determined from all genes in *A. fumigatus* and *A. nidulans* annotated to the GO term "developmental process" (GO:0032502) in AmiGO [52]. This data was accessed on 2014-07-19.

Strains and culture conditions

The strains used in this study include *A. fumigatus* CEA10, TSD1.15(Δ*veA*) and TTDS4.1(Δ*mtfA*) [41,53] and *A. nidulans* TRV50.2 [54], TXFp2.1(Δ*veA*) generated in this study, and TRVpΔ*mtfA* [40]. Many *A. nidulans* studies have used a *veA* partial deletion [55]. For the present study we generated a strain with a complete deletion of the *veA* coding region, TXFp2.1(Δ*veA*). This strain was generated as follows. First, The *veA* deletion cassette was obtained by fusion PCR as previously described [56]. A 1.4 kb 5' UTR and a 1 kb 3' UTR *veA* flanking regions were PCR amplified from wild type FGSC4 genomic DNA with primers veA_comF and AnidveA_p2, and ANVeASTagP3 and ANVeASTagP4 primers sets, respectively. The *A. fumigatus pyrG* ($pyrG^{A.fum}$) selectable marker was amplified with AnidveA_p5 and ANVeASTagP6 primers from plasmid p1439. The 5' and 3' UTR fragments were then PCR fused to $pyrG^{A.fum}$ to generate the *veA* replacement construct using primers AnidveA_P7 and AnidveA_P8. The deletion cassette was transformed into *A. nidulans* RJMP1.49 strain [57]. The resulting transformants were then transformed with the pSM3 plasmid containing the *A. nidulans pyroA* to generate prototrophs, obtaining the Δ*veA* strain. This strain was confirmed by DNA analysis (data not shown) and designated as TXFp2.1.

All strains were grown in liquid stationary cultures in Czapek-Dox medium (Difco) in the dark. The experiments were carried out with two replicates. After 72 hours of incubation at 37^{o}C mycelia samples were harvested, immediately frozen in liquid nitrogen and lyophilized.

RNA extraction

Total RNA was isolated from lyophilized mycelia using the directzol RNA MiniPrep Kit (Zymo) according to the manufacturer's instructions. RNA then was quantified using a nanodrop instrument. Expression patterns of *veA* and *mtfA* were verified in the *A. fumigatus* and *A.*

nidulans wild types as well as in the deletion mutants by qRT-PCR prior to RNA sequencing (not shown), conforming the absence of transcripts in the deletion mutants.

RNA sequencing

RNA-Seq libraries were constructed and sequenced at Vanderbilt Technologies for Advanced Genomics using the Illumina Tru-seq RNA sample prep kit as previously described [39,58,59]. In brief, total RNA quality was assessed via Bioanalyzer (Agilent). Upon passing quality control, poly-A RNA was purified from total RNA and the second strand cDNA was synthesized from mRNA. cDNA ends were then blunt repaired and given an adenylated 3' end. Next, barcoded adapters were ligated to the adenylated ends and the libraries were PCR enriched, quantified, pooled and sequenced an on Illumina HiSeq 2500 sequencer. Two biological replicates were generated for each strain sequenced.

RNA-seq read alignment and differential gene expression

Raw RNA-seq reads were trimmed of low-quality reads and adapter sequences using Trimmomatic using the suggested parameters for single-end read trimming [60]. Trimmed reads were aligned to *A. nidulans* and *A. fumigatus* genomes using Tophat2 using the reference gene annotation to guide alignment and without attempting to detect novel transcripts (parameter –no-novel-juncs) [61]. Reads aligning to each gene were counted using HTSeq-count with the intersection-strict mode [62]. Differential expression between Δ*veA* and WT and Δ*mtfA* and WT strains of *A. fumigatus* and *A. nidulans* were determined using DESeq2 [63]. Genes were considered differentially expressed if their adjusted *P*-value was less than 0.1 and their $\log_2$ fold change was greater than 1 or less than -1.

Statistical analyses

GO term enrichment was determined for over- and under-expressed genes in all four conditions tested (*A. nidulans* and *A. fumigatus* Δ*veA* vs. WT and Δ*mtfA* vs. WT) using the Cytoscape plugin Bingo [64,65]. To allow for a high-level view of the types of differentially expressed gene sets, the Aspergillus GOSlim term subset developed by AspGD was used. The Benjamani-Hochberg multiple testing correction was applied, and terms were considered significantly enriched if the adjusted *P*-value was less than 0.05.

Fisher's exact tests were performed using the R function fisher.test with a two-sided alternative hypothesis [66]. *P*-values were adjusted for multiple comparisons using the R function p.adjust with the Benjamini-Hochberg multiple testing correction [67] Figures were created using the R plotting system ggplot2 [68] and circos [69].

Results

The majority of SM gene clusters in Aspergillus *are species-specific*

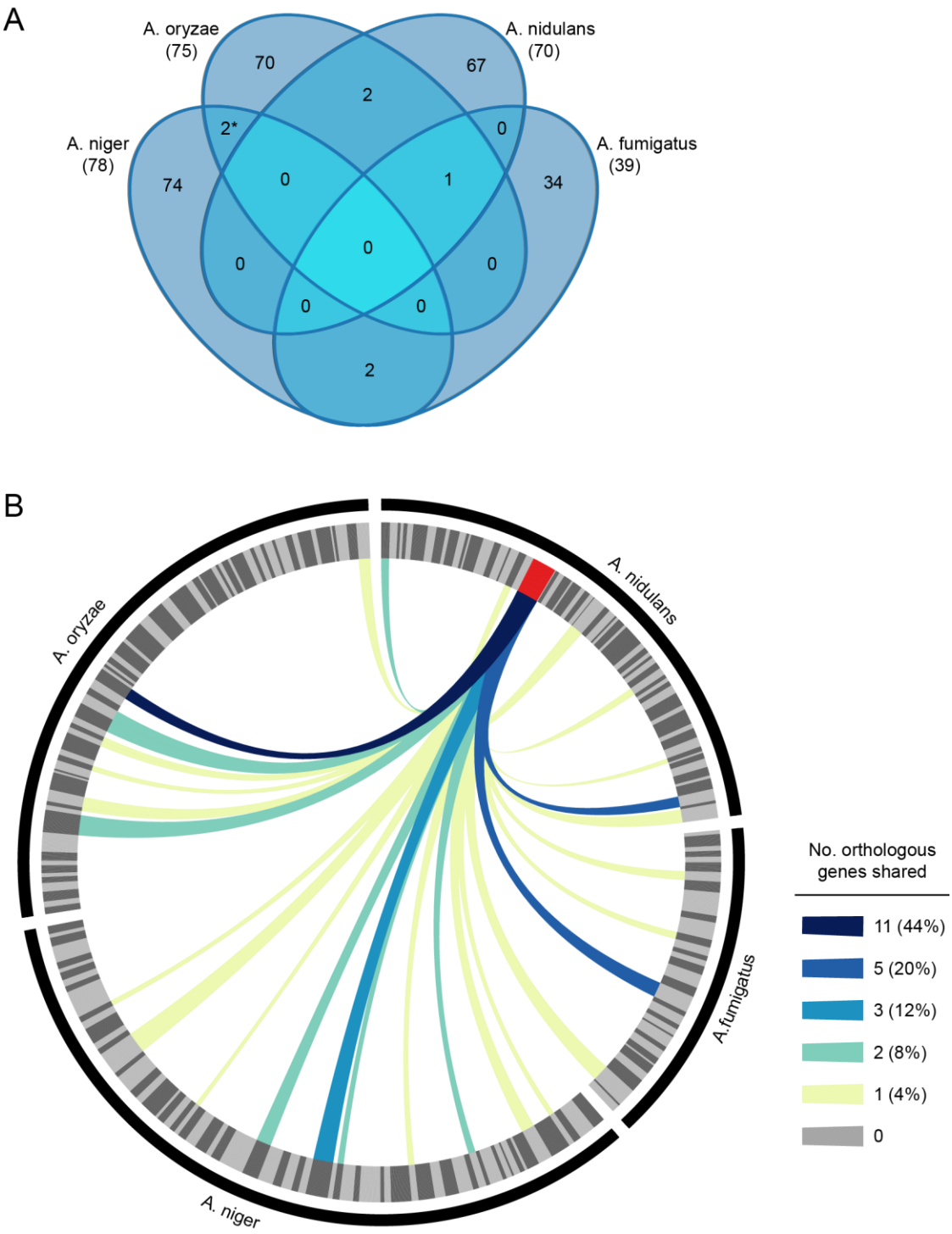

Figure 1. SM gene clusters in *Aspergillus* show minimal evolutionary conservation. A. Venn diagram showing homologous SM gene clusters between *A. fumigatus*, *A. nidulans*, *A. niger*, and *A. oryzae*. Two SM gene clusters were

considered homologous if greater than 50% of their genes were orthologs. Numbers in parenthesis indicate the total number of SM gene clusters present in each species. The asterisk (*) is to clarify that that two SM gene clusters in *A. oryzae* are homologous to one gene cluster in *A. niger*. **B.** Circos plot showing all SM gene clusters in *A. fumigatus*, *A. nidulans*, *A. niger*, and *A. oryzae*. The outer black track shows the relative SM gene counts in each of the four species. SM clusters are indicated by the alternating light and dark grey wedges of the inner track; wedge thickness is proportional to number of clustered genes. The sterigmatocystin gene cluster in *A. nidulans* is colored red. Links indicate SM clusters containing one or more genes assigned to the same orthogroup as gene(s) in the sterigmatocystin gene cluster; link color indicates the number of shared genes.

The genomes of *A. fumigatus*, *A. nidulans*, *A. oryzae*, and *A. niger* contain 317, 498, 725, and 584 secondary metabolic genes, respectively, which are organized in 37, 70, 75, and 78 corresponding secondary metabolic gene clusters [28]. We considered SM gene clusters to be conserved between species if greater than half of the genes in the larger gene cluster were orthologous to greater than half of the genes in the smaller gene cluster. Even with this very liberal definition of gene cluster conservation, we found that no SM gene clusters were conserved across all four species. Moreover, 91.9-96.1% of SM gene clusters were specific to each species, with only 7 SM gene clusters conserved between any species (Figure 1a). As none of these SM gene clusters have chemically characterized products, little can be inferred about the similarity or differences of the products of these conserved gene clusters.

While very few conserved SM gene clusters can be identified between these four species, SM gene clusters do contain genes whose orthologs are parts of other, non-homologous, SM gene clusters. For example, the 25 genes in the sterigmatocystin gene cluster in *A. nidulans*, one of the largest SM gene clusters present in the genomes analyzed, have orthologs in 25 SM gene clusters in the other three species as well as inparalogs in 8 other *A. nidulans* SM gene clusters (Figure 1b). However, in all but one case, less than 20.0% (5 genes) of the sterigmatocystin gene cluster is present in the other gene cluster. The only exception is the truncated aflatoxin gene cluster of *A. oryzae*, which shares 11 orthologs with the ST gene cluster. Although the *A. oryzae* aflatoxin gene cluster is non-functional [58,70,71], the evolutionary conservation between the aflatoxin and sterigmatocystin gene clusters is reflected in the fact that sterigmatocystin is the penultimate precursor product of the aflatoxin biosynthetic pathway [72].

Aspergillus *SM genes are significantly less conserved than genes for primary metabolism*

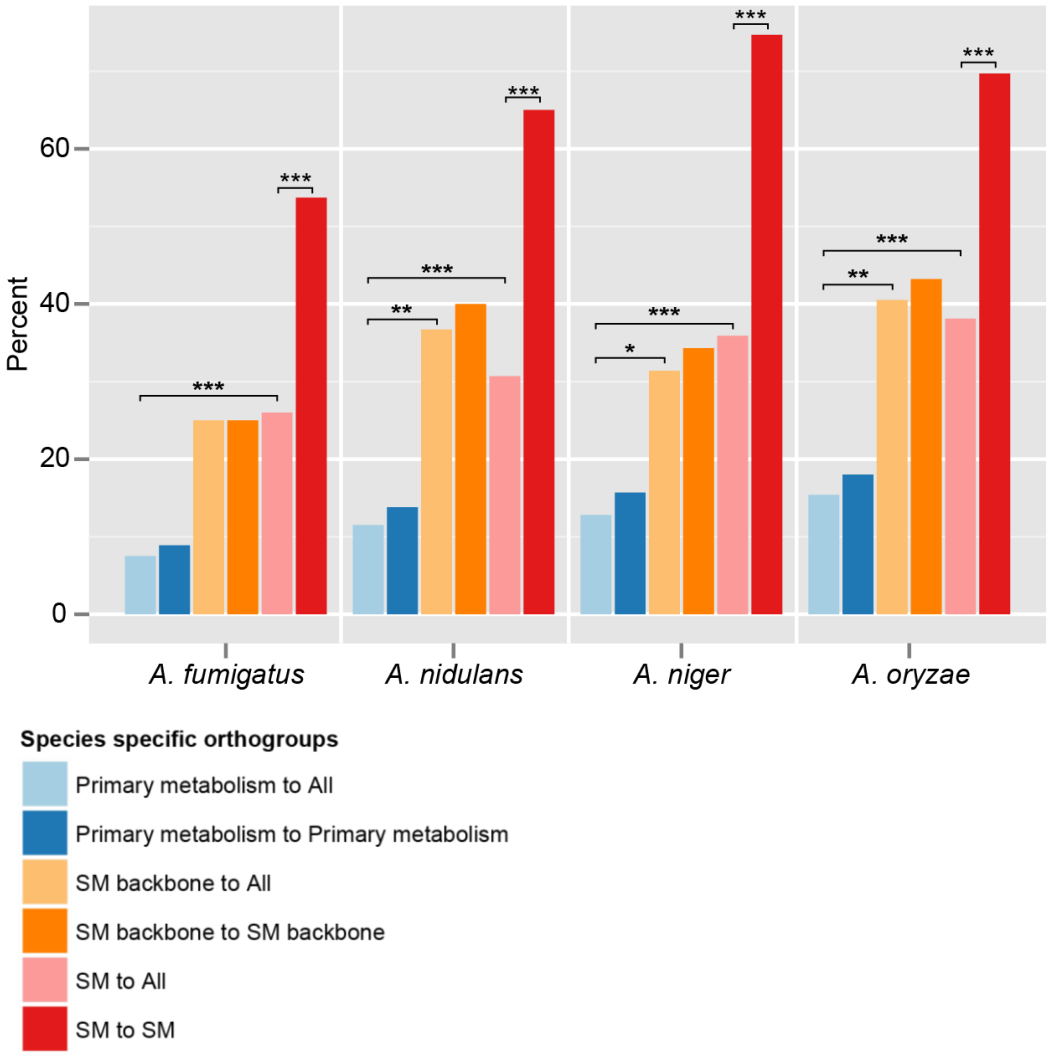

Figure 2 SM genes in *A. fumigatus*, *A. nidulans*, *A. niger*, and *A. oryzae* are less conserved than genes involved in primary metabolism. For each species, dark blue bars indicate the percentage of primary metabolism orthogroups that is species-specific when compared to all genes similarly annotated as participating in primary metabolism in the other three species; light blue bars indicate the percentage of each species' primary metabolism orthogroups that is species-specific compared to all genes, irrespective of their annotation, in the other three species. Similarly, light orange bars indicate the percentage of each species' SM backbone synthesis orthogroups that is species-specific

when compared to all genes similarly annotated as SM backbone synthesis genes in the other three species; dark orange bars indicate the percentage of each species' SM backbone synthesis orthogroups that is species-specific compared to all other genes. Light red bars indicate the percentage of each species' SM orthogroups that is species-specific when compared to all genes similarly annotated as SM genes in the other three species; dark red bars indicate the percentage of each species' SM orthogroups that is species-specific compared to all other genes. Asterisks indicate statistically significant differences based on a *P*-value ≤ 0.01 (*), ≤ 0.001 (**), or ≤ 0.0001 (***) in a two-tailed Fisher's exact test (Appendix C).

To determine the percentage of lineage-specific orthogroups, we determined the number of orthogroups annotated to a particular GOSlim term with at least one gene present in at least one other genome as well as the number of orthogroups with at least one gene annotated to the same functional category in at least one other genome. We found that SM orthogroups were significantly far less conserved than primary metabolic orthogroups in all four genomes examined (adjusted $P < 1e^{-10}$ for all combinations; Appendix C). No more than 18% of primary metabolic orthogroups were lineage-specific in any *Aspergillus* species; this low percentage was observed both for comparisons between just primary metabolic genes as well as across all genes (Figure 2). In contrast, SM orthogroups as well as orthogroups containing just SM backbone genes were more likely to be lineage-specific (Figure 2). For example, in *A. fumigatus,* the smallest genome in our analysis, 8.9% (117/1322) of primary metabolic orthogroups were lineage-specific versus 19.0% (4/21) of SM backbone orthogroups and 26.0% (73/281) of SM orthogroups. Strikingly, genes involved in secondary metabolism in at least 2 genomes were by

far the least conserved in our analysis. Between 53.7% (in *A. fumigatus*, with 37 SM gene clusters) to 74.7% (in *A. niger*, with 77 SM gene clusters) of SM orthogroups had no SM ortholog in any of the other species examined. SM backbone genes were conserved at a similar rates when compared to all genes and all SM backbone genes, which reflects the accurate prediction of polyketide synthase and non-ribosomal peptide synthase genes in the organisms under study.

VeA regulates the same biological processes as well as the same fraction of the genome in both A. nidulans *and* A. fumigatus

Table 1. Differentially expressed genes in Δ*veA* vs. WT and Δ*mtfA* vs. WT comparisons for *A. fumigatus* and *A. nidulans*

Condition	Species	Total dif. expressed[a]	Per. diff. expressed[a]	Under-expressed[a]	Over-expressed[a]
ΔveA	*A. fumigatus*	3,101	31.7%	1,555	1,546
	A. nidulans	2,836	26.5%	1,671	1,165
ΔmtfA	*A. fumigatus*	97	0.9%	63	34
	A. nidulans	968	9.0%	568	400

[a]Number of differentially expressed genes relative to wild type

We next examined the function of the conserved secondary metabolic regulator VeA by performing RNA sequencing [27,58,59] of Δ*veA* and wild-type (WT) *A. fumigatus* strains TSD1.15 [53] and CEA10 and *A. nidulans* strains TXFp2.1 and TRV50.2 [54] and analyzing the data to identify genes that are differentially regulated in Δ*veA* vs WT in the two species. Of the 9,783 transcribed genes in the *A. fumigatus* genome, 1,546 (15.8%) were over-expressed and 1,555 (15.9%) were under-expressed in the Δ*veA* vs WT analysis in *A. fumigatus* (Table 1). We observed very similar numbers of genes differentially regulated in the *A. nidulans* Δ*veA* vs WT analysis; out of 10,709 genes in the *A. nidulans* genome, were 1,165 (10.9%) were over-

expressed and 1,671 genes (15.6%) were under-expressed. In total, approximately 32% and 26% of protein coding genes were differentially regulated in Δ*veA* compared to WT in *A. fumigatus* and *A. nidulans*, respectively.

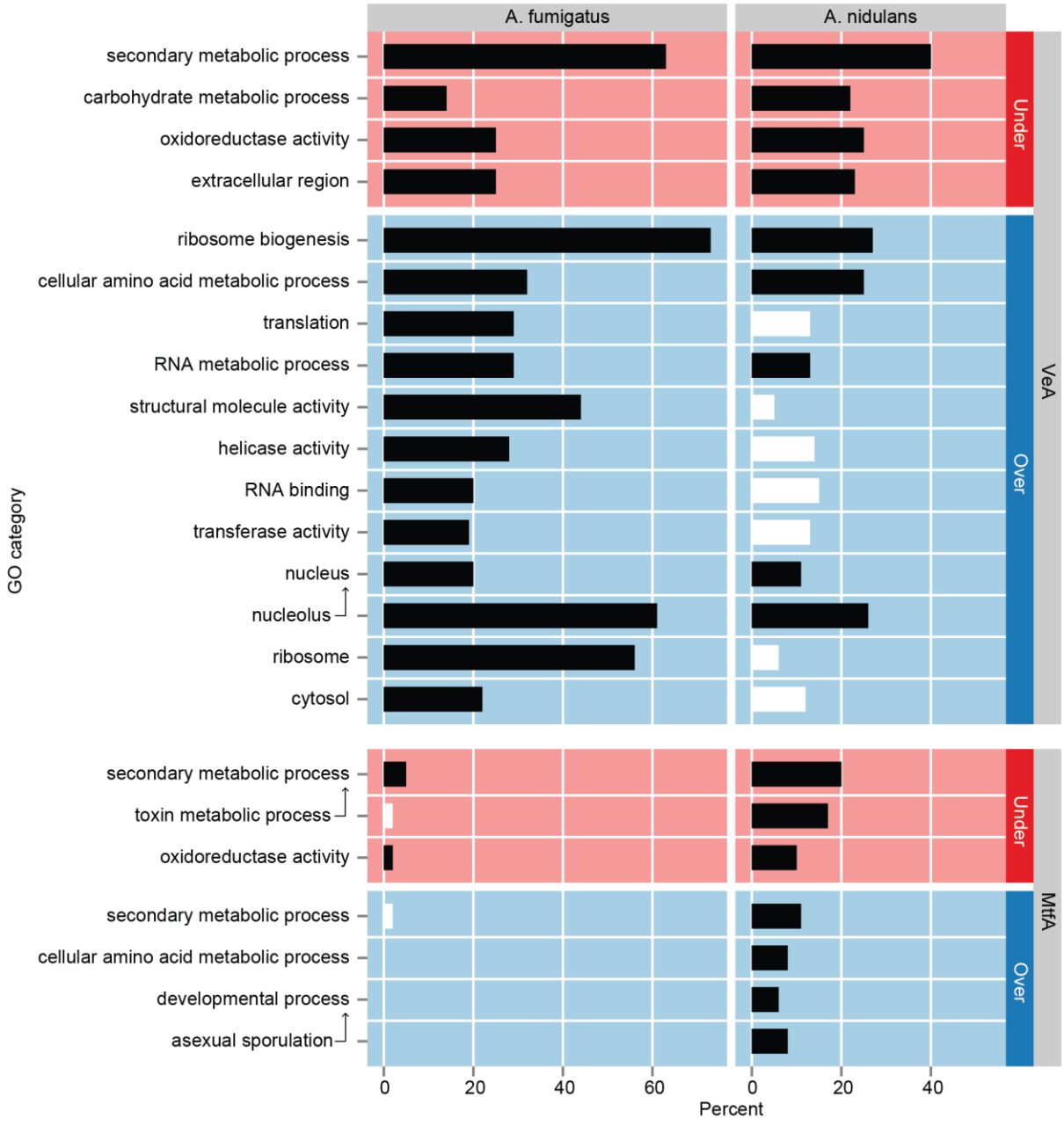

Figure 3 GO term enrichment analysis of genes differentially expressed in Δ*veA* and Δ*mtfA* in *A. fumigatus* and *A. nidulans*. Gene ontology (GO) categories statistically

overrepresented in under-expressed (red) and over-expressed (blue) gene sets in Δ*veA* and Δ*mtfA* relative to wild type. Arrows point to GO term ancestors. Horizontal bars show the percentage of each gene set assigned to a particular GO term with black bars indicating significant enrichment (Benjamini & Hochberg adjusted P-value ≤ 0.05 in a hypergeometric test; Appendix D); white bars indicate no significant enrichment.

To characterize the broad functional categories of these differentially regulated genes, we performed GO term enrichment analysis using the Aspergillus GOSlim term hierarchy [48,73]. Four GO terms, namely SECONDARY METABOLIC PROCESS, CARBOHYDRATE METABOLIC PROCESS, OXIDOREDUCTASE ACTIVITY, and EXTRACELLULAR REGION, are significantly enriched in under-expressed genes in both *A. nidulans* and *A. fumigatus*, showing that VeA is a positive regulator of similar processes in both species (Figure 3). Over-expressed genes in *A. fumigatus* were significantly enriched for twelve GO terms potentially related to cell growth, namely RIBOSOME BIOGENESIS, CELLULAR AMINO ACID METABOLIC PROCESS, TRANSLATION, RNA METABOLIC PROCESS, STRUCTURAL MOLECULE ACTIVITY, HELICASE ACTIVITY, RNA BINDING, TRANSFERASE ACTIVITY, NUCLEUS, NUCLEOLUS, RIBOSOME, and CYTOSOL. Five of these twelve terms were also significantly enriched in *A. nidulans* (RIBOSOME BIOGENESIS, CELLULAR AMINO ACID METABOLIC PROCESS, RNA METABOLIC PROCESS, NUCLEUS, and NUCLEOLUS). Over-expressed genes were present in the remaining seven terms in *A. nidulans* but did not show statistically significant enrichment (Appendix D).

MtfA's regulatory role is smaller in scope in A. fumigatus *compared to* A. nidulans

We next examined the role of the recently identified SM regulator MtfA [40,41] in *A. fumigatus* and *A. nidulans* by performing RNA sequencing and differential gene expression analysis of Δ*mtfA* vs WT strains of both species (*A. fumigatus* tTDS4.1 Δ*mtfA* [41] and CEA10, *A. nidulans* TRVp Δ*mtfA* and TRV50.2 [40]). In contrast to our findings with *veA*, we found a striking difference in the percentage of genes regulated in both species (Table 1). Thirty-six genes were over-expressed (0.4%) and 63 (0.6%) were under-expressed in the *A. fumigatus* Δ*mtfA* vs WT analysis, whereas in the *A. nidulans* Δ*mtfA* vs WT analysis 400 genes were over-expressed (3.7%) and 568 were under-expressed (5.3%).

To determine the functional categories impacted by *mtfA* deletion both species, we performed GO term enrichment analysis on the genes differentially expressed between Δ*mtfA* and WT strains. Under-expressed as well as over-expressed genes in *A. nidulans* were significantly enriched for SECONDARY METABOLIC PROCESS, TOXIN METABOLIC PROCESS and OXIDOREDUCTASE ACTIVITY, suggesting that MtfA is involved in positive and negative regulation of different secondary metabolites (Figure 3). Over-expressed genes in *A. nidulans* were also significantly enriched for asexual developmental processes, namely DEVELOPMENTAL PROCESS and ASEXUAL SPORULATION.

Under-expressed genes in *A. fumigatus* were significantly enriched for two of the three processes as in *A. nidulans,* namely SECONDARY METABOLIC PROCESS and OXIDOREDUCTASE ACTIVITY. However, over-expressed genes in *A. fumigatus* were not significantly enriched for any GO terms; some over-expressed genes were present in the SECONDARY METABOLIC PROCESS term, though this was not statistically significant (Appendix D).

Regulation of similar processes regardless of gene conservation

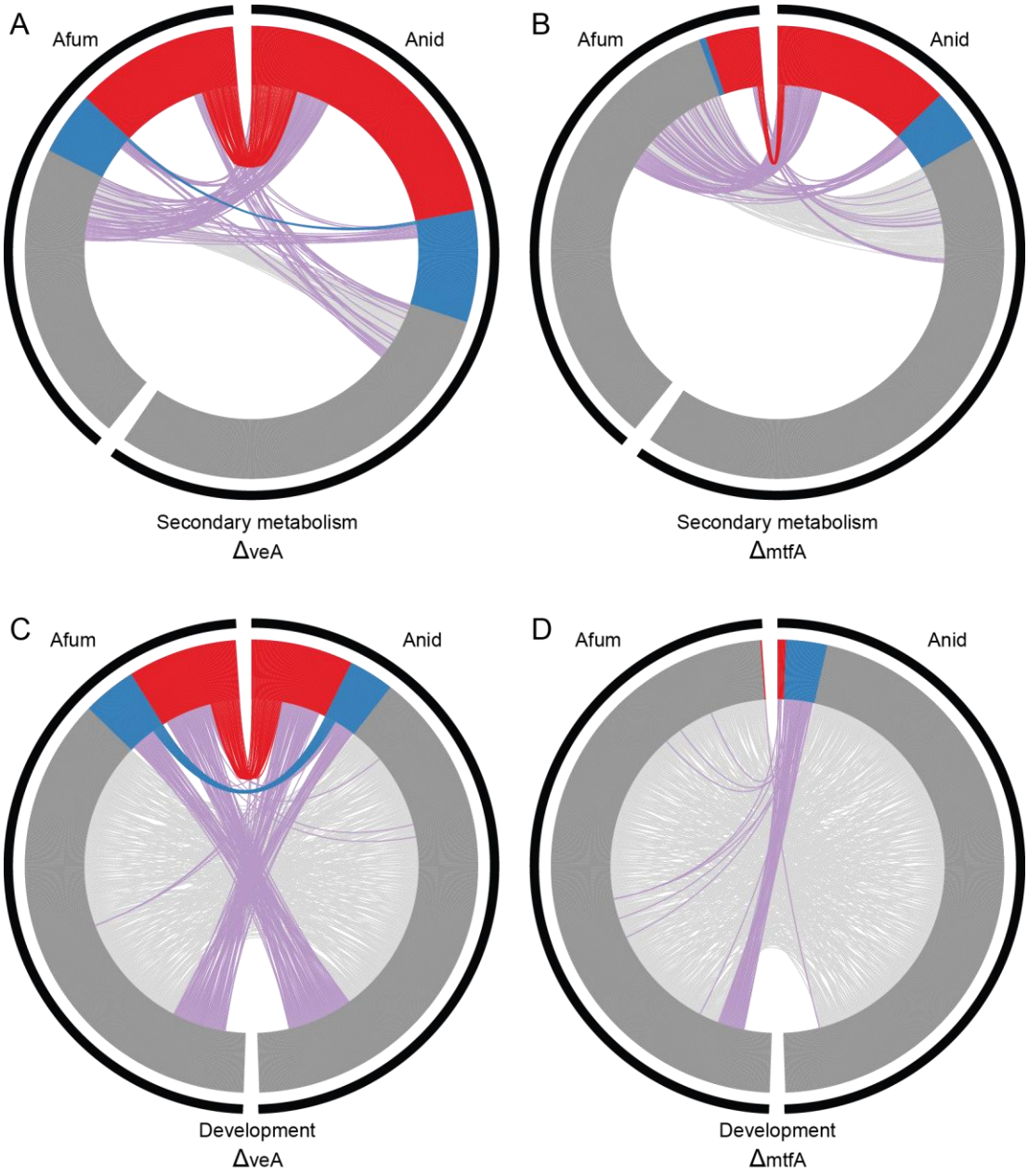

Figure 4 Orthology of SM and development genes differentially expressed in Δ*veA* and Δ*mtfA* in *A. fumigatus* and *A. nidulans*. Circos plots of SM genes **(A,B)** and developmental genes **(C,D)** showing change in gene expression patterns under Δ*veA* **(A,C)** and Δ*mtfA* **(B,D)** conditions. Outer black track shows the relative gene counts in *A. nidulans* (right) and *A. fumigatus* (left). Inner track shows the relative number of under-expressed genes (red), over-expressed genes (blue) and not differentially expressed genes

(grey). Links indicate orthologous genes between the two species that are both under-expressed (red links), both over-expressed (blue links) and both not differentially expressed (light grey links); purple links indicate that the orthologous genes have conflicting expression patterns.

To examine whether SM gene conservation correlated with conservation of regulation by VeA and MtfA, we examined whether orthologous genes in *A. nidulans* and *A. fumigatus* showed the same responses in Δ*veA* vs WT and Δ*mtfA* vs WT analyses (Figure 4). SM gene expression in Δ*veA A. nidulans* and *A. fumigatus* was similar in terms of numbers of differentially expressed genes despite the large amount of genes without orthologs between these two species (Figure 4a; Figure 1a). Of the 184 under-expressed SM genes in Δ*veA A. nidulans*, only 64 genes (34.8%) had an ortholog in *A. fumigatus* (Appendix E). Of these 64 conserved genes, 45 (70.3%) had at least one differentially expressed ortholog in *A. fumigatus*, and 37 (57.8%) had at least one similarly under-expressed ortholog in *A. fumigatus* (Appendix E). Fewer SM genes were over-expressed in either Δ*veA A. nidulans* or *A. fumigatus*; of the 67 over-expressed genes in *A. nidulans*, 14 (20.9%) had orthologs in *A. fumigatus*. Of these 14 conserved genes, 5 had at least one differentially expressed ortholog in *A. fumigatus*, and 3 had at least one similarly over-expressed ortholog in *A. fumigatus* (Appendix E).

When *mtfA* was deleted, fewer SM genes were differentially expressed in *A. fumigatus* than in *A. nidulans.* Of the 107 under-expressed genes in Δ*mtfA A. nidulans*, 36 (33.6%) had an ortholog in *A. fumigatus* (Figure 4b; Appendix E). Unlike *veA*, however, only 6 of these conserved genes had differentially expressed orthologs in *A. fumigatus*. Finally, of the 32 over-

expressed genes in Δ*mtfA A. nidulans*, 2 of the 12 genes with orthologs in *A. fumigatus* had orthologs that were differentially expressed.

Apart from their involvement in the global regulation of SM, both VeA and MtfA are also involved in the regulation of asexual and sexual development. In contrast to genes involved in SM, genes involved in asexual and sexual development in *Aspergillus* have been shown to be highly conserved across the genus [4]. Of the 490 genes annotated to the GO term DEVELOPMENTAL PROCESS, 462 have at least one ortholog among the 478 genes annotated to this term in *A. fumigatus.* In Δ*veA* A. *nidulans*, 72 developmental genes are under-expressed and 32 are over-expressed. Of the 72 under-expressed genes, 66 (91.7%) have an ortholog in *A. fumigatus*, 30 of which are have a differentially expressed ortholog (Figure 4c; Appendix E). There are fewer over-expressed developmental genes in Δ*veA A. nidulans*, but they show similar trends; 31 of the 32 over-expressed genes have an ortholog, 15 of which have differentially expressed orthologs in *A. fumigatus* and 11 of which have over-expressed orthologs. In contrast with *veA*, many more developmental genes were differentially expressed in *A. nidulans* Δ*mtfA* (35) than in *A. fumigatus* Δ*mtfA* (1). While 4 of the 6 under-expressed genes and 28 of the 29 over-expressed genes in *A. nidulans* had orthologs in *A. fumigatus*, none of these orthologs were differentially expressed (Figure 4d; Appendix E).

Discussion

Here, we examined the interplay between secondary metabolites with narrow taxonomic distributions and their broadly conserved SM global regulators in four *Aspergillus* species. We found remarkably few conserved SM gene clusters *A. fumigatus*, *A. nidulans*, *A. oryzae*, and *A. niger* (Figure 1a). Further, the genes comprising these clusters were significantly more species-

specific than genes involved in primary metabolism (Figure 2), and those SM genes that were conserved were assigned to non-homologous pathways (e.g., genes in the sterigmatocystin gene cluster; Figure 1b). Despite the high level of divergence in SM pathways, regulators of SM production are conserved throughout filamentous fungi [12]. We assessed the conservation of roles of two of these regulators, *veA* and *mtfA*, in *A. fumigatus* and *A. nidulans* by comparing genome-wide gene expression of deletion mutants of *veA* and *mtfA* in both species with wild-type strains. We found that the role of *veA* in controlling secondary metabolism and development was conserved in both species (Figure 3), though the regulated genes were often different (Figure 4). In contrast, we found that while deleting *mtfA* negatively impacted SM gene expression in both *A. fumigatus* and *A. nidulans*, developmental genes were only impacted in *A. nidulans* (Figure 3).

Aspergillus *secondary metabolic genes and gene clusters are largely species-specific*

Previous studies have described relatively small numbers of conserved SM gene clusters between the closely related species *Aspergillus fumgiatus*, *fisherianus*, and *clavatus* [7]. These studies considered SM gene clusters to be conserved if 80% or more of their genes were shared. Since the four *Aspergillus* species under study here are much less closely related [6], we used a lower threshold of 50% for conservation. Even using this relaxed threshold, we found no clusters that were conserved in all four species, one cluster conserved in three species, and a small number conserved between pairs of species (Figure 1a). SM gene clusters have been described as not only species specific but sometimes strain specific; two isolates of *A. fumigatus* differ in one putative SM gene cluster [74], and multiple SM gene clusters vary in their presence and absence in isolates of *A. niger* [75].

In addition to the non-conservation of SM gene clusters, we found that the genes comprising SM genes are much more likely to be lineage specific than those involved in primary metabolic functions (Figure 2). Surprisingly, many genes involved in SM clusters had orthologs in other *Aspergilli* that were not in an SM gene cluster themselves. This observation may support the hypothesis that SM gene clusters can be formed or altered by incorporating non-SM genes through genomic rearrangements and possible co-regulation with other genes in the biosynthetic pathway [76]. However, many of the SM genes were species-specific even when compared against all genes in the other four *Aspergilli* (Figure 1), and as many as 21.7% of SM genes were not present in any other sequenced *Aspergillus* species. This high number of species-specific genes involved in *Aspergilli* may be explained by extensive gene duplication and loss, *de novo* gene emergence, very high sequence divergence driven by selection, or horizontal gene transfer.

The evolution of the circuit regulating secondary metabolism and development

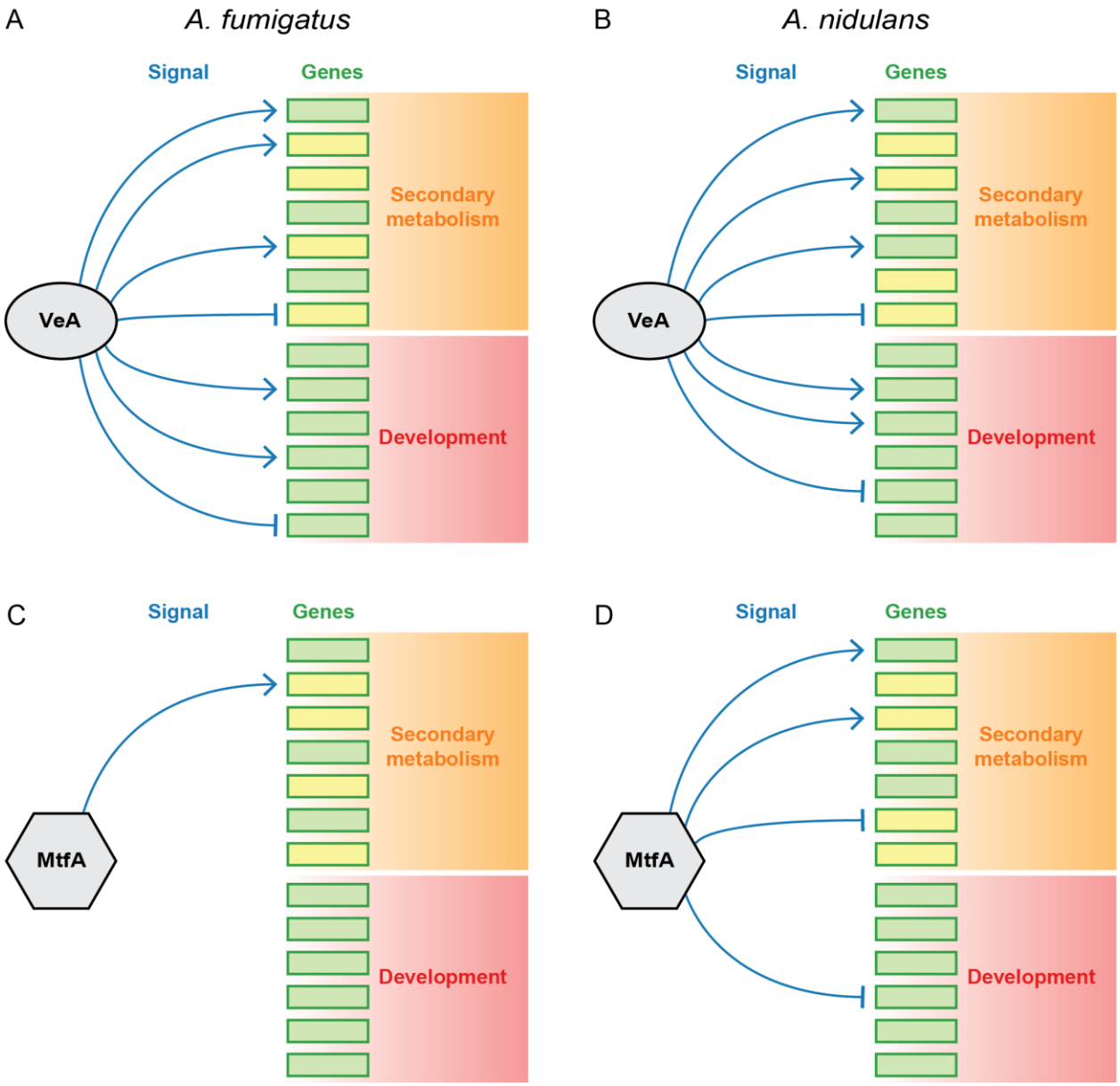

Figure 5. Model of gene regulatory network evolution in *Aspergillus*. Generalized gene regulatory networks for VeA **(A,B)** and MtfA **(C,D)** in *A. fumigatus* **(A,C)** and *A. nidulans* **(B,D)**. As master transcriptional regulators, VeA and MtfA can both promote the expression of a gene (indicated by an arrow →), or inhibit gene expression (indicated by a bar ⊣). Target genes are either present in both species (green fill) or species-specific (yellow fill).

Our findings with respect to the genes differentially expressed in the absence of *veA* support a conserved role for *veA* in regulating secondary metabolism and development in *A. fumigatus* and *A. nidulans*; however, the downstream genes regulated by VeA are different between the two species. Whether lineage-specific or conserved, SM genes are differentially expressed in *veA*'s absence in both *A. fumigatus* and *A. nidulans.* Interestingly, conserved genes differentially expressed in one species are often not differentially expressed in the other. We propose that the transcriptional circuit by which VeA regulates secondary metabolism and development in both fungi has diverged at the level of both the target genes and the regulatory signal (Figure 5). VeA is known to have many interacting partners [44]; among these, it is responsible for transporting the Velvet family protein VelB from the cytoplasm to the nucleus, where both proteins interact with LaeA, forming a trimeric complex that regulates secondary metabolism production and development [35]. However, each protein has functions in the cell outside of this complex, which can be seen through the different effects of individual gene deletion on gene expression and morphogenesis [35,39,77]. VeA interacts with red light-sensing proteins in the nucleus, and it is speculated that VeA may act as a scaffold protein recruiting additional transcriptional regulators [47]. Finally, recent analysis has shown that the Velvet domain is a DNA-binding domain, and that Velvet family proteins may act as direct transcriptional regulators [78]. The number and complexity of VeA's interacting partners, and its putative transcription factor function, offers many degrees of freedom for changes in specific gene regulation in both *A. fumigatus* and *A. nidulans*, while preserving its important ecological role in coordinating sexual development and secondary metabolism production in response to dark conditions.

Much less is known about the regulatory partners of the putative C2H2 zinc finger transcription factor MtfA; however, we suggest that MtfA acts downstream of VeA in *A. nidulans*, but not in *A. fumigatus*, as its expression is decreased in Δ*veA A. nidulans* but not Δ*veA A. fumigatus* . Our results show that MtfA is involved in regulating secondary metabolism in both *A. fumigatus* and *A. nidulans*, though it regulates fewer clusters in *A. fumigatus* than in *A. nidulans*. As was the case with *veA*, deleting *mtfA* results in differentially expressed lineage-specific and conserved SM genes, though differentially expressed conserved genes were not necessarily differentially expressed in the other species, indicating a divergence in the signal that targets these genes for regulation by MtfA or its interacting partners. Unlike *veA*, however, deleting *mtfA* resulted in developmental gene expression changes exclusively in *A. nidulans*, suggesting that there has been a loss in *A. fumigatus* or gain in *A. nidulans* of the regulatory signal that directs MtfA or its downstream targets to regulate developmental processes. Taken together, our results suggest extensive rewiring in the regulatory circuit governing secondary metabolism and development between *A. nidulans* and *A. fumigatus*.

CHAPTER III

SUMMARY

This study enhances our understanding of the scope of regulatory rewiring in *Aspergillus*. The well-conserved protein VeA, which has been linked to developmental changes and SM production in many species, regulates genes involved in SM and development in both *A. nidulans* and *A. fumigatus*, but specific genes that are regulated in each species differ dramatically. While SM genes are very divergent between *A. fumigatus* and *A. nidulans*, developmental genes are well conserved, suggesting that the change in VeA's gene regulation between these two species has happened through extensive regulatory rewiring that maintains the biological processes under VeA's control. Analysis of the regulator MtfA revealed that it is entirely decoupled from regulating developmental processes in *A. fumigatus*, while it regulates many developmental genes in *A. nidulans*. MtfA is involved in regulating SM genes in both species, though fewer genes are impacted by its deletion in *A. fumigatus* than in *A. nidulans*. Finally, *mtfA* is under-expressed Δ*veA A. nidulans* but not in *A. fumigatus*, suggesting that MtfA acts downstream of VeA in *A. nidulans* but not *A. fumigatus*. These results taken together suggest that the regulatory circuit governed by *veA* has undergone extensive rewiring that has changed which genes are regulated but preserves regulation of biological processes, and that this may be proceeding through changes in VeA's interacting partners.

Our understanding of regulatory rewiring in *Aspergillus* could be further refined by adding species to the analysis presented here. *A. fumigatus* and *A. nidulans* are relatively distantly related organisms, and including a more closely related species (such as *A. fischerianus*

or *A. clavatus*, close relatives of *A. fumigatus*) in this analysis would shed light on the rate of regulator rewiring. Additional possibilities for future work on comparative gene regulation in *Aspergillus* would be examining the interplay of different environmental factors (such as pH or temperature) with *veA* and *mtfA*, as well as examining gene expression changes in different fungal tissue types.

REFERENCES

1. Geiser DM, Klich MA, Frisvad JC, Peterson SW, Varga J, et al. (2007) The current status of species recognition and identification in Aspergillus. Stud Mycol 59: 1–10. doi:10.3114/sim.2007.59.01.

2. Patterson TF, Kirkpatrick WR, White M, Hiemenz JW, Wingard JR, et al. (2000) Invasive aspergillosis. Disease spectrum, treatment practices, and outcomes. Medicine (Baltimore) 79: 250–260.

3. Eaton DL, Gallagher EP (1994) Mechanisms of aflatoxin carcinogenesis. Annu Rev Pharmacol Toxicol 34: 135–172. doi:10.1146/annurev.pa.34.040194.001031.

4. Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, et al. (2005) Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae. Nature 438: 1105–1115. doi:10.1038/nature04341.

5. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, et al. (2007) Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88. Nat Biotechnol 25: 221–231. doi:10.1038/nbt1282.

6. Gibbons JG, Rokas A (2012) The function and evolution of the Aspergillus genome. Trends Microbiol 21: 14–22. doi:10.1016/j.tim.2012.09.005.

7. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, et al. (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol 47: 736–741. doi:10.1016/j.fgb.2010.06.003.

8. Bennett J, Bentley R (1989) What's in a name?—Microbial secondary metabolism. Adv Appl Microbiol 34.

9. Rohlfs M, Albert M, Keller NP, Kempken F (2007) Secondary chemicals protect mould from fungivory. Biol Lett 3: 523–525. doi:10.1098/rsbl.2007.0338.

10. Yim G, Wang HH, Davies J (2007) Antibiotics as signalling molecules. Philos Trans R Soc Lond B Biol Sci 362: 1195–1200. doi:10.1098/rstb.2007.2044.

11. Scharf DH, Heinekamp T, Remme N, Hortschansky P, Brakhage A a, et al. (2012) Biosynthesis and function of gliotoxin in Aspergillus fumigatus. Appl Microbiol Biotechnol 93: 467–472. doi:10.1007/s00253-011-3689-1.

12. Brakhage A a (2013) Regulation of fungal secondary metabolism. Nat Rev Microbiol 11: 21–32. doi:10.1038/nrmicro2916.

13. Tuch BB, Li H, Johnson AD (2008) Evolution of eukaryotic transcription circuits. Science (80-) 319: 1797–1799. doi:10.1126/science.1152398.

14. Keller NP, Turner G, Bennett JW (2005) Fungal secondary metabolism - from biochemistry to genomics. Nat Rev Microbiol 3: 937–947. doi:10.1038/nrmicro1286.

15. Brakhage AA (1998) Molecular Regulation of beta -Lactam Biosynthesis in Filamentous Fungi. Microbiol Mol Biol Rev 62: 547–585.

16. Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, et al. (1999) Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science 284: 1368–1372.

17. Weber G, Schörgendorfer K, Schneider-Scherzer E, Leitner E (1994) The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame. Curr Genet 26: 120–125. doi:10.1007/BF00313798.

18. Chang PK, Ehrlich KC, Yu J, Bhatnagar D, Cleveland TE (1995) Increased expression of Aspergillus parasiticus aflR, encoding a sequence-specific DNA-binding protein, relieves nitrate inhibition of aflatoxin biosynthesis. Appl Environ Microbiol 61: 2372–2377.

19. Brown DW, Butchko RAE, Busman M, Proctor RH (2007) The Fusarium verticillioides FUM gene cluster encodes a Zn(II)2Cys6 protein that affects FUM gene expression and fumonisin production. Eukaryot Cell 6: 1210–1218. doi:10.1128/EC.00400-06.

20. Proctor RH, Brown DW, Plattner RD, Desjardins AE (2003) Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in Gibberella moniliformis. Fungal Genet Biol 38: 237–249.

21. Vining LC (1990) Functions of secondary metabolites. Annu Rev Microbiol 44: 395–427. doi:10.1146/annurev.mi.44.100190.002143.

22. Demain AL, Fang A (2000) The natural functions of secondary metabolites. Adv Biochem Eng Biotechnol 69: 1–39.

23. Rohlfs M, Churchill ACL (2011) Fungal secondary metabolites as modulators of interactions with insects and other arthropods. Fungal Genet Biol 48: 23–34. doi:10.1016/j.fgb.2010.08.008.

24. Keller NP, Hohn TM (1997) Metabolic Pathway Gene Clusters in Filamentous Fungi. Fungal Genet Biol 21: 17–29.

25. Gardiner DM, Howlett BJ (2005) Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of Aspergillus fumigatus. FEMS Microbiol Lett 248: 241–248. doi:10.1016/j.femsle.2005.05.046.

26. Hoffmeister D, Keller NP (2007) Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat Prod Rep 24: 393–416. doi:10.1039/b603084j.

27. Lin H-C, Chooi Y-H, Dhingra S, Xu W, Calvo AM, et al. (2013) The Fumagillin Biosynthetic Gene Cluster in Aspergillus fumigatus Encodes a Cryptic Terpene Cyclase Involved in the Formation of β-trans-Bergamotene. J Am Chem Soc 135: 4616–4619. doi:10.1021/ja312503y.

28. Inglis DO, Binkley J, Skrzypek MS, Arnaud MB, Cerqueira GC, et al. (2013) Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of Aspergillus nidulans, A. fumigatus, A. niger and A. oryzae. BMC Microbiol 13: 91. doi:10.1186/1471-2180-13-91.

29. Frisvad JC, Rank C, Nielsen KF, Larsen TO (2009) Metabolomics of Aspergillus fumigatus. Med Mycol 47 Suppl 1: S53–S71. doi:10.1080/13693780802307720.

30. Christensen M, Frisvad JC, Tuthill D (1999) Taxonomy of the Penicillium miczynskii group based on morphology and secondary metabolites. Mycol Res 103: 527–541. doi:10.1017/S0953756298007515.

31. McCowen MC, Callender ME, Lawlis JF (1951) Fumagillin (H-3), a new antibiotic with amebicidal properties. Science 113: 202–203.

32. Asami Y, Kakeya H, Onose R, Chang Y-H, Toi M, et al. (2004) RK-805, an endothelial-cell-growth inhibitor produced by Neosartorya sp., and a docking model with methionine aminopeptidase-2. Tetrahedron 60: 7085–7091. doi:10.1016/j.tet.2003.09.104.

33. Wiemann P, Guo C-J, Palmer JM, Sekonyela R, Wang CCC, et al. (2013) Prototype of an intertwined secondary-metabolite supercluster. Proc Natl Acad Sci U S A 110: 17065–17070. doi:10.1073/pnas.1313258110.

34. Frisvad JC, Andersen B, Thrane U (2008) The use of secondary metabolite profiling in chemotaxonomy of filamentous fungi. Mycol Res 112: 231–240. doi:10.1016/j.mycres.2007.08.018.

35. Bayram Ö, Krappmann S, Ni M, Bok JW, Helmstaedt K, et al. (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science (80-) 320: 1504–1506. doi:10.1126/science.1155888.

36. Kato N, Brooks W, Calvo AM (2003) The expression of sterigmatocystin and penicillin genes in Aspergillus nidulans is controlled by veA, a gene required for sexual development. Eukaryot Cell 2: 1178–1186.

37. Stinnett SM, Espeso EA, Cobeño L, Araújo-Bazán L, Calvo AM (2007) Aspergillus nidulans VeA subcellular localization is dependent on the importin alpha carrier and on light. Mol Microbiol 63: 242–255. doi:10.1111/j.1365-2958.2006.05506.x.

38. Purschwitz J, Müller S, Kastner C, Schöser M, Haas H, et al. (2008) Functional and physical interaction of blue- and red-light sensors in Aspergillus nidulans. Curr Biol 18: 255–259. doi:10.1016/j.cub.2008.01.061.

39. Dhingra S, Lind AL, Lin H-C, Tang Y, Rokas A, et al. (2013) The fumagillin gene cluster, an example of hundreds of genes under veA control in Aspergillus fumigatus. PLoS One 8: e77147. doi:10.1371/journal.pone.0077147.

40. Ramamoorthy V, Dhingra S, Kincaid A, Shantappa S, Feng X, et al. (2013) The Putative C2H2 Transcription Factor MtfA Is a Novel Regulator of Secondary Metabolism and Morphogenesis in Aspergillus nidulans. PLoS One 8: e74122. doi:10.1371/journal.pone.0074122.

41. Smith TD, Calvo AM (2014) The mtfA Transcription Factor Gene Controls Morphogenesis, Gliotoxin Production, and Virulence in the Opportunistic Human Pathogen Aspergillus fumigatus. Eukaryot Cell 13: 766–775. doi:10.1128/EC.00075-14.

42. Tsai HF, Chang YC, Washburn RG, Wheeler MH, Kwon-Chung KJ (1998) The developmentally regulated alb1 gene of Aspergillus fumigatus: its role in modulation of conidial morphology and virulence. J Bacteriol 180: 3031–3038.

43. Calvo AM, Wilson RA, Bok JW, Keller NP (2002) Relationship between secondary metabolism and fungal development. Microbiol Mol Biol Rev 66: 447–459, table of contents.

44. Calvo AM (2008) The VeA regulatory system and its role in morphological and chemical development in fungi. Fungal Genet Biol 45: 1053–1061. doi:10.1016/j.fgb.2008.03.014.

45. Roze L V, Chanda A, Linz JE (2011) Compartmentalization and molecular traffic in secondary metabolism: a new understanding of established cellular processes. Fungal Genet Biol 48: 35–48. doi:10.1016/j.fgb.2010.05.006.

46. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci U S A 100: 15670–15675. doi:10.1073/pnas.2532165100.

47. Bayram O, Braus GH (2012) Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. FEMS Microbiol Rev 36: 1–24. doi:10.1111/j.1574-6976.2011.00285.x.

48. Arnaud MB, Chibucos MC, Costanzo MC, Crabtree J, Inglis DO, et al. (2010) The Aspergillus Genome Database, a curated comparative genomics resource for gene, protein and sequence information for the Aspergillus research community. Nucleic Acids Res 38: D420–D427. doi:10.1093/nar/gkp751.

49. Crabtree J, Angiuoli S V, Wortman JR, White OR (2007) Sybil: methods and software for multiple genome comparison and visualization. Methods Mol Biol 408: 93–108. doi:10.1007/978-1-59745-547-3_6.

50. Enright AJ, Van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res 30: 1575–1584.

51. Wisecaver JH, Slot JC, Rokas A (2014) The Evolution of Fungal Metabolic Pathways. PLoS Genet.

52. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, et al. (2009) AmiGO: online access to ontology and annotation data. Bioinformatics 25: 288–289. doi:10.1093/bioinformatics/btn615.

53. Dhingra S, Andes D, Calvo AM (2012) VeA Regulates Conidiation, Gliotoxin Production and Protease Activity in the Opportunistic Human Pathogen Aspergillus fumigatus. Eukaryot Cell 11: 1531–1543. doi:10.1128/EC.00222-12.

54. Ramamoorthy V, Shantappa S, Dhingra S, Calvo AM (2012) veA-dependent RNA-pol II transcription elongation factor-like protein, RtfA, is associated with secondary metabolism and morphological development in Aspergillus nidulans. Mol Microbiol 85: 795–814. doi:10.1111/j.1365-2958.2012.08142.x.

55. Kim H-S, Han K-Y, Kim K-J, Han D-M, Jahng K-Y, et al. (2002) The veA gene activates sexual development in Aspergillus nidulans. Fungal Genet Biol 37: 72–80. doi:10.1016/S1087-1845(02)00029-4.

56. Szewczyk E, Nayak T, Oakley CE, Edgerton H, Xiong Y, et al. (2006) Fusion PCR and gene targeting in Aspergillus nidulans. Nat Protoc 1: 3111–3120. doi:10.1038/nprot.2006.405.

57. Shaaban M, Palmer JM, El-Naggar WA, El-Sokkary MA, Habib E-SE, et al. (2010) Involvement of transposon-like elements in penicillin gene cluster regulation. Fungal Genet Biol 47: 423–432. doi:10.1016/j.fgb.2010.02.006.

58. Gibbons JG, Salichos L, Slot JC, Rinker DC, McGary KL, et al. (2012) The evolutionary imprint of domestication on genome variation and function of the filamentous fungus Aspergillus oryzae. Curr Biol 22: 1403–1409. doi:10.1016/j.cub.2012.05.033.

59. Gibbons JG, Beauvais A, Beau R, McGary KL, Latgé J-P, et al. (2012) Global transcriptome changes underlying colony growth in the opportunistic human pathogen Aspergillus fumigatus. Eukaryot Cell 11: 68–78. doi:10.1128/EC.05102-11.

60. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics: btu170 – . doi:10.1093/bioinformatics/btu170.

61. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14: R36. doi:10.1186/gb-2013-14-4-r36.

62. Anders S, Pyl PT, Huber W (2014) HTSeq - A Python framework to work with high-throughput sequencing data. bioRxiv. doi:10.1101/002824.

63. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Cold Spring Harbor Labs Journals. doi:10.1101/002832.

64. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498–2504. doi:10.1101/gr.1239303.

65. Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21: 3448–3449. doi:10.1093/bioinformatics/bti551.

66. R Core Team (2014) R: A Language and Environment for Statistical Computing.

67. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57: pp. 289–300.

68. Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer New York.

69. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19: 1639–1645. doi:10.1101/gr.092759.109.

70. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, et al. (2005) Genome sequencing and analysis of Aspergillus oryzae. Nature 438: 1157–1161. doi:10.1038/nature04300.

71. Kiyota T, Hamada R, Sakamoto K, Iwashita K, Yamada O, et al. (2011) Aflatoxin non-productivity of Aspergillus oryzae caused by loss of function in the aflJ gene product. J Biosci Bioeng 111: 512–517. doi:10.1016/j.jbiosc.2010.12.022.

72. Brown DW, Yu JH, Kelkar HS, Fernandes M, Nesbitt TC, et al. (1996) Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in Aspergillus nidulans. Proc Natl Acad Sci U S A 93: 1418–1422.

73. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29. doi:10.1038/75556.

74. Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, et al. (2008) Genomic islands in the pathogenic filamentous fungus Aspergillus fumigatus. PLoS Genet 4: e1000046. doi:10.1371/journal.pgen.1000046.

75. Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJI, Culley D, et al. (2011) Comparative genomics of citric-acid-producing Aspergillus niger ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21: 885–897. doi:10.1101/gr.112169.110.

76. Wong S, Wolfe KH (2005) Birth of a metabolic gene cluster in yeast by adaptive gene relocation. Nat Genet 37: 777–782. doi:10.1038/ng1584.

77. Perrin RM, Fedorova ND, Bok JW, Cramer RA, Wortman JR, et al. (2007) Transcriptional regulation of chemical diversity in Aspergillus fumigatus by LaeA. PLoS Pathog 3: e50. doi:10.1371/journal.ppat.0030050.

78. Ahmed YL, Gerke J, Park H-S, Bayram Ö, Neumann P, et al. (2013) The Velvet Family of Fungal Regulators Contains a DNA-Binding Domain Structurally Similar to NF-κB. PLoS Biol 11: e1001750. doi:10.1371/journal.pbio.1001750.

APPENDIX A

ROLE OF THE STUDENT

I was responsible for determining the level of conservation of metabolic gene sets and SM gene clusters in *A. nidulans*, *A. fumigatus*, *A. oryzae*, and *A. niger*. Additionally, I performed RNA-seq read alignment, differential expression analysis, GOSlim enrichment analysis, and ortholog differential expresson calculations for *A. nidulans* and *A. fumigatus* Δ*veA*, Δ*mtfA*, and wild-type strains.

APPENDIX B. METABOLISM ORTHOGROUP CONSERVATION PERCENTAGES

Genome	Orthogroup comparison	Total number orthogroups	Orthogroups shared in at least one genome	Species-specific orthogroups	Percent species-specific orthogroups
Anid	All to All	9577	7356	2221	23.2%
Anid	Primary to All	1359	1203	156	11.5%
Anid	Primary to Primary	1359	1172	187	13.8%
Anid	PKS-NRPS-Hybrid to All	30	19	11	36.7%
Anid	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	30	18	12	40.0%
Anid	Synthase to All	47	30	17	36.2%
Anid	Synthase to Synthase	47	25	22	46.8%
Anid	SM to All	417	289	128	30.7%
Anid	SM to SM	417	146	271	65.0%
Afum	All to All	8810	7234	1576	17.9%
Afum	Primary to All	1322	1223	99	7.5%
Afum	Primary to Primary	1322	1205	117	8.9%
Afum	PKS-NRPS-Hybrid to All	20	15	5	25.0%
Afum	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	20	15	5	25.0%
Afum	Synthase to All	25	19	6	24.0%
Afum	Synthase to Synthase	25	18	7	28.0%
Afum	SM to All	281	208	73	26.0%
Afum	SM to SM	281	130	151	53.7%
Aniger	All to All	12386	7170	5216	42.1%
Aniger	Primary to All	1410	1229	181	12.8%
Aniger	Primary to Primary	1410	1188	222	15.7%
Aniger	PKS-NRPS-Hybrid to All	35	24	11	31.4%
Aniger	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	35	23	12	34.3%
Aniger	Synthase to All	50	37	13	26.0%
Aniger	Synthase to Synthase	50	28	22	44.0%
Aniger	SM to All	612	392	220	35.9%
Aniger	SM to SM	612	155	457	74.7%
Aory	All to All	10440	7051	3389	32.5%
Aory	Primary to All	1502	1270	232	15.4%
Aory	Primary to Primary	1502	1231	271	18.0%
Aory	PKS-NRPS-Hybrid to All	37	22	15	40.5%
Aory	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	37	21	16	43.2%
Aory	Synthase to All	50	31	19	38.0%
Aory	Synthase to Synthase	50	26	24	48.0%
Aory	SM to All	515	319	196	38.1%
Aory	SM to SM	515	156	359	69.7%

APPENDIX C. METABOLISM ORTHOGROUP CONSERVATION STATISTICAL SIGNIFICANCE

Genome	Category 1	Category 2	Species-specific ogroups in category one	Shared ogroups in category one	Species-specific ogroups in category two	Shared ogroups in category two	Fisher's exact p-value	False discovery rate	Sig. level
Anid	Primary to All	Primary to Primary	156	1203	187	1172	8.30E-02	1.29E-01	
Anid	Primary to All	PKS-NRPS-Hybrid to All	156	1203	11	19	3.86E-04	1.11E-03	*
Anid	Primary to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	156	1203	12	18	8.16E-05	2.69E-04	**
Anid	Primary to All	Synthase to All	156	1203	17	30	1.55E-05	5.44E-05	***
Anid	Primary to All	Synthase to Synthase	156	1203	22	25	4.60E-09	2.34E-08	***
Anid	Primary to All	SM to All	156	1203	128	289	1.48E-18	1.03E-17	***
Anid	Primary to All	SM to SM	156	1203	271	146	1.05E-99	2.35E-98	***
Anid	Primary to Primary	PKS-NRPS-Hybrid to All	187	1172	11	19	1.71E-03	4.17E-03	*
Anid	Primary to Primary	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	187	1172	12	18	4.37E-04	1.19E-03	*
Anid	Primary to Primary	Synthase to All	187	1172	17	30	1.43E-04	4.44E-04	**
Anid	Primary to Primary	Synthase to Synthase	187	1172	22	25	1.06E-07	5.14E-07	***
Anid	Primary to Primary	SM to All	187	1172	128	289	5.13E-14	3.19E-13	***
Anid	Primary to Primary	SM to SM	187	1172	271	146	1.66E-88	3.11E-87	***
Anid	PKS-NRPS-Hybrid to All	PKS-NRPS-Hybrid to PKS-NRPS-	11	19	12	18	1.00E+00	1.00E+00	

		Hybrid							
Anid	PKS-NRPS-Hybrid to All	Synthase to All	11	19	17	30	1.00E+00	1.00E+00	
Anid	PKS-NRPS-Hybrid to All	Synthase to Synthase	11	19	22	25	4.80E-01	6.48E-01	
Anid	PKS-NRPS-Hybrid to All	SM to All	11	19	128	289	5.41E-01	6.97E-01	
Anid	PKS-NRPS-Hybrid to All	SM to SM	11	19	271	146	2.87E-03	6.42E-03	*
Anid	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to All	12	18	17	30	8.11E-01	9.60E-01	
Anid	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to Synthase	12	18	22	25	6.41E-01	7.98E-01	
Anid	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to All	12	18	128	289	3.11E-01	4.41E-01	
Anid	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to SM	12	18	271	146	9.70E-03	1.98E-02	
Anid	Synthase to All	Synthase to Synthase	17	30	22	25	4.03E-01	5.64E-01	
Anid	Synthase to All	SM to All	17	30	128	289	5.07E-01	6.68E-01	
Anid	Synthase to All	SM to SM	17	30	271	146	2.01E-04	6.09E-04	**
Anid	Synthase to Synthase	SM to All	22	25	128	289	3.20E-02	5.19E-02	
Anid	Synthase to Synthase	SM to SM	22	25	271	146	1.69E-02	3.15E-02	
Anid	SM to All	SM to SM	128	289	271	146	2.72E-23	2.34E-22	***

Afum	Primary to All	Primary to Primary	99	1223	117	1205	2.27E-01	3.35E-01	
Afum	Primary to All	PKS-NRPS-Hybrid to All	99	1223	5	15	1.54E-02	2.97E-02	
Afum	Primary to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	99	1223	5	15	1.54E-02	2.97E-02	
Afum	Primary to All	Synthase to All	99	1223	6	19	1.01E-02	2.03E-02	
Afum	Primary to All	Synthase to Synthase	99	1223	7	18	2.25E-03	5.25E-03	*
Afum	Primary to All	SM to All	99	1223	73	208	2.31E-16	1.52E-15	***
Afum	Primary to All	SM to SM	99	1223	151	130	3.46E-66	5.53E-65	***
Afum	Primary to Primary	PKS-NRPS-Hybrid to All	117	1205	5	15	2.91E-02	4.86E-02	
Afum	Primary to Primary	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	117	1205	5	15	2.91E-02	4.86E-02	
Afum	Primary to Primary	Synthase to All	117	1205	6	19	2.13E-02	3.67E-02	
Afum	Primary to Primary	Synthase to Synthase	117	1205	7	18	5.56E-03	1.20E-02	
Afum	Primary to Primary	SM to All	117	1205	73	208	1.63E-13	9.58E-13	***
Afum	Primary to Primary	SM to SM	117	1205	151	130	3.75E-60	5.26E-59	***
Afum	PKS-NRPS-Hybrid to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	5	15	5	15	1.00E+00	1.00E+00	
Afum	PKS-NRPS-Hybrid to All	Synthase to All	5	15	6	19	1.00E+00	1.00E+00	
Afum	PKS-NRPS-Hybrid to All	Synthase to Synthase	5	15	7	18	1.00E+00	1.00E+00	
Afum	PKS-NRPS-	SM to All	5	15	73	208	1.00E+00	1.00E+00	

	Hybrid to All								
Afum	PKS-NRPS-Hybrid to All	SM to SM	5	15	151	130	1.88E-02	3.39E-02	
Afum	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to All	5	15	6	19	1.00E+00	1.00E+00	
Afum	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to Synthase	5	15	7	18	1.00E+00	1.00E+00	
Afum	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to All	5	15	73	208	1.00E+00	1.00E+00	
Afum	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to SM	5	15	151	130	1.88E-02	3.39E-02	
Afum	Synthase to All	Synthase to Synthase	6	19	7	18	1.00E+00	1.00E+00	
Afum	Synthase to All	SM to All	6	19	73	208	1.00E+00	1.00E+00	
Afum	Synthase to All	SM to SM	6	19	151	130	5.81E-03	1.23E-02	
Afum	Synthase to Synthase	SM to All	7	18	73	208	8.15E-01	9.60E-01	
Afum	Synthase to Synthase	SM to SM	7	18	151	130	2.01E-02	3.52E-02	
Afum	SM to All	SM to SM	73	208	151	130	2.30E-11	1.23E-10	***
Aniger	Primary to All	Primary to Primary	181	1229	222	1188	3.13E-02	5.15E-02	
Aniger	Primary to All	PKS-NRPS-Hybrid to All	181	1229	11	24	3.92E-03	8.60E-03	*
Aniger	Primary to All	PKS-NRPS-Hybrid to PKS-NRPS-	181	1229	12	23	1.16E-03	3.01E-03	*

		Hybrid							
Aniger	Primary to All	Synthase to All	181	1229	13	37	1.69E-02	3.15E-02	
Aniger	Primary to All	Synthase to Synthase	181	1229	22	28	1.22E-07	5.68E-07	***
Aniger	Primary to All	SM to All	181	1229	220	392	6.25E-31	7.00E-30	***
Aniger	Primary to All	SM to SM	181	1229	457	155	2.75E-164	3.08E-162	***
Aniger	Primary to Primary	PKS-NRPS-Hybrid to All	222	1188	11	24	1.91E-02	3.40E-02	
Aniger	Primary to Primary	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	222	1188	12	23	8.17E-03	1.70E-02	
Aniger	Primary to Primary	Synthase to All	222	1188	13	37	7.46E-02	1.18E-01	
Aniger	Primary to Primary	Synthase to Synthase	222	1188	22	28	3.54E-06	1.37E-05	***
Aniger	Primary to Primary	SM to All	222	1188	220	392	1.07E-22	8.52E-22	***
Aniger	Primary to Primary	SM to SM	222	1188	457	155	2.40E-145	1.34E-143	***
Aniger	PKS-NRPS-Hybrid to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	11	24	12	23	1.00E+00	1.00E+00	
Aniger	PKS-NRPS-Hybrid to All	Synthase to All	11	24	13	37	6.30E-01	7.93E-01	
Aniger	PKS-NRPS-Hybrid to All	Synthase to Synthase	11	24	22	28	2.67E-01	3.88E-01	
Aniger	PKS-NRPS-Hybrid to All	SM to All	11	24	220	392	7.17E-01	8.64E-01	
Aniger	PKS-NRPS-Hybrid to All	SM to SM	11	24	457	155	2.53E-07	1.09E-06	***
Aniger	PKS-NRPS-Hybrid to PKS-NRPS-	Synthase to All	12	23	13	37	4.72E-01	6.45E-01	

	Hybrid								
Aniger	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to Synthase	12	23	22	28	5.00E-01	6.67E-01	
Aniger	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to All	12	23	220	392	1.00E+00	1.00E+00	
Aniger	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to SM	12	23	457	155	1.35E-06	5.62E-06	***
Aniger	Synthase to All	Synthase to Synthase	13	37	22	28	9.28E-02	1.42E-01	
Aniger	Synthase to All	SM to All	13	37	220	392	1.69E-01	2.56E-01	
Aniger	Synthase to All	SM to SM	13	37	457	155	8.29E-12	4.64E-11	***
Aniger	Synthase to Synthase	SM to All	22	28	220	392	2.86E-01	4.10E-01	
Aniger	Synthase to Synthase	SM to SM	22	28	457	155	1.13E-05	4.21E-05	***
Aniger	SM to All	SM to SM	220	392	457	155	4.67E-43	5.81E-42	***
Aory	Primary to All	Primary to Primary	232	1270	271	1231	6.32E-02	1.01E-01	
Aory	Primary to All	PKS-NRPS-Hybrid to All	232	1270	15	22	2.67E-04	7.86E-04	**
Aory	Primary to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	232	1270	16	21	6.82E-05	2.32E-04	**
Aory	Primary to All	Synthase to All	232	1270	19	31	1.26E-04	4.04E-04	**
Aory	Primary to All	Synthase to Synthase	232	1270	24	26	1.28E-07	5.76E-07	***
Aory	Primary to All	SM to All	232	1270	196	319	4.53E-25	4.62E-24	***
Aory	Primary to All	SM to SM	232	1270	359	156	7.62E-114	2.84E-112	***

Aory	Primary to Primary	PKS-NRPS-Hybrid to All	271	1231	15	22	1.83E-03	4.36E-03	*
Aory	Primary to Primary	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	271	1231	16	21	4.18E-04	1.17E-03	*
Aory	Primary to Primary	Synthase to All	271	1231	19	31	1.24E-03	3.17E-03	*
Aory	Primary to Primary	Synthase to Synthase	271	1231	24	26	2.11E-06	8.44E-06	***
Aory	Primary to Primary	SM to All	271	1231	196	319	3.18E-19	2.37E-18	***
Aory	Primary to Primary	SM to SM	271	1231	359	156	1.36E-100	3.81E-99	***
Aory	PKS-NRPS-Hybrid to All	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	15	22	16	21	1.00E+00	1.00E+00	
Aory	PKS-NRPS-Hybrid to All	Synthase to All	15	22	19	31	8.28E-01	9.66E-01	
Aory	PKS-NRPS-Hybrid to All	Synthase to Synthase	15	22	24	26	5.20E-01	6.78E-01	
Aory	PKS-NRPS-Hybrid to All	SM to All	15	22	196	319	8.61E-01	9.94E-01	
Aory	PKS-NRPS-Hybrid to All	SM to SM	15	22	359	156	4.45E-04	1.19E-03	*
Aory	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to All	16	21	19	31	6.63E-01	8.16E-01	
Aory	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	Synthase to Synthase	16	21	24	26	6.71E-01	8.17E-01	
Aory	PKS-NRPS-Hybrid to	SM to All	16	21	196	319	6.00E-01	7.64E-01	

	PKS-NRPS-Hybrid								
Aory	PKS-NRPS-Hybrid to PKS-NRPS-Hybrid	SM to SM	16	21	359	156	1.59E-03	3.95E-03	*
Aory	Synthase to All	Synthase to Synthase	19	31	24	26	4.19E-01	5.80E-01	
Aory	Synthase to All	SM to All	19	31	196	319	1.00E+00	1.00E+00	
Aory	Synthase to All	SM to SM	19	31	359	156	1.31E-05	4.72E-05	***
Aory	Synthase to Synthase	SM to All	24	26	196	319	1.75E-01	2.61E-01	
Aory	Synthase to Synthase	SM to SM	24	26	359	156	2.46E-03	5.63E-03	*
Aory	SM to All	SM to SM	196	319	359	156	1.70E-24	1.59E-23	***

[a]<0.01 is *,
<0.001 is **,
<0.0001 is ***

APPENDIX D. GO TERM ENRICHMENT OF GENES DIFFERENTIALLY REGULATED BY VEA AND MTFA IN *AFUM* AND *ANID*

Condition	Species	Gene set	GOSlim category	No. genes in set and category	Total genes category	percentage	adjusted p-value
VeA	Anid	under expressed	carbohydrate metabolic process	113	511	22.1%	0.00065313
VeA	Anid	under expressed	secondary metabolic process	113	280	40.4%	1.0159E-22
VeA	Anid	under expressed	extracellular region	54	234	23.1%	0.032509
VeA	Anid	under expressed	oxidoreductase activity	274	1109	24.7%	1.7295E-15
VeA	Anid	over expressed	ribosome biogenesis	46	171	26.9%	5.1145E-08
VeA	Anid	over expressed	cellular amino acid metabolic process	78	309	25.2%	1.1425E-11
VeA	Anid	over expressed	translation	27	201	13.4%	0.98454
VeA	Anid	over expressed	RNA metabolic process	84	665	12.6%	0.072582
VeA	Anid	over expressed	nucleolus	52	200	26.0%	2.2256E-08
VeA	Anid	over expressed	cytosol	167	1428	11.7%	1
VeA	Anid	over expressed	nucleus	244	2174	11.2%	1
VeA	Anid	over expressed	ribosome	9	156	5.8%	1
VeA	Anid	over expressed	helicase activity	13	90	14.4%	0.49953
VeA	Anid	over expressed	transferase activity	109	859	12.7%	0.19287
VeA	Anid	over expressed	structural molecule activity	10	199	5.0%	1
VeA	Anid	over expressed	RNA binding	37	243	15.2%	0.11892
VeA	Afum	under expressed	carbohydrate metabolic process	66	186	35.5%	5.1791E-07
VeA	Afum	under expressed	secondary metabolic process	117	469	24.9%	3.2906E-10
VeA	Afum	under expressed	extracellular region	57	226	25.2%	0.0016833
VeA	Afum	under expressed	oxidoreductase activity	219	865	25.3%	8.0805E-14
VeA	Afum	over expressed	ribosome biogenesis	122	168	72.6%	1.1842E-61
VeA	Afum	over expressed	cellular amino acid metabolic process	82	257	31.9%	1.1985E-10
VeA	Afum	over expressed	translation	118	407	29.0%	8.2226E-12
VeA	Afum	over expressed	RNA metabolic process	182	630	28.9%	4.9301E-18
VeA	Afum	over expressed	nucleolus	123	202	60.9%	3.5094E-49

VeA	Afum	over expressed	cytosol	319	1426	22.4%	1.2379E-13
VeA	Afum	over expressed	nucleus	411	2091	19.7%	6.2564E-09
VeA	Afum	over expressed	ribosome	85	152	55.9%	3.1963E-30
VeA	Afum	over expressed	helicase activity	27	97	27.8%	0.013603
VeA	Afum	over expressed	transferase activity	159	846	18.8%	0.022743
VeA	Afum	over expressed	structural molecule activity	82	186	44.1%	1.0106E-19
VeA	Afum	over expressed	RNA binding	89	445	20.0%	0.026205
MtfA	Anid	under expressed	toxin metabolic process	9	52	17.3%	0.019476
MtfA	Anid	under expressed	secondary metabolic process	56	280	20.0%	4.5508E-17
MtfA	Anid	under expressed	oxidoreductase activity	113	1109	10.2%	4.76E-11
MtfA	Anid	over expressed	cellular amino acid metabolic process	25	309	8.1%	3.83E-03
MtfA	Anid	over expressed	secondary metabolic process	32	280	11.4%	4.89E-07
MtfA	Anid	over expressed	asexual sporulation	15	190	7.9%	4.27E-02
MtfA	Anid	over expressed	developmental process	28	440	6.4%	4.27E-02
MtfA	Afum	over expressed	toxin metabolic process	1	64	1.6%	1
MtfA	Afum	over expressed	secondary metabolic process	9	186	4.8%	2.89E-05
MtfA	Afum	over expressed	oxidoreductase activity	17	865	2.0%	1.28E-04
MtfA	Afum	over expressed	cellular amino acid metabolic process	1	257	0.4%	1
MtfA	Afum	over expressed	secondary metabolic process	3	186	1.6%	0.10974
MtfA	Afum	over expressed	asexual sporulation	0	184	0.0%	1
MtfA	Afum	over expressed	developmental process	0	437	0.0%	1

APPENDIX E. GENE EXPRESSION ORTHOLOG COUNTS.

Gene set	condition	species	relative expression[a]	no. genes	Genes with ortholog(s)[b]	Genes with no ortholog(s) differentially expressed[c]	Genes with ortholog(s) differentially expressed[d]	Genes with ortholog(s) showing matching expression[e]	Genes with ortholog(s) showing opposite expression[f]
SM	vea	Anid	-1	184	64	19	45	37	13
SM	vea	Anid	1	67	14	9	5	3	3
SM	vea	Anid	0	247	49	29	20	na	na
SM	vea	Afum	-1	98	36	6	30	29	3
SM	vea	Afum	1	38	11	2	9	3	7
SM	vea	Afum	0	181	53	18	35	na	na
dev	vea	Anid	-1	72	66	36	30	29	1
dev	vea	Anid	1	32	31	16	15	11	4
dev	vea	Anid	0	385	355	289	66	na	na
dev	vea	Afum	-1	77	73	41	32	28	4
dev	vea	Afum	1	38	35	24	11	11	1
dev	vea	Afum	0	363	344	291	53	na	na
dev	mtfa	Anid	-1	6	4	4	0	0	0
dev	mtfa	Anid	1	29	28	28	0	0	0
dev	mtfa	Anid	0	454	420	419	1	na	na
dev	mtfa	Afum	-1	1	1	1	0	0	0
dev	mtfa	Afum	1	0	0	0	0	0	0
dev	mtfa	Afum	0	477	451	418	33	na	na
SM	mtfa	Anid	-1	107	36	30	6	6	0
SM	mtfa	Anid	1	32	12	10	2	0	2
SM	mtfa	Anid	0	359	79	68	11	na	na
SM	mtfa	Afum	-1	32	9	4	5	5	1
SM	mtfa	Afum	1	4	0	0	0	0	0
SM	mtfa	Afum	0	281	91	38	53	na	na
all	mtfa	Anid	-1	568	310	291	19	17	2
all	mtfa	Anid	1	400	301	292	9	6	3
all	mtfa	Anid	0	9741	6985	6901	84	na	na
all	mtfa	Afum	-1	63	33	19	14	13	2
all	mtfa	Afum	1	34	21	13	8	6	2
all	mtfa	Afum	0	9686	7433	6615	818	na	na
all	vea	Anid	-1	1671	1065	431	634	498	212

all	vea	Anid	1	1165	825	335	490	383	143
all	vea	Anid	0	7873	5706	4015	1691	na	na
all	vea	Afum	-1	1555	1131	572	559	461	134
all	vea	Afum	1	1546	1221	705	516	388	169
all	vea	Afum	0	6682	5135	4002	1133	na	na

[a]gene expression category relative to wild type. Under expressed (-1), over expressed (1), or not differentially expressed (0)
[b]number of genes with one or more orthologs that belongs to the same geneset (e.g. if in SM gene set, ortholog must also function in SM)
[c]number of genes with no ortholog that is differentially expressed
[d]number of genes with one or more orthologs that is differentially expressed
[e]number of genes with one or more orthologs has matching expression relative to wild type
[f]number of genes with one or more othologs has opposite expression relative to wild type