Phylogenomic Analyses of Non-Dikarya Fungi Supports Horizontal Gene Transfer Driving Diversification of Secondary Metabolism in the Amphibian Gastrointestinal Symbiont, *Basidiobolus*.
Phylogenomic Analyses of Non-Dikarya Fungi Supports Horizontal Gene Transfer Driving Diversification of Secondary Metabolism in the Amphibian Gastrointestinal Symbiont, Basidiobolus

Javier F. Tabima,*1 Ian A. Trautman,* Ying Chang,* Yan Wang,†1 Stephen Mondo,§ Alan Kuo,§ Asaf Salamov,§ Igor V. Grigoriev,** Jason E. Stajich,†‡ and Joseph W. Spatafora*

*Department of Botany and Plant Pathology, College of Agricultural Sciences, Oregon State University, Corvallis, †Department of Microbiology and Plant Pathology, ‡Institute for Integrative Genome Biology, University of California—Riverside, California §US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, California, and **Department of Plant and Microbial Biology, University of California-Berkeley, California

ABSTRACT Research into secondary metabolism (SM) production by fungi has resulted in the discovery of diverse, biologically active compounds with significant medicinal applications. The fungi rich in SM production are taxonomically concentrated in the subkingdom Dikarya, which comprises the phyla Ascomycota and Basidiomycota. Here, we explore the potential for SM production in Mucoromycota and Zoopagomy- cota, two phyla of nonflagellated fungi that are not members of Dikarya, by predicting and identifying core genes and gene clusters involved in SM. The majority of non-Dikarya have few genes and gene clusters involved in SM production except for the amphibian gut symbionts in the genus Basidiobolus. Basidiobolus genomes exhibit an enrichment of SM genes involved in siderophore, surfactin-like, and terpene cyclase production, all these with evidence of constitutive gene expression. Gene expression and chemical assays also confirm that Basidiobolus has significant siderophore activity. The expansion of SMs in Basidiobolus are partially due to horizontal gene transfer from bacteria, likely as a consequence of its ecology as an amphibian gut endosymbiont.

KEYWORDS Basidiobolus zygomycetes secondary metabolism horizontal gene transfer siderophores

Fungi produce a wealth of biologically active small molecules – secondary or specialized metabolites – that function in interactions with other organisms, environmental sensing, growth and development, and numerous other processes (Rokas et al. 2020). Several of these compounds have led to the successful development of pharmaceuticals (e.g., antibiotics, immunosuppressants, statins, etc.) that have had dramatic and positive impacts on human health. Understanding the evolution of fungal secondary metabolites and linking them with their ecological and physiological functions in nature can inform searches for compounds with applications in human society.

Secondary metabolism (SM) is imprecisely defined but can be characterized generally as the production of bioactive compounds that are not part of primary metabolism and that are not required for growth and survival in the laboratory (Keller et al. 2005; Brakhage 2013; Rokas et al. 2020). In fungi, the genes responsible for the synthesis of secondary metabolites are frequently co-located in biosynthetic gene clusters (Smith et al. 1990; Brakhage 2013), which contain the genes that control regulation of expression, biosynthesis, tailoring, and transport of these compounds out of the cell (Smith et al. 1990; Keller et al., 2005; Osbourn 2010). In the kingdom Fungi, the diversity of products synthesized via SM is substantial and...
primarily includes alkaloids, peptides, polyketides, and terpenes (Collemare et al. 2008; Helaly et al. 2018). Each of these groups of compounds are synthesized by core genes that are characteristic of the pathways and include, but are not limited to, dimethylallyl tryptophan synthases (DMAT), non-ribosomal peptide synthetases (NRPS), polyketide synthetases (PKS), and terpene cyclases (TC). These bioactive compounds fulfill various roles that are hypothesized to increase the fitness of the fungus by promoting better recognition and adaptation to environmental cues.

Bio synthesis of secondary metabolites is heterogeneous across the fungal tree of life, but the vast majority of discovered and predicted secondary metabolites are reported within the fungal phyla Ascomycota and Basidiomycota of the subkingdom Dikarya (Wisecaver et al. 2014; Rokas et al. 2018). Filamentous ascomycetes are the major producers of polyketide and peptidic secondary metabolites (Rokas et al. 2018) (e.g., penicillin, cyclosporin, etc.), with the majority of genes, gene clusters, and enzymes involved with fungal SM found in the subphylum Pezizomycotina (Collemare et al. 2008; Wisecaver et al. 2014; Helaly et al. 2018; Rokas et al. 2018). Although less than Ascomycota, Basidiomycota is also a prominent producer of SM, including diverse terpene genes (e.g., sesquiterpenes, Quin et al. 2014), some of the better-known hallucinogens (e.g., psilocybin of Psilocybe; Reynolds et al. 2018), and compounds toxic to humans (e.g., amanitin of Amanita; Luo et al. 2010).

For reasons that are unclear, the remainder of kingdom Fungi is characterized by a paucity of secondary metabolites (Voigt et al. 2016). This includes the zoosporic fungi and relatives classified in Blastocladiomycota, Chytridiomycota and Rozellomycota, and the nonflagellated, zoomyocyte fungi of Mucoromycota and Zoopagomycota. This pattern of SM diversity supports the hypothesis that diversification of secondary metabolism is a characteristic of Ascomycota and Basidiomycota (subkingdom Dikarya), which share a more recent common ancestor relative to the other phyla. Recent genome sampling efforts have focused on increased sequencing of non-Dikarya species (Nagy et al. 2014; Kohler et al. 2015; Spatafora et al. 2016; Quandt et al. 2017; Ahrendt et al. 2018). These efforts have provided a better understanding of the relationships of the phyla of kingdom Fungi (e.g., Spatafora et al. 2016), and processes and patterns that shaped the evolution of morphologies (e.g., Nagy et al. 2014) and ecologies (e.g., Chang et al. 2019; Quandt et al. 2017) within the kingdom. The availability of a diversity of these genomes provides an opportunity to characterize and focus on the secondary metabolism composition of non-Dikarya taxa, which have remained relatively unexplored.

While the majority of non-Dikarya taxa have low SM diversity, genomic sequencing of the genus Basidiobolus (Phylum Zoopagomycota) revealed that it possesses an unusually large composition of SM gene clusters. Species of Basidiobolus have complex life cycles, that involve the production of multiple spore types that occur in various environmental niches. Species have been found in the digestive tracts and feces of fish, reptiles, and amphibians (Nickerson and Hutchison 1971), where their function and impact on the host remains unknown. They have also been isolated from cadavers of mites and collembola species (Manning et al. 2007; Werner et al. 2012, 2018), centipedes (Fonseca et al. 2019), spiders (Henriksen et al. 2018), and associated with the gut microbiota of rove beetles (Stefani et al. 2016). Basidiobolus is known from leaf litter (Smith and Callaghan 1987; Manning and Callaghan 2008) and it has been identified as part of the microbiome of aquatic carnivorous plants (Sirová et al. 2018). Finally, species of the genus have also been implicated in opportunistic infections of humans and other mammals (Vikram et al. 2012; Geramizadeh et al. 2015; Okada et al. 2015). Across these environments Basidiobolus is dispersed by both forcibly discharged asexual spores (blastoconidia) and passively dispersed asexual spores (capilloconidia) that adhere to exoskeletons of small insects. Basidiobolus also reproduces sexually through the production of zygospores (meiospores) either by selfing (homothallic) or outcrossing (heterothallic) according to species. Based on these diverse ecologies, Basidiobolus must have adapted for survival in numerous environmental niches including the digestive systems of diverse animal species, feces, insect phoresis, and on decaying plant matter or leaf litter.

In this study we demonstrate that the genomes of Basidiobolus contain a larger number of genes related to SM than predicted by phylogeny and that in several cases the evolution of many of these SM genes is inconsistent with vertical evolution. Our objectives were to: i) characterize the diversity of SM in Basidiobolus, ii) identify the phylogenetic sources of this diversity, and iii) determine which classes of SM gene clusters are functional and may predict the secondary metabolites produced by species of Basidiobolus. Finally, we propose a model in which the amphibian gastrointestinal system is an environment that promotes noncanonical evolution of its fungal inhabitants.

MATERIALS AND METHODS

Data collection
Annotated genome and amino-acid translation of predicted gene model sequences for three isolates of two species within the genus Basidiobolus were used in this study: Basidiobolus meristosporus CBS 931.73 (Mondo et al. 2017), isolated from gecko dung in the locality of Lamco, Ivory Coast; B. meristosporus B9252 (Chibucos et al. 2016) isolated from human eye in Saudi Arabia; and B. heterosporus B8920 (Chibucos et al. 2016) isolated from plant debris in India. The genomic sequence of B. meristosporus CBS 931.73 was sequenced with PacBio and annotated by Mondo et al. (2017) and obtained from the US Department of Energy Joint Genome Institute MycoCosm genome portal (https://mycocosm.jgi.doe.gov; Grigoriev et al. 2014). Genomic sequences and annotation of B. meristosporus B9252 and B. heterosporus B8920 sequenced by Chibucos et al. (2016) were obtained directly from the authors. The raw reads for these two species are available in GenBank (Accession numbers GCA_000697375.1 and GCA_000697455.1). Additional genomes of 66 Mucoromycota and Zoopagomycota species were used in this study (Ma et al. 2009, Tisserant et al. 2013, Wang et al. 2013, Schwartz et al. 2014, Chang et al. 2015, Chibucos et al. 2016, Corrochano et al. 2016, Lastovetsky et al. 2016, Wang et al. 2016, Mondo et al. 2017, Uehling et al. 2017, Ahrendt et al. 2018, Chen et al. 2018). Data sources for the genomes included in these analyses are available in Table S1.

Secondary metabolite gene cluster prediction
Predicted SM proteins were retrieved for genomes of 66 Mucoromycota and Zoopagomycota species (not including Basidiobolus) based on the Secondary Metabolite Unique Regions Finder (SMURF) (Khalidi et al. 2010) predictions available at MycoCosm for NRPS and PKS, and using local prediction with AntiSMASH v4.2.0 (Blin et al. 2017) for Terpene Cyclases. Secondary Metabolite gene clusters for the three genomes of Basidiobolus were predicted with AntiSMASH v4.2.0 and SMURF from the annotated genomes of the three isolates used in this study. The AntiSMASH prediction was performed on local HPC, while SMURF predictions were obtained by
submission of the genomes to SMURF web server (http://smurf.jcvi.org/). Predictions were contrasted manually to include all clusters of secondary metabolites from both sources. Orthologous sets of core SM genes across the three genomes of Basidiobolus were identified using OrthoFinder (Emms and Kelly 2015).

**Secondary metabolite expression analysis**

To assess the expression of predicted SM proteins, we calculated summarized counts of RNA transcript per million (TPM) for genes in B. meristosporus CBS 931.73 isolate by aligning RNA-Seq reads to the assembled B. meristosporus CBS 931.73 genome with HiSat v2.1.0 (Kim et al. 2019). The aligned sequence reads were processed with HTSeq-counts (Anders et al. 2015) to generate the counts of overlapping reads found for each gene and the normalized TPM for the genes was calculated using the cpm function in edgeR (Robinson et al. 2010). A distribution of RNA-Seq read counts per gene was plotted using the ggplot2 package in the R statistical framework (R Core Team 2018).

**Phylogenomic analyses**

Phylogenetic analyses were used to assess the evolutionary relationships of the NRPS, PKS, and terpene cyclase/synthase predicted proteins. For the NRPS genes, the adenylation domains (A-domains) were identified by hmmsearch from HMMer 3.0 suit (Eddy 2004), using the A-domain profile reported by Bushley and Turegon (2010) as a reference profile HMM. The predicted A-domains were extracted from the resulting HMMER table into a FASTA file using the esl-reformat program included in the HMMER suite. The predicted A-domains for all Basidiobolus, Mucoromycota and Zoopagomycota species were added to an A-domain amino-acid alignment reported by Bushley and Turegon (2010) with MAFFT v7 7 (Katoh et al. 2017) (File S1). This reference alignment contains A-domains from NRPS proteins from nine major subfamilies of fungal and bacterial NRPS proteins. The phylogenetic domain tree was constructed using a maximum likelihood approach implemented in RAxML v. 8.2.11 (Stamatakis 2014) with the JTT amino acid substitution matrix, after protein selection for the analysis. To screen for PKS proteins in Dikarya, OrthoFinder was used to build orthologous clusters of genes between Basidiobolus, Mucoromycota and Zoopagomycota. To include additional fungal TC orthologs in the phylogenetic reconstruction, we identified TC proteins from 58 published genomes of Dikarya isolates (21 Basidiomycetes and 37 Ascomycetes, Table S3 (Armaleo et al. 2019, Burmester et al. 2011, DiGuistini et al. 2011, Yang et al. 2011, Zuccaro et al. 2011, Floudas et al. 2012, Morin et al. 2012, Ohm et al. 2012, Padamsee et al. 2012, Xiao et al. 2012, Gosticnar et al. 2014, Pendleton et al. 2014, Riley et al. 2014, Terfehr et al. 2014, Wang et al. 2014, Kroken et al. 2015, Okagaki et al. 2015, Nagy et al. 2016, Nguyen et al. 2017, Teixeira et al. 2017, Martino et al. 2018, Vesth et al. 2018, Crouch et al. 2020, Haridas et al. 2020)), available in DOE-JGI MycoCosm, each from a different family. One isolate per family was randomly selected for the analysis. To screen for TC proteins in Dikarya, OrthoFinder was used to build orthologous clusters of genes between Basidiobolus, Mucoromycota and Zoopagomycota species. Basidiobolus isolates were clustered within orthologous groups that contain TC of Basidiobolus, Mucoromycota and Zoopagomycota were considered valid TC orthologs and used in downstream analyses. Finally, we identified bacterial TC to evaluate the potential for HGT into Basidiobolus, Mucoromycota or Zoopagomycota species. Bacterial TCs were identified by screening each protein from the RefSeq, release 87 (May 2018, O’Leary et al. 2015) FASTA dataset against a BLAST database (Altschul et al. 1997), one from each of the zygomycete orthologous groups identified by OrthoFinder in the previous step, for a total of four protein sequences in the database. The BLASTP program was used to perform the searches, using an e-value of 1e-10 and the BLOSUM62 substitution matrix. Positive matches from the BLASTP assay were used as bacterial TC for subsequent analyses. A multi-sequence alignment containing all predicted TC from Basidiobolus, Mucoromycota and Zoopagomycota species, the orthologous TC from additional fungal species, and the TC proteins from reference bacterial predicted gene models result of the BLASTP search was performed in MAFFT v7 using the G-INS-I algorithm for progressive global alignment (Katoh et al. 2017) (File S5). A phylogenetic tree was reconstructed (Acaromycetes ingoldii MCA 4198 v1.0 (Kipponmyyngan et al. 2018), Fibroporia radiculosa TFFF 294 (Tang et al. 2012), Fomitiporia mediterranea v1.0 (Floudas et al. 2012), Gloeophyllum trabeum v1.0 (Floudas et al. 2012), Gymnopus luxurians v1.0 (Kohler et al. 2015), Laccaria bicolor v2.0 (Martin et al. 2008), Microbotryum lychnidis–dioicae p1A1 Lamole (Perlin et al. 2015), Piloderma croceum F 1398 v1.0 (Kohler et al. 2015), Pisolithus tinctorius Marx 270 v1.0 (Kohler et al. 2015)), four Neocallichlamydocymycota isolates (Anaerozyma robustus v1.0 (Haitjema et al. 2017), Orpinomyces sp. (Youssef et al. 2013), Piromyces fennis v3.0 (Haitjema et al. 2017), Piromyces sp. E2 v1.0 (Haitjema et al. 2017)) and one Chytriomyzomycota species (Spizellomyces punctatus DAOM 1117 (Russ et al. 2016)). The predicted PKS proteins from additional species were all obtained from the DOE-JGI MycoCosm genome portal by searching for all genes with “PKS” on the SM annotation from MycoCosm. The KS domains were identified for the subset of PKS proteins using a HMMER KS profile as mentioned above. A phylogenetic tree was reconstructed using maximum likelihood using all KS domains using similar parameters described in the NRPS step. To determine if the novel PKS candidates had the common domains of PKS proteins (AT, KR, KS, DH and PP), an identification of these profiles were performed based on HMMER models from PFAM (Finn et al. 2009) and SMART (Schultz et al. 2000) databases (KS (PF001009), AT (SM00827), KR (SM00822), DH (SM00826), and PP (SM00823)).

The terpene cyclase (TC) proteins were predicted in AntiSMASH for all 69 assembled genome sequences from Basidiobolus, Mucoromycota and Zoopagomycota. To include additional fungal TC proteins in the phylogenetic reconstruction, we identified TC proteins from 58 published genomes of Dikarya isolates (21 Basidiomycetes and 37 Ascomycetes, Table S3 (Armaleo et al. 2019, Burmester et al. 2011, DiGuistini et al. 2011, Yang et al. 2011, Zuccaro et al. 2011, Floudas et al. 2012, Morin et al. 2012, Ohm et al. 2012, Padamsee et al. 2012, Xiao et al. 2012, Gosticnar et al. 2014, Pendleton et al. 2014, Riley et al. 2014, Terfehr et al. 2014, Wang et al. 2014, Kroken et al. 2015, Okagaki et al. 2015, Nagy et al. 2016, Nguyen et al. 2017, Teixeira et al. 2017, Martino et al. 2018, Vesth et al. 2018, Crouch et al. 2020, Haridas et al. 2020)), available in DOE-JGI MycoCosm, each from a different family. One isolate per family was randomly selected for the analysis. To screen for TC proteins in Dikarya, OrthoFinder was used to build orthologous clusters of genes between Basidiobolus, Mucoromycota and Zoopagomycota species. Basidiobolus isolates were clustered within orthologous groups that contain TC of Basidiobolus, Mucoromycota and Zoopagomycota were considered valid TC orthologs and used in downstream analyses. Finally, we identified bacterial TC to evaluate the potential for HGT into Basidiobolus, Mucoromycota or Zoopagomycota species. Bacterial TCs were identified by screening each protein from the RefSeq, release 87 (May 2018, O’Leary et al. 2015) FASTA dataset against a BLAST database (Altschul et al. 1997), one from each of the zygomycete orthologous groups identified by OrthoFinder in the previous step, for a total of four protein sequences in the database. The BLASTP program was used to perform the searches, using an e-value of 1e-10 and the BLOSUM62 substitution matrix. Positive matches from the BLASTP assay were used as bacterial TC for subsequent analyses. A multi-sequence alignment containing all predicted TC from Basidiobolus, Mucoromycota and Zoopagomycota species, the orthologous TC from additional fungal species, and the TC proteins from reference bacterial predicted gene models result of the BLASTP search was performed in MAFFT v7 using the G-INS-I algorithm for progressive global alignment (Katoh et al. 2017) (File S5). A phylogenetic tree was reconstructed
Horizontal gene transfer in *Basidiobolus* species

Identification of genic regions with evidence for horizontal gene transfer (HGT) from bacteria was performed by searching all translated proteins from the predicted gene models of the *Basidiobolus meristosporus* isolates against a custom BLAST protein database. This database included all amino-acid sequences available in the NCBI RefSeq proteomics database and all available amino-acid sequences for *Mucoromycomycota* and *Zoopagomycota* species at the DOE-JGI MycoCosm genome portal. The proteins were searched with BLASTP against the combined RefSeq/Mucoromycomycota/Zoopagomycota custom database, using an e-value cutoff of $1e^{-10}$ and the BLOSUM62 substitution matrix. A summary of the taxonomy identifier for all best hits was obtained by identifying whether the best hit was a *Mucoromycomycota/Zoopagomycota* protein, or a RefSeq protein. We used the rentrez package (Winter 2017) in the R statistical framework to extract the top-ten taxonomic identifiers from the NCBI database when hits did not correspond to *Mucoromycomycota/Zoopagomycota*. Proteins that had no hits to a fungal protein and only hits to a bacterial protein were considered candidate HGT genes.

To increase the accuracy of the prediction of genes product of HGT, we tested whether HGT candidate genes showed signs of errant assembly or in silica incorporation into the *Basidiobolus* genomes. The mean genomic read coverage of each candidate gene was calculated and compared to the mean coverage of the scaffold harboring the HGT candidate gene for all three *Basidiobolus* isolates. A z-score was calculated in R to determine the number of standard deviations of the HGT candidate from the mean coverage of the harboring scaffold. All HGT candidates with standard deviations greater or less than two were removed from the analysis. The genomic reads were mapped to each reference genome using BWA (Li and Durbin 2009). Coverage per HGT and harboring scaffold was estimated using samtools depth (Li et al. 2009). A Kruskal-Wallis test was performed to determine significant differences across the siderophore activity of the strains.

**Measuring siderophore activity in *Basidiobolus***

To measure the siderophore activity (chelation of ferric ions) of *Basidiobolus*, an assay of detection of siderophore activity based on the universal chrome azurul S (CAS) assay (Andrews et al. 2016) was performed for the strain *Basidiobolus meristosporus* CBS 931.73. This colorimetric assay uses a complex of Fe(III) – CAS – DDAPS (Surfactant). When this complex is combined with acetate yeast agar (AY agar), it results in a greenish-blue color, where the color changes to yellow upon the removal of the iron. A total of 450 ml of AY agar (Andrews et al. 2016) was mixed with 50 ml of autoclaved 10X CAS assay (20 ml of 10 mM Fe(NO₃)₃, 40 ml of 10 mM chrome azuro S, and 100 ml of 10 mM DDAPS). A 1 cm layer of the AY-CAS media was poured in small petri dishes and cooled. An upper layer of 10 cm of AY agar was poured after cooling, and the media was left to diffuse overnight and stored at 4°C for 24 hr. *B. meristosporus, Conidiobolus thomboides* FSU 785 (negative control), and *Cladosporium sp.* from the *herbarum* species complex PE-07 (positive control) were transferred into the AY-CAS plates by transferring a small amount of mycelium via a sterile toothpick and piercing the media in the center. The assay was performed in triplicate for each isolate used. Cultures were grown at room temperature for 12 days. Siderophore activity was measured as the area of the plate that has changed to yellow color, when compared to a negative control and an empty petri dish with AY-CAS agar. Pictures of the plates were taken. These images were imported into Adobe Photoshop CC 2019 and size-corrected to 5.5 cm (diameter of the small petri dish). All images were concatenated in the same file. The Color Range tool of Adobe Photoshop CC tool was used to measure the yellow area for all plates. The area measurements were exported into R, where an analysis of variance was performed to determine significant differences across the siderophore activity of the strains.

**Data availability**

File S1 includes the ortholog groups of SM for *Basidiobolus* species. File S2 includes the alignment of all A-domains from NRPS sequences used in the study. File S3 includes the alignment of A-domains from *Zoopagomycomycota* and *Mucoromycomycota* NRPS sequences. File S4 includes the KS alignment for all PKS sequences used in this study. File S5 includes all terpene cyclase core genes used in this study. Figure S1 contains maximum likelihood reconstruction of NRPS-A domains. Figure S2 contains graphical representation of cluster 5 from *B. meristosporus* CBS 931.73. Figure S3 contains phylogenetic sources and abundance of KS domains. Figure S4 contains maximum likelihood reconstruction of PKS KS domains. Figure S5 contains presence and absence of domains characteristic of PKS core gene models in non-zygomycete fungal species. Figure S6 contains presence and absence of domains characteristic of PKS core gene models in zygomycete fungal species. Figure S7 contains phylogenetic sources and abundance of terpene cyclases.

The 5-mer analysis was conducted in all predicted coding sequences from the annotated gene models using the oligonucleotideFrequency function of the Biostring R package (Pagès et al. 2019). Codon usage was estimated using the uco function of the seqinr R package (Charif and Lobry 2007). Both these indices were divided by gene length to normalize the nucleotide composition by the effect of gene length. Principal component analyses were performed in R for both 5-mer and codon usage analysis to compare the HGT candidates to the fungal candidate genes. Lastly, the putative functions of the genes with evidence for HGT were summarized using the functional annotations available in the gene format file (GFF) of *B. meristosporus* CBS 931.73.
comparisons for HGT genes. Figure S10 includes a PCA 5-mer analysis for HGT genes. Figure S11 includes a PCA for codon usage analysis for HGT genes. Figure S12 includes siderophile activity assays. Table S1 includes all Mucoromycota and Zoopagomycota isolates used in the manuscript. Table S2 is a summary of secondary metabolite gene models predicted for Basidiobolus species. Table S3 includes the isolates of Dikarya used for the terpene cyclase assay. Table S4 includes the summary of hits of HGT genes against NCBI taxonomy. Table S5 is a summary of gene annotations from HGT candidates. Supplementary files, tables and figures have been deposited to figshare. All additional files included as supplementary materials in this manuscript and custom scripts used for this analysis can be found in the ZyGoLife GitHub repository (https://github.com/zygolife/Basidiobolus_SM_repo). Supplemental material available at figshare: https://doi.org/10.25387/g3.12691859.

RESULTS

Secondary metabolite gene cluster prediction

A total of 834 SM gene clusters were predicted for the 69 Mucoromycota and Zoopagomycota genomes including 117 non-ribosomal peptide synthetases (NRPS), 204 NRPS-like, 103 polyketide syntheses (PKS), 97 PKS-like, one PKS-NRPS hybrid, and 312 terpene cyclase (TC) gene models across both phyla (Figure 1, Table S1). For Basidiobolus species, 3874 total SM gene models were predicted for any member of Mucoromycota or Zoopagomycota, 300 SM gene models were predicted, including 45 NRPS, 143 NRPS-like, 103 polyketide synthases (PKS), 97 PKS-like, and 59 TC gene models. In Mucoromycota, 534 total SM gene models were predicted, including 45 NRPS, 143 NRPS-Like, 43 PKS, 50 PKS-Like, and 253 TC gene models.

The highest number of SM gene clusters were predicted for Basidiobolus. A total of 38 gene clusters and 44 SM core genes were predicted in the B. meristosporus CBS 931.73 genome, 40 SM gene clusters and 44 SM core genes for B. meristosporus B9252, and 23 SM gene clusters and 23 SM core genes for B. heterosporus B8920 (Table 1; Table S2). Seventy-eight percent of the SM gene models predicted were found to be shared across the Basidiobolus isolates. Ten SM core genes were found to be unique to B. meristosporus CBS 931.73, 10 SM gene models unique to B. meristosporus B9252, and 5 SM genes unique to B. heterosporus B8920 (File S1).

The next three isolates with the most numerous predicted SM gene clusters, not including Basidiobolus genomes, were Dimargaris cristalligena RSA 468 with 33 predicted SM proteins (21 NRPS, 7 NRPS-Like, 2 PKS-Like, 3 TC), Linderia pennsinsula ATCC 12442 V 1.0 with 23 SM predicted (1 NRPS, 15 PKS, 3 PKS-Like, 4 TC), and Martensiomyces pterosporus CBS 209.56 v1.0 with 23 SM proteins predicted (1 NRPS, 5 PKS, 14 PKS-Like, 3 TC). No DMAT gene models were predicted for any member of Mucoromycota or Zoopagomycota.

Expression of core SM genes

A total of 83.45% of the RNA sequenced reads were mapped uniquely to the reference genome of B. meristosporus CBS 931.73, while 12.08% of the reads were mapped in more than one location. Only 4.47% of the RNA sequenced reads did not map to the reference genome. The majority of predicted SM for B. meristosporus CBS 931.73 were expressed at the same or higher levels than constitutive housekeeping genes, such as Beta-tubulin, Elongation Factor 1, Actin, and Ubiquitin (Figure 2). The highest expressed SM core genes per SM group were: NRPS – gene model 387529 (Cluster 5) with 74.03 TPM mapped; NRPS-like – gene model 221915 (Cluster 20) with 45.58 TPM mapped; PKS – gene model 290138 (Cluster 37) with 687.55 TPM mapped; PKS-like – gene model 207695 (Cluster 38) with 26.15 TPM mapped, and Terpene cyclase – gene model 301341 (Cluster 13) with 78.26 TPM mapped.

Phylogenetic analysis of NRPS/NRPS-Like A-domains

The phylogenetic reconstruction of A-domains was performed with 951 A-domains from a combined dataset including the A-domain dataset from Bushley and Turgeon (2010) and the predictions of NRPS/NRPS-like A-domains from Mucoromycota and Zoopagomycota genome sequences (File S2). A total of 395 NRPS A-domains were predicted for Basidiobolus, Mucoromycota and Zoopagomycota genome sequences (File S3). The phylogenetic analyses recovered the nine major families reported by Bushley and Turgeon (2010) with the addition of the two new clades, surfactin-like and ChNSP 12-11-like, reported here. The total number of Mucoromycota/Zoopagomycota A-domains and their distribution across these clades are as follows:

74 to the AAR clade; 115 to the major bacterial clade (MBC), including 104 A-domains with the surfactin-like clade with and an additional 11 A-domains scattered elsewhere in the MBC; the CYCLO clade with eight A-domains; 65 A-domains to the ChNSP 12-11-like clade; 25 A-domains to the ChNSP 12-11 clade; 11 A-domains to the PKS/NRPS clade; 16 A-domains to the SID clade; 76 A-domains to the SIDE clade; and 5 A-domains to the EAS clade (Figure 3).

Of fungi classified in Mucoromycota or Zoopagomycota, Basidiobolus genomes contained the most A-domains clustered in the EAS, SIDE, and CYCLO clades. The A-domains clustered within EAS represent two domains from B. meristosporus CBS 931.73, and one A-domain of Rhizopagus irregularis DAO1 197198 v2.0. The A-domains clustered within SIDE represent 76 domains, 44 from B. meristosporus CBS 931.73, 19 from B. meristosporus B9252, and 13 from B. heterosporus. The A-domains clustered within the CYCLO clade represent six domains, four from B. meristosporus CBS 931.73 and two from B. meristosporus B9252. The NRPS A-domains grouped in the CYCLO clade correspond to cluster 5 NRPS from B. meristosporus CBS 931.73 (Protein ID 387529) and cluster 9 of B. meristosporus B9252 (Protein ID N161_4477). Gene model 387529 was predicted as a tetra-modular protein, with two N-methyltransferase domains and a thioesterase domain. In addition, AntiSMASH analyses predict a six gene clade (cluster 5) that includes gene model 387529, a zinc finger transcription factor (312194), carrier protein (277549), transporter (277552), peptidase (277554), and a SNARE associated protein (334759) (Figure S2).

All of these gene models have evidence of gene expression (Figure 2).

Of fungi classified in Mucoromycota or Zoopagomycota, Basidiobolus genomes contained the most A-domains clustered in the EAS, SIDE, and CYCLO clades. The A-domains clustered within EAS represent two domains from B. meristosporus CBS 931.73, and one A-domain of Rhizopagus irregularis DAO1 197198 v2.0. The A-domains clustered within SIDE represent 76 domains, 44 from B. meristosporus CBS 931.73, 19 from B. meristosporus B9252, and 13 from B. heterosporus. The A-domains clustered within the CYCLO clade represent six domains, four from B. meristosporus CBS 931.73 and two from B. meristosporus B9252. The NRPS A-domains grouped in the CYCLO clade correspond to cluster 5 NRPS from B. meristosporus CBS 931.73 (Protein ID 387529) and cluster 9 of B. meristosporus B9252 (Protein ID N161_4477). Gene model 387529 was predicted as a tetra-modular protein, with two N-methyltransferase domains and a thioesterase domain. In addition, AntiSMASH analyses predict a six gene clade (cluster 5) that includes gene model 387529, a zinc finger transcription factor (312194), carrier protein (277549), transporter (277552), peptidase (277554), and a SNARE associated protein (334759) (Figure S2). All of these gene models have evidence of gene expression (Figure 2).

The Basidiobolus NRPS A-domains clustered in the surfactin-like clade included 32 A-domains from eight gene models from B. meristosporus CBS 931.73. These included gene model 375475 (Cluster 19) with eight A-domains, gene models 307892 (Cluster 33) and 343011 (Cluster 35) with seven A-domains each; gene model 368581 (Cluster 8) with three A-domains; gene models 337511 (Cluster 16), 372991 (Cluster 17), and 298977 (Cluster 7) each with two A-domains; and gene model 322666 (Cluster 2) with one A-domain. B. meristosporus B9252 contained two surfactin-like A-domains from one gene model (N161_8304; Cluster 18). B. heterosporus possessed 13 A-domains from seven gene models, including N168_07733 (Cluster 6), N168_06479 (Cluster 13), N168_02885 (Cluster 1)
and N168_00034 (Cluster 14) with two A-domains each, and gene models N168_05934 (Cluster 8), N168_04239 (Cluster 12) and N168_07140 (Cluster 9) with one A-domain each.

Evolutionary relationships of predicted PKSs
A total of 388 KS domains were included in the phylogenetic reconstruction (File S4). A total of 253 domains were predicted here including: two KS domains from Chytridiomycota genomes, 21 from Neocallimastigomycota, 46 from Basidiomycota, 78 from Ascomycota, 44 from Mucoromycota, and 61 from Zoopagomycota (File S4). The 136 additional KS domains were obtained from Kroken et al. (2003). For Mucoromycota and Zygomycota, the genomes with the highest number of KS domains were Linderina pennispora ATCC 12442 v1.0 (15 KS domains), Coemansia reversa NRRL 1564 v1.0 (6 KS domains), Coemansia mojavensis RSA 71 v1.0 (6 KS domains), Martensiozymes pterosporus CBS 209.56 v1.0 (6 KS domains). The KS domain numbers for the Basidiobolus genomes included three for B. meristosporus CBS 931.73, two for B. meristosporus B9252, and one for B. heterosporus. All of the predicted PKS for non-Basidiobolus species were obtained from the DOE-JGI MycoCosm genome portal predictions, as AntiSMASH only predicted PKS for the Basidiobolus species.

The major clades of PKS proteins reported by Kroken et al. (2003) were recovered by this study, including the fungal reducing (R) PKS, animal fatty acid synthases (FAS), fungal non-reducing (NR) PKS, and bacterial PKS, as well as fungal FAS I and II clades (Figure S3 and S4). The Fungal R PKS1 clade included domains from B. meristosporus

Figure 1 Genome size, number of genes, proportion of secondary metabolism (SM) gene clusters and number of predicted SM gene clusters for 69 Zoopagomycota and Mucoromycota sequenced genomes. Color code represents the category of SM predicted per isolate. NRPS: Non-ribosomal peptide synthetases. PKS: Polyketide synthases. Glom.: Glomeromycotina. Mort.: Morteriellomycotina. Ent.: Entomophtoromycotina. Kickx.: Kickxellomycotina. Zoo.: Zoopagomycotina.
CBS 931.73 (2 KS domains) and B. meristosporus B9252 (2 KS domains). No KS domains of Mucoromycota or Zoopagomycota genomes were clustered within the animal FAS clade, the bacterial PKS clade, nor the fungal FAS II clade. The Fungal NR PKS clade contained a new clade, “non-reducing PKS proteins clade IV”, comprising KS domains from Basidiomycota, Neocallimastigomycota, and one KS domain of B. meristosporus B9252. The fungal FAS clade I comprised all the remaining KS domains of Mucoromycota and Zoopagomycota genomes, while the fungal FAS II clade comprised mostly KS domains from Neocallimastigomycota, Basidiomycota and one Chytridiomycota representative (Figure S3 and S4).

To better understand the predicted PKSs of Mucoromycota and Zoopagomycota that clustered in the fungal FAS clade, the patterns of domains that comprise fungal PKS protein sequences were analyzed. Predicted PKS of Ascomycota contained AT, KS and PP domains in all sequences (Figure S5). Predicted PKS from Basidiomycota contained AT and KS domains in all sequences, while KR and PP domains were present on 36% of the sequences. All PKS that contain a KS domain in the FAS clade are missing the PP domain. For Chytridiomycota, the predicted PKS with the KS domain clustered in the Fungal (R) clade contained AT, KS and DH domains, but no PP domain. Chytridiomycota PKS with KS domains associated to FAS clade only possessed AT and KS domains. For Neocallimastigomycota, 100% of PKS clustered in fungal R and NR clades contained AT and KS domains, however, only 5 PKS contained the PP domain. All Neocallimastigomycota PKS that clustered within the FAS clade contained only AT and KS domains. For Mucoromycota and Zoopagomycota, 100% of PKS contained either AT and/or KS domain, but no other domains (Figure S6). Some Basidiobolus PKS were differentiated from the other Mucoromycota and Zoopagomycota PKS as they possessed one or more of the DH or PP domains (Figure S6).

**Evolutionary relationships of predicted terpene cyclase gene models**

A total of 1,108 terpene cyclase or terpene cyclase-like gene models were predicted and used for phylogenetic analyses (Additional file 5). These include 253 identified in the Mucoromycota species genomes, 59 in the Zoopagomycota genomes, and 401 ortholog proteins from 58 fungal genomes of Basidiomycota and Ascomycota. An additional 395 TC candidates were identified in bacterial genomes from RefSeq.
via BLASTP. The phylogenetic reconstruction of TC resulted in two main clades: an outgroup clade that comprised predicted TC annotated as tRNA threonylcarbamoyladenosine dehydratase and phytoene synthases, and a main ingroup TC core clade (Figure 5, Figure S7). The tRNA threonylcarbamoyladenosine dehydratase subclade contained mostly bacterial sequences plus one TC gene model of \textit{B. meristosporus} B9252 and \textit{B. meristosporus} CBS 931.73 each. The phytoene synthases clade comprised two subclades, one clade that grouped most bacterial gene models, and a second clade clustering bacterial and Mucoromycota predicted TC gene models, but no \textit{Basidiobolus} TC core gene models were found in this clade.

The TC core clade (Figure 5 and Figure S7) comprised a bacterial exclusive clade (Bacteria I), followed by four highly supported subclades containing TC core genes from Mucoromycota, Zoopagomycota, and Dikarya, and are referred to here as Fungi TC clades I – IV.

The majority of TC genes from Mucoromycota and Zoopagomycota clustered within clade Fungi TC II, which included no other fungal or bacterial TC. These genes were annotated as Squalene synthetases by the MycoCosm genome portal. The Fungi TC I, III and IV clades included mostly Mucoromycota TCs. The Fungi TC I clade contained TC genes from Mucoromycota isolates, as well as TC genes from the ascomycete isolates \textit{Fusarium verticillioides} 7600 and \textit{Hypoxylon sp.} EC38 v1.0, and the basidiomycete \textit{Suillus luteus} UH-Slu-Lm8-n1 v2.0. These gene models were annotated as associated with the Ubiquitin C-terminal hydrolase UCHL1 according to the MycoCosm portal. Fungi TC III clade also included TC genes from Ascomycota sequenced genomes. The majority of these genes were annotated as containing a terpene synthase family, metal binding domain, and a polypropenyl synthetase domain in the MycoCosm genome portal.

Finally, Fungi TC IV clade included TC genes from Ascomycota sequenced genomes. The majority of these genes were annotated as containing a terpene synthase family, metal binding domain, and a polypropenyl synthetase domain in the MycoCosm genome portal. The terminal clades of TC clusters include the Bacteria II + \textit{Basidiobolus} clade that included bacterial TC and six TC SM core genes predicted from \textit{Basidiobolus} (two from each genome), and the clade comprising TC core genes solely from Dikarya and one gene from \textit{Mortierella verticillata} NRRL 6337 and \textit{Rhizopus microsporus} ATCC11559 v1.0 (Figures 5 and Figure S7).

\textbf{Signatures for HGT in Basidiobolus genomes}

The identity search for genes with evidence for HGT identified 934 genes in \textit{B. meristosporus} CBS 931.73, 620 genes of \textit{B. meristosporus} B9252, and 382 genes of and \textit{B. heterosporus} B8920 with zero BLASTP hits to fungal proteins. These genes were used for the

\textbf{Figure 3}

Phylogenetic sources and abundance of A-domains from NRPS predicted gene models for Zoopagomycota and Mucoromycota species. The maximum likelihood phylogenetic tree (left) represent a simplification of the reconstructed tree which includes the clades that include more than one A-domain from Mucoromycota/Zoopagomycota genomes. The heatmap (right) represents the abundances A-domains predicted for each domain clustered within each clade, including the dikarya species included in the analysis. \textit{Basidiobolus} species are enriched in NRPS from the major bacterial clade (MBC), cyclosporin (CYCLO), surfactin-like, and siderophore (SIDE, SID) clades.
coverage assay to identify gene model coverage deviation from the harboring scaffold median coverage (Figure S8). This assay resulted in 810 genes of B. meristosporus CBS 931.73, 532 genes of B. meristosporus B9252, and 301 genes of and B. heterosporus B8920 with z-scores under 2 standard deviations from the harboring scaffold median coverage. Finally, 4 HGT genes from B. meristosporus CBS 931.73, 106 genes from B. meristosporus B9252, and 89 genes from B. heterosporus B8920 were not located in scaffolds surrounded by genes with best fungal hits. All other HGT genes were located in scaffolds surrounded by fungal genes: 806 from from B. meristosporus CBS 931.73, 426 genes from, B. meristosporus B9252, and 212 genes of of and B. heterosporus B8920. These candidate genes with evidence for HGT represented 5%, 3% and 2% of the gene content of B. meristosporus CBS 931.73, B. meristosporus B9252, and B. heterosporus B8920, respectively. The HGT candidates showed significant differences in GC content (F-value = 6.395, df = 2, p-value = 0.00167) when contrasted to all genes considered of fungal origin (Mean fungal GC content = 0.4865, Mean HGT GC content = 0.4911, Tukey HSD p-value = 0.001), but no differences were detected in GC content between HGT candidates and fungal genes that co-occur on common scaffolds (Mean surrounding gene GC content = 0.488, Tukey HSD p-value = 0.30). Similarly, no differences in codon usage, or in 5-mer nucleotide composition (Figures S10 and S11, respectively), were found between HGT candidates and genes of fungal origin. Differences were found, however, in intron number and normalized intron length between genes HGT candidate genes and fungal genes, where the distribution of intron number showed a median of 0 introns for HGT candidates and 2 for fungal genes (Kruskal-Wallis test; \( \chi^2 = 272.88, df = 1, p-value < 2.2e-16 \); Figure S9). The majority of HGT candidate genes (61% of the genes) have no introns, compared to 36% of genes from fungal origin with no introns (Figure S9). The candidate HGT genes have a significantly smaller normalized intron length than the genes with a fungal origin (Kruskal-Wallis test; \( \chi^2 = 272.88, df = 1, p-value < 2.2e-16 \)).

A large percentage of SM core genes appeared to be the product of HGT from bacterial species into Basidiobolus. The SM core genes identified as candidate HGT genes is 61%, 27% and 30% for B. meristosporus CBS 931.73, B. meristosporus B9252, and B. heterosporus B8920, respectively (Table 2). NRPS/NRPS-like SM core genes represent the largest percentage of HGT evidence, while TC have the lowest percentage of HGT evidence (Table 2, Table 3). The identification of taxonomic sources for HGT into Basidiobolus indicated that most HGT comes from bacteria in the phylum Proteobacteria with 232, 135 and 69 gene models in B. meristosporus CBS 931.73, B. meristosporus B9252, and B. heterosporus B8920, respectively. Firmicutes (112, 68, and 35 gene models, respectively) and Actinobacteria/High Gram C+ (127, 53, and 25 gene models, respectively) (Figure 6, Table S4) were the second and third most abundant source of HGT from bacteria into Basidiobolus.

Most HGT candidate genes resulted in no gene ontology (GO) annotation (655 gene models) and/or no InterPro domain annotation (140 gene models, Table S5). The ten top GO terms were oxidation-reduction process (59 gene models), protein binding (34 gene models), phosphorelay sensor kinase activity (32 gene models), N-acetyltransferase activity (28 gene models), catalytic activity (26 gene models), extracellular space (25 gene models), hydrolase activity (24 gene models), oxidoreductase activity (23 gene models), hydrolase activity (24 gene models), hydrolyzing O-glycosyl compounds (23 gene models), and NRPS (18 gene models) (Table S5).

Siderophore activity in Basidiobolus meristosporus

The siderophore activity assay resulted in visible activity for the three replicates for B. meristosporus and Cladosporium sp., and a small halo for one of the replicates of C. thromboides. No visible siderophore activity was detected for the other replicates of C. thromboides or for the empty AY-CAS plates (Figure 7, Figure S12). The analysis of
variance for the area of siderophore activity measured showed significant differences across isolates/empty plates (ANOVA, F value = 19.62, \(P = 0.0004\)). Post-hoc tests showed no differences between *B. meristosporus* and *Cladosporium sp.* siderophore activity (Tukey HSD, \(P = 0.9801489\)) nor between *C. thromboides* and the empty AY-CAS plates (Tukey HSD, \(P = 0.9396603\)), but significant differences were found for all other comparisons (Tukey HSD, \(P < 0.001\) for all remaining comparisons).

**DISCUSSION**

Secondary, or specialized, metabolite (SM) production is an important element of fungal metabolism. It has resulted in numerous natural products with human health implications, such as mycotoxins in our food supply, and medical applications, including antibiotics, immunosuppressants, and antitumor agents. The majority of the known genetic and chemical diversity of fungal SM has been described from the phyla Ascomycota and Basidiomycota with few SM gene clusters, and limited SM production, reported for fungi in Mucoromycota and Zoopagomycota. This observation has led to the dogma that ‘zygomycete’ species are depauperate of these chemical pathways (Voigt et al. 2016). Here we report gene clusters involved in SM production in the largest survey to date of Mucoromycota and Zoopagomycota species using genomics approaches and estimate the SM potential of these fungi.

Genome sequencing of a total of 69 isolates across diverse lineages of Mucoromycota and Zoopagomycota enabled detailed identification of SM clusters. These results support the hypothesis that zygomycete fungi have a low abundance of secondary metabolism (Figure 1, Table S1) and agree with previous reports (Voigt et al. 2016). Outliers to this pattern exist, however, and are particularly true of the genus *Basidiobolus* (Zoopagomycota), which possesses a large number of SM gene clusters predicted for the NRPS, PKS and TC families. This discovery of abundant candidate genes for production of secondary metabolite in *Basidiobolus* is novel and its presence is most consistent with a signal of horizontal gene transfer from bacteria to fungi, a phenomenon we propose is facilitated by living in the amphibian gut environment.

**Distribution and evolution of NRPS genes across Mucoromycota and Zoopagomycota**

A deeper examination of *Basidiobolus* SM gene clusters indicates that this genus surpasses the number of expected NRPS genes for zygomycete species (Figure 1). Most of these SM genes also show evidence for transcription, indicating that the majority of these genes are expressed constitutively under laboratory conditions (Figure 2). Several of these core genes, such as the NRPS gene model 387529, appear to be expressed at a higher rate than several housekeeping genes.

**Table 2 Number of predicted secondary metabolite core genes with evidence for HGT in *Basidiobolus* genomes.** NRPS: Non-ribosomal peptide synthetases, PKS: Polyketide synthases, TC: Terpene cyclases

| Isolate          | Total SM | NRPS/NRPS-like | PKS/PKS-Like | NRPS-PKS hybrids | TC    |
|------------------|----------|----------------|--------------|------------------|-------|
| *B. meristosporus* CBS 931.73 | 26/44 (61%) | 22/30 (60%) | 2/4 (50%) | 0/0 (0%) | 2/10 (2%) |
| *B. meristosporus* B9252 | 12/44 (27%) | 9/30 (30%) | 1/7 (14%) | 0/1 (0%) | 2/6 (1.4%) |
| *B. heterosporus* B8920 | 7/23 (30%) | 7/18 (40%) | 0/1 (0%) | 0/0 (0%) | 0/4 (0%) |
A more detailed census of core genes reveals unique patterns of evolution of NRPS genes across Mucoromycota and Zoopagomycota when compared to Dikarya fungi (Figure 3, Table 1). The only NRPS A-domains found throughout the two phyla are members of the AAR clade (with the exception of *Rhizopus delemar* 99-80 and *Piptoccephalis cylindrospora* RSA 2659). These A-domains are from the genes that encode α-aminoadipate reductases (AAR), an enzyme responsible for the reduction of α-aminoadipic acid, which is essential for the lysine biosynthesis pathway and is present in all fungal phyla (Bushley and Turgeon 2010). In contrast, the remaining NRPS genes, and their respective A-domains, show discontinuous and patchy distributions across Mucoromycota and Zoopagomycota (Figure 3).

The most pronounced NRPS diversification in *Basidiobolus* are for genes that are predicted to encode for siderophores, iron chelating metabolites. A-domains for predicted siderophores are distributed throughout four clades including the three clades of major bacterial genes and the fungal SID and SIDE clades (Figure 3, Figure S1). Major bacterial clades (MBC) exclusively comprise bacterial siderophore synthases, such as pyoverdine, yersiniabactin and pyochelin (Bushley and Turgeon 2010), with the exception of the surfactin-like clade. Our results show that genomes of *Mortierella* and *Basidiobolus* contain A-domains that are members of the MBC, and that they are the only fungal representatives of these clades. SID clade contains all NRPS associated with siderophore production from Ascomycota and Basidiomycota species. All *Basidiobolus* isolates contain one NRPS A-domain in this clade, as well as three A-domains from three NRPS gene models of *Conidiobolus coronatus* NRRL28638. Finally, SIDE, a clade comprising NRPS genes responsible for the production of siderophores in filamentous ascomycetes is expanded in *Basidiobolus* (Figure 3), which is also the only zygomycete with A-domains

Table 3 Summary of secondary metabolite core genes with HGT evidence from *Basidiobolus* isolates

| Isolate | Gene   | Z-score of gene coverage | SM class       | Taxonomy for best NCBI hit |
|---------|--------|--------------------------|----------------|---------------------------|
| B. heterosporus | N168_02885 | -0.104 | NRPS | b-proteobacteria |
| B. heterosporus | N168_05934 | -0.183 | NRPS | b-proteobacteria |
| B. heterosporus | N168_07140 | 0.260 | NRPS | b-proteobacteria |
| B. heterosporus | N168_06479 | 0.328 | NRPS | b-proteobacteria |
| B. heterosporus | N168_00176 | 0.013 | NRPS | cyanobacteria |
| B. heterosporus | N168_05966 | 0.672 | NRPS | d-proteobacteria |
| B. heterosporus | N168_08580 | 0.047 | NRPS-Like | d-proteobacteria |
| B. meristosporus B9252 | N161.mRNA.1431.1 | -0.350 | NRPS | cyanobacteria |
| B. meristosporus B9252 | N161.mRNA.1485.1 | -0.258 | NRPS | d-proteobacteria |
| B. meristosporus B9252 | N161.mRNA.4115.1 | 0.058 | NRPS | enterobacteria |
| B. meristosporus B9252 | N161.mRNA.4324.1 | 0.213 | NRPS | CFB group bacteria |
| B. meristosporus B9252 | N161.mRNA.8304.1 | 0.358 | NRPS | b-proteobacteria |
| B. meristosporus B9252 | N161.mRNA.11289.1 | 0.596 | NRPS | cyanobacteria |
| B. meristosporus B9252 | N161.mRNA.1486.1 | 0.662 | NRPS-Like | firmicutes |
| B. meristosporus B9252 | N161.mRNA.6486.1 | 0.760 | NRPS-Like | firmicutes |
| B. meristosporus B9252 | N161.mRNA.8699.1 | -0.201 | NRPS-Like | firmicutes |
| B. meristosporus B9252 | N161.mRNA.6146.1 | 0.999 | PKS | high GC Gram+ |
| B. meristosporus CBS 931.73 | 366903 | 0.043 | NRPS | a-proteobacteria |
| B. meristosporus CBS 931.73 | 368581 | -0.232 | NRPS | b-proteobacteria |
| B. meristosporus CBS 931.73 | 349800 | 0.208 | NRPS | high GC Gram+ |
| B. meristosporus CBS 931.73 | 351909 | -0.120 | NRPS | cyanobacteria |
| B. meristosporus CBS 931.73 | 372749 | -0.088 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 372991 | -0.260 | NRPS | b-proteobacteria |
| B. meristosporus CBS 931.73 | 373247 | 0.533 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 373940 | -0.142 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 375475 | 0.522 | NRPS | enterobacteria |
| B. meristosporus CBS 931.73 | 338397 | 0.033 | NRPS | d-proteobacteria |
| B. meristosporus CBS 931.73 | 375580 | 0.033 | NRPS | cyanobacteria |
| B. meristosporus CBS 931.73 | 375613 | -0.369 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 377413 | -0.245 | NRPS | g-proteobacteria |
| B. meristosporus CBS 931.73 | 387529 | -0.157 | NRPS | b-proteobacteria |
| B. meristosporus CBS 931.73 | 307892 | 0.404 | NRPS | a-proteobacteria |
| B. meristosporus CBS 931.73 | 298977 | -0.201 | NRPS | b-proteobacteria |
| B. meristosporus CBS 931.73 | 343011 | 0.147 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 363930 | -0.358 | NRPS | firmicutes |
| B. meristosporus CBS 931.73 | 300898 | -0.165 | NRPS-Like | firmicutes |
| B. meristosporus CBS 931.73 | 322666 | -0.174 | NRPS-Like | CFB group bacteria |
| B. meristosporus CBS 931.73 | 146993 | -0.200 | NRPS-Like | cyanobacteria |
| B. meristosporus CBS 931.73 | 382467 | -0.264 | NRPS-Like | a-proteobacteria |
| B. meristosporus CBS 931.73 | 340613 | -0.136 | NRPS-Like | GNS bacteria |
| B. meristosporus CBS 931.73 | 237744 | -0.001 | PKS | bacteria |
| B. meristosporus CBS 931.73 | 292783 | -0.429 | PKS | high GC Gram+ |
| B. meristosporus CBS 931.73 | 301341 | -0.210 | Terpene | CFB group bacteria |
| B. meristosporus CBS 931.73 | 304520 | 0.389 | Terpene | cyanobacteria |
clustered in this clade. These findings are consistent with enrichment of both bacterial and fungal siderophores and are suggestive of the importance of iron metabolism in *Basidiobolus*.

The CYCLO clade contains the A-domains for core genes associated with biosynthesis of cyclic peptides, such as beauvericin and cyclosporin (Bushey and Turgeon 2010; Bushey et al. 2013). Sister to all other CYCLO clade A-domains are the A-domains that comprise the NRPS core gene model 387529 from *B. meristosporus* (Figure S2). 387529 is expressed at the highest rate of any SM gene under laboratory conditions (Figure 2). It is annotated as a tetra-modular gene model, that includes two N-methylation domains, four adenylation domains, four condensation domains, and a TE domain. When compared to *simA*, the NRPS responsible for biosynthesis of cyclosporin, structural similarities can be found, such as the presence of the N-methylation domains and the TE terminator domain. Its phylogenetic and structural similarities to *simA* suggest that 387529 results in the synthesis of a cyclic peptide with methylated amino acid residues. Cyclopentapeptides were recently described from *B. meristosporus* (Gene model N161_14278) clustered with the *B. subtilis* SrfA-A domain (Gene model N161_14278) clustered with the *B. subtilis* SrfA-A domain. The placement of this NRPS-PKS A-domain is interesting because surfactins are lipopeptides, which contain a hydrophobic fatty acid chain, whose biosynthesis is consistent with an NRPS-PKS hybrid. The A-domains from the remaining Mucoromycota and Zoopagomycota lack PKS diversity

**Figure 7** Siderophore activity of *Basidiobolus meristosporus* CBS 931.73 in a universal CAS assay using layered AY-CAS plates after 12 days. Bars represent mean siderophore activity measured per strain as the yellow area in AY-CAS plates for three replicates. Error bars represent the standard deviation for each replicate. Cladosporium sp. PE-07 represents the positive control. Conidiobolus thromboideus FSU 785 represents a zygomycete with no evidence for siderophore NRPS expansion.

including the gene models 298977, 368581 and 372991 from *B. meristosporus* CBS 931.73. Gene model 298977 showed no evidence for gene expression, while gene models 368581 and 372991 show high rates of expression (Figure 2). This is the first report of the prediction of a surfactin-like gene in fungi, but surfactant production was recently reported in *Mortierella alpina* (Baldeweg et al. 2019). These include malpinins, amphiphilic acetylated hexapeptides that function as natural emulsifiers during lipid secretion, and malpibaldins, hydrophobic cyclopentapeptides. This finding is consistent with the genomic data and reveals that *Mortierella*, in addition to *Basidiobolus*, possesses homologs in the surfactin-like clade that are phylogenetically different from *B. subtilis* surfactins.

Surfactins, encoded by the *srfA* gene cluster in *Bacillus subtilis*, are functionally active as surfactants, as well as toxins and antibiotics. However, the surfactin genes from *B. subtilis* (*SrfA-AA, SrfA-AB and SrfA-AC*) were included in our analysis and clustered in a different clade within the MBC. A-domains from single module NRPS-like protein from *B. meristosporus* CBS 931.73 (Gene model 146993) and from a NRPS-PKS hybrid A-domain from of *B. meristosporus* B9252 (Gene model N161_14278) clustered with the *B. subtilis* SrfA genes. The placement of this NRPS-PKS A-domain is interesting because surfactins are lipopeptides, which contain a hydrophobic fatty acid chain, whose biosynthesis is consistent with an NRPS-PKS hybrid. The A-domains from the remaining Mucoromycota and Zoopagomycota lack PKS diversity

Polyketide synthases (PKS) are abundant SM of Ascomycota and Basidiomycota and are involved in antibiotic production, carotenoid biosynthesis and other functional roles. In contrast, literature on PKS diversity for Mucoromycota and Zoopagomycota is limited. Our analyses update the phylogenetic reconstruction reported by Kroken et al. (2013) by adding genomic information from Chytridiomycota, Mucoromycota and Zoopagomycota (Figures S3 and S4). Overall, we
report the discovery of two new clades: A clade of non-reducing PKS proteins (clade IV) comprising Neocastimastigomycota and Basidiobolus, and a novel FAS II clade consisting of Neocastimastigomycota, Chytridiomycota and Basidiomycota KS domains. The results of our PKS prediction revealed a number of potential PKS core genes in zygomycete species (Figures S3 and S4), but the results of our phylogenetic reconstruction show that the KS domain of the predicted PKS gene models are fungal fatty acid synthases (FAS). Only Basidiobolus meristosporus genomes possessed KS domains associated with fungal PKS, either reducing or non-reducing, and only B. meristosporus CBS 931.73 possessed a PKS in the Fungal FAS clade.

A domain-by-domain presence/absence analysis of PKS gene models (Figures S5 and Figure S6) shows that in addition to AT and KT domains, a third domain (either KR, DH or PP) is found in the majority of fungal PKSs (Figure S5). Conversely, the zygomycete genes in the FAS clade only possess the AT and/or KT domains, which are domains common between FAS and PKS. However, Basidiobolus KS domains are found in fungal reducing PKS clades. Gene models 292783 and 237744 from both isolates of B. meristosporus cluster with fungal reducing PKS II and possess the DH domain (Figure S4), and the expression levels of 292783 is consistent with an actively transcribed gene (Figure 2).

**Terpene cyclase-like genes are expanded in zygomycetes**

Terpene cyclases (TC) are the most common SM predicted for zygomycete species (Figure 1, Table S1). Phylogenetic reconstruction shows that zygomycete TC core genes are clearly distinct from Ascomycota and Basidiomycota TC, where at least 4 new clades of predominantly zygomycete TC are found (Figure 5, Figure S7). Fungi II TC clade comprises TC from all zygomycte genomes analyzed with the exception of Piptocephalis cylindrospora RSA 2659 single—cell v3.0, Phycymyes blakesleeanus NRRL1555 v2.0, and five genomes of Rhizopogus irregularis, which show no prediction of TC in their genomes. No Dikarya TC genes cluster within this clade. Fungi I, III and IV group TC genes are found almost exclusively in Mucoromycota species, as well as some Dikarya species. TC genes from Fungi I are present only in Mucoromycotina genomes. Functional annotations of TC genes from this clade indicate that these TC genes code for proteins associated with squalene and phytoene synthases and are part of the synthesis of carotenoids. Carotenoids are important compounds for the synthesis pathway of trisporic acid, the main molecule responsible for initiating sexual reproduction in zygomyctes (Burmester et al. 2007). Finally, Basidiobolus is the only non-Dikarya genus with TC genes clustered within the Bacteria I clade of TC. Both these SM core genes show evidence for expression comparable to housekeeping genes or other SM core genes (Figure 2). The presence of bacterial-like TC genes in Basidiobolus present more evidence on the plasticity of the genome of Basidiobolus and its ability to integrate and possibly express foreign SM associated DNA.

**Horizontal gene transfer of SM genes to Basidiobolus**

Basidiobolus is a genus with a complex biology, alternating ecologies, and multiple spore types. This complex biology is also reflected at the genomic scale. The sequenced genomes show a larger genome size than other zygomycetes, as well as a higher number of genes than any other Zoopagomycota genomes sequenced to date. Our SM prediction assay is concordant with these patterns in which Basidiobolus has an excess of SM gene clusters when compared to other zygomycetes. Evidence points to HGT as a main driver of SM diversity in Basidiobolus as supported by the phylogenetic reconstructions of NRPS and TC gene clusters with bacterial homologs. Basidiobolus and Mortierella are the only fungi with genes associated with the bacterial clades in each of these SM phylogenetic reconstructions. Moreover, these HGT candidates are integrated into the Basidiobolus genome assembly and do not show evidence of artificial assembly as evidenced by discontinuous coverage (Figure S8). The most abundant functional group of SM core genes overall are siderophores and their overall functionality is supported by both the RNA expression analyses (Figure 2) and siderophore plate assays (Figure 7).

In one stage of the Basidiobolus life cycle, the fungus lives as a gut endosymbiont where it co-occurs with bacteria and other organisms that comprise the gut microbiome. Animal gut environments can facilitate HGT between bacteria and fungi, as previously reported for the zoosporic species of Neocallimastigomycetes (Chytridiomycota), which live in the ruminant gut environment and whose genomes exhibit a 2–3.5% frequency of genes with HGT evidence (Wang et al. 2019; Murphy et al. 2019). Phylogenomic analyses of the Basidiobolus NRPS A-domains support a phylogenetic affinity with A-domains from bacterial taxa and more rarely other fungi, a pattern most consistent with HGT. Our HGT survey comprised an extensive search of reference genomes across the tree of life available in NCBI RefSeq, as well as all of the gene models predicted for the Mucoromycota and Zoopagomycota genomes. We find that 2–5% of all predicted gene models present in Basidiobolus genomes are consistent with signatures of HGT from bacteria. However, the percentages change from 27 to 61% of predicted SM core genes models with bacterial HGT evidence. It should be noted that the two genomes with the lower percentage of predicted HGT SM genes, B. meristosporus B9252 (27%) and B. heterosporus B8920 (30%), are more fragmented, Illumina-only genomes, which complicates assembly of biosynthetic gene clusters. The SM gene models with bacterial signatures are highly abundant, with NRPS and NRPS-like genes comprising the top 25 ontology categories of HGT genes in B. meristosporus (Table S5). These results are consistent with the life history of Basidiobolus, where the fungus lives in close proximity with other microorganisms associated with the amphibian gut environment.

The additional analyses to discover distinct genetic features for HGT candidates showed that the largest differences found are in intron number and intron length, but not in nucleotide composition or by codon usage. A significantly smaller number of introns and smaller normalized intron length in HGT candidates provide more support to the HGT hypothesis, where we expected that bacterially transferred genes would maintain a smaller number and length of introns. Intronic expansion of transferred genes into fungal species after an HGT event appears to be rapid in order to reflect the genetic makeup across the genome (Da Lage et al. 2013) and can explain the introns in some of the HGT candidates. However, up to 60% of the HGT candidates still maintain absence of introns as expected for genes of bacterial origin. Finally, the nucleotide composition of HGT candidates and fungal genes were indistinguishable. Reports show that foreign genes with similar codon usage are more likely to become fixed on the receiving genome (Medrano-Soto et al. 2004; Amorós-Moya et al. 2010; Tuller 2011). We interpreted these results to indicate that horizontally transferred genes are evolving toward a similar nucleotide composition of the fungal genome based on the 5-mer/codon usage assay, but still maintain high protein similarity to and group with donor lineage copy in phylogenetic reconstructions.

The taxonomic survey of our HGT analysis shows that a diverse array of bacteria may have consistently contributed genetic information into the Basidiobolus genomes (Figure 6 and Table S4).
The most abundant bacterial taxonomic groups associated with HGT are the Proteobacteria, Firmicutes, Actinobacteria/high GC gram positive bacteria, and Bacteroidetes (Figure 6, Table S4). The proportion of HGT for each taxonomic group appears to be consistent among the three Basidiobolus genomes, and there are consistencies between the most common taxonomic groups responsible for HGT in Basidiobolus and the reported composition of bacteria associated with the gut microbiome in amphibians (Bletz et al. 2016; Kohl et al. 2013) and reptiles (Colston and Jackson 2016; Costello et al. 2010).

CONCLUSIONS

Our results confirm that the majority of zygomycete fungi classified in Mucoromycota and Zoopagomyxota do not possess a large genomic potential for secondary metabolism. Significant departures from this pattern exist, however, as exemplified by Basidiobolus, a genus with a complex genomic evolution and potential for considerable and diverse secondary metabolite production. First, it possesses larger than average genome with less than 8% content of repetitive regions, but a diverse secondary metabolite production. First, it possesses larger than average genome with less than 8% content of repetitive regions, but a genetic plasticity to integrate and express extrinsic DNA. Second, the incorporation of extrinsic DNA is consistent with selection for increased SM production, especially gene models that are related to the acquisition of these SM core genes via HGT from co-inhabiting bacterial species. More information is needed to further test these hypotheses, including sequencing of additional Basidiobolus species with long read technologies; more accurate characterization of amphibian microbiomes that test positive for Basidiobolus; and Liquid Chromatography – Tandem Mass Spectrometry (LC – MSMS) characterization of the Basidiobolus metabolome.

ACKNOWLEDGMENTS

We would like to thank Dr. Vincent Bruno and Dr. Ashraf S. Ibrahim for kindly sharing the files of the genome annotation, assembly and gene predictions of B. meristosporus B9252 and B. heterosporus B8920; Dr. Marc Cubeta and Dr. Gregory Bonito for discussions and recommendations for additional experiments regarding siderophore assays and surfactin-like genes; Dr. Gregory Bonito for allowing us to use the unpublished genome sequences of Mortierella humilis PMI_1414, Mortierella sp. GBAus27b; Dr. Olafur S. Andrésen for allowing us to use the unpublished genome sequences of Lobaria palmonaria; Dr. Paul Dryer for allowing us to use the unpublished genome sequence of Xanthoria parieta 46-1-SA22; Dr. Francis Martin and Dr. Pierre Gladeix for allowing us to use the unpublished genome sequence of Achatostromium strumarium CBS333.67; Dr. Francis Martin for allowing us to use the unpublished genome sequence of Boletus edulis V1 and Amylostereum chailletii DWAg2; Dr. Marie-Noelle Rosso for allowing us to use the unpublished genome sequences of Abortiporus biennis CIRM-BRFM1778; Dr. Francis Martin and Dr. Gregory Bonito for allowing us to use the unpublished genome sequence of Atractiella rhizopha v2.0; the 1000 Fungal Genomes Project for the use of the dikarya unpublished genomes; and Sabrina Heitmann, Dr. Ed Barge, Dr. Devin Leopold and Dr. Posy Busby for the Cladosporium sp. from the herbarum species complex PE-07 sample. This work was supported by the National Science Foundation (DEB-1441604 to JWS; DEB-1557110 and DEB-1441715 to JES). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

LITERATURE CITED

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang et al., 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Amorós-Moya, D., S. Bedhomme, M. Hermann, and I. G. Bravo, 2010 Evolution in regulatory regions rapidly compensates the cost of nonoptimal codon usage. Mol. Biol. Evol. 27: 2141–2151. https://doi.org/10.1093/molbev/msq103

Andrews, S., P. T. Pyl, and W. Huber, 2015 HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, M. Y., C. M. Santelli, and O. W. Duckworth, 2016 Layer plate CAS assay for the quantitation of siderophore production and determination of exudation patterns for fungi. J. Microbiol. Methods 121: 41–43. https://doi.org/10.1016/j.mimet.2015.12.012

Ahrendt, S. R., C. A. Quandt, D. Giobanu, A. Clun, A. Salamov et al., 2018 Leveraging single-cell genomics to expand the fungal tree of life. Nat. Microbiol. 3: 1417–1428. https://doi.org/10.1038/s41564-018-0261-0

Armaleo, D., O. Müller, F. Lutzon, O. S. Andrésson, G. Blanc et al., 2019 The lichen symbiosis re-viewed through the genomes of Cladonia grayi and its algal partner Achorosclerosis glomerata. BMC Genomics 20: 605. https://doi.org/10.1186/s12864-019-5629-x

Baldeweg, F., P. Warncke, D. Fischer, and M. Gressler, 2019 Fungal Biosurfactants from Mortierella alpina. Org. Lett. 21: 1444–1448. https://doi.org/10.1021/acs.orglett.9b00193

Bletz, M. C., D. J. Goeldbroed, E. Sanchez, T. Reinhardt, C. C. Tebbe et al., 2016 Amphibian gut microbiota shifts differentially in community structure but converges on habitat-specific predicted functions. Nat. Commum. 7: 13699. https://doi.org/10.1038/ncomms13699

Blin, K., T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen et al., 2017 antSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. 45: W36–W41. https://doi.org/10.1093/nar/gkw319

Brakhage, A. A., 2013 Regulation of fungal secondary metabolism. Nat. Rev. Microbiol. 11: 21–32. https://doi.org/10.1038/nrmmicro2916

Burmeister, A., M. Richter, K. Schultze, K. Voelz, D. Schachtschabel et al., 2007 Cleavage of β-carotene as the first step in sexual hormone synthesis in zygomycetes is mediated by a trisporic acid regulated β-carotene oxygenase. Fungal Genet. Biol. 44: 1096–1108. https://doi.org/10.1016/j.fgb.2007.07.008

Burmeister, A., E. Shelest, G. Glöckner, C. Heddergott, S. Schindler et al., 2011 Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. Genome Biol. 12: R7. https://doi.org/10.1038/gb-2011-12-1-r7

Bushley, K. E., and B. G. Turgeon, 2010 Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evol. Biol. 10: 26. https://doi.org/10.1186/1471-2148-10-26

Bushley, K. E., R. Raja, P. Jaiswal, J. S. Cumbie, M. Nonogaki et al., 2013 The genome of Tolypocladium inflatum: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. PLoS Genet. 9: e1003496. https://doi.org/10.1371/journal.pgen.1003496

Chang, Y., S. Wang, S. Sekimoto, A. L. Aerts, C. Choi et al., 2015 Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. Genome Biol. Evol. 7: 1590–1601. https://doi.org/10.1093/gbe/evv090

Chang, Y., A. Desirò, H. Na, L. Sandor, A. Lipzen et al., 2019 Phylogenomics of Endogonaceae and evolution of mycorrhizas within Mucoromycota. New Phytol. 222: 511–525. https://doi.org/10.1111/nph.15613
Luo, F., Y. He, J. Wei, C. Zhao, X. Zhou et al., 2020 Basidiosins A and B: Cyclopentapeptides from the entomophthoralean fungus Basidiobolus meristosporus. Fitoterapia 146: 104671. https://doi.org/10.1016/j.fitote.2020.104671

Ma, L. J., A. S. Ibrahim, C. Skory, M. G. Grubberr, G. Burger et al., 2019 Comparative genome analysis of Trichophyton rubrum and related dermatophytes reveals candidate genes involved in infection. MBio 3: e00259–12. https://doi.org/10.1128/mBio.00259–12

Mardon-Soto, A., G. Moreno-Hagelsieb, P. Vinuesa, J. A. Christen, and J. Collado-Vides, 2004 Successful lateral transfer requires codon usage compatibility between foreign genes and recipient genomes. Mol. Biol. Evol. 21: 1884–1894. https://doi.org/10.1093/molbev/msh202

Murphy, C., N. Youssef, R. A. Hanafy, M. B. Couger, J. E. Stajich et al., 2019 Horizontal gene transfer as an indispensible driver for Neocallimastigomycota evolution into a distinct gut-dwelling fungal lineage. Appl. Environ. Microbiol. 85: e00988–19. https://doi.org/10.1128/AEM.00988–19

Nagy, G., A. Farkas, Á. Csernekits, O. Bencsik, A. Szekeres et al., 2014 Transcription of the three HMG-CoA reductase genes of Mucor circinelloides. BMC Microbiol. 14: 93. https://doi.org/10.1186/s12864-014-1894-3

Nagy, L. G., R. Riley, A. Tritt, C. Adam, C. Daum et al., 2016 Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. Mol. Biol. Evol. 33: 959–970. https://doi.org/10.1093/molbev/msw237

Nickerson, M. A., and J. A. Hutchison, 1971 The distribution of the fungus Basidiobolus ranarium Eidam in fish, amphibians and reptiles. Am. Midl. Nat. 86: 500–502. https://doi.org/10.2307/2423642

Nguyen, T. A., O. H. Cissé, J. Y. Wong, P. Zheng, D. Hewitt et al., 2017 Innovation and constraint leading to complex multicellularity in the Ascomycota. Nat. Commun. 8: 14444. https://doi.org/10.1038/s41478-017-06453-z

Ohm, R. A., N. Feau, B. Henriassat, C. L. Schoch, B. A. Horwitz et al., 2012 Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. PLoS Pathog. 8: e1003037. https://doi.org/10.1371/journal.ppat.1003037

Okada, K., S. Amano, Y. Kawamura, and Y. Kagawa, 2015 Gastrointestinal basidiobolomycosis in a dog. J. Vet. Med. Sci. 77: 1311–1313. https://doi.org/10.1292/jvms.15-0177

Okagaki, L. H., C. C. Nunes, J. Sailsbery, B. Clay, D. Brown et al., 2015 Genome sequences of three phytopathogenic species of the Magnaporthaceae family of fungi. G3 (Bethesda) 5: 2539–2545.

O’Leary, N. A., M. W. Wright, J. R. Brister, S. Cuño, D. Haddad et al., 2015 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44: D733–D745. https://doi.org/10.1093/nar/gkv1189

Osbour, A., 2010 Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiol. 154: 531–535. https://doi.org/10.1104/pp.110.161315

Pagès, H., P. Aboyoun, R. Gentleman and R. DeB Roy, 2019 Biostrings: Efficient manipulation of biological strings. R package version 2.52.0.

Padamsee, M., T. A. Kumar, R. Riley, M. Binder, A. Boyd et al., 2012 The genome of the xenotolerant mold Wallomia sebi reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. Fungal Genet. Biol. 49: 217–226. https://doi.org/10.1016/j.fgb.2012.01.007

Pendleton, A. L., K. E. Smith, N. Feau, F. M. Martin, J. V. Grigoriev et al., 2014 Duplications and losses in gene families of rust pathogens highlight putative effectors. Front. Plant Sci. 5: 299. https://doi.org/10.3389/fpls.2014.00299

Perlin, M. H., J. Anselem, E. Fontanillas, S. San Toh, Z. Chen et al., 2015 Sex and parasites: genomic and transcriptomic analysis of Microbotryum lycnidis-dioicae, the biotrophic and plant-castrating anther smut fungus. BMC Genomics 16: 461. https://doi.org/10.1186/s12864-015-1660-8

Peypoux, F., J. M. Bonmatin, and J. Wallach, 1999 Recent trends in the biochemistry of surfactin. Appl. Microbiol. Biotechnol. 51: 553–563. https://doi.org/10.1007/s002530051432

Quandt, C. A., D. Beaudet, D. Corsaro, J. Walochnik, R. Michel et al., 2017 The genome of an intracelluar parasite, Paramicrosporidium saccamoeae, reveals alternative adaptations to obligate intracelllar parasitism. eLife 6: e29594. https://doi.org/10.7554/eLife.29594

Quin, M. B., C. M. Flynn, and C. Schmidt-Dannert, 2014 Traversing the fungal terpenome. Nat. Prod. Rep. 31: 1449–1473. https://doi.org/10.1039/C3NP0075G

Reynolds, H. T., V. Vijayakumar, E. Gluck Thaler, H. B. Korotkin, P. B. Matheny et al., 2018 Horizontal gene cluster transfer increased hallucinogenic mushroom diversity. Evol. Lett. 2: 88–101. https://doi.org/10.1002/evl3.42

Robinson, M. D., D. J. McCarthy, and G. K. Smyth, 2010 edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rokas, A., J. H. Wisecaver, and A. L. Lind, 2018 The birth, evolution and death of metabolic gene clusters in fungi. Nat. Rev. Microbiol. 16: 731–744. https://doi.org/10.1038/s41579-018-0075-3

Rokas, A., M. E. Mead, J. L. Steenwyk, H. A. Raja, and N. H. Oberlies, 2020 Biosynthetic gene clusters and the evolution of fungal chemo-diversity. Nat. Prod. Rep. 17: 167–178.

Riley, R., A. A. Salamov, D. W. Brown, L. G. Nagy, D. Floudas et al., 2014 Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. USA 111: 9923–9928. https://doi.org/10.1073/pnas.140092111

Russ, C., B. F. Lang, Z. Chen, S. Guija, T. Shea et al., 2016 Genome sequence of Spizellomyces punctatus. Genome Announc. 4: e00849–16. https://doi.org/10.1128/genomeA.00849-16

Schultz, J., R. R. Copley, T. Döerks, C. P. Ponting, and P. Bork, 2000 SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 28: 231–234. https://doi.org/10.1093/nar/28.1.231

Schwarzt, V. U., S. Winter, E. Shelest, M. Maréc-Houben, F. Horn et al., 2014 Gene expansion shapes genome architecture in the human pathogen Litchthemia corymbifera: an evolutionary genomics analysis in the ancient terestrial mucorales (Mucoromycotina). PLoS Genet. 10: e1004496. https://doi.org/10.1371/journal.pgen.1004496

Sirová, D., J. Bártá, K. Šimek, T. Posch, J. Pech et al., 2018 Hunters or farmers? Microbiome characteristics help elucidate the diet composition in an aquatic carnivorous plant. Microbiome 6: 225. https://doi.org/10.1186/s40168-018-0600-7
