Integrative activity of mating loci, environmentally responsive genes, and secondary metabolism pathways during *Chaetomium globosum* sexual development

Zheng Wang[1], Francesc López-Giráldez[2], Junrui Wang[1], Frances Trail[5], Jeffrey P. Townsend[1,4,5]

[1] Department of Biostatistics, Yale University, New Haven, CT 06510, USA

[2] Yale Center for Genome Analysis (YCGA) and Department of Genetics, Yale University, New Haven, CT 06511, USA

[3] Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA

[4] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA

[5] Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06511, USA

Running title: *Chaetomium* sexual development

Abstract

The origins and maintenance of the rich morphological and ecological diversity of fungi has been a longstanding question in evolutionary biology. To investigate how differences in expression regulation contribute to divergences in development and ecology among closely related species, comparative transcriptomics was applied to *Chaetomium globosum* and previously studied model species of *Neurospora* and *Fusarium*, which represent diversity from saprotrophic to pathogenetic biology, from post-fire terrestrial to highly humid ecology, and from heterothallic, pseudo-homothallic to homothallic lifestyles. Gene expression was quantified in perithecia at nine distinct morphological stages during nearly synchronous sexual development. Unlike *N. crassa*, expression of all mating loci in *C. globosum* was highly correlated. Key regulators of the initiation of sexual development in response to light stimuli—including orthologs of *N. crassa sub-1*, *sub-1*-dependent gene NCU00309, and *asl-1*—showed regulatory dynamics matching between *C. globosum* and *N. crassa*. Among 24 secondary-metabolism gene clusters in *C. globosum*, 11—including the cochliodones biosynthesis cluster—exhibited highly coordinated expression across the perithecial development. *C. globosum* exhibited coordinately up-regulated expression of histidine kinases in hyperosmotic response pathways—consistent with ecological adaptation to high humidity as previously demonstrated in pathogenic *F. graminearum*. Bayesian networks indicated that gene interactions during sexual development have diverged in concert with the capability to reproduce asexually and homothallic *vs.* heterothallic life cycle in *N. crassa* and *C. globosum*, shifting the hierarchical roles of genes associated with conidiation and heterokaryon incompatibility. This divergence supports an evolutionary history of loss of conidiation due to unfavorable combinations of heterokaryon incompatibility in homothallic species.

Importance:

Introduction

Fungi encompass an enormous morphological diversity that is distributed across almost every kind of environment on earth (Stajich et al. 2009). Understanding the evolution of such a rich diversity has become one of the central questions in evolutionary developmental biology (Jedd 2011; Wang et al. 2018; Minelli 2016). Sordariomycetes fungi include the genetic model *Neurospora crassa*, pathogenic model *Fusarium graminearum*, and developmental model *Sordaria macrospora*, which have together enabled the identification of numerous genes critical for sexual development via comparative genomics and transcriptomics (Lehr et al. 2014; Wang et al. 2014; Trail et al. 2017; Wang et al. 2016; Teichert et al. 2014; Nygren et al. 2012). Among Sordariomycetes are numerous examples of heterothallic (self-incompatible) and homothallic (self-compatible) species. We have yet to understand how heterothallism and homothallism affect the developmental process that generates perithecia in diverse ecologies—from underwater substrates to leaves in the air, and from strict pathogens of plants, animals or fungi to opportunistic species.

The draft genome sequence of the conditional human pathogen *Chaetomium globosum*—a close relative of the model fungus *N. crassa*—was recently made available (Cuomo et al. 2015). The genome of homothallic *C. globosum* strain CBS 148.51 (ATCC 6205) comprised 11,124 predicted protein-coding genes (Cuomo et al. 2015). Under laboratory conditions, colonies of *C. globosum* can be quickly established from ascospores or hyphal elements. Production of perithecia can be quickly induced, usually forming concentrically around the point of inoculation and maturing in a gradient towards the edge of the colony. The genus *Chaetomium* (Chaetomiaceae, Sordariales) includes about 100 species identified from various substrates,

primarily occupying environments of high humidity, including aquatic niches (Kirk et al. 2011; Domsch et al. 1993).

Morphologically, *Chaetomium* perithecia develop in parallel forms to previously-studied fungi such as *Neurospora* and *Fusarium*, with a few features that have been suggested to be adaptive to their highly humid habitats. *Chaetomium* perithecia feature a characteristic membranaceous wall, covered by conspicuously flexuous long hairs (Lentz and Seth 1973). Single-celled, smooth, pigmented ascospores are released from asci inside the perithecium, then squeezed out through the ostiole and trapped by the coiled hairs (Stevenson and Ames 1963). *Chaetomium globosum*, the type species of the genus, occupies diverse substrates (aerial, terrestrial, aquatic, marine, pathogenic; Wang et al. 2016; Yamada et al. 2011). Whereas most pathogenic fungi rely on asexual reproduction of a large number of simple conidia as a major distribution strategy, asexual spores have not yet been observed in *C. globosum*, implying a key role of sexual reproduction in propagation and dispersal. The light-colored hairs have been suggested to serve as an air-trap, facilitating dispersal of perithecia and ascospores by water droplets (Dixon 1961). Perithecial hairs of *Chaetomium* species have also been attributed defensive value as physical barriers against predatory insects (Wicklow 1979) and to function as mucilage-filled sacs that enable flotation-based water transport during the maturation of the perithecia (Ellis 1981).

Biologically, *C. globosum* has received considerable scientific attention as a consequence of its pathogenesis of diverse animal hosts and humans (Miller and McMullin 2014; Naidu et al. 1991; Kim et al. 2013; Awad et al. 2014) its production of secondary metabolites, including toxic chaetoglobosins, chaetomugilins, chaetoviridins, and cochliodones (Abdel-Azeem et al. 2016; Wang et al. 2017; Li et al. 2008; Winter et al. 2012). These chemicals also have potential for

applications as the biocontrol of pests and the treatment of cancer because of their high cytotoxicity (Nielsen et al. 1999; Pieckova 2003; Miller and McMullin 2014; McMullin et al. 2013; Došen et al. 2017; Jiang et al. 2017; Luo et al. 2013). *C. globosum* has been found to infect humans (Mackenzie 1979; Hassett et al. 2009; Vesper et al. 2007; Green et al. 2014), in whom it causes symptoms of rhinitis and asthma. *C. globosum* can also infect finger- and toe-nail beds causing onychomycosis—the incidence of which is increasing worldwide over recent decades (Naidu et al. 1991; Kim et al. 2013; Shi et al. 2016). Additional interest has accrued because of its cellulolytic ability, which could facilitate critical biodegradation efforts and might enable it to be a bioenergy powerhouse (Kim et al. 2015; Wanmolee et al. 2016). Consequent attention regarding these fungal properties has mainly been devoted to asexual growth. Recently, molecular evidence suggested that the heterotrimeric Gα protein-cAMP/PKA signaling pathway leads to co-regulation of melanin biosynthesis and secondary metabolism during sexual reproduction (Hu et al. 2018; Hu et al. 2012).

Reproductively, *C. globosum* follows a classic homothallic life cycle: successful sexual development and production of ascospores does not require mating and crossing between partners of different genetic backgrounds. This homothallic life cycle has frequently evolved as a derived feature within diverse fungal lineages (Wilson et al. 2015; Nagel et al. 2018). Research on mating activity in fungi has primarily focused on mating type genes and their evolution (Heitman 2015; Lee et al. 2010; Lu et al. 2011; Wilson et al. 2015; Martin et al. 2011; Nygren et al. 2012; Klix et al. 2010). There is a lack of knowledge about how divergence in mating behavior between heterothallic and homothallic fungi differentially affects downstream sexual development. The closely-related genus *Neurospora* includes a pseudo-homothallic model *N. tetrasperma*, and two heterothallic models *N. crassa* and *N. discreta*, for which two mating

types are required for successful crossing and subsequent sexual development. Gene expression studies within these individual species have illuminated gene expression dynamics associated with pigmentation and meiotic sporulation (Wang et al. 2014; Lehr et al. 2014), perithecial development (Lehr et al. 2014; Trail et al. 2017), and responses to genetic and environmental signals (Wang et al. 2014). Several comparative studies have shed light on the evolution of gene expression and on the genetic basis of perithecium production (Lehr et al. 2014; Wang et al. 2014; Trail et al. 2017; Wang et al. 2018; Sikhakolli et al. 2012; Kim et al. 2018). Studies of several filamentous fungi (Kim et al. 2015; Pöggeler and Kück 2000; Coppin et al. 1997; Ait Benkhali et al. 2013; Rodenburg et al. 2018) have investigated the genome-wide impacts of activities of mating-type and other sex-associated loci during post-crossing sexual development—activities that are associated with the evolution of divergent development of homothallic and heterothallic lifestyles in filamentous fungi. However, these studies all focused only the early stages of sexual development with an emphasis on the crossing process.

To investigate how differences in expression regulation may contribute to dramatic divergences in development and ecology among closely related species, we conducted comparative transcriptomic sequencing of *Chaetomium globosum, Neurospora spp.* (Lehr et al. 2014; Wang et al. 2014; Wang et al. 2019), and *Fusarium spp.* (Trail et al., 2017; Sikhakoli et al 2012). We compared expression of mating-type loci between *C. globosum* and *N. crassa* from both mating types across the entirety of perithecial development. We identified orthologous genes responsive to environmental signals among *C. globosum*, *N. crassa*, and *F. graminearum*, and assessed their ancestrally retained, convergent, or divergent expression patterns and their unique habitat adaptation. We also analyzed expression during sexual development of all 41 predicted secondary metabolism pathways, evaluating the regulation of secondary metabolite

synthesis in coordination with the maturation of asci and ascospores. Unannotated genes exhibiting similar expression patterns between *C. globosum* and *N. crassa* with peak expression during asci and ascus development were selected, and their knockout phenotypes were characterized in *N. crassa*.

Materials and Methods

Induction of synchronous perithecial development—To compare sexual development of different species fungi under a single growth condition, the genome-sequenced strain of *Chaetomium globosum* (CBS 148.51) was cultured on Carrot Agar (CA; Klittich and Leslie 1988), enabling comparison to related studies conducted on *Neurospora* and *Fusarium* species (Klittich and Leslie 1988; Wang et al. 2014; Lehr et al. 2014; Sikhakolli et al. 2012). In contrast with species in *Neurospora* and *Fusarium*, *C. globosum* produces no conidia. Culture of *C. globosum* on a medium with a low concentration of simple sugars—as in CA medium—represses germination of *C. globosum* ascospores, which otherwise can be used to induce a large amount of synchronic growth. To culture enough tissue exhibiting synchronized perithecial development, hyphae of *C. globosum* were inoculated in 200 ml liquid CA in a 500 ml flasks. The flask cultures were incubated at 27 C on a 100 rpm shaker under constant light. 10-day liquid cultures were filtered with a sterilized single-layer miracloth (Calbiochem™), and abundant hyphal elements were harvested from the filtrate (**Fig. S1**). Two ml of the filtrate were plated out on a cellophane membrane covering solid CA in petri dish (9 cm in diameter), then incubated at 27 C under constant white light in a refrigerated incubator (VWR Signature™ Diurnal Growth Chamber). Light-colored protoperithecia became visible 5 days after inoculation, and perithecia with melanized ascospores were fully developed 8 days after appearance of protoperithecia. Plates

with apparent protoperithecia were gently disturbed with a glass microbiological spreader to mimic crossing protocols applied to other heterothallic systems in previous studies; such a disturbance is known to be critical for setting a synchronous start-time for perithecial development in *F. graminearum* (Trail et al. 2017). Fungal tissues were collected by removal of cellophane membranes at the protoperithecial stage (0 time point, at disturbance), as well as at 2, 24, 48, 72, 96, 120, 144, and 168 h after disturbance. Tissue samples were flash frozen in liquid nitrogen and stored at -80 C. Tissues collected from a single plate were used as one biological replicate. At least three biological replicates were prepared for every sampled timepoints.

RNA isolation and transcriptome profiling—Total RNA was extracted from homogenized tissue with TRI REAGENT (Molecular Research Center) as in Clark et al. (2008). Messenger RNA was purified using Dynabeads oligo(dT) magnetic separation (Invitrogen). Preparation of cDNA for sequencing followed the Illumina mRNA Sequencing Sample Preparation Guide. Complementary DNA was primed for reverse transcription using N_6 primers. The quality of cDNA samples was verified with a bioanalyzer (Agilent Technologies); then 22 qualified samples—including 5 purely technical replicates—were sequenced at the Yale Center for Genomics Analysis (YCGA).

Data acquisition and analysis—17 sequencing libraries were produced from purified total RNA samples using the Illumina TruSeq stranded protocol. The libraries underwent 76-bp single-end sequencing using Illumina HiSeq 2500 according to Illumina protocols, generating an average of 22 million single-end reads per library. Adapter sequences, empty reads, and low-quality

sequences were removed. For each read, we trimmed the first 6 nucleotides, and trimmed the last nucleotides at the point where the Phred score of an examined base fell below 20 using in-house scripts. If, after trimming, the read was shorter than 45 bp, the entire read was discarded. Trimmed reads were aligned to the *C. globosum* genome from the NCBI with its genome annotation using `Tophat` v.2.1.1 (Trapnell et al. 2009), applying the very-sensitive preset and providing the corresponding gene model annotation. Only the reads that mapped to a single unique location within the genome, with a maximum of two mismatches in the anchor region of the spliced alignment, were reported in these results. We used the default settings for all other `Tophat` options. We tallied reads by aligning to exons of genes with the program `HTSeq` v0.6.1p1. A tally of the number of the reads that overlapped the exons of a gene, was calculated using aligned reads and the gene structure annotation file for the reference genome. In addition to the 17 libraries representing biological replicates, five technical replicates were sequenced to ensure quality control among serial sequencing runs. Data from the five technical replicates exhibited very high consistency. All sequence data were deposited to the Gene Expression Omnibus database (**GSE131190**). Statistical analysis of gene expression levels based on the tallies of reads for each gene for the 17 libraries representing biological replicates was conducted with `LOX` v1.6 (Zhang et al. 2010). Raw reads that mapped ambiguously or to multiple loci were excluded from `LOX` input. Previously generated data on gene expression during sexual development in *N. crassa* (GSE41484; Wang et al. 2014) and in *F. graminearum* (GSE61865; Sikhakolli et al. 2012) were compared to the new *C. globosum* gene expression data.

Gene orthology assessment and ancestral gene expression estimation—To identify orthologs for comparison of gene expression, protein and nucleotide sequences for *N. crassa* and *C. globosum*

genomes were downloaded from the JGI genome database (Grigoriev et al. 2014). Predicted protein sequences were used to identify single copy orthologs (cluster) with `ReMark` (Kim et al. 2011), specifying the BLOSUM62 amino-acid transition matrix and an inflation factor of 1.6. The ortholog set was compared to those reported in the latest (8th) InParanoid database (Sonnhammer and Östlund 2015) for missing or misidentified clusters. Any contrasting results as well as genes belonging to large gene families were further verified by manually conducting phylogenetic analyses on sequences obtained from exhaustive reciprocal `BLAST` searches. Orthologs of interest were also identified in *N. tetrasperma* and *F. graminearum* with the same approach. In this research, we did not further consider the orthologous genes that were multi-copy in either *N. crassa* (OR74A) or *C. globosum* (CBS 148.51). For phylogenetic analyses, amino acid sequences were aligned using `SATé-II` (Liu et al. 2012) specifying `MAFFT` as the aligner, `MUSCLE` as the merger, and `RaxML` as the tree estimator under WAG model. Alignment scores were estimated using 10 iterations of `SATé-II-ML` for three independent runs, and the alignment with the best score was retained for further analysis with `MrBayes` 3.2 (Ronquist et al. 2012). Bayesian analyses were performed with the Metropolis-coupled Markov-chain Monte Carlo method, and the prior for the amino-acid model was set to be mixed. Trees were sampled every 1000th generation over four chains for 2,000,000 generations. 1000 trees obtained prior to convergence were discarded before computing a 50% majority-rule consensus of the remaining trees. Clades with Posterior Probability (PP) higher than 0.95 were deemed significantly supported. Gene expression patterns that were ancestrally retained, convergent, or divergent were classified with ancestral expression reconstructed with taxa to be compared with orthologous genes as in (Trail et al. 2017). Briefly, the fold-change between stages and the molecular evolutionary tree of selected species were supplied as input files to the Continuous Ancestral

Character Estimation (Schluter et al. 1997; CACE, Paradis et al. 2004) tool in the Discovery Environment Application list in CyVerse (Matasci et al. 2002; Goff et al. 2011; Paradis 2012; Chougule et al. 2018), which provided ancestral changes in expression across adjacent stages at all internal nodes for every ortholog set.

Bayesian network reconstruction—Biological networks were modeled using the Bayesian Network Web Server (Ziebarth et al. 2013) supplied with perithecial development expression data for *N. crassa* and *C. globosum* separately. To scale changes between serial sample points appropriately for Bayesian Network inference, they were quantified as

$$\begin{cases} \left((x_{t+1} - x_t)/min[x_t, x_{t+1}]\right)/2, & x_{t+1} - x_t < 2 \\ \log_2\left((x_{t+1} - x_t)/min[x_t, x_{t+1}]\right), & x_{t+1} - x_t \geq 2, \end{cases}$$

where x_t is the relative expression level as quantified by LOX (Zhang et al. 2010) at time point (stage) t, $t \in 1 .. 8$ for *N. crassa* and $t \in 1 .. 9$ for *C. globosum*. Global structure learning settings were retained at default settings. Each network depicted is the 50% majority consensus of the 100 highest-scoring models that retain edges exceeding a selection threshold of 0.5, performed without imposing any structural constraints.

Functional enrichment and secondary metabolic gene clusters—Functional annotation of statistically significantly differentially expressed genes in metabolic pathways was gathered via the biochemical pathway and annotation data from the Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa and Goto 2000). GO Enrichment analysis was performed with Panther provided by Gene Ontology Resource (Mi et al. 2017). Functional annotation was further checked for genes of interest using the FungiDB database (Stajich et al. 2012). Forty-one

secondary-metabolic gene clusters indicated within the *C. globosum* genome by entries in the JGI database were further confirmed using `AntiSMASH` (Blin et al. 2019); gene clusters with more than three genes were further analyzed for their expression patterns. Genome-wide expression patterns were clustered with a hierarchical algorithm of euclidean distance using `Morpheus` (https://software.broadinstitute.org/morpheus).

Nucleic acid manipulation and genetic transformation—Potential functions of *C. globosum* genes during sexual development were identified by examining phenotypes of strains with the orthologous gene knocked out in *N. crassa*. Knockout strains were constructed for more than 9600 genes via the *Neurospora crassa* knockout project (Dunlap et al. 2007; Colot et al. 2006), obtained from the Fungal Genetic Stock Center (FGSC; McCluskey et al. 2010). Deletion of target genes in the knockout mutants was verified by PCR (Fu et al. 2011; Chinnici et al. 2014; Wang et al. 2014). Perithecium formation was examined by stereomicroscopic analysis of shape, size, number formed, and stereomicroscopic examination of beak developmental morphology. Additionally, squash mounts of perithecia in water were examined using a compound microscope to ascertain the presence and maturity of asci, ascospore shape, number per ascus, and presence of paraphyses. Each knockout strain that exhibited a significant morphological phenotype was first verified as a knockout by PCR, then crossed with the wild type strain. Cosegregation of the observed phenotype with deletion of the gene in the offspring was verified to ensure that the intended deletion was responsible for the mutant phenotype (Fu et al. 2011; Chinnici et al. 2014; Wang et al. 2014) .

Results

A total of 5784 single-copy orthologs were identified by comparison of *C. globosum* CBS 148.51 and *N. crassa mat **A*** OR74A genomes (**Table S1**). The ortholog of mating type gene *mta-1* (coding MAT1-2-1) in *C. globosum* was identified as corresponding to the *mat **a*** gene from heterothallic *N. crassa* trp-3 and pseudohomothallic *N. tetrasperma* P0656.

Morphology and development—Our 0-hour sample was timed such that the majority (>50%) of expected protoperithecia were well-developed and had brown-colored protrusion with prolonged stiff hairs. Our 2-hour sample was selected to characterize gene expression of those protoperithcia after the disturbance that was used to facilitate greater synchronization of development. At 24, 48, 72, 96, 120, 144 and 168 h after disturbance, the majority of sexual reproductive structures had reached respective developmental stages characterized by young perithecia (abundant hairs), young perithecia (dark colored), double-sized perithecia with thin-walled ascogenous cells, asci containing condensed spore content, asci with light-colored young ascospores, mature asci with dark ascospores, and released ascospores. Five days after hyphal elements from liquid cultures were plated on carrot medium, *C. globosum* developed abundant pale protoperithecia with a few hairs. Dark perithecia were covered with brownish hairs, which became curly after 72 h of perithecial development. Differentiation of ascogenous tissue manifested at 72 h, and young asci were present at 96 h. Asci with 8 ascospores were evident in 120-hour perithecia. Nearly 70% of perithecia matured and released ascospores within 8 days after appearance of protoperithecia (**Figs. 1, S1**). Perithecial development between *C. globosum* and species of *Neurospora* are highly similar in terms of timing in appearance of major morphological characters, including size expansion of perithecia, ascogenous tissues, asci, and ascospore maturation (Wang et al. 2014; Trail et al. 2017). The transcriptome was assayed at the additional stage of ascospore release (168 h) in *C. globosum*.

Gene expression across perithecial development—LOX yielded well-measured relative expression levels across more than one sampled stage for 11170 of the 11232 genes in the *C. globosum* genome (**Table S2, S3**). Sixty-two genes without orthologs in *N. crassa* exhibited no measurable expression across *C. globosum* sexual reproduction. 2023 genes showed no detectable expression in at least one out of the nine sampled stages; among these genes only 231 have identified orthologs in *N. crassa*, with predicted functions ranging from metabolic processes to meiotic regulation. Genome-wide gene expression of *C. globosum* across sexual development generally followed two frequent patterns (**Fig. 2**). One frequent pattern starts with up-regulation for the two hours after disturbance, then a general down-regulation toward the last stage of spore release. The other frequent pattern starts with a down-regulation, followed by up-regulation towards the end of the perithecial development. Within each of these frequent patterns, gene expression can be further clustered into four sub-patterns (**Fig. 2**, panels **B–I**; **Table S3**). Few single-copy orthologs were detected that belong to group B. *C. globosum* gene CHGG_02344 (**Fig. 2-B2**) exhibited nearly identical gene expression dynamics to its *N. crassa* ortholog NCU05882 .

Many single-copy orthologs between *N. crassa* and *C. globosum* were detected in groups C–I. In general, genes that are critical for development in *N. crassa* showed similar expression regulation between *C. globosum* and *N. crassa*. Examples include two late light-responsive genes (NCU04510 and 08677) and their orthologs (CHGG_04293 and 05004; **Fig. 2-C2**), a check-point kinase and a 3'-5' exonuclease in replication and recombination (NCU01940 and CHGG_06604; NCU02961 and CHGG_05895; **Fig. 2-E2**), as well as basal hyphal growth and asexual/sexual development regulator *adv-1* (NCU07392 and CHGG_00993; **Fig. 2-G2**).

Interestingly, numerous genes with direct roles in metabolism exhibited divergent expression between *C. globosum* and *N. crassa*. Examples include a likely secondary-metabolism-related reductase and a dehydrogenase (NCU01904 and CHGG_06640; NCU01905 and CHGG_06639; **Fig. 2-D2**), chitinase (NCU04603 and CHGG_00523; **Fig. 2-F2**), as well as a kinase-activating protein-coding gene (NCU01242 and CHGG_02675; **Fig 2-H2**). *C. globosum* genes CHGG_06511 and 06512 (**Fig. 2-I2**) exhibited similar up-regulation at the end of perithecial development to the upregulation observed for their homolog *acu-9* (NCU10007, malate synthase) in *N. crassa*.

Regulation of dark pigment production—Many species in Sordariomycetes produce dark perithecia and ascospores. Dark pigmentation formed by melanin synthesis is a phenotypic marker of sexual development in these fungi. Orthologs of 4 genes encoding melanin synthesis enzymes in *N. crassa*, including polyketide synthase (*per-1*), scytalone dehydratase (*scy-1*), and tetrahydroxynaphthalene reductase 1 and 2 (*tnr-1*, and *tnr-2*), were identified in *C. globosum* and *N. tetrasperma*. Expression of these four melanin synthesis genes across perithecial development were found to be highly coordinately regulated within species (previously reported in N. crassa, Wang et al. 2014). However, in contrast to the two-phase up-regulation in *Neurospora* species (in response to crossing—the 2-hour sampling point—and later during ascospores development), expression of these genes exhibited a monotonic up-regulation in *C. globosum* (**Fig. 3**). The lack of effect of the 0-hour disturbance on expression of melanin synthesis in *C. globosum* contrasted with the significant disruption of expression observed in many other metabolic pathways.

Expression of mating loci and pheromone genes—Orthologs of genes coding for all four mating-type proteins in *N. crassa*, including Mat A-1, Mat A-2 and Mat A-3 for the **A** strain and Mat a-1 for the **a** strain, coexist in the homothallic genome of *C. globosum*. Orthologs of *N. crassa* genes that code for **A**-specific pheromone precursor *ccg-4*, and for two pheromone receptors *pre-1* (responsive to *ccg-4*) and *pre-2* (responsive to *mfa-1*), have been annotated in *C. globosum* genes. *N. crassa* **a**-specific pheromone precursor *mfa-1* encodes a very short protein; no ortholog of *mfa-1* or similar precursor protein was identified or annotated in *C. globosum* (**Fig. 4**). In *C. globosum*, All four mating-type loci were highly coordinately regulated, especially during the late stages of sexual development (> 48 h; **Fig. 4A**). Expression among the three mating type **A** genes *mtA-1*, *mtA-2* and *mtA-3* and between them and the mating type **a** gene *mta-1* were inconsistent during sexual development in *N. crassa* (**Fig. 4B**). Expression of *mat a-1* was not detectable in **A** protoperithecia, and was almost undetectable before 48 h, but after 48 h its expression increased monotonically, peaking at the last sampling stage of ascospore maturation in *N. crassa* at 107-fold above 0 h. In *C. globosum*, expression of *ccg-4* and *pre-1* exhibited similar up-regulation toward spore maturation and release (**Fig. 4C**). In *N. crassa*, *ccg-4* experienced a monotonic, dramatic increase of expression culminating in 45-fold upregulation above 0-hour expression by the end of perithecial development. However, expression levels of *ccg-4* do not necessarily reflect the pheromone level within the fungal culture: the pheromone genes it upregulates encode pre-pro-pheromone that require further post-transcriptional processing to yield mature pheromones.

Regulators of sexual development in response to environmental signals—As has been demonstrated in several fungal model species, light serves as a key environmental signal of

fungal development (Wang et al. 2017; Wang et al. 2016; Chen et al. 2009). Three light-responsive genes—submerged protoperithecia (*sub-1*), *sub-1*-responsive gene NCU00309 (Sancar et al. 2015), and ascospore lethal (*asl-1*), are required for normal sexual development in *N. crassa*. Interestingly, their orthologs are present in *C. globosum*. Expression of *sub-1* and NCU00309 orthologs in *C. globosum* exhibited similar expression regulation as in *N. crassa*, only slightly different in scale (**Fig. 5**). Gene *asl-1* is characterized by an undulating expression across sexual development in *N. crassa*. Its ortholog in *C. globosum* exhibits a similar dynamic, but appears phase-shifted 24 h earlier: it peaked in expression at 2 h, whereas in *N. crassa*, *asl-1* expression dropped to its lowest point at the 2-hour crossing time-point. Crossing in *N. crassa* is controlled by mating type loci, the functions of which are not known in homothallic *C. globosum*. Aside from developmental guidance obtained from environmental light signals (Wang et al. 2016; Wang et al. 2017; Rodriguez-Romero et al. 2010; Chen et al. 2009), sexual reproduction in many fungal species is regulated by numerous other environmental stress factors (Rodriguez-Romero et al. 2010). Catalase genes are found in aerobically respiring organisms, and function to protect cells from the toxic effects of hydrogen peroxide; three *N. crassa* catalase genes—*cat-1*, *cat-2*, and *cat-3*—are known to be highly expressed during asexual reproduction. Orthologs of these genes exhibited down-regulated expression during perithecial development in *C. globosum*, in contrast to the coordinate up-regulation (especially as perithecia and ascospores mature) observed in *N. crassa* (**Fig. S2**), for which air flow is a critical component of the successful dispersal of forcibly released ascospores.

Expression regulation for genes critical for asexual reproduction and heterokaryon incompatibility—Asexual reproduction has never been observed for *C. globosum* (Wang et al.

2016). Nevertheless, orthologs of some *N. crassa* conidiation genes were identified in the *C. globosum* genome, including *aconidiate-2* and *-3* (*acon-2* and *acon-3*), *conidiation-3* and *-10* (*con-3* and *con-10*), *non-repressor of conidiation-1* and *-2* (*nrc-1, nrc-2*), and *conidia separation-1* (*csp-1*). In *N. crassa*, genes *acon-2* and *acon-3* are required for macroconidiation (Bailey-Shrode and Ebbole 2004). In *N. crassa,* their expression was up-regulated during ascospore formation (48–96 hours post-crossing); in *C. globosum*, expression of their homologs was upregulated across sexual development as well as during ascospore maturation and release. Conidiation gene *con-10* is known to exhibit a dramatic expression response to light in conidiating tissues (Lee and Ebbole 1998). However, proliferative expression of *con-10* has also been reported during perithecial development of *N. crassa* (Liu et al. 2017; Wang et al. 2014; Chen et al. 2009). Gene *csp-1* is a global circadian repressor that regulates membrane formation via ergosterol synthesis (Sancar et al. 2011), while *nrc-2* is a serine/threonine kinase that is required to repress conidiation (Kothe and Free 1998). In *N. crassa*, *csp-1* expression was up-regulated during sexual development and *nrc-2* expression was consistently high. Expression of their orthologs in *C. globosum* contrasted: *con-10* and *csp-1* were both down-regulated (**Fig. 6**). No orthologs of *N. crassa con-6*, *con-8, con-13* or *fluffy* (*fl*) were identified in the *C. globosum* genome. The genome of *N. crassa* hosts many heterokaryon incompatibility (*het*) genes annotated based on the characteristic HET domain, which is fairly common in ascomycetes genomes (Espagne et al. 2002). Evolutionary analyses of the well studied *het* genes in *N. crassa*—including two copies of *tol*, two copies of *het-C*, *het-6*, *-13*, *-14* and *-15*—distinguish three gene families, HET-6/13/15, HET-C, and HET14-TOL, composed of22, 4, and 21 genes with clear homology between *N. crassa* and *C. globosum* (**Fig. S3A**). An ortholog of *N. crassa het-13* and two copies of *het-C* were identified as CHGG_03992, _00869, and 03493 in

C. globosum, and single-copy homologs of *N. crassa tol* and *het-6* were identified as CHGG_01991 and _08450 (**Fig. S3B, C**). Along with more than half of the genome being up-regulated (4983 / 9717 genes), expression of 15 (out of 24, $P = 0.31$, Fisher Exact Test) predicted *het* genes exhibited coordinated up-regulation at 24 h in *N. crassa*, during which nuclei from opposite mating types fused with the phase shift from dikaryotic to diploid, although the overall increase was not significant ($P = 0.053$, one sample *t*-test *vs.* 0) among these *het* genes. A second coordinated but insignificant ($P = 0.343$, one sample *t*-test *vs.* 0) up-regulation at 48 h was observed for 16 predicted *het* genes, which is significant ($P = 0.0021$, Fisher Exact Test) in contrast to about a third of the genome (3425/9717) that was up-regulated, during meiotic sporulation, in *N. crassa* (**Fig. S3D**). Such coordination was not observed among most predicted *het* genes in *C. globosum* (**Fig. S3E**). For the selected *N. crassa het* genes with orthologs or single-copy homologs identified in *C. globosum*, a peak of expression during ascospore maturation at 120 h was observed for *het-C*, *het-6* and *het-13* orthologs in *C. globosum* (**Fig. S3F**).

Bayesian networks for associations among conidiation, heterokaryon incompatibility, and sexual development in N. crassa and C. globosum—Bayesian networks (Ziebarth et al. 2013) based on time series expression data convey co-regulatory posterior probability on each edge and the degree at each vertex representing the associations among a set of genes. Note that in BN statistical model output, the direction of the edge in a BN network is not necessarily the regulatory direction. However, evidence of the centrality of gene function is strengthened with dense direct network connections, presumably because multiple genes are co-regulated for a specific developmental purpose. Our previous studies demonstrated that genes inferred to play a

role via BN analysis do play central roles in the network model, tend to be placed in the top-tier modules, and have multiple edge connections (Wang et al. 2014; Wang et al. 2019)—consistent with a conception that gene hubs in the network represent critical points (Jeong et al. 2001; Yu et al. 2004). BN networks herein were constructed to examine interactions of selected conidiation, *het*, and sexual development genes that have been functionally characterized in *N. crassa* and their homologs in *C. globosum* during perithecial development (**Fig. 7, Table S4**). As expected, sexual development genes were positioned as core hubs of the process in BNs for both species. These core regulators were interacting with both *con* and *het* genes. However, *con* genes were inferred as central in *C. globosum* **(Fig. 7A)**, while *het* genes were inferred as central in *N. crassa* **(Fig. 7B)**. Expression of *het* genes was highly coordinated during the pairing and fusion of two haploid nuclei from opposite mating type separately in early sexual development of *N. crassa*. Consistent with that coordination, *het-C2*, and *het-6* were attributed central roles during the early development and positioned as top-tier regulators of the *N. crassa* sexual development network, with dense direct associations with asci and ascospore developmental genes and other *het* genes. Expression of conidiation genes *con-10* and *acon-2* appear co-regulated with major sexual development genes in *N. crassa*. Associations with expression of other conidiation genes, *het* genes, and sexual development were minimal. Associations among genes *acon-3*, *nrc-1*, and *con-3* and the two mating-type genes *mat A1* and *mat a1*—which have known roles in crossing and regulate *het* genes during early perithecial development in *N. crassa* (Glass et al. 1990)—were not inferred as central regulatory modules—and were positioned as lower-tier elements in the Bayesian network. The central roles of *het* genes in the *het-con*-sexual development network in *N. crassa* contrasts with the less critical regulatory roles of *het* genes observed in *C. globosum* network. The *het* genes—except for *het-C2*—were positioned low with

presumably less regulatory roles in sexual development in the *C. globosum* network. Another contrast between the *N. crassa* network and the *C. globosum* network is consistent with the life cycle divergence between the two species: mating type loci *mat A1* and *mat a1* were placed as top-tier regulators in the *C. globosum* sexual development BN along with *asd-1* and *con-10*. This position and the gene associations indicated suggest a more centralized post-crossing role for mating-type loci in homothallic *C. globosum* than in heterothallic *N. crassa*—in which mating-type loci function to regulate the crossing process before the initiation of the perithecial development.

Expression of osmolarity responsive genes and secondary metabolism clusters—Chaetomium species prefer high-humidity habitats, and actuate a variety of secondary metabolism via secondary metabolism clusters (SMCs). Phylogenetic analyses of the cellular hyperosmotic responsive histidine kinases (HKs) revealed seven *C. globosum* HKs that are homologous with osmotic-1 (*os-1*), osmolarity two-component system protein *sln-1*, development and carotenogenesis control-1 protein (*dcc-1*), and two-component system protein A (*hcp-1*; **Fig. 8A**). Expression of these HKs was highly coordinately up-regulated in *C. globosum* (**Fig. 8B**) as it is in *F. graminearum* (**Fig. 8C**) during the perithecial development process (with the exception of a coordinately down-regulation in expression in response to disturbance in *F. graminearum*). In comparison, expression of these HKs in *N. crassa* sexual development is patternless (**Fig. 8D**). A total of 41 SMCs were predicted within the JGI database to be present in the genome of *C. globosum* CBS 148.51 (**Table S5**), and only 16 clusters reported in *N. crassa* OR74A. Among the 41 SMCs, 28 are composed of three or more genes that are physically linked on chromosomes, and 12 of these multigene SMCs exhibited highly synchronized expression

across sexual development (**Fig. S4**, panels **A–L**). Generally a down-regulation was observed for most multilocus SMCs, but biosynthetic clusters for aureonitol (CHGG_00239-00246), chaetoglocin (CHGG_10645-10649), and cochliodones (CHGG_10019-10029) exhibited peak expression in mature perithecia—especially for cochliodones synthetic clusters, in which expression was up-regulated only after 48 h of perithecial development, as asci and ascospores developed (**Fig. S4G**). By searching for genes orthologous to those identified in SMCs of other fungal genomes, thirteen genes identified in *C. globosum* as orthologous members of SMCs contain only a single gene; nine of these apparently solitary secondary metabolism genes exhibited up-regulation during later perithecial development (**Fig. S4M–O**).

Knockout phenotypes—In comparison of *C. globosum* gene expression to *N. crassa*, 46 genes exhibited highly similar expression levels across sexual development. Among these 46 genes, 15 are hypothetical proteins that have yet to be annotated with functions (**Table S6**). Ten of these hypothetical proteins showed at least once with higher than 3-fold expression change in *N. crassa* and much higher expression changes in *C. globosum*. These 10 genes were selected for knockout phenotyping during sexual development in *N. crassa*; one (NCU06316) exhibited a knockout phenotype of early-stage arrested perithecial development (**Fig. 9A–C**). Homozygotes of NCU06316 knockouts (FGSC20345 *mat* ***a*** and FGSC20346 *mat* ***A***) produced dark and enlarged perithecia; however, these perithecia failed to produce asci and ascospores (c.f. Lehr et al. 2014). An additional 12 late-responding light-induced genes and 4 early-responding light-induced genes (Chen et al. 2009) exhibiting differential expression regulation between the two species were also investigated for their regulatory roles in sexual development with *N. crassa* knockout mutants. Of the 16 selected light-responsive genes, one (NCU07441; in the cross of **a**

strain FGSC15502 and **A** strain FGSC15503) exhibited a knockout phenotype of arrest during protoperithecial development (**Fig. 9D–F**).

Discussion

Here we have studied transcriptomics during sexual reproduction for *Chaetomium globosum* and compared it with closely related model fungi based on their shared similarity in sexual morphological development and response to environmental signals for initiation of sexual development. We have revealed integrative activities among mating loci, environmentally responsive genes and secondary metabolism pathways during *C. globosum* sexual development. Highly coordinated regulation of mating loci suggests that these genes contribute to the further regulation of post-crossing development in homothallic *C. globosum*. Up-regulated expression of the cellular hyperosmotic responsive histidine kinases potentially reflects that sexual development of *C. globosum* has become highly adapted to humid environments. The active regulation of clusters of secondary-metabolism genes across sexual development suggests that active fungal defence of fruiting bodies from predation is an important aspect of sexual development in *C. globosum*.

Although *C. globosum* and *N. crassa* are adapted to very different ecologies (highly humid, even aquatic substrates vs. heat-killed post-fire vegetation), the morphology and sexual development of *C. globosum* and the genetic model *N. crassa* are highly similar. Under the same laboratory settings on carrot medium, the sexual development of *C. globosum* proceeds in nearly perfectly parallelism with that of *N. crassa* in terms of sequence and timing of major development characters, including protoperithecia, initiation of perithecia, ascogenous center appearance, asci and ascospores formation, and ascospores maturation and release. Consistent

with this observation, it is not surprising that expression regulation of many gene markers associated with morphology development showed similar patterns between *C. globosum* and *N. crassa*. For example, genes in the melanin synthesis pathways that are critical for ascospore pigment production (Ao et al. 2019) exhibited highly similar up-regulation during ascospore development in *Chaetomium* and *Neurospora* species.

Unfortunately many genes in *C. globosum* genome are not functionally annotated, and high non-homologous random recombination against foreign DNA in *C. globosum* make it a hard system to introduce precise gene knockout approach for genetic purposes (Nakazawa et al. 2013). Taking advantage of the well annotated *N. crassa* genome and *N. crassa* knockout mutants available from a systematic knockout program, we were able to assess possible functions of some of these genes in *C. globosum*. Some unannotated genes exhibited similar expression regulation between *C. globosum* and *N. crassa*, and two of them showed knockout phenotypes in sexual development in *N. crassa* making them interesting regarding genetics of sexual development in *C. globosum*. However, differences in expression dynamics for some genes were also revealed by comparative transcriptomics between *C. globosum* and other fungal models, and such expression divergences likely reflect the genetic divergence and ecological adaptation during sexual development among these fungi, including mating type loci activities, pheromone regulation, osmolarity response, secondary metabolites production, as well as asexual elements expression during sexual development.

Although genetic background in mating type loci has been intensively studied for both heterothallic and homothallic lifestyle (Heitman 2015), little is known about how such divergence in genetic settings play different roles in sexual development regulation between the two lifestyles. Comparative transcriptomics between *C. globosum* and *N. crassa* showed

divergent regulatory behaviors of mating type genes between homothallic and heterothallic life styles. Mating type loci, from both mating types, behave highly coordinated through the whole sexual reproduction process in homothallic *C. globosum*, and high coordination was also observed for pheromone receptors and precursor, which are regulated by mating type loci and regulate the mating anc crossing process (Wilson et al. 2019). Nevertheless, current knowledge about mating loci and pheromone regulatory pathways and their function during mating processes for ascomyceteous fungi are mainly and ironically based on homothallic yeast models (Lin and Heitman 2007; Tsong et al. 2007). One of the filamentous fungi that has provided great insights into the genetics of homothallic species is *Sordaria macrospora*—a species closely related to *N. crassa* and *C. globosum*. In *S. macrospora*, mating-type genes are functional genes that are required for the initiation of sexual reproduction (protoperithecia) and likely regulate fruiting body development and ascosporogenesis via tight interaction with other genes (Poggeler 2007). In the heterothallic and pseudo-homothallic speciess of *Neurospora*, a neither significant upregulation of the *mat **A*** mating-type loci nor coordinated expression of the *mat **A*** and *mat **a*** loci has been observed during the perithecial development. The exception is a dramatic increase of *mat a-1* expression in *N. crassa* *(Lehr et al. 2014; Wang et al. 2014)*, arising as a consequence of a cross with *mat **A*** as the protoperithecial partner and *mat **a*** as the filamental partner in *N. crassa*. The divergence in expression activity of the mating-type loci between these homothallic and heterothallic species likely arises as a shift in the major roles of mating-type genes in homothallic species: a loss of role in the regulation of mating and crossing in heterothallic species, perhaps accompanied by an enhanced role in the regulation of development of fruiting bodies in homothallic species. In the reconstructed BN including both mating type loci and sexual development genes, the mating-type loci in *C. globosum* tightly associated with sexual

developmental genes, whereas in the *N. crassa* BN, mating-type loci were loosely associated with *het* and *con* genes, suggesting loss of regulatory roles during asci and ascospore development. Experimental research perturbing the mating type loci and MAPK signaling pathways in post-crossing stages of perithecial and ascospore development from both species would further clarify the genetic divergence in mating-type gene regulatory action and the concomitant changes in fungal life history.

There are many commonalities among filamentous fungi in how they respond to environmental stresses by production of resistant sexual reproductive structures and ascospores (Rodriguez-Romero et al. 2010). It would be of great interest to better understand how genome-wide gene expression underlying fungal sexual development evolved to maintain these commonalities while simultaneously adapting to diverse species-specific habitats. Genes responsive to light, oxygen, and humidity exhibited dissimilar expression patterns among *C. globosum*, *F. graminearum*, and *N. crassa* during sexual reproduction, indicating that potentially common processes of perithecial development are responsive to different environmental conditions among these fungi. Whole-genome expression data across the complexity of fruiting body development also captured how fungi respond differently to environments during sexual reproduction. While we did not assay expression in differing environmental conditions in this study, we did maintain a homogeneous environment in which species with differing genetics were cultured; interestingly, genes that regulate fungal response to environmental factors such as hydrogen peroxide or osmolarity were expressed in a species-specific manner. Cellular hyperosmotic responsive histidine kinases (HKs) are PAS domain proteins that have long been recognized as transducers of diverse environmental signals (Zhao et al. 2018), and our evolutionary analyses revealed additional copies of these HKs in *C. globosum* that are

homologous to those in *N. crassa* and *F. graminearum*. Coordinately up-regulated expression of all hyperosmotic-responsive histidine kinases in *C. globosum* and *F. graminearum* is consistent with observation of preferences for high humidity and the demonstrated role of water in spore release in these fungi (David et al. 2016; Dixon 1961). In contrast, *N. crassa* has a fully developed beak, through which ascospores are forcibly released;in natural settings. We observed low fold-changes that composed mixed expression patterns for these HKs, including *os-1*, *hcp-1*, *sln-1*, and *dcc-1*, during sexual development in *N. crassa*, consistent with a conception that humidity is not as critical for the propagation of the post-fire, air-dispersed fungus *N. crassa* as it is in water-dispersed *C. globosum* and *F. graminearum*.

Nevertheless, humidity affects conidiation in *N. crassa* (Guignard et al. 1984), and expression of HKs does respond coordinately to environmental humidity in some stages of life-history. Two of the cellular hyperosmotic responsive histidine kinases—*hcp-1* (Sun et al. 2012) and *dcc-1* (Barba-Ostria et al. 2011)—are conidiation-related. The expression of these HKs exhibited divergent expression during *N. crassa* conidia germination between different media (Wang et al. 2019). All these HKs exhibited similar patterns of down-regulation when *N. crassa* was cultured on Bird medium that was specifically formulated to induce conidiation. In contrast, these HKs—except *hcp-1*—were generally up-regulated for cultures on a natural, carbon-rich nitrogen-poor Maple sap medium that supports both asexual and sexual reproduction. These findings indicate that these genes may be key components of the network that regulates the environmentally-mediated asexual-sexual switch in fungi.

Another exciting aspect of sexual development in *C. globosum* are the activities of its numerous pathways devoted to synthesis of secondary metabolites. Our results provide insight into how expression of secondary metabolism clusters (SMCs)—which are highly enriched in

C. globosum—were regulated during sexual development. Secondary metabolites are part of fungal defense system (Künzler 2018). Thus, it is not surprising that many SMCs were up-regulated in response to the disturbance step implemented to synchronize perithecial development. Eleven SMCs that were composed of more than 3 physically linked loci exhibited highly coordinated expression dynamics across sexual development. Ten of these SMCs were expressed at a high level during early perithecial development before the fungus has equipped with physical defences such as thick-walled cells and setae. One exception was gene clusters that encoded enzymes for the synthesis of cochliodones. Expression of these genes increased markedly only after 48 h of post-cross perithecial development—concomitant with the development of asci and ascospores. Our observation of coordinated up-regualtion of mating loci, melanin synthesis pathways, and many secondary metabolism pathways during *C. globosum* sexual development is consistent with recent studies reporting highly associated activities among mating loci and secondary metabolism pathways during sexual development of pathogenic fungi (Yu et al. 2018; Hu et al. 2018). These secondary metabolites of *Chaetomium* are considered harmful to human health (Došen et al. 2017); Our discovery of stage-specific expression in the majority of SMCs in *C. globosum* will be useful to the development of strategies to manipulate the growth and development of these fungi for high production of secondary metabolites.

Lack of asexual reproduction in *C. globosum* remains a mystery—especially when it is noted that some members of the genus *Chaetomium*—and even the *C. globosum* species complex—do produce abundant conidia (Wang et al. 2016). Our finding that orthologs of major conidiation genes exist and are actively expressed during sexual development in *C. globosum*—a Sordariomycete without asexual spores—calls for further investigation of the ecology, evolution,

and genetics of fungal condidiation and other sporulation pathways. In natural settings, fungi quickly occupy substrate by inducing rapid growth and producing myriad fragile, genetically uniform asexual spores (conidia). In contrast, sexual reproduction is usually a slow process, during which fruiting bodies facilitate the production and dissemination of hardy propagules that have diversified genetics due to recombination (Wang et al. 2009). However, conidiation has never been observed for many fungi, including many species in the *C. globosum* complex (Wang et al. 2016) as well as the homothallic fungal model *Sordaria macrospora* (Nowrousian et al. 2010). As in the *S. macrospora* genome, the *C. globosum* genome hosts homologs of most conidiation-related genes as annotated in *N. crassa*, but orthologs of conidiation genes *con-6*, *con-8* and *con-13* were not found in *C. globosum* genome. Mutants of *con-8* and *con-13* exhibited defective growth of aerial hyphae and conidia production but normal vegetative growth in *N. crassa* (Berlin and Yanofsky 1985; Hager and Yanofsky 1990; Roberts and Yanofsky 1989). Many homologs of *N. crassa* conidiation genes are expressed in both asexual growth and sexual reproduction in *S. macrospora*. Three hypotheses, including yet-to-be-discovered asexual reproduction, loss of key conidiation genes or other mutation leading to loss of conidiation, and unfavorable combination of heterokaryon incompatibility genes, were posed to explain the lack of conidiation in *S. macrospora* (Nowrousian et al. 2010). In *N. crassa,* expression of *con* genes has long been known in all three sporulation pathways, including the production of macroconidia, microconidia, and ascospores, suggesting that their roles are typically not strictly relegated to any single sporulation process (Springer and Yanofsky 1992), and that they are independently activated during each sporulation. This conception is supported by the observation of divergent expression of conidiation genes during ascospore production between *C. globosum* and *N. crassa*.

It has been hypothesized that the loss of conidiation in homothallic *C. globosum* and *S. macrospora* occurs as a consequence of the conflict caused by heterokaryon incompatibility (Nowrousian et al. 2010). Investigation of the behavior of *het* genes during conidiation for these fungi that produce no conidia is challenging. However, interactions among conidiation genes and heterokaryon incompatibility during sexual spore production could provide clues regarding possible conflicts between the two functional groups. Indeed, heterokaryon incompatibility proteins with *het*-domains are abundant in these fungi, and expression of many heterokaryon incompatibility genes was up-regulated during karyotic fusion and ascospore production during *N. crassa* sexual reproduction (Wang et al. 2014), but homologs of these genes in *C. globosum* did not exhibit any consistent expression pattern, except for a trend toward up-regulation during ascospore maturation of some homologs of *N. crassa het* genes. Reconstructed Bayesian networks relating the associations of expression among sexual development, conidiation, and heterokaryon incompatibility genes suggest that conidiation and heterokaryon incompatibility relate differently to sexual development in *N. crassa* and *C. globosum*. Namely, heterokaryon incompatibility was tightly associated with regulation of sexual development of asci and ascospores in *N. crassa*, whereas it was not involved in the regulation of sexual development in *C. globosum*, indicated with lack of direct connections with sexual regulators. However, the preservation and frequent up-regulation of these *het* genes during ascospore maturation suggest that these genes operate in unknown but important roles in sexual development in both *N. crassa* and *C. globosum*.

Interestingly, expression of conidiation genes was highly associated with the expression of sexual regulators during asci and ascospore development in *C. globosum*, and some of these associations—including *con-3* and *asd-1*, *spo11* and *acon-2*, *asd-1* and *acon-3*, and *acon-2* and

asl-1—are conserved between the *N. crassa* and *C. globosum* BNs. Apparently, as the production of ascospores commences, these conidiation genes are regulated by divergent developmental mechanisms between *C. globosum* and *N. crassa*. In *C. globosum*, conidiation-associated genes and heterokaryon-incompatibility loci might have lost their co-regulation along with losing heterothallism. There has been speculation of a shared genetic basis among different sporulation pathways (Springer and Yanofsky 1992). Nevertheless, the roles of conidiation genes in ascospore production are barely characterized in *N. crassa*. Because conidia are essential to starting the fertilization process, it is challenging to study the role of *con* genes during sexual development in heterothallic *N. crassa*. Our observation would promote *C. globosum* as an alternative model to study shared regulatory mechanism among different sporulation pathways and conidiation genes' roles in ascospore production in these fungi. Additional genetic data for these genes and how they interact with heterokaryon incompatibility gene during asexual growth in homothallic and no-conidium producing *C. globosum* and *S. macrospora*, homothallic and conidium producing *F. graminearum*, and heterothallic and conidium producing *N. crassa* would be highly informative. Comparative genomics between conidiation and non-conidiation species in the species complex of *C. globosum* could also illuminate why some of these fungi have lost their proclivity or their ability to reproduce asexually.

Conclusion

Comparative transcriptomics of the sexual development between *C. globosum* and closely related fungi revealed that mating type loci are expressed highly coordinately in *C. globosum*, and appear to play divergent roles in regulating post-crossing sexual development from their roles as regulators of mating in *N. crassa*. This study calls for further investigation of the means by

which conidiation genes have evolved to interact with heterokaryon incompatibility genes in diverse fungal models to understand why conidiation has not yet been observed for some of the fungi. We have shown that environmental responses to humidity and secondary-metabolite synthesis pathways are actively regulated during *C. globosum* sexual reproduction. Some pathways are highly expressed during sexual reproduction producing resistant perithecia and ascospores in *C. globosum*, providing useful information for diagnostic and treatment purposes of this pathogenic fungus.

Acknowledgements: We thank the Broad Institute and JGI for making *Chaetomium globosum* gene and genomic data available for oligonucleotide prediction. All authors have declared that no competing interests exist. This study was supported by funding to JPT from The National Institutes of Health P01 grant GM068067, by funding from the National Science Foundation (Grant number IOS 1457044) to JPT, and (IOS 1456482) and Michigan AgBioResearch to FT. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Figure 1. Stages of sexual development of *Chaetomium globosum*. Dissection microscope images (scale bars: 1 mm) of protoperithecia A) prior (0 h) and B) subsequent to disturbance, and C–E) perithecial development sampled at 2 h, 24 h, and 48 h. F1–J1) Dissection microscope images (scale bars: 1 mm) of cultures and F2–I2) light microscope images (scale bars: 50 μm) of perithecial squashes mounted in water showing inner tissue development, performed at 72, 96,

120, 144 and 168 h. J2) ascospores released at 168 h . Arrows identify protoperithecia (Pp), perithecia (Pt), ascogenous tissues (Ag), ascii (Ac) and ascospores (As).

Figure 2. Genome-wide gene expression of *Chaetomium globosum* across sexual development, clustered with Euclidean distance and representative expression patterns. A) Hierarchical gene clustering and corresponding expression patterns over sexual development, for which expression was compared in fold changes. B–I) Representative clusters showing predominant expression patterns. B1–I1) Gene expression profiles (relative to lowest stage-specific expression) for genes in each selected cluster. B2–I2) Comparison of expression between *C. globosum* (blue) and *Neurospora crassa* (orange) for selected orthologous gene pairs in each cluster, suggesting ancestrally retained or recently divergent expression between the two fungi during sexual reproduction.

Figure 3. Gene expression of melanin synthases (in fold change relative to the stage of lowest expression) exhibited up-regulation in A) *Chaetomium globosum*, B) *Neurospora tetrasperma*, and C) *Neurospora crassa* as ascospores matured. A local peak of expression during early perithecial development in *Neurospora* species is concomitant with the melanization of perithecia; in *C. globosum,* perithecial wall cells are generally hyaline to lightly tinted. Gene names (*per-1*, *scy-1*, *tnr-1*, and *tnr-2*) correspond to those used in *N. crassa*; orthologs in *N. tetrasperma* and *C. globosum* are identified by the name of their *N. crassa* ortholog. Error bars convey 95% confidence intervals.

Figure 4. Expression of mating type A genes (*mtA-1*, blue line;, *mtA-2*, red line; and *mtA-3*, green line) the mating type a gene (*mta-1*, purple dashed line), and pheromone-related genes (*pre-1*, blue line; *pre-2*, red line; and *ccg-4*, green line) across sexual development in *Chaetomium globosum* and *Neurospora crassa*, where expression has been quantified relative to the lowest stage-specific expression level. A) In *C. globosum*, expression of mating-type loci was highly coordinated and was up-regulated during ascospore development (starting at 48 h after initiation of perithecial development). B) In *N. crassa*, expression of *mat A* was very low (not shown; note large credible intervals) and changes in expression across sexual development were not statistically significant. In contrast, expression of *mta-1* was dramatically upregulated. C) In *C. globosum*, expression of pheromone genes generally followed a pattern of up-regulation across sexual development similar to that of the *N. crassa* mating-type loci. D) In *N. crassa,* pheromone genes were not coordinately expressed; *matA*-specific pheromone precursor *ccg-4* was highly up-regulated during sexual development. Note the second *y*-axis (purple dashed) provided to quantify relative expression of *mta-1* occurring on a different overall scale. Error bars convey 95% confidence intervals.

Figure 5. Three *Chaetomium globosum* genes and their orthologs in *Neurospora crassa* that are light-responsive and critical to the initiation of *N. crassa* sexual reproduction exhibit similar expression dynamics in A) *C. globosum* and B) *N. crassa* during perithecial development. Error bars convey 95% confidence intervals.

Figure 6. Divergent expression of genes critical for conidiation during sexual development, comparing orthologs in A) *Chaetomium globosum* to genes of B) *Neurospora crassa*. Note the

second *y*-axis for *con-10* (blue dashed) in *C. globosum* and *con-3* (orange dashed) in *N. crassa*, provided to quantify relative expression occurring on a different overall scale. Error bars convey 95% confidence intervals.

Figure 7. Bayesian networks during sexual development in A) *C. globosum* and B) *N. crassa*, relating associations among homologs of sexual development (red), conidiation (blue), and heterokaryon incompatibility (black) . Edge connections represented in 50% or more of the likely models between the two networks (green).

Figure 8. Phylogeny and comparative expression of hyperosmotic responsive histidine kinases (HKs) and related proteins in *Chaetomium globosum*, *Fusarium graminearum*, and *Neurospora crassa*. A) Phylogeny of the cellular HKs and related proteins. Thick branches received significant support (Bayesian Posterior Probability > 0.95). B) Expression profiles of cellular HKs across sexual development in C) *C. globosum*, D) *F. graminearum*, and E) *N. crassa* (relative to the lowest stage-specific expression). Error bars convey 95% confidence intervals. Note the second *y*-axes (purple dashed axis for *C. globosum* gene CHGG_06323; blue dashed axis for *N. crassa* gene NCU07221), provided to quantify relative expression for one gene occurring on a different overall scale from other genes.

Figure 9. Knockout strains of unannotated genes exhibit mutant phenotypes on synthetic crossing medium, indicating critical functions in *N. crassa* perithecal development. Knockout strain ΔNCU06316 A) formed bands of perithecia along the line of confluence after *A* and *a* strains were inoculated on the opposite sides of plates, B) produced dark perithecia along these

crossing bands, and C) arrested perithecial development before production asci and ascospores. Knockout strain ΔNCU07441 D) did not form bands of perithecia along the line of confluence after *A* and *a* strains were inoculated on the opposite sides of plates, E) produced abundant, light-colored protoperithecia on the plate surface, and F) arrested protoperithecial development, exhibiting an abortive ascogenous center.

Figure S1. Synchronization of perithecial development in *Chaetomium globosum* was aided via inoculation of solid carrot agar cultures by hyphal fragments from A) filtered liquid cultures. Synchronized development of *C. globosum* perithecia on CA plates is sampled at B) 5, C) 8, and D) 11 days after inoculation.

Figure S2. Expression of catalase genes *cat-1*, *cat-2*, and *cat-3* exhibited divergent regulation between A) *Chaetomium globosum* and B) *Neurospora crassa*, suggesting perithecial development in these species have adapted to differing microenvironmental conditions such as oxygen level and air flow intensity. Note the second *y*-axis (blue dashed) provided to quantify relative expression of *cat-1* occurring on a different overall scale.

Figure S3. Molecular phylogenies and sexual development expression profiles of heterokaryon incompatibility genes in *N. crassa* and *C. globosum*. A) RAxML tree based on the best-scored alignment by SATe II identifies three major clades of *het* homologs between *N. crassa* and *C. globosum*, including orthologs of two *N. crassa* HET-C genes, members of B) the TOL-HET14 Bayesian phylogeny based on the best-supported SATe II alignment with focused sequences, and members of C) the HET-6/13/15 Bayesian phylogeny, based on the best-supported SATe II alignment with focused sequences (thick branches are supported with a posterior probability ≥ 0.95). D) Expression profiles of *het* genes during sexual development in *N. crassa* (average expression, dashed black curve), E) expression profiles for *het* genes during sexual development in *C. globosum* (average expression, dashed black curve), and F) comparison between *N. crassa* and *C. globosum* of expression of selected het genes during

sexual development (relative to the lowest stage-specific expression). Error bars convey 95% confidence intervals.

Figure S4. Expression profiles of genes within fifteen secondary-metabolism gene clusters across nine stages of sexual development, where expression has been quantified relative to the lowest stage-specific expression level for each gene.

Table S1. Single-copy orthologs identified between *Chaetomium globosum* and *Neurospora crassa* genomes and, where available, KEGG functional annotation.

Table S2. Sequencing and mapping quality statistics for all RNAseq samples.

Table S3. Genome-wide gene expression levels across sexual development in *Chaetomium globosum*.

Table S4. Input data matrix for the Bayesian network reconstruction, the model-averaged scores of edges, and the structure matrix underlying the Bayesian networks depicted in Figure 7.

Table S5. Forty-one secondary-metabolism clusters predicted within the *Chaetomium globosum* genome at the JGI Mycocosm (Grigoriev et al. 2014) database.

Table S6. Forty-six genes exhibiting highly similar gene expression patterns between *Chaetomium globosum* and *Neurospora crassa* across sexual development; 15 of the 46 are currently annotated as hypothetical proteins.

References: