UNIVERSITY OF CALIFORNIA
RIVERSIDE

Functional Genomics and G-Protein Signaling in *Neurospora crassa*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Microbiology

by

Alexander J Carrillo

September 2021

Dissertation committee:
Dr. Katherine Borkovich, Chairperson
Dr. Patricia Manosalva
Dr. Howard Judelson

The Dissertation of Alexander J Carrillo is approved by:

_____

_____

_____

Committee Chairperson

Acknowledgements

The road to obtaining a Ph.D. is a long and difficult one. There were many people who have helped me through this journey towards my Ph.D. So many professors were important for helping me learn how generate hypothesis, conduct experiments, and ultimately contribute to our understanding of cellular signaling. I am grateful to all the people I have met in my Ph.D. journey for helping me become who am I today.

First and foremost, I must thank my major professor, Dr. Katherine Borkovich, for being a fantastic mentor and giving me the opportunity to work in her lab. My work in her lab as both an undergraduate and graduate student solidified my desire for pursuing research. With her wealth of knowledge supervision and guidance I have become a better researcher, teacher, writer, collaborator, and leader.

I thank all my lab mates past and present for their guidance, wisdom and great attitudes. They are: Ilva Cabrera, Arit Ghosh, Alexander Michkov, and Shouqiang Ouyang. I want to sincerely thank Ilva Cabrera, who was the first graduate student that I worked with as an undergraduate and she helped guide me into my graduate career. Additionally, I want to thank Logan Collier for being there to discuss data and bounce ideas off. I will miss us supporting each other throughout our time in graduate school.

I also want to thank all of the undergraduate students that I mentored including Chathushya Keerthisinghe, Emily Bossard, and Veronica Batallones. Berenise Lopez-Lopez for being such a great student and having patience with me as I learned how to be a mentor. I also want to thank is Lida Halilovic for being an amazing student that really stepped up and took charge of the projects I gave to her.

Finally, I would like to thank my family and my partner Clarissa Gomez who has been there and helped me through this journey. I am forever grateful of Clarissa's understanding of the long hours and weekends that I would spend in lab. We both have grown so much during this time, and I am glad you have been there for me when I needed you.

The text of this dissertation, in part, is a reprint of the material as it appears in Functional profiling of Transcription Factor genes in *Neurospora crassa* (2017), Genetic relationships between the RACK1 homolog *cpc-2* and heterotrimeric G protein subunit genes in *Neurospora crassa* (2019) and Clustering analysis of large-scale phenotypic data in the model filamentous fungus *Neurospora crassa* (2020). The co-author Katherine Borkovich listed in these publications directed and supervised the research which forms the basis for this dissertation.

ABSTRACT OF THE DISSERTATION

Functional Genomics and G-Protein Signaling in *Neurospora crassa*

by

Alexander J Carrillo

Doctor of Philosophy, Graduate Program in Microbiology
University of California, Riverside, September 2021
Dr. Katherine Borkovich, Chairperson

Functional genomics is a powerful tool for identifying genes that function in the same pathway in an organism. Environmental sensing mechanisms, and G protein signaling in particular, are important for relaying information from outside the cell to allow it to generate the proper response in eukaryotes. The main objectives of this dissertation are to 1. Generate and utilize phenotypic data that may reveal potential gene pathways in *N. crassa*, 2. Decipher the ways in which G-protein signaling is regulated and 3. Determine the role G-protein signaling has in regulating metabolism.

In Chapter 2, *N. crassa* transcription factors were annotated and characterized for growth and development phenotypes. Publicly available RNAseq datasets were mined to determine possible correlations between transcription factor phenotype and transcript abundance during sexual development. We identified a total of 312 transcription factors in *N. crassa.* Complete phenotypic data were obtained for 242 strains using a combination of publicly available data and new analysis of gene deletion mutants

generated during the study. Of the 242 transcription factor gene deletion strains, 64% had at least one defect in growth and development. The combination of RNAseq analysis with phenotypic data revealed several transcription factor genes with sexual development phenotypes that correlated with transcript abundance in wild type.

In Chapter 3, I took all available phenotypic data that has been generated for gene deletion mutants in *N. crassa* (data for nearly 1,300 strains) and tested the ability of several statistical clustering methods to group mutants based on their growth and development phenotypes, with the goal of identifying potential pathways. Analysis of several clustering methods showed that using a weighted partitioning around medoids approach generated the most biologically relevant grouping of mutants. Publicly available RNAseq datasets were used to determine if there is any correlation between gene expression and phenotype. Most phenotypic clusters contained multiple expression profiles, suggesting that co-expression is not generally observed for genes with shared phenotypes. Yeast ortholog data for genes that co-clustered with MAPK signaling cascade genes were mined and revealed potential networks of interacting proteins in *N. crassa*.

In Chapter 4, we investigated genetic interactions between the Receptor for Activated C Kinase-1 (RACK1) homolog *cpc-2*, the Gβ subunit *gnb-1* and other G protein signaling components in *N. crassa*. We showed that CPC-2 is a cytosolic protein via cell fractionation and fluorescent microscopy. We observed genetic epistasis between *cpc-2* and *gna-2* for basal hyphae growth rate and aerial hyphae height. Mutational activation of *gna-3* alleviated the submerged conidiation defect observed in the Δ*cpc-2*

mutant. *cpc-2* and *gnb-1* exhibited a largely synergistic relationship, with mutants lacking both genes showing more severe defects for all phenotypic traits.

In Chapter 5, I used a combination of RNAseq and liquid chromatography-mass spectrometry to profiles the transcriptomes and metabolomes of wild type, Δ*gna-1*, Δ*gna-3* and Δ*ric8 N. crassa* strains. We observed large transcriptional differences between mutants and wild type. Many of the differentially expressed genes encode metabolic enzymes, and the electron transport chain was impacted some strains. Metabolome analysis revealed changes in levels of several primary metabolites for all mutants. Comparing the RNAseq and metabolomics data provided evidence for both transcriptional and post-transcriptional regulation of certain metabolic proteins in the various mutants.

Table of Contents

List of tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 5

Chapter 1
Introduction

Neurospora crassa as a model organism

Neurospora crassa is a multicellular eukaryote, filamentous fungus in the phylum Ascomycota. The ease of obtaining mutants led to its use as an early model organism for molecular genetics and biology and had profound effects on biochemistry (1). *N. crassa* was the organism used to generate some of the foundational theories in modern genetics, such as the one-gene-one-enzyme hypothesis (2). *N. crassa* was also used to study the first conditional mutants (3), compartmentation of metabolic pools and pathways within vacuoles (4) and connecting biochemical defects to altered vegetative morphologies (5). Later, Neurospora became a well-known model for the analysis of the molecular basis for the eukaryotic circadian oscillatory system (6) and is now potentially a new model for study of mycoviruses (7).

The long history of laboratory studies that use *N. crassa* means that there is a vast amount of biochemical/genetic information available which leads to many biological and bioinformatics tools being available for *N. crassa*. The complete genome was released in 2003 and the latest assembly is publicly available at FungiDB (http://fungidb.org/fungidb/ (8,9)). Methods to quickly create knock-out mutants and tagged strains are well defined. With these tools, targeted DNA insertions, deletions and replacements in the genome can be quickly created and studied (10). With the combination of available gene deletion mutants and the wide array of phenotypic assays for which protocols have been developed for high-throughput analysis, many gene deletion mutants have now been characterized in *N. crassa* (11).

Life cycle of *N. crassa*

N. crassa is a heterothallic (self-sterile) fungus, which grows vegetatively by apical extension of basal hyphae (12). Hyphae are tube-like structures that increase surface area and perform specialized functions during growth. Hyphae have chitinous cell walls, and undergo fusion, branching and linear growth by addition of material at the hyphal tip to make up the hyphal network (12-14). There are six specialized hyphal types: leading hyphae, trunk hyphae, aerial hyphae, enveloping hyphae, trichogynes and ascogenous hyphae (13, 15-23).

Environmental cues, such as carbon deprivation and exposure to oxygen can stimulate formation of aerial hyphae that grow perpendicular to the trunk hyphae (12, 24). Repeated apical budding on the aerial hyphae results in formation of conidiophores, eventually giving rise to the mature spores or macroconidia (12). Macroconidia are spread throughout the environment via wind dispersal or mechanical agitation (25). Upon acquisition of appropriate nutrient conditions, conidia can establish a new colony, thereby completing the asexual phase of the life cycle (25).

N. crassa has two different mating types (*mat a* and *mat A*) that are encoded by one locus (14). When starved for nitrogen, *N. crassa* enters sexual development, which starts with the formation of female reproductive structures known as protoperithecia (17). Next, female hyphae (trichogynes) of one mating type (e.g., *mat a*) will grow chemotropically from a protoperithecium towards a male cell (conidium) of opposite mating type (e.g., *mat A*) resulting in cell and nuclear fusion. Thereafter, meiosis occurs and the protoperithecium matures into a perithecium (26). During the process of

maturation from a protoperithecium to perithecium, asci develop and undergo a pair of meiotic divisions followed by single mitotic division, forming eight sexual spores termed ascospores (17). Ascospores are dormant until exposed to heat, after which they will germinate to produce hyphae (20).

Phenomics

N. crassa has a rich history of forward genetics, with more than 1000 loci identified using classical genetics approaches (27). The Neurospora Genome project sought to mutate all genes in the *N. crassa* genome using reverse genetics (28,29). With the nearly 9000 gene deletion mutants in *N. crassa* being available to the community, another goal of the Neurospora Genome Project was to perform phenotypic characterization of all knockout mutants. An undergraduate research program at the University of California, Los Angeles, pioneered methods for phenotyping mutants, with ~1000 mutants analyzed (11). The first large group of mutants analyzed were 103 transcription factors, out of a total of 182 genes annotated at the time (28,30). With only 44 of the 103 (43%) of the TFs exhibiting at least one phenotype (28).

The Borkovich laboratory has used the phenotypic methods developed in collaboration with UCLA to analyze additional groups of mutants lacking serine/threonine protein kinases, serine/threonine/tyrosine protein phosphatases and G protein coupled receptors (31-33). These studies analyzed 10 phenotypes across vegetative growth, asexual, and sexual development for large groups of gene deletion mutants (11). Serine/Threonine (S/T) kinases are important enzymes in signaling in eukaryotes (34). There are 86 serine/threonine(S/T) protein kinases in *N. crassa*, with 77

viable gene deletion mutants. When analyzed for the 10 phenotypes, 57% exhibited at least one growth or development defect, with 40% of all mutants possessing a defect in more than one trait (31). Another integral part of signaling pathways are the protein phosphatases. In *N. crassa* there are 30 protein phosphatases, with 24 viable gene deletion mutants. The protein phosphatases have 91% of the gene deletion mutants exhibiting at least one phenotype, and 29% overall having at least one defect in growth, asexual and sexual development (32). The third group of genes analyzed were the G-protein coupled receptors (GPCRs). There are 45 predicted GPCRs, with 36 available gene deletion mutants. The GPCRs had fewer phenotypes as compared to the previous two groups, with only 47% possessing at least one defect in growth or development (33).

With a large number of mutants and several traits measured for each mutant, organizing the data into relevant groups can become unwieldy. One method of organizing large amounts of data is through clustering (35). Clustering is the process of grouping items so that those in the same cluster are as similar as possible, while at the same time items in different clusters are divergent from those in other clusters (36). Clustering, as applied to phenotypes, groups mutants with similar defects, revealing genes that possibly act in the same or related pathways. To accomplish this, there are several different methods and programs. Broadly, there are two categories of clustering: partitioning and hierarchy, with several subcategories therein (37). The most common methods of clustering are K-means, Partitioning Around Medoids (PAM), complete linkage hierarchical clustering, and Pearson's Correlation (37).

Large-scale phenomics clustering analysis using Pearson's Correlation has previously been performed in several other fungal organisms: *Saccharomyces cerevisiae*, *Fusarium graminearum*, and *Cryptococcus neoformans* (38-40). In *S. cerevisiae*, individual homozygous deletion mutants lacking nonessential genes were exposed to a broad range of cytotoxic and cytostatic agents, generating a quantitative sensitivity profile for 4756 strains. The data for these strains was then clustered using Pearson's correlation (38). In *F. graminearum*, 17 different phenotypes were generated from analysis of vegetative growth, asexual development, sexual development, virulence, and stress. A total of 657 transcription factor mutants were analyzed, leading to ~11,000 different data points. The available data were both categorical and quantitative, and for this study they converted both the categorical and quantitative data to a single numerical scale and used Pearson's correlation to generate clusters (39). In *C. neoformans*, 30 different quantitative and categorical phenotypes were determined for 129 kinase mutants. As with the study in *F. graminearum*, the phenotypes were converted to the same numeric scale and Pearson's correlation was applied to generate phenotypic clusters (40).

Heterotrimeric G protein signaling

General background. Heterotrimeric G proteins play a critical role in transmitting extracellular stimuli into changes in gene expression (41). Chemical and sensory stimuli are recognized by specialized cell-surface proteins called G Protein Coupled Receptors (GPCRs). GPCRs are seven-helix membrane proteins that when

bound to a ligand act as Guanine nucleotide Exchange Factors (GEFs) for Gα proteins (Figure 1.1) (42). Thus, the receptors transduce information from environmental stimuli to intracellular signaling pathways via heterotrimeric G proteins. (41). The heterotrimer consists of a Gα, a Gβ and a Gγ subunit. The Gα subunit can be bound to either guanosine di-phosphate (GDP) or guanosine triphosphate (GTP) (Figure 1.1). When bound to GTP, the Gα dissociates from the βγ dimer, and both may regulate downstream signaling (41, 43). When the Gα is in its GDP-bound state, it is inactive and sequestered at the membrane with the tightly bound Gβγ heterodimer with some GPCR/G-protein complexes preassembled in the absence of any ligand (41, 42). The Gα protein has native GTPase activity and eventually hydrolyzes a phosphate, converting its GTP to GDP, and then returns to the membrane where it re-associates with the Gβγ dimer (44). In addition, Regulators of G Protein Signaling (RGS) proteins stimulate the native GTPase activity of the Gα subunit, accelerating its return to the inactive GDP bound form by more than 2000 times (Figure 1.1) (45).

G proteins in *N. crassa*. *N. crassa* possesses three Gα subunits, termed GNA-1, GNA-2, and GNA-3, one Gβ subunit (GNB-1), another possible Gβ subunit (CPC-2), one Gγ subunit (GNG-1), 45 putative GPCR proteins, and a non-GPCR guanine nucleotide exchange factor (RIC8) (30, 33, 46). Considering the Gα subunits, deletion of *gna-1* or *gna-3* results in defects in growth and development. Female sterility and slow hyphal growth rate were observed in Δ*gna-1* mutants (47, 48). The Δ*gna-3* mutant has defects in asexual development, with shorter aerial hyphae and premature conidiation on solid medium and inappropriate conidiation in submerged cultures (49). The Δ*gna-2* mutation,

when paired with deletions of either *gna-1* or *gna-3*, leads to phenotypes more severe than observed in the individual gene deletion mutants (48, 50). One of the few known phenotypes for the Δ*gna-2* single mutant is reduced growth rate when grown with glycerol as the sole carbon source. (51). Finally, both Δ*gna-1* Δ*gna-3* double mutants or a mutant that lacks all three Gα genes have severe defects in hyphal growth, asexual (short aerial hyphae) and sexual (female-sterile) development (48).

The Gβ subunit GNB-1 is an import regulator of Gα subunit stability, as deletion of *gnb-1* results in reduced levels of all three Gα subunits (52). The Δ*gnb-1* mutant also has defects during asexual and sexual differentiation, exhibiting inappropriate conidiation in submerged cultures and female sterility (52, 53). However, Δ*gnb-1* mutants do not have significant hyphal growth rate defects (53).

CPC-2, a possible alternative Gβ subunit. Receptor for Activated C Kinase 1 (RACK1) is a scaffolding protein that binds soluble, and membrane bound proteins and is known to shuttle activated Protein Kinase C (PKC) (54, 55). RACK1 is comprised of seven WD40 repeats that assemble into a seven-blade-β-propeller structure that facilitates interaction with other proteins (56). RACK1 has been shown to be involved in many roles during cellular function, from development, to cell migration, apoptosis, and circadian rhythms, as well as brain functions (56, 57). The homolog of RACK1in *S. cerevisiae*, Acs1p, has been implicated as alternative Gβ subunit through direct interaction with Gα subunits (58). Asc1p acts on the Gα Gpa2p, functioning as a Guanine nucleotide Dissociation Inhibitor (GDI) inhibiting the spontaneous exchange of GTP for GDP that is intrinsic to the Gα. Additionally, Asc1p is involved in regulating glucose

responsiveness through its binding to the effector enzyme adenylyl cyclase and lowering the amount of cAMP in response to glucose stimulation (58). In *Cryptococcus neoformans,* there is evidence for a physical interaction between the RACK1 homolog Gib2 and the protein kinase C homolog Pkc1 (59). In *Magnaporthe oryzae*, the RACK1 ortholog, MoMip11 interacts with both the Gα protein MoMagA and RGS protein MoRgs7 to regulate pathogenicity (60, 61). In *N. crassa*, the RACK1 homolog Cross Pathway Control-2 (CPC-2) regulates global derepression of amino acid biosynthetic genes during amino acid starvation under amino acid limited conditions (62). Possible interactions between CPC-2 and Gα subunits have not been studied.

RIC8, a non-GPCR GEF. RIC8 has been implicated as a positive regulator of Gα proteins in several animal species, including *C. elegans, Drosophila melanogaster* and mammalian cells (44). RIC8 is required for both asymmetric cell division in zygotes and priming of synaptic vesicles in *Caenorhabditis elegans* (44). Invertebrates such as *C. elegans* and Drosophila have one RIC8 protein, while vertebrates have two homologs, Ric-8A and Ric-8B (63). In Drosophila, RIC8 is essential for responses to extracellular ligands, for maintenance of polarity during asymmetric cell division in embryogenesis and for the stability of Gα and Gβ proteins (64). In mice, attempts were made to make *Ric8-A* or *Ric8-B* knockouts; however, these efforts failed, as the knockout mice died early during embryogenesis (65). Additionally, RIC8 is not conserved throughout all eukaryotes. RIC8 is found in several filamentous fungi, but there are no homologs in baker's yeast or plants (46).

In *N. crassa* there is one predicted RIC8 protein (46). Unlike in mice, *ric8* knock out mutants are viable in *N. crassa* (46). The Δ*ric8* strain displays severe defects during sexual and asexual development, as well as during basal growth. The phenotypes observed in Δ*ric8* are similar to those observed in Δ*gna-1*, Δ*gna-3* double gene deletion strain (46). Protein levels for the three Gα proteins (GNA-1, GNA-2, and GNA-3) and the Gβ GNB-1 are all decreased in the Δ*ric8* mutant (46). The defects were partially rescued by introduction of predicted constitutively activated (no intrinsic GTPase activity) and GTP-bound *gna-1*^{Q204L} and *gna-3*^{Q208L} alleles into the Δ*ric8* background (46). *In vitro* GTPγS-binding assays showed that in the presence of RIC8, both GNA-1 and GNA-3 had increased binding of GTPγS, demonstrating that RIC8 acts as a GEF towards GNA-1 and GNA-3 (46). This evidence taken together indicates that RIC8 may play a role in the regulation of Gα subunits, specifically GNA-1 and GNA-3. These results support the hypothesis that RIC8 regulates G protein signaling similarly in *N. crassa* and higher eukaryotes, and that discoveries in *N. crassa* will be relevant for progressing work on RIC8 in animal systems.

Metabolomics

One way we can understand the impact of G-protein signaling is to explore the effects of G protein mutations on the metabolite profile using metabolomics. Metabolomics is the process of determining the metabolites present in cells, tissues, and biofluids using advanced analytical chemistry techniques with modern statistical methods

(66, 67). Metabolites are substrates and products of metabolism that drive essential cellular functions, such as energy production, signal transduction and apoptosis (66). Metabolites have varied and important roles beyond just energy production; compounds such as ATP, NAD^+, and S-adenosyl methionine have been shown to act as co-substrates, regulating post-translational modification (68, 69). Generally, the collection of all small molecule chemicals in an organism is referred to as the metabolome (67, 70). Metabolite profiling encompasses two different areas: intracellular metabolites (endometabolome – metabolic fingerprinting) and extracellular metabolites (exometabolome – metabolic footprinting) (71).

Metabolite profiling strategies involve scanning for detectable metabolites using any of several techniques: Gas-chromatography-mass-spectrometry (GC-MS), gas-chromatography-time-of-flight-mass-spectrometry (GC-TOF-MS), liquid-chromatography-mass-spectrometry (LC-MS) and nuclear magnetic resonance (NMR). These are the principle analytical methods used for metabolite profiling due to their high sensitivity and/or resolution (71-73). One method of metabolite profiling is untargeted, global investigations, which uses wide polarity gradients aiming to cover as much of the chromatographic "space" available as possible, combined with full scan MS and generation of peak tables using peak-picking algorithms (74). The other general method used in metabolomics is targeted analysis, in which quantification and unambiguous detection of predefined metabolites is accomplished (71).

Several metabolomics studies have been performed in fungi. For example, in *S. cerevisiae*, GC-MS was used to determine the metabolic profiles of three different yeast strains involved in alcohol production and a laboratory reference strain at two different time points in the fermentation process. The laboratory reference strain was the most different metabolically as it did not produce as much acetate, ethyl esters and alcohols (75). In *Phialemonium curvatum* the metabolomes of the fungus cultivated on crude palm oil versus glucose were compared using targeted metabolomics for metabolites in the tricarboxylic acid cycle (76). On glucose, all metabolites in the TCA cycle were detected, while metabolites such as α-ketoglutarate and fumaric acid were undetectable on palm oil. Additionally, higher abundance of glyoxylic acid was detected, suggesting the glyoxylate cycle is the key metabolic pathway on palm oil (76). RNA-seq analysis was combined with metabolomics in *Fusarium graminearum* to assign metabolites to genes in a strain lacking the H3K27 methyltransferase. A total of 22 fungal metabolites were identified, with 10 compounds having not been previously reported in *F. graminearum* (77). In *N. crassa*, NMR was used to characterize the metabolome in mutant lacking the heterotrimeric Ga subunit *gna-3* (78). A difference was observed between Δ*gna-3* and wild type in response to low carbon conditions, with the Δ*gna-3* appearing to possess a carbon-sensing defect (78).

Hypothesis and Objectives

Signaling pathways are the mechanism by which cells receive outside information and induce biochemical changes inside the cell. Determining all of the proteins involved in a signaling cascade as either a regulator or a part of the cascade is important to having a complete understanding of the pathway. One way of determining the members of a pathway is by using reverse genetics to identify genes that have similar functions (79). Beyond signaling, understanding the relationship between gene expression and protein feature with phenotype can inform on the mechanics of how these genes influence phenotype. The first hypothesis of this thesis is that clustering the available mutant phenotypic data from *N. crassa* gene deletion mutants will reveal new potential members of known cellular pathways and uncover relationships between gene expression, protein feature and phenotype. I utilize phenotypic screening and statistical clustering to determine the influence of genotype and protein features on phenotype.

One important cellular pathway is G-protein signaling, playing a critical role in growth and development in all eukaryotic organisms. An important regulator of G-protein signaling is the non-GPCR GEF RIC8. The roles that the Gα subunits and RIC8 play in metabolism are not well-characterized. The second hypothesis of this thesis is RIC8 affects metabolism through its regulation of GNA-1 and GNA-3, with some possible Gα-independent regulation of metabolism. To determine the effects on metabolism, I utilized the latest techniques in metabolomics and RNA sequencing to analyze strains lacking *gna-1*, *gna-3* and *ric8*.

Aim 1: Identify and perform phenotypic analysis on all available transcription factor mutants in *Neurospora crassa*

Transcription factors are the proteins that are directly responsible for activating and deactivating gene expression. They typically function at the end of complex signaling pathways that respond to environmental cues. The Colot et al. study described above included less that 50% of the transcription factor genes in *N. crassa*. In order to provide more complete information for an important group of regulatory genes, the phenotypes of all 212 available transcription factor mutants were investigated in *N. crassa*. In this study, the complete set of transcription factors were catalogued, and the corresponding available mutants were analyzed for phenotypes using methods developed in (11).

Aim 2: Apply clustering algorithms to all available phenotypic data to sort genes into clusters and to identify potential pathways

With phenotypic data available for ~1200 gene *N. crassa* gene deletion mutants, large scale clustering of phenotypes could potentially reveal new members of cellular pathways. Using statistical analysis, genes that yield similar phenotypes can be grouped together in an unbiased way. The clusters could also reveal some correlations between phenotype and protein features. Several different clustering algorithms were compared to determine the method that would best produce phenotypically similar groups. The clustered data was then used to investigate correlations between phenotypes and protein features, protein phosphorylation data and RNA expression data. A novel correlation between aerial hyphae height and membrane-associated proteins was found. It was also determined that gene expression does not necessarily correlate with phenotype. Yeast

ortholog data for the *N. crassa* genes that co-clustered MAPK signaling cascades were analyzed, revealing additional potential networks of interacting proteins in *N. crassa*.

Aim 3: Determine the effects on gene expression and metabolism in Δ*gna-1*, Δ*gna-3* and Δ*ric8* mutants

As mentioned above, there is evidence that Gα subunits influence *N. crassa* metabolism (78). In addition, preliminary RNA-seq analysis of a Δ*ric8* mutant showed that >100 metabolic genes are mis-regulated (Arit Ghosh and Katherine Borkovich, unpublished). To obtain more understanding of the role of G-protein signaling in metabolite regulation, state-of-the-art mass spectrometry equipment for targeted polar metabolomics at the UCR Metabolomics Core Facility was utilized to analyze the metabolomes of Δ*gna-1*, Δ*gna-3* and Δ*ric8* mutants. A total of 120 metabolites were detected. All three gene deletion mutants have mis-regulated metabolites, with Δ*ric8* having the most abnormal metabolome. To determine whether control over metabolism is at the transcriptional level, RNA samples from Δ*ric8* and the two Gα mutants were submitted for RNA-seq analysis. Comparison of the results from the metabolomics and transcriptomics experiments will allow us to determine whether any mis-regulated metabolites are being controlled at a transcriptional or post-transcriptional level.

References

1. Kasbekar DP, McCluskey K. Neurospora genomics and molecular biology. Norfolk, UK: Caister Academic Press; 2013.

2. Beadle GW, Tatum EL. Genetic Control of Biochemical Reactions in Neurospora. Proc Natl Acad Sci U S A. 1941;27:499–506.

3. Stokes JL, Larsen A, Woodward CR, Foster JW. A NEUROSPORA ASSAY FOR PYRIDOXINE. Journal of Biological Chemistry. 1943;150:17–24.

4. Weiss RL. Intracellular localization of ornithine and arginine pools in Neurospora. J Biol Chem. 1973;248:5409–13.

5. Mishra NC. Genetics and biochemistry of morphogenesis in Neurospora. Adv Genet. 1977;19:341–405.

6. Loros JJ, Dunlap JC. Genetic and molecular analysis of circadian rhythms in Neurospora. Annu Rev Physiol. 2001;63:757–94.

7. Honda S, Eusebio-Cope A, Miyashita S, Yokoyama A, Aulia A, Shahi S, et al. Establishment of Neurospora crassa as a model organism for fungal virology. Nat Commun. 2020;11:5627.

8. Basenko EY, Pulman JA, Shanmugasundram A, Harb OS, Crouch K, Starns D, et al. FungiDB: An Integrated Bioinformatic Resource for Fungi and Oomycetes. J Fungi (Basel). 2018;4.

9. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus Neurospora crassa. Nature. 2003;422:859–68.

10. Roche CM, Loros JJ, McCluskey K, Glass NL. Neurospora crassa: looking back and looking forward at a model microbe. Am J Bot. 2014;101:2022–35.

11. Turner GE. Phenotypic analysis of Neurospora crassa gene deletion strains. Methods Mol Biol. 2011;722:191–8.

12. Springer ML. Genetic control of fungal differentiation: the three sporulation pathways of Neurospora crassa. Bioessays. 1993;15:365–74.

13. Glass NL, Rasmussen C, Roca MG, Read ND. Hyphal homing, fusion and mycelial interconnectedness. Trends Microbiol. 2004;12:135–41.

14. Pöggeler S, Kück U. Comparative analysis of the mating-type loci from Neurospora crassa and Sordaria macrospora: identification of novel transcribed ORFs. Mol Gen Genet. 2000;263:292–301.

15. Riquelme M, Yarden O, Bartnicki-Garcia S, Bowman B, Castro-Longoria E, Free SJ, et al. Architecture and development of the Neurospora crassa hypha -- a model cell for polarized growth. Fungal Biol. 2011;115:446–74.

16. Robertson NF. The mechanism of cellular extension and branching. New York: Academic Press; 1965.

17. Raju NB. Meiosis and ascospore genesis in Neurospora. Eur J Cell Biol. 1980;23:208–23.

18. Bistis GN. Chemotropic Interactions between Trichogynes and Conidia of Opposite Mating-Type in Neurospora crassa. Mycologia. 1981;73:959–75.

19. Read ND. A scanning electron microscopic study of the external features of perithecium development in Sordaria humana. Can J Bot. 1983;61:3217–29.

20. Raju NB. Genetic control of the sexual cycle in Neurospora. Mycological Research. 1992;96:241–62.

21. Read ND. Cellular nature and multicellular morphogenesis pf higher fungi. London: Academic Press; 1994.

22. Bistis GN, Perkins DD, Read ND. Different cell types in Neurospora crassa. Fungal Genetics Reports. 2003;50:17–9.

23. Roca MG, Arlt J, Jeffree CE, Read ND. Cell biology of conidial anastomosis tubes in Neurospora crassa. Eukaryot Cell. 2005;4:911–9.

24. Toledo I, Hansberg W. Protein oxidation related to morphogenesis inNeurospora crassa. Experimental Mycology. 1990;14:184–9.

25. Metzenberg RL. The Sexual Cycle in Neurospora. In: Stocchi V, Bonfante P, Nuti M, editors. Biotechnology of Ectomycorrhizae: Molecular Approaches. Boston, MA: Springer US; 1995. p. 85–98. doi:10.1007/978-1-4615-1889-1_8.

26. Nelson MA, Metzenberg RL. Sexual development genes of Neurospora crassa. Genetics. 1992;132:149–62.

27. Perkins DD, Radford A, Sachs MS. The Neurospora Compendium: Chromosomal Loci. Academic Press; 2001.

28. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, et al. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A. 2006;103:10352–7.

29. Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, et al. Enabling a Community to Dissect an Organism: Overview of the Neurospora Functional Genomics Project. Advances in Genetics. 2007;Volume 57:49–96.

30. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, et al. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev. 2004;68:1–108.

31. Park G, Servin JA, Turner GE, Altamirano L, Colot HV, Collopy P, et al. Global analysis of serine-threonine protein kinase genes in Neurospora crassa. Eukaryotic Cell. 2011;10:1553–64.

32. Ghosh A, Servin JA, Park G, Borkovich KA. Global analysis of serine/threonine and tyrosine protein phosphatase catalytic subunit genes in Neurospora crassa reveals interplay between phosphatases and the p38 mitogen-activated protein kinase. G3 (Bethesda). 2014;4:349–65.

33. Cabrera IE, Pacentine IV, Lim A, Guerrero N, Krystofova S, Li L, et al. Global Analysis of Predicted G Protein-Coupled Receptor Genes in the Filamentous Fungus, Neurospora crassa. G3 (Bethesda). 2015;5:2729–43.

34. Edelman AM, Blumenthal DK, Krebs EG. Protein serine/threonine kinases. Annu Rev Biochem. 1987;56:567–613.

35. Jain AK, Murty MN, Flynn PJ. Data clustering: a review. ACM Comput Surv. 1999;31:264–323.

36. Xu D, Tian Y. A Comprehensive Survey of Clustering Algorithms. Ann Data Sci. 2015;2:165–93.

37. Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, et al. A review of clustering techniques and developments. Neurocomputing. 2017;267:664–81.

38. Brown JA, Sherlock G, Myers CL, Burrows NM, Deng C, Wu HI, et al. Global analysis of gene function in yeast by quantitative phenotypic profiling. Mol Syst Biol. 2006;2:2006.0001.

39. Son H, Seo Y-S, Min K, Park AR, Lee J, Jin J-M, et al. A phenome-based functional analysis of transcription factors in the cereal head blight fungus, Fusarium graminearum. PLoS Pathog. 2011;7:e1002310.

40. Lee K-T, So Y-S, Yang D-H, Jung K-W, Choi J, Lee D-G, et al. Systematic functional analysis of kinases in the fungal pathogen Cryptococcus neoformans. Nat Commun. 2016;7:12766.

41. Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 2007;61:423–52.

42. Syrovatkina V, Alegre KO, Dey R, Huang X-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. J Mol Biol. 2016;428:3850–68.

43. Chakravorty D, Assmann SM. G protein subunit phosphorylation as a regulatory mechanism in heterotrimeric G protein signaling in mammals, yeast, and plants. Biochem J. 2018;475:3331–57.

44. Wilkie TM, Kinch L. New roles for Galpha and RGS proteins: communication continues despite pulling sisters apart. Curr Biol. 2005;15:R843-854.

45. Ross EM, Wilkie TM. GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins. Annu Rev Biochem. 2000;69:795–827.

46. Wright SJ, Inchausti R, Eaton CJ, Krystofova S, Borkovich KA. RIC8 is a guanine-nucleotide exchange factor for Galpha subunits that regulates growth and development in Neurospora crassa. Genetics. 2011;189:165–76.

47. Ivey FD, Kays AM, Borkovich KA. Shared and independent roles for a Galpha(i) protein and adenylyl cyclase in regulating development and stress responses in Neurospora crassa. Eukaryot Cell. 2002;1:634–42.

48. Kays AM, Borkovich KA. Severe impairment of growth and differentiation in a Neurospora crassa mutant lacking all heterotrimeric G alpha proteins. Genetics. 2004;166:1229–40.

49. Kays AM, Rowley PS, Baasiri RA, Borkovich KA. Regulation of conidiation and adenylyl cyclase levels by the Galpha protein GNA-3 in Neurospora crassa. Mol Cell Biol. 2000;20:7693–705.

50. Baasiri RA, Lu X, Rowley PS, Turner GE, Borkovich KA. Overlapping functions for two G protein alpha subunits in Neurospora crassa. Genetics. 1997;147:137–45.

51. Li L, Borkovich KA. GPR-4 is a predicted G-protein-coupled receptor required for carbon source-dependent asexual growth and development in Neurospora crassa. Eukaryot Cell. 2006;5:1287–300.

52. Yang Q, Poole SI, Borkovich KA. A G-protein beta subunit required for sexual and vegetative development and maintenance of normal G alpha protein levels in Neurospora crassa. Eukaryot Cell. 2002;1:378–90.

53. Krystofova S, Borkovich KA. The heterotrimeric G-protein subunits GNG-1 and GNB-1 form a Gbetagamma dimer required for normal female fertility, asexual

development, and galpha protein levels in Neurospora crassa. Eukaryot Cell. 2005;4:365–78.

54. McCahill A, Warwicker J, Bolger GB, Houslay MD, Yarwood SJ. The RACK1 scaffold protein: a dynamic cog in cell response mechanisms. Mol Pharmacol. 2002;62:1261–73.

55. Ron D, Jiang Z, Yao L, Vagts A, Diamond I, Gordon A. Coordinated movement of RACK1 with activated betaIIPKC. J Biol Chem. 1999;274:27039–46.

56. Adams DR, Ron D, Kiely PA. RACK1, A multifaceted scaffolding protein: Structure and function. Cell Commun Signal. 2011;9:22.

57. Gandin V, Senft D, Topisirovic I, Ronai ZA. RACK1 Function in Cell Motility and Protein Synthesis. Genes Cancer. 2013;4:369–77.

58. Zeller CE, Parnell SC, Dohlman HG. The RACK1 ortholog Asc1 functions as a G-protein beta subunit coupled to glucose responsiveness in yeast. J Biol Chem. 2007;282:25168–76.

59. Palmer DA, Thompson JK, Li L, Prat A, Wang P. Gib2, a novel Gbeta-like/RACK1 homolog, functions as a Gbeta subunit in cAMP signaling and is essential in Cryptococcus neoformans. J Biol Chem. 2006;281:32596–605.

60. Li G, Zhang X, Tian H, Choi Y-E, Tao WA, Xu J-R. MST50 is involved in multiple MAP kinase signaling pathways in Magnaporthe oryzae. Environ Microbiol. 2017;19:1959–74.

61. Yin Z, Zhang X, Wang J, Yang L, Feng W, Chen C, et al. MoMip11, a MoRgs7-interacting protein, functions as a scaffolding protein to regulate cAMP signaling and pathogenicity in the rice blast fungus Magnaporthe oryzae. Environ Microbiol. 2018;20:3168–85.

62. Kruger D, Koch J, Barthelmess IB. cpc-2, a new locus involved in general control of amino acid synthetic enzymes in Neurospora crassa. Curr Genet. 1990;18:211–5.

63. Tall GG, Gilman AG. Resistance to inhibitors of cholinesterase 8A catalyzes release of Galphai-GTP and nuclear mitotic apparatus protein (NuMA) from NuMA/LGN/Galphai-GDP complexes. Proc Natl Acad Sci U S A. 2005;102:16584–9.

64. Hampoelz B, Hoeller O, Bowman SK, Dunican D, Knoblich JA. Drosophila Ric-8 is essential for plasma-membrane localization of heterotrimeric G proteins. Nat Cell Biol. 2005;7:1099–105.

65. Gabay M, Pinter ME, Wright FA, Chan P, Murphy AJ, Valenzuela DM, et al. Ric-8 Proteins Are Molecular Chaperones That Direct Nascent G Protein α Subunit Membrane Association. Sci Signal. 2011;4:10.1126/scisignal.2002223.

66. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016;17:451–9.

67. Wishart DS. Metabolomics for Investigating Physiological and Pathophysiological Processes. Physiol Rev. 2019;99:1819–75.

68. Wellen KE, Hatzivassiliou G, Sachdeva UM, Bui TV, Cross JR, Thompson CB. ATP-citrate lyase links cellular metabolism to histone acetylation. Science. 2009;324:1076–80.

69. Nakahata Y, Kaluzova M, Grimaldi B, Sahar S, Hirayama J, Chen D, et al. The NAD+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. Cell. 2008;134:329–40.

70. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007;35 Database issue:D521-526.

71. Jewett MC, Hofmann G, Nielsen J. Fungal metabolite analysis in genomics and phenomics. Curr Opin Biotechnol. 2006;17:191–7.

72. Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. Metabolite profiling: from diagnostics to systems biology. Nat Rev Mol Cell Biol. 2004;5:763–9.

73. Zhang S, Nagana Gowda GA, Ye T, Raftery D. Advances in NMR-based biofluid analysis and metabolite profiling. Analyst. 2010;135:1490–8.

74. Theodoridis GA, Gika HG, Want EJ, Wilson ID. Liquid chromatography-mass spectrometry based global metabolite profiling: a review. Anal Chim Acta. 2012;711:7–16.

75. Mendes I, Sanchez I, Franco-Duarte R, Camarasa C, Schuller D, Dequin S, et al. Integrating transcriptomics and metabolomics for the analysis of the aroma profiles of Saccharomyces cerevisiae strains from diverse origins. BMC Genomics. 2017;18:455.

76. Zamani AI, Barig S, Ibrahim S, Mohd Yusof H, Ibrahim J, Low JYS, et al. Comparative metabolomics of Phialemonium curvatum as an omnipotent fungus cultivated on crude palm oil versus glucose. Microb Cell Fact. 2020;19:179.

77. Adpressa DA, Connolly LR, Konkel ZM, Neuhaus GF, Chang XL, Pierce BR, et al. A metabolomics-guided approach to discover Fusarium graminearum metabolites after removal of a repressive histone modification. Fungal Genet Biol. 2019;132:103256.

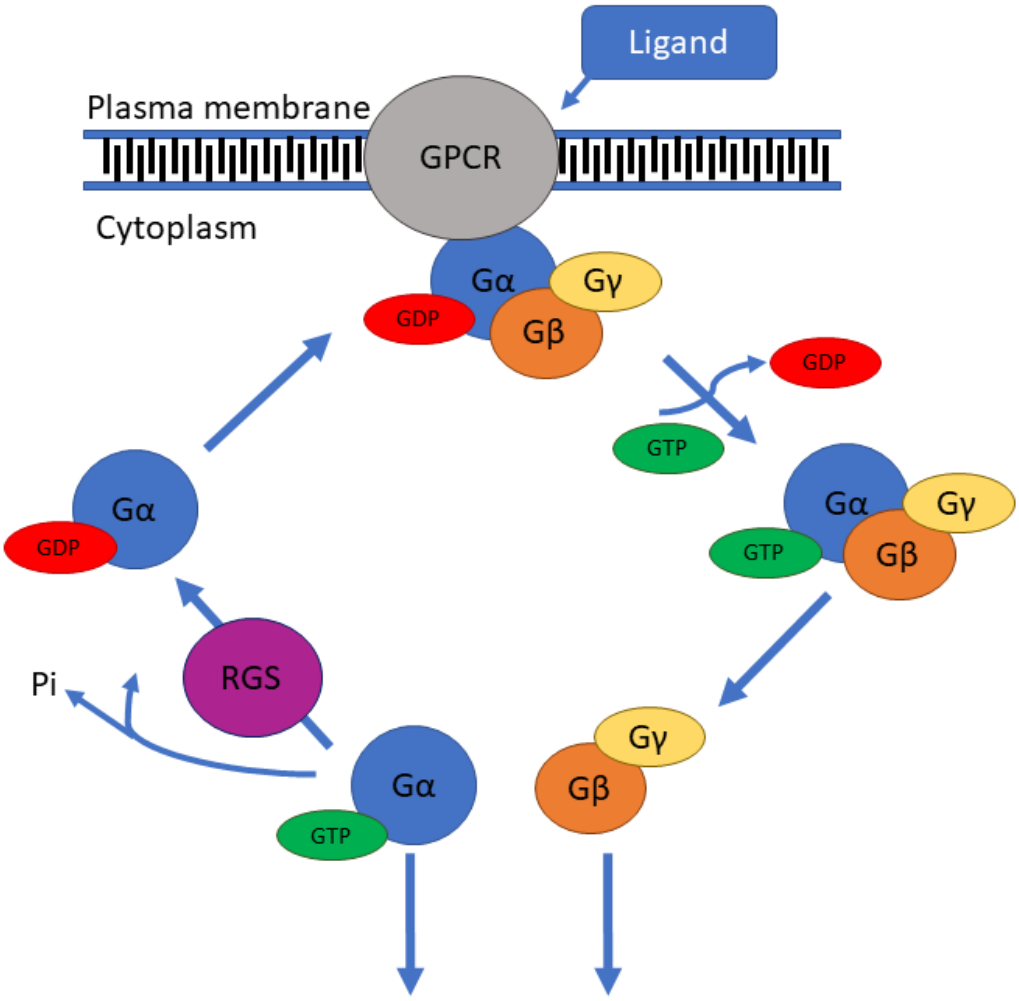
78. Kim JD, Kaiser K, Larive CK, Borkovich KA. Use of 1H nuclear magnetic resonance to measure intracellular metabolite levels during growth and asexual sporulation in Neurospora crassa. Eukaryotic Cell. 2011;10:820–31.

79. Mutlu N, Kumar A. Mapping paths: new approaches to dissect eukaryotic signaling circuitry. F1000Res. 2016;5:F1000 Faculty Rev-1853.

Figure legends

Figure 1.1 Heterotrimeric G protein cycle in N. crassa.

An extracellular ligand binds to the G protein coupled receptor (GPCR), facilitating exchange of GDP for GTP on the Gα subunit. The Gα bound to GTP then dissociates from the Gβγ dimer and both act on downstream effectors. Gα subunits have intrinsic GTPase activity which with hydrolyze the GTP to GDP or an RGS can stimulate the GTPase activity increasing hydrolyzation by up to 2000 times. The GDP-bound Gα will then reassociate with the Gβγ dimer and the GPCR


Ligand
Plasma membrane
GPCR
Cytoplasm
Gα
Gγ
GDP
Gβ
GDP
GTP
Gα
GDP
Gα
Gγ
GTP
Gβ
RGS
Pi
Gγ
Gβ
Gα
GTP

Chapter 2

Functional Profiling of Transcription Factor Genes in Neurospora *crassa*

Abstract

Regulation of gene expression by DNA-binding transcription factors is essential for proper control of growth and development in all organisms. In this study, we annotate and characterize growth and developmental phenotypes for transcription factor genes in the model filamentous fungus *Neurospora crassa* We identified 312 transcription factor genes, corresponding to 3.2% of the protein coding genes in the genome. The largest class was the fungal-specific Zn_2Cys_6 (C6) binuclear cluster, with 135 members, followed by the highly conserved C2H2 zinc finger group, with 61 genes. Viable knockout mutants were produced for 273 genes, and complete growth and developmental phenotypic data are available for 242 strains, with 64% possessing at least one defect. The most prominent defect observed was in growth of basal hyphae (43% of mutants analyzed), followed by asexual sporulation (38%), and the various stages of sexual development (19%). Two growth or developmental defects were observed for 21% of the mutants, while 8% were defective in all three major phenotypes tested. Analysis of available mRNA expression data for a time course of sexual development revealed mutants with sexual phenotypes that correlate with transcription factor transcript abundance in wild type. Inspection of this data also implicated cryptic roles in sexual development for several co-transcribed transcription factor genes that do not produce a phenotype when mutated.

Introduction

Transcriptional control by sequence-specific DNA-binding proteins is a major regulatory mechanism in all organism (1). It has been estimated that there are >90 types of transcription factors in eukaryotes (1). Major structural classes of eukaryotic transcription factors include zinc coordinating (C2H2 and C4 zinc fingers and Zn_2Cys_6/C6 zinc clusters), helix-turn-helix (HTH), β-scaffold, and proteins with basic domains (basic leucine zipper/bZIP and basic helix-loop-helix/bHLH) (1). Although zinc finger proteins predominate in eukaryotic genomes, the HTH group is the most widely conserved transcription factor group across evolution, as it comprises the majority of transcription factors in bacterial and archaeal genomes (1,2).

Early studies proposed that transcription factors were 0.5–8% of the gene content in the genome, and that the number of genes was roughly proportional to the size of the genome (3). In recent years, the widespread availability of genome sequences has made it possible to annotate the actual number of transcription factor genes from numerous eukaryotic species. There are estimated to be ∼19,000 protein-coding genes in humans (4). Various groups have annotated DNA-binding transcription factor genes over the years and identified >1300, with most discrepancies resulting from transcript variants for some genes (5,6). For example, Vaquerizas et al., 2009 annotated 1391 transcription factors, comprising 7.3% of total protein coding genes in the human genome. They found that three classes of transcription factor genes predominate in humans: C2H2 zinc finger (675 genes), homeodomain (257 genes), and bHLH (87 genes). A later study identified 1558 transcription factor genes in the human genome (8.2% of protein-coding genes),

comprising 111 families in 40 classes and 10 superclasses (6). C2H2 transcription factors were reported as the largest structural class at 53%, followed by HTH at 26%, and basic domain transcription factors at 11%. In the fruit fly Drosophila melanogaster, there are 708 genes out of 13,929 protein-coding genes (5%) that encode predicted transcription factors (7,8). C2H2 zinc fingers comprise the largest group with 255 members, followed by homeobox (101 genes), and HLH (58 genes) proteins. The model plant Arabidopsis thaliana has 1717 predicted transcription factors out of the 27,029 protein-coding genes in the genome (6.4% of genes) (9,10). These genes are organized in 58 families, with the largest classes being HLH (225 genes), MYB (168 genes), AP2/ERF (139 genes), the plant-specific NAC family (138 genes), bZIP (127 genes), and C2H2 zinc fingers (116 genes).

Fungi contain several types of transcription factors not found in animals or plants, including the Zn_2Cys_6 (C6) zinc cluster, APSES (*Neurospora crassa* ASM1, *Saccharomyces cerevisiae* Phd1p, *Aspergillus nidulans* StuA, *Candida albicans* Efg1, and *S. cerevisiae* Sok2p), copper fist, STE (sterile) and Velvet classes (11–14). In the model yeast S. cerevisiae, there are currently 301 predicted transcription factors out of the 6604 protein-coding genes in the genome [4.6% of genes; (14–16)]. The largest class of transcription factors in the yeast genome is the C6 zinc cluster (57 genes), followed by C2H2 zinc fingers (41 proteins), bZIP (15 proteins), homeodomain (12 proteins), GATA factors (10 proteins), and bHLH (eight proteins) (14).

N. crassa is multicellular fungus and eukaryotic model system that has been studied for >75 yr (17,18). A genome-wide project has resulted in a nearly complete gene

knockout mutant collection for the almost 10,000 genes in the genome (19–21). We had previously annotated 182 transcription factor genes in the *N. crassa* genome (17), and attempted mutation of 103 of these genes (19). We were able to isolate viable knockout mutants for 99 genes, and these were analyzed for growth and developmental phenotypes (19). The results demonstrated that 43% of the transcription factor mutants had at least one phenotype, with greater than half of these possessing multiple defects (19).

In this study, we have cataloged functions for the majority of transcription factors in a filamentous fungus. We annotate 130 more transcription factor genes in the genome, bringing the total number in *N. crassa* to 312. We combine the data from our earlier study Colot et al., 2006 with that for the additional genes, presenting phenotypes for a total of 242 available viable knockout mutants. We mine an RNAseq dataset for expression of the 312 genes during different phases of sexual development. Our results demonstrate that the majority of genes yield at least one phenotype, and that multiple transcription factors are required to control different aspects of growth and development in *N. crassa*.

Materials and Methods

Transcription factor gene annotation: Data from the 182 predicted transcription factor genes identified previously (17,19) were included in this study. This list was augmented with additional genes obtained through searches at the FungiDB or Broad Institute Neurospora Genome databases (*Neurospora crassa* Genome Project 2015; (22)), the CIS-BP database at http://cisbp.ccbr.utoronto.ca (23), and from a list that was generously shared with us by Luis Larrondo (Pontificia Catholic University of Chile, Santiago, Chile). All entries were evaluated in our laboratory using BLAST homology

searches (24), followed by analysis of PFAM domains using Interpro (25). Genes encoding domain(s) established to be DNA binding transcription factors, and with *E*-values $<10^{-5}$, were included in our final list. An exception to the DNA binding requirement was made for WD40 domain-containing transcriptional adapters, in light of their importance to growth, development, and environmental sensing in fungi.

Media and mutant construction: Vogel's minimal medium (VM) (26) was used to support asexual growth and development, while synthetic crossing medium (SCM) (27) was utilized to assess sexual development. Formation of tight colonies on plates was facilitated by growth on sorbose-containing medium plates (28). Hygromycin B (Calbiochem, San Diego, CA) was used at a concentration of 200 μg/ml in media where indicated. Inoculations were performed using macroconidia harvested from VM agar slants (28).

Wild-type strains ORS-SL6a (FGSC 4200; *mat a*) and 74-OR23-IVA (FGSC 2489; *mat A*) were obtained from the Fungal Genetics Stock Center (FGSC; Kansas State University, Manhattan, KS). Available transcription factor mutants (homokaryons or heterokaryons) were obtained from the FGSC, or produced in our laboratory. All *N. crassa* gene numbers are preceded by the prefix "NCU." We attempted purification of homokaryons from 28 heterokaryotic transformants that were available from the FGSC or our laboratory stocks using sexual crosses (as described in Colot *et al.*, 2006) or streaking of macroconidia in the vegetative phase (28). All putative homokaryons were checked for the presence of the knockout cassette using *hph* and gene flank-specific primers by polymerase chain reaction as described (29). Those strains purified through streaking of

macroconidia were also screened for the absence of the wild-type open reading frame using gene-specific primers. This step was not needed for the cross progeny, since the ascospore meiotic products are homokaryons. We were able to successfully purify 15 mutants using this approach. Two additional mutants (NCU07430 /*mad-1* and NCU09496) were purified using this method, but were inviable after storage. Another seven mutants were removed due to the observation in another study that the strains carry a secondary mutation that results in female sterility (30). In summary, a total of 70 mutants were not available for analysis at the time of this study, either due to failures in construction of knockout cassettes; changes in gene annotation that made the knockout mutant unusable, or resulted in the knockout mutant not being made; incorrect insertion of knockout cassettes in the genome; the inability to purify homokaryotic mutants from transformants, secondary mutations; or because the original purified knockout mutant was inviable after storage. This left us with 242 viable homokaryotic mutants for study.

Phenotypic analysis: The 242 viable mutants were analyzed for phenotypes using methods described in (31–34). Data were obtained from the Broad Institute *Neurospora* website (*Neurospora crassa* Genome Project 2015) or from experiments performed for this study. Some of the data at the Broad Institute website was previously published (35). We have omitted measurements of pigmentation and aerial hyphae height on yeast extract-containing medium from our analysis (31). Near-isogenic wild type strains FGSC4200 and/or FGSC2489 were used as a control for all determinations. Race tubes made of glass or prepared from disposable plastic pipets were

used for quantitative analysis of hyphal growth rates (36,37). Multiple race tubes were analyzed for each mutant, with a minimum of four independent replicates with $R^2 > 0.95$ used to obtain the average growth rate. The wild-type growth rate range for data obtained from Colot *et al.*, 2006 and the Broad Institute *Neurospora* Database was 70–85 mm/d, while that from this study was 75–85 mm/d. Binned data from Colot *et al.*, 2006 and/or the Broad Database were averaged to allow comparison to actual growth rate measurements.

Aerial hyphae height was assessed after incubation of 2-ml VM standing liquid cultures for 3 d in the dark at room temperature, with at least six replicates/strain. Mutants that displayed a growth rate <70 or >85 mm/d, or an aerial hyphae height measurement <30 or >45 mm (over 3 d), were considered significantly different than wild type. Macroconidia production was qualitatively assessed by visual inspection of VM agar slants incubated for 3 d in the dark at 25°, followed by 4 d in light at room temperature, with at least four replicates/strain. Three stages of sexual development were qualitatively analyzed using SCM agar slants cultured in constant light at room temperature: formation of protoperithecia after 7 d; development of perithecia from protoperithecia 7 d after fertilization with opposite mating type wild-type conidia; and ascospore shooting from mature perithecia 14 d after fertilization. The quantity and size of protoperithecia, perithecia, and ascospores were observed using a S8APO stereomicroscope with a DFC280 digital camera (Leica Microsystems, Buffalo Grove, IL), or an Olympus SZX9 stereomicroscope with a C-4040 digital camera (Olympus, Lake Success, NY). At least four replicates were analyzed for each strain.

Clustering of transcription factor gene expression data and heatmap generation: A sexual development time course RNAseq dataset (38) was mined for expression of the 284 transcription factor genes essentially as described (31). To allow visualization of the expression data, heat maps were produced using pheatmap (V1.0.2) (39) in R V3.1.1; (R Development Core Team 2014). The scaling function in pheatmap was used to normalize expression data, with the reads per kilobase of transcript per million mapped reads (RPKM) values at each time point for a given gene normalized to the RPKM at the time point with lowest expression for that gene.

Naming of transcription factor genes: In keeping with the *N. crassa* convention, and the naming used in our previous study (35), unnamed transcription factor genes whose mutations revealed phenotypes received names that reflected the defects. Strains showing abnormalities in all three phenotypes, basal hyphal extension during vegetative growth, asexual development, and sexual development are known as *all development altered* (*ada*). Strains showing altered hyphal growth and asexual development are named as *vegetative asexual development* (*vad*). Mutants with normal hyphal growth but altered sexual and vegetative development are referred to as *sexual and vegetative development* (*svd*), while those with defects in vegetative hyphal growth and sexual development are called *vegetative and sexual development* (*vsd*). Strains showing slower basal hyphal extension, but normal asexual and sexual development are known as *colonial* (*col*) or *slower growth rate* (*sgr*), while mutants with hyphal growth rates greater than wild type are named *faster growth rate* (*fgr*). Mutants with longer aerial hyphae than wild type are named *long aerial hyphae* (*lah*), while those with shorter aerial

hyphae are called *short aerial hyphae* (*sah*). Genes whose loss results in a block in female fertility are designated female fertility (*ff*), while those with more subtle sexual cycle defects are known as *defective sexual development* (*dsd*). Mutants that produce submerged perithecia are named as *sub*, while those with defective beaks are referred to as *bek*.

Results

Annotation of additional transcription factor genes in the *N. crassa* genome. In this study, we annotated an additional 130 transcription factor genes in *N. crassa*, bringing the total number to 312 (Table 2.1). This corresponds to 3.2% of the 9760 protein-coding genes in the genome (22), a proportion that is comparable to S. cerevisiae. We identified 25 classes of transcription factors with a single domain, and seven groupings with two distinct domains (Figure 2.1 and Table 2.1). The largest group of transcription factors is the C6 binuclear cluster, with 130 genes having this as the only domain (42% of the transcription factors), and another five genes possessing this domain in combination with a second motif (135 genes total; 43% of the transcription factors) (Table 2.1). This is more than twice the number identified in *S. cerevisiae* (57), despite both organisms having a similar total number of transcription factors (14).

The second largest group in *N. crassa* is the C2H2 zinc finger, with 54 genes possessing only C2H2, and seven also containing a second domain, for a total of 61 genes (20% of transcription factors)—a number greater than observed for *S. cerevisiae* [41 genes; (14)]. The third, fourth, and fifth largest groups of transcription factors in *N.*

crassa are the bZIP (23 genes; 7.4% of total), MYB (16 genes with single domain; one gene with two domains; 5.4% of total), and bHLH (13 genes; 4.2% of total) (Figure 2.1 and Table 2.1). *N. crassa* has significantly more single domain bZIP (23 vs. 15), bHLH (13 vs. 8), and MYB factors (16 vs. 6) (14). Finally, we identified a total of nine genes with two distinct domains, corresponding to 2.9% of the annotated transcription factors (Table 2.1). As noted above, the most common domain in the genes with two domains was the C2H2 zinc finger, with seven genes (Table 2.1).

Mutant production and analysis of hyphal growth phenotypes. All mutants used in this study were produced during the Neurospora Genome Project (20), which included a high throughput gene knockout project (20,40–42). Each mutant carries a gene replacement mutation, with insertion of a hygromycin resistance gene cassette in place of the open reading frame (19). There were instances where gene annotation changes or issues with primer, knockout cassette or strain construction rendered some mutants unusable, or unable to be produced (see Materials and Methods for details). In the end, we were able to assemble a group of 242 viable mutants for phenotypic analysis (78% of the total genes; Table 2.1). The mutants included homokaryons for the four putative essential genes described in our earlier study (19), which were isolated by streak-purifying in vegetative phase or by screening many ascospores: NCU00340/*pp-1*, *asl-1*/NCU01345, *ts* (formerly *asl-2*; NCU01459), and *cpc-1*/NCU04050.

We analyzed the 242 viable mutants for an array of growth and developmental phenotypes, beginning with the linear growth rate on minimal medium (19,43). *N.*

crassa grows by polar extension, branching, and anastomosis (fusion) of tube-like structures called hyphae to form the web-like multicellular structure termed the mycelium (reviewed in 1–4). Hyphae contain incomplete crosswalls (septa) that separate cellular compartments, but allow distribution of metabolites and organelles throughout the mycelium (45).

We assessed the hyphal growth rate of the mutants on minimal medium using race tubes (37,46). The results demonstrated that 105 mutants had growth rates significantly different from wild type (43% of viable mutants), and altered growth rate was the predominant phenotype in the transcription factor mutants (Figure 2.2 and Table 2.1). Consistent with its size, the transcription factor class (present alone or with a second domain) with the largest absolute number of mutants with hyphal growth defects was the C6 family (31/103 viable mutants; 30%; Table 2.1). The C2H2 group had a slightly higher proportion of mutants with hyphal growth phenotypes (19/51 mutants; 37%; Table 2.1). However, there were several transcription factor classes with 65–100% of the mutants exhibiting a hyphal growth phenotype. For example, 100% of the mutants with CP2, HSF, and WD40 domains had a hyphal growth defect. Similarly, 10/12 BHLH genes yielded phenotypic data, and nine of these had a growth rate phenotype (75% of viable mutants). For genes with the MYB domain, 13/17 were represented as viable mutants, and 8/13 (62%) had a hyphal growth defect (Table 2.1).

Less than half of the mutants possessed a growth rate significantly slower than wild type (range of 70–84 mm/d; Figure 2.3A). Several mutants displayed a growth rate within the 60–64 and 65–69 mm/d increments, while far fewer were found in the groups

with the slowest growth rates (Figure 2.3A). The <59 mm/d groups contained representation from several transcription factor classes, with loss of the WD40 gene, *rco-1*/NCU06205 (47), resulting in the slowest growth rate observed in our study (Figure 2.3A). Only two mutants grew faster than wild type; one lacking the C6 zinc finger gene *fgr-1*/NCU10597, and one deleted for the APSES gene *vsd-5*/NCU07587 (Figure 2.3A).

Asexual development defects. ***N. crassa*** produces two types of asexual spores, microconidia, and macroconidia (44,48,49). Microconidia are small, uninucleate, and relatively nonabundant, and are difficult to observe under laboratory conditions (44,49). In contrast, the multinucleated macroconidia are produced profusely in wild-type cultures (48). The macroconidiation (hereafter referred to as conidiation) pathway is regulated by oxygen/reactive oxygen species, carbon and nitrogen availability, high temperatures, blue light, and the circadian rhythm (44,48,50–52). Due to these multiple layers of regulation, wild type does not form conidia in submerged liquid cultures unless subjected to heat shock, carbon stress, or nitrogen stress (53–57). The conidiation pathway begins with adhesion of basal hyphae, followed by growth of aerial hyphae that rise perpendicular to the growth surface (48,50,44). Aerial hyphae form branches, with some transitioning from hyphal growth to apical budding as constrictions form between cellular compartments at the hyphal tip (48). With time, the constrictions tighten to separate the mature multinucleated conidia.

Several transcriptional regulatory proteins have been identified that regulate aspects of conidiation in *N. crassa*. For example, mutants lacking the transcription factor fluffy [*fl*/NCU08726; (58)] and the transcriptional adapter *rco-1*/NCU06205 (47)) are blocked at distinct points in the process. Other work has demonstrated transcription factors that influence conidiation under different environmental conditions or the extent of conidiation, including *vad-5*/NCU06799 (59), *chc-1*/NCU00749 (60) and *hsf-2*/NCU08480 (61). The white collar complex (WCC) containing the transcription factors *wc-1*/NCU02356 and *wc-2*/NCU00902 has been demonstrated to directly control expression of 24 transcription factor genes under blue light regulation (62,63).

In this study, we identified 92 transcription factor mutants (38% of total mutants) with a defect in aerial hyphae height and/or conidia production (Figure 2.2 and Table 2.1). Of these 92 strains, the majority (87; 95%) displayed an aerial hyphae phenotype, either singly (69 mutants; 75%) or in combination with a conidiation defect (18 mutants; 20%). A total of five mutants (5%) had a conidiation defect with normal aerial hyphae height. Similar to what was observed for hyphal growth rate, and in keeping with the large number of genes in this family, the C6 transcription factor class had the largest number of mutants with a phenotype (37; 36% of all mutants with C6 domain; Table 2.1). However, other domain classes had a larger proportion of viable mutants with an asexual development defect, including HSF (2/2; 100%), GATA (5/7; 71%), WD40 (2/3; 67%), HMG-box (2/3; 67%), and MYB (7/13; 54%) (Table 2.1).

Of the mutants with an aerial hyphae height defect, the majority had shorter aerial hyphae (60 mutants; 65%; Figure 2.3B). However, in comparison to the hyphal growth

rate analysis described above, there were more mutants with aerial hyphae height greater than wild type (27 mutants; 29%; Figure 2.3, A and B). The large C6 family dominated the mutants at both ends of the aerial hyphae height spectrum, while other groups appeared in only one section. For example, the Far1 (*tah-9*/NCU06551) and CP2 (*csp-2*/NCU06095) class transcription factor mutants were all taller than wild type, with the *csp-2* mutant the tallest in our study. The other end of the spectrum includes mutants lacking genes with two transcription factor domains: the C2H2 + STE mutant *pp-1*/NCU00340 , the ARID + MYB mutant *svd-1*/NCU04079 and the ARID + APSES mutant *ada-9*/NCU01238, which were all shorter than wild type. Similar to observations for hyphal growth rate, the WD40 mutant *rco-1*/NCU06205 had the shortest aerial hyphae overall.

Inspection of the qualitative data for conidia production in agar slants showed that a total of 23 mutants was affected. Three mutants were increased (lacking the C2H2 *cre-1*/NCU08807, C2H2+STE *pp-1*/NCU00340, and MYB *ada-22*/NCU08003), and 16 reduced, relative to wild type. Three mutants (lacking the C6 gene *fl*/NCU08726, the BZIP *ada-1*/NCU00499, and the ARID/BRIGHT *ada-20*/NCU05891) did not form any conidia, and the *kal-1*/NCU03593 homeodomain mutant had an abnormal conidiation pattern (Colot et al. 2006). The mutant lacking the MYB gene *rcn-1*/NCU07834 had a conidiation defect as its only phenotype.

Sexual development phenotypes. *N. crassa* is a hermaphrodite, in that a single colony produces female and male gametes. However, *N. crassa* is not self-fertile, requiring that

the female and male cells be of opposite mating type [*mat* A and *mat* a; (64)]. Nitrogen starvation induces differentiation of female reproductive structures (protoperithecia) (64) from the basal hyphae of the mycelium (see wild-type image in Figure 2.4B). Protoperithecial development begins with coiling, extension, adhesion, and then septation and branching of hyphae to form a coil (65). Enveloping hyphae encircle the coil, and then grow and branch to form the protoperithecium. Mating occurs when male-receptive hyphae (trichogynes) from the mature female structure grows toward a male (typically an asexual spore; macro- or microconidium) of opposite mating type in a process involving a pheromone response (66). After fusion of the trichogyne and conidium, the nuclei from the male and female divide synchronously to form the ascogenous hyphae within the developing perithecium (64) (see wild-type image in Figure 2.4B). Nuclear fusion and meiosis takes place in a specialized cell type (crozier). The perithecium enlarges, melanizes, and forms a beak at the tip as the meiotic progeny (ascospores) mature within. Ascospores are then forcibly ejected from a hole (ostiole) in the tip of the beak (64). The entire sexual cycle can be completed within 2.5 wk (28). Blue light is an important environmental cue for sexual development, regulating the abundance of protoperithecia.

In the budding yeast *S. cerevisiae*, only a few genes necessary for sexual development and meiosis are transcription factors (67–71). Our earlier study revealed 15 transcription factor genes that regulate aspects of sexual development in *N. crassa*, and mutation of 13 of these genes resulted in a complete block in ascospore production (19). Work from other laboratories has identified genes encoding the transcription factors *asm*-

1/NCU01414 (72), *asd-4*/NCU15829 (73), *pp-1*/NCU00340 (74), and tan spore [*ts*/NCU01459; (75)] all of which influence sexual development.

The collection of 242 viable transcription factor mutants was screened for the number, size, and other properties of protoperithecia (Figure 2.4A). For perithecia, the mutants were assayed for abundance, relative size, and the presence of beaks. Ascospore ejection was scored by visual inspection of spores on the sides of the glass culture tube. We identified 47 mutants with at least one type of sexual cycle phenotype, corresponding to 19% of the viable mutants (Figure 2.1, Figure 2.4A, and Table 2.1). The C2H2 class had the largest number of genes with a sexual cycle phenotype (12 mutants; Table 2.1). Other transcription factor classes with a large proportion of mutants possessing a sexual development defect include WD40 (2/3 with phenotypes; 67%) and MYB (8/13 with phenotypes; 62%; Table 2.1).

Of the 47 mutants with a defect, 26 exhibited a complete block at a step in sexual development, with no development of ascospores (Figure 2.4A). A total of nine mutants were completely blocked in protoperithecial development (Figure 2.4A). Another 10 mutants produced protoperithecia, but no perithecia or ascospores, for a total of 19 mutants that did not form perithecia (Figure 2.4A). An example of a mutant that produces protoperithecia, but no perithecia, is *fmf-1*/NCU09387 (Figure 2.4B). Seven mutants elaborated protoperithecia and perithecia, but no ascospores (Figure 2.4A). In addition to those with a complete block at a step in sexual development, we identified numerous mutants with a reduction in the relative quantity of sexual structures (Figure 2.4A). There were 16 mutants affected in the number of protoperithecia, 12 in perithecial production,

and 15 with fewer ascospores (Figure 2.4A). A majority of ARID (2/3; 67%), WD40 (2/3; 67%), and MYB (8/13; 62%) class mutants are affected in sexual development (Table 2.1).

Several mutants displayed more unique defects during the sexual cycle. Similar to previous results, the *ts*/NCU01459 mutant (75) produced tan ascospores. The GATA mutant sub-1/NCU01154 produced protoperithecia that were submerged in the agar. Six different transcription factor families contribute to proper perithecial beak development; the homeodomain mutant *bek-1*/NCU00097 (76), the C6 mutant *bek-2*/NCU07139 , the BHLH mutant *vsd-3*/NCU08999 and the BZIP mutant *ada-1*/NCU00499 completely lacked perithecial beaks, while another four mutants had a reduced number of perithecial beaks (C2H2 genes *cre-1*/NCU08807 and *vsd-9*/NCU07952, BHLH *dsd-3*/NCU05970, and APSES *vsd-5*/NCU07587). The beak defect of *vsd-5* is presented in Figure 2.4B.

Correlation between phenotypes and gene expression during sexual development. We took advantage of a publicly available dataset for a time course during perithecial and ascospore development (38) to investigate expression of transcription factor genes. This dataset includes eight time points (0, 2, 24, 48, 72, 96, 120, and 144 hr) after fertilization with opposite mating type wild type. Time = 0 corresponds to nitrogen-starved vegetative hyphae and unfertilized protoperithecia just before application of wild type conidia. Perithecia are obvious at the 24 hr time point, while croziers (evidence of meiosis) appear at 48–72 hr. Asci containing the meiotic progeny are formed after 96 hr, and perithecial beaks at 120–144 hr.

We first interrogated the data for transcription factor genes that yielded a sexual cycle phenotype (Figure 2.4C). There were several genes for which phenotypes and gene expression showed a correlation (Figure 2.4C). For example, Group 1 genes, with expression peaking early from 0 to 2 hr, included 6/9 mutants with defects in protoperithecial development, and 100% with perithecial phenotypes. The expression pattern and high proportion of perithecial defects suggests roles in fertilization for this group. Group 3 genes peak at 48 hr (time of meiosis), with 100% of the mutants possessing an ascospore defect. Group 4 genes are highly expressed from 72 to 144 hr and 4/5 of the corresponding mutants have a defect in perithecial formation and ascospore development/shooting. The large Group 6 is generally highly expressed late (120–144 hr), with 11/14 and 12/14 of the mutants exhibiting defects in perithecial and ascospore development, respectively. The more heterogeneously expressed Group 2 and 5 genes did not follow strict expression patterns, but all but one gene yielded a protoperithecial phenotype.

We next mined the expression data to identify co-transcribed genes in the same class of transcription factors (Figure 2.5). This analysis revealed several instances of similarly expressed genes where one produced a moderate (nonblocking) sexual cycle phenotype, and the second a nonblocking defect or no phenotype, suggesting possible redundancy. The MYB genes *rca-1*/NCU01312 (no sexual phenotype) and *svd-4*/NCU09329 (severely reduced ascospores) are expressed at lower levels from 0 to 48 hr, with steady, elevated expression from 72 to 144 hr (Figure 2.5A), suggesting redundant roles late during sexual development. The C2H2 genes *svd-3*/NCU06487 (reduced

number of ascospores) and *vad-12*/NCU05035 (no sexual phenotype) are highly expressed at 120–144 hr (Figure 2.5B), the time of ascospore production. The C2H2 genes NCU02487 (no sexual phenotype) and *svd-3*/NCU06487 (reduced number of ascospores) are highly expressed in unfertilized protoperithecia (t = 0), and from 120 to 144 hr, when ascospores are produced (Figure 2.5B). The C2H2 genes *vsd-9*/NCU07952 (reduced ascospores) and NCU03699 (no sexual phenotype) are coordinately expressed, with a dip at 48 hr (time of meiosis) and 144 hr (ascospore maturation) (Figure 2.5B). The C6 class genes *vsd-4*/NCU01243 (reduced number ascosopores) and *acu-15*/NCU06656 (normal sexual development) are both highly expressed in protoperithecia, and then peak again from 24 to 48 hr, at the time of meiosis (Figure 2.5C).

Associations between phenotypes during different phases of the lifecycle. We identified a total of 154 mutants with at least one phenotype during growth or development (64% of viable mutants; Table 2.1). Inspection of mutants with two defects revealed 34 with growth and asexual development phenotypes, 12 with growth and sexual development defects, and six with phenotypes in asexual and sexual development, totaling 21% of the viable mutants (Figure 2.2 and Table 2.1). There were 19 strains with defects in all three stages assayed (growth, asexual development, and sexual development), comprising 4.1% of the mutants (Figure 2.2 and Table 2.1). Of interest, the majority of mutants with a conidiation defect also exhibited reduced hyphal growth rate (27/92 mutants; 82%). A lesser proportion also possessed sexual cycle defects (17/28

mutants; 61%). These observations suggest a possible linkage between the ability to produce conidia and growth rate and/or sexual development.

As our quantitative data for growth rate and aerial hyphae height were both obtained using vegetative (asexual) cultures, we investigated a possible relationship between these two phenotypes in our entire population of mutants (Figure 2.3C). Visual inspection of the plot did not reveal a clear correlation, and regression analysis of a line drawn through the global data had an R^2 value of only 0.165, not supporting a linear relationship between growth rate and aerial hyphae height. When we restricted the data to mutants with a defect in at least one trait, the R^2 value dropped to 0.144 (data not shown). Therefore, the growth rate of basal hyphae and aerial hyphae height appear to be independent traits in our group of transcription factor mutants.

We also investigated phenotypes for the group of 24 WCC-regulated transcription factor genes (62). We were most interested in whether two major phenotypes that are known to be light regulated—asexual and sexual development—were prevalent in this group of genes. Three of the genes did not pass our requirements to be classified as transcription factors (NCU00275, NCU06534, and NCU07846). Of the remaining 21 genes, we have complete phenotypic data for 18 mutants, while phenotypes for another mutant (*ve-1*/NCU05964) have been published (77). Nine of these 19 mutants have defects in asexual sporulation, two have both asexual and sexual sporulation phenotypes, and three mutants have only sexual cycle defects, for a total of 14 mutants (74%) with a sporulation phenotype (Figure 2.6A). Of interest, two of the mutants that lack phenotypes in asexual or sexual sporulation (NCU07705 /*clr-1* and NCU08000) are similarly

expressed during sexual development (Figure 2.6B), suggesting possible redundancy during this process. Taken together, our results mesh well with the findings of the earlier study (62), and support these transcription factors as potential second-level regulators in the blue light gene expression hierarchy that regulates asexual and sexual development in *N. crassa*.

Discussion

We have analyzed phenotypes in available mutants for annotated transcription factor genes in *N. crassa*. Overall, 64% of the mutants analyzed possessed at least one growth or developmental phenotype. The largest proportion of mutants exhibited defects in hyphal growth rate, either alone or in combination with another defect. This result underscores the importance of transcriptional regulation to the various steps of hyphal growth, including spore germination, hyphal growth and polarity, branching, and anastomosis. The latter two phenomena are often accentuated in slow-growing mutants (so-called colonials; (78,79)) and this may contribute to the slow growth rate of some of the mutants in our study.

Inspection of data for individual transcription factor classes revealed several with a large number of genes with defects and/or skewing toward specific defects. For transcription factor classes represented by more than two mutants, 100% of the ARID/BRIGHT, NDT80 and WD40 mutants possessed phenotypes. The majority (83%) of GATA factor mutants possessed defects in asexual development, while 75% of the BHLH and 62% of the MYB domain mutants had a growth rate phenotype. There was

less concentration of sexual development defects in certain classes, with WD40 (67%) and MYB (58%) having the largest proportion of mutants with sexual cycle phenotypes.

Comparison of our current data with our original analysis of 99 viable transcription factor mutants revealed several phenotypes that were greatly under-represented in the earlier study. For example, there has been a fivefold increase in mutants with only a basal hyphae growth rate phenotype, and a fourfold increase in mutants with defects in both growth rate and asexual development. There were zero and one mutants with hyphal growth/sexual development and asexual development/sexual development phenotypes, respectively (19); there are now 12 mutants in the first group and six mutants in the second (Figure 2.2). In general, the more than doubling of the number of mutants analyzed has resulted in discovery of more genes that affect basal hyphae growth and asexual development. This difference also contributes to the increase in the relative number of genes that yield phenotypes, from 40% (19) to 64% (this study).

We have previously applied our phenotypic screening platform to other large gene families in *N. crassa*, including serine-threonine protein kinases [77 mutants; (33)], serine-threonine protein phosphatases [24 mutants; (29)] and G protein coupled receptors [GPCRs; 36 mutants; (31)]. The percentage of transcription factor mutants with at least one growth/developmental phenotype (64%) is less than that for the protein phosphatases (91%), but greater than that observed for GPCRs (47%) and protein kinases (57%). Comparison between these groups reveals that the proportion of mutants with hyphal growth rate defects is relatively low for GPCRs (14%), but similar for the protein kinase (42%), transcription factor (43%), and protein phosphatase (50%) genes. In the case of

asexual sporulation, the fraction of affected mutants was identical/near identical for transcription factors (38%), GPCRs (39%) and protein kinases (40%), but significantly higher for protein phosphatases (58%). For sexual development, GPCRs and transcription factors had a much lower proportion of genes with phenotypes (17 and 19%, respectively) than the protein kinases (42%) and protein phosphatases (63%). There were also striking differences in the percentage of mutants with defects in all three categories among the four groups, from GPCRs with no such mutants, to 8% of transcription factor, 26% of protein kinase, and 29% of protein phosphatase mutants. The lower number of transcription factor and GPCR mutants with sexual development defects or phenotypes in all three categories analyzed may reflect greater gene redundancy in transcription factors and GPCRs. We hypothesize that there may be greater functional redundancy in genes at the opposite ends of the environmental sensing spectrum (receptors and transcription factors), than with those more involved in signal transduction and integration (kinases and phosphatases). This may be more obvious during sexual development due to the large number of cell types involved (49), and the need to coordinate cell morphogenesis with meiosis. Alternatively, the greater apparent redundancy may reflect functions for receptors and transcription factors that are not currently being analyzed in our phenotypic assays. Testing of these and other alternative hypotheses will require further investigation.

Transcription factor genes have been annotated in several other filamentous fungal species. In the genus *Trichoderma*, the biological control agents and mycoparasites *Trichoderma virens* and *Trichoderma atroviride*, and the efficient

cellulose degrader and industrial species *T. reesei* have 641, 592, and 448 transcription factor genes, respectively (80) (5.1, 5.0, and 4.9% of the genes in each genome). The two mycoparasitic species exhibited a large expansion in the C6 zinc cluster transcription factor family, with 382 genes for *T. atroviride* and 422 for *T. virens*, compared to 258 in *T. reesei*. In *Aspergillus nidulans*, there are 490 annotated transcription factors, corresponding to 4.6% of the predicted genes (81). Similar to *Trichoderma* species, *A. nidulans* has significantly more C6 genes than *N. crassa* [330 genes; (81)]. A more recent study reported the number of C6 zinc cluster transcription factor genes in three Aspergilli species, with 180 in A. clavatus, 276 in *A. nidulans*, and 306 in *A. flavus* (82). In contrast to C6 proteins, the number of C2H2 proteins reported for other species is similar to, or even less than in, *N. crassa*, with 60 in *A. nidulans* (81), 61 in *T. virens*, 53 in *T. atroviride*, and 49 in *T. reesei* (80). Thus, the larger number of C6 proteins in these other fungi is not due to a proportional increase in all transcription factor families. Rather, it appears that the expansion of the fungal-specific C6 class is a major contributor to the greater number of transcription factors overall in *Trichoderma* species and *A. nidulans* relative to *N. crassa*. The reduced size of the C6 class in *N. crassa* may result, at least in part, from Repeat-Induced Point Mutation (RIP), a mechanism that mutates duplicated DNA sequences during the sexual cycle and that has been hypothesized to limit the size of gene families in the *N. crassa* genome (83,84).

In this study, we investigated functions for 242 predicted transcription factor genes during growth and asexual and sexual development in *N. crassa*. Considering the significant sample size and the accompanying penchant for gene redundancy, our results

revealed a surprisingly large proportion of mutants with at least one phenotype. We identified numerous genes that lack sexual cycle phenotypes that are coordinately expressed with other genes with nonblocking phenotypes during sexual development. This suggests that these genes possess overlapping functions during sexual differentiation, a hypothesis that can be addressed in future experiments through construction of mutants lacking multiple transcription factor genes. This work has augmented our knowledge of the functions of transcription factors during growth and development in filamentous fungi.

References

1. A Catalogue of Eukaryotic Transcription Factor Types, Their Evolutionary Origin, and Species Distribution - PubMed. Available at: https://pubmed.ncbi.nlm.nih.gov/21557078-a-catalogue-of-eukaryotic-transcription-factor-types-their-evolutionary-origin-and-species-distribution/ [Accessed December 4, 2019]

2. The Many Faces of the Helix-Turn-Helix Domain: Transcription Regulation and Beyond - PubMed. Available at: https://pubmed.ncbi.nlm.nih.gov/15808743-the-many-faces-of-the-helix-turn-helix-domain-transcription-regulation-and-beyond/ [Accessed December 4, 2019]

3. Transcription Regulation and Animal Diversity - PubMed. Available at: https://pubmed.ncbi.nlm.nih.gov/12853946-transcription-regulation-and-animal-diversity/ [Accessed December 4, 2019]

4. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, Vazquez J, Valencia A, Tress ML. Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum Mol Genet* (2014) **23**:5866–5878. doi:10.1093/hmg/ddu309

5. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* (2009) **10**:252–263. doi:10.1038/nrg2538

6. Wingender E, Schoeps T, Dönitz J. TFClass: an expandable hierarchical classification of human transcription factors. *Nucleic Acids Res* (2013) **41**:D165-170. doi:10.1093/nar/gks1123

7. Gramates LS, Marygold SJ, Santos GD, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al. FlyBase at 25: looking to the future. *Nucleic Acids Res* (2017) **45**:D663–D671. doi:10.1093/nar/gkw1016

8. Hammonds AS, Bristow CA, Fisher WW, Weiszmann R, Wu S, Hartenstein V, Kellis M, Yu B, Frise E, Celniker SE. Spatial expression of transcription factors in Drosophila embryonic organ development. *Genome Biol* (2013) **14**:R140. doi:10.1186/gb-2013-14-12-r140

9. Jin J, He K, Tang X, Li Z, Lv L, Zhao Y, Luo J, Gao G. An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol Biol Evol* (2015) **32**:1767–1773. doi:10.1093/molbev/msv058

10. Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, et al. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res* (2008) **36**:D1009-1014. doi:10.1093/nar/gkm965

11. Iyer LM, Koonin EV, Aravind L. Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSES transcription factors. *Genome Biol* (2002) **3**:RESEARCH0012. doi:10.1186/gb-2002-3-3-research0012

12. Weirauch MT, Hughes TR. A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. *Subcell Biochem* (2011) **52**:25–73. doi:10.1007/978-90-481-9069-0_3

13. Ahmed YL, Gerke J, Park H-S, Bayram Ö, Neumann P, Ni M, Dickmanns A, Kim SC, Yu J-H, Braus GH, et al. The velvet family of fungal regulators contains a DNA-binding domain structurally similar to NF-κB. *PLoS Biol* (2013) **11**:e1001750. doi:10.1371/journal.pbio.1001750

14. Hughes TR, de Boer CG. Mapping yeast transcriptional networks. *Genetics* (2013) **195**:9–36. doi:10.1534/genetics.113.153262

15. de Boer CG, Hughes TR. YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. *Nucleic Acids Res* (2012) **40**:D169-179. doi:10.1093/nar/gkr993

16. Engel SR, Dietrich FS, Fisk DG, Binkley G, Balakrishnan R, Costanzo MC, Dwight SS, Hitz BC, Karra K, Nash RS, et al. The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3 Bethesda Md* (2014) **4**:389–398. doi:10.1534/g3.113.008995

17. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, et al. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. *Microbiol Mol Biol Rev MMBR* (2004) **68**:1–108. doi:10.1128/mmbr.68.1.1-108.2004

18. Selker EU. Neurospora. *Curr Biol CB* (2011) **21**:R139-140. doi:10.1016/j.cub.2011.01.006

19. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. *Proc Natl Acad Sci U S A* (2006) **103**:10352–10357. doi:10.1073/pnas.0601456103

20. Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, McCluskey K, Plamann M, Galagan JE, Birren BW, et al. Enabling a community to dissect an

organism: overview of the Neurospora functional genomics project. *Adv Genet* (2007) **57**:49–96. doi:10.1016/S0065-2660(06)57002-6

21. Park G, Colot HV, Collopy PD, Krystofova S, Crew C, Ringelberg C, Litvinkova L, Altamirano L, Li L, Curilla S, et al. High-throughput production of gene replacement mutants in Neurospora crassa. *Methods Mol Biol Clifton NJ* (2011) **722**:179–189. doi:10.1007/978-1-61779-040-9_13

22. Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, Kissinger JC, Li W, Nayak V, Pinney DF, et al. FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Res* (2012) **40**:D675-681. doi:10.1093/nar/gkr918

23. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* (2014) **158**:1431–1443. doi:10.1016/j.cell.2014.08.009

24. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* (1990) **215**:403–410. doi:10.1016/S0022-2836(05)80360-2

25. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, et al. InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* (2012) **40**:D306-312. doi:10.1093/nar/gkr948

26. Vogel HJ. Distribution of lysine pathways among fungi: Evolutionary implications. *Am Nat* (1964) **98**:435–446.

27. Westergaard M, Mitchell HK. Neurospora V. A synthetic medium favoring sexual reproduction. *Amer J Bot* (1947) **34**:573–577.

28. Davis RH, deSerres FJ. Genetic and microbiological research techniques for Neurospora crassa. *Methods Enzym* (1970) **71A**:79–143.

29. Ghosh A, Servin JA, Park G, Borkovich KA. Global analysis of serine/threonine and tyrosine protein phosphatase catalytic subunit genes in Neurospora crassa reveals interplay between phosphatases and the p38 mitogen-activated protein kinase. *G3 Bethesda Md* (2014) **4**:349–365. doi:10.1534/g3.113.008813

30. Fu C, Ao J, Dettmann A, Seiler S, Free SJ. Characterization of the Neurospora crassa cell fusion proteins, HAM-6, HAM-7, HAM-8, HAM-9, HAM-10, AMPH-1 and WHI-2. *PLoS One* (2014) **9**:e107773. doi:10.1371/journal.pone.0107773

31. Cabrera IE, Pacentine IV, Lim A, Guerrero N, Krystofova S, Li L, Michkov AV, Servin JA, Ahrendt SR, Carrillo AJ, et al. Global Analysis of Predicted G Protein-

Coupled Receptor Genes in the Filamentous Fungus, Neurospora crassa. *G3 Bethesda Md* (2015) **5**:2729–2743. doi:10.1534/g3.115.020974

32. Ghosh A, Servin JA, Park G, Borkovich KA. Global analysis of serine/threonine and tyrosine protein phosphatase catalytic subunit genes in Neurospora crassa reveals interplay between phosphatases and the p38 mitogen-activated protein kinase. *G3 Bethesda* (2014) **4**:349–65. doi:10.1534/g3.113.008813

33. Park G, Servin JA, Turner GE, Altamirano L, Colot HV, Collopy P, Litvinkova L, Li L, Jones CA, Diala F-G, et al. Global analysis of serine-threonine protein kinase genes in Neurospora crassa. *Eukaryot Cell* (2011) **10**:1553–1564. doi:10.1128/EC.05140-11

34. Turner GE. Phenotypic analysis of Neurospora crassa gene deletion strains. *Methods Mol Biol* (2011) **722**:191–8. doi:10.1007/978-1-61779-040-9_14

35. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. *Proc Natl Acad Sci U A* (2006) **103**:10352–7. doi:10.1073/pnas.0601456103

36. Dharmananda S, Feldman JF. Spatial Distribution of Circadian Clock Phase in Aging Cultures of Neurospora crassa. *Plant Physiol* (1979) **63**:1049–54.

37. White B, Woodward D. A simple method for making disposable race tubes. *Fungal Genet Rep* (1995) **42**:79. doi:10.4148/1941-4765.1357

38. Wang Z, Lopez-Giraldez F, Lehr N, Farré M, Common R, Trail F, Townsend JP. Global Gene Expression and Focused Knockout Analysis Reveals Genes Associated with Fungal Fruiting Body Development in Neurospora crassa. *Eukaryot Cell* (2014) **13**:154–169. doi:10.1128/EC.00248-13

39. Kolde R. *pheatmap: Pretty Heatmaps*. (2019). Available at: https://CRAN.R-project.org/package=pheatmap [Accessed December 6, 2019]

40. Colot, Hildur VPark, GyungsoonTurner, Gloria ERingelberg, CarolCrew, Christopher MLitvinkova, LiubovWeiss, Richard LBork.

41. Park G, Servin JA, Turner GE, Altamirano L, Colot HV, Collopy P, Litvinkova L, Li L, Jones CA, Diala F-G, et al. Global analysis of serine-threonine protein kinase genes in Neurospora crassa. *Eukaryot Cell* (2011) **10**:1553–1564. doi:10.1128/EC.05140-11

42. Park G, Colot HV, Collopy PD, Krystofova S, Crew C, Ringelberg C, Litvinkova L, Altamirano L, Li L, Curilla S, et al. High-throughput production of gene replacement mutants in Neurospora crassa. *Methods Mol Biol* (2011) **722**:179–89. doi:10.1007/978-1-61779-040-9_13

43. Turner GE. Phenotypic analysis of Neurospora crassa gene deletion strains. *Methods Mol Biol Clifton NJ* (2011) **722**:191–198. doi:10.1007/978-1-61779-040-9_14

44. Springer ML. Genetic control of fungal differentiation: the three sporulation pathways of Neurospora crassa. *BioEssays News Rev Mol Cell Dev Biol* (1993) **15**:365–374. doi:10.1002/bies.950150602

45. Glass NL, Rasmussen C, Roca MG, Read ND. Hyphal homing, fusion and mycelial interconnectedness. *Trends Microbiol* (2004) **12**:135–141. doi:10.1016/j.tim.2004.01.007

46. Dharmananda S, Feldman JF. Spatial Distribution of Circadian Clock Phase in Aging Cultures of Neurospora crassa. *Plant Physiol* (1979) **63**:1049–1054. doi:10.1104/pp.63.6.1049

47. Yamashiro CT, Ebbole DJ, Lee BU, Brown RE, Bourland C, Madi L, Yanofsky C. Characterization of rco-1 of Neurospora crassa, a pleiotropic gene affecting growth and development that encodes a homolog of Tup1 of Saccharomyces cerevisiae. *Mol Cell Biol* (1996) **16**:6218–6228. doi:10.1128/mcb.16.11.6218

48. Springer ML, Yanofsky C. A morphological and genetic analysis of conidiophore development in Neurospora crassa. *Genes Dev* (1989) **3**:559–571. doi:10.1101/gad.3.4.559

49. Bistis GN, Perkins DD, Read ND. Different cell types in Neurospora crassa. *Fungal Genet Rep* (2003) **50**:17–19. doi:10.4148/1941-4765.1154

50. Hansberg W, de Groot H, Sies H. Reactive oxygen species associated with cell differentiation in Neurospora crassa. *Free Radic Biol Med* (1993) **14**:287–293. doi:10.1016/0891-5849(93)90025-p

51. Greenwald CJ, Kasuga T, Glass NL, Shaw BD, Ebbole DJ, Wilkinson HH. Temporal and spatial regulation of gene expression during asexual development of Neurospora crassa. *Genetics* (2010) **186**:1217–1230. doi:10.1534/genetics.110.121780

52. Baker CL, Loros JJ, Dunlap JC. The circadian clock of Neurospora crassa. *FEMS Microbiol Rev* (2012) **36**:95–110. doi:10.1111/j.1574-6976.2011.00288.x

53. Cortat M, Turian G. Conidiation of Neurospora crassa in submerged culture without mycelial phase. *Arch Mikrobiol* (1974) **95**:305–309. doi:10.1007/bf02451771

54. That TC, Turian G. Ultrastructural study of microcyclic macroconidiation in Neurospora crassa. *Arch Microbiol* (1978) **116**:279–288. doi:10.1007/bf00417852

55. Plesofsky-Vig N, Light D, Brambl R. Paedogenetic conidiation inNeurospora crassa. *Exp Mycol* (1983) **7**:283–286. doi:10.1016/0147-5975(83)90049-X

56. Guignard R, Grange F, Turian G. Microcycle conidiation induced by partial nitrogen deprivation in Neurospora crassa. *Can J Microbiol* (2011) **30**:1210–1215. doi:10.1139/m84-192

57. Madi L, McBride SA, Bailey LA, Ebbole DJ. rco-3, a gene involved in glucose transport and conidiation in Neurospora crassa. *Genetics* (1997) **146**:499–508.

58. Bailey LA, Ebbole DJ. The fluffy gene of Neurospora crassa encodes a Gal4p-type C6 zinc cluster protein required for conidial development. *Genetics* (1998) **148**:1813–1820.

59. Sun J, Tian C, Diamond S, Glass NL. Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in Neurospora crassa. *Eukaryot Cell* (2012) **11**:482–93. doi:10.1128/EC.05327-11

60. Sun X, Zhang H, Zhang Z, Wang Y, Li S. Involvement of a helix-loop-helix transcription factor CHC-1 in CO(2)-mediated conidiation suppression in Neurospora crassa. *Fungal Genet Biol FG B* (2011) **48**:1077–1086. doi:10.1016/j.fgb.2011.09.003

61. Thompson S, Croft NJ, Sotiriou A, Piggins HD, Crosthwaite SK. Neurospora crassa heat shock factor 1 Is an essential gene; a second heat shock factor-like gene, hsf2, is required for asexual spore formation. *Eukaryot Cell* (2008) **7**:1573–1581. doi:10.1128/EC.00427-07

62. Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF, et al. Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. *Eukaryot Cell* (2010) **9**:1549–1556. doi:10.1128/EC.00154-10

63. Chen C-H, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ. Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora. *EMBO J* (2009) **28**:1029–1042. doi:10.1038/emboj.2009.54

64. Raju NB. Genetic control of the sexual cycle in Neurospora. *Mycol Res* (1992) **96**:241–262. doi:10.1016/S0953-7562(09)80934-9

65. Lord KM, Read ND. Perithecium morphogenesis in Sordaria macrospora. *Fungal Genet Biol FG B* (2011) **48**:388–399. doi:10.1016/j.fgb.2010.11.009

66. Bistis GN. Evidence for diffusible, mating-type-specific trichogyne attractants inNeurospora crassa. *Exp Mycol* (1983) **7**:292–295. doi:10.1016/0147-5975(83)90051-8

67. Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO, Herskowitz I. The transcriptional program of sporulation in budding yeast. *Science* (1998) **282**:699–705. doi:10.1126/science.282.5389.699

68. Kassir Y, Granot D, Simchen G. IME1, a positive regulator gene of meiosis in S. cerevisiae. *Cell* (1988) **52**:853–862. doi:10.1016/0092-8674(88)90427-8

69. Kassir Y, Adir N, Boger-Nadjar E, Raviv NG, Rubin-Bejerano I, Sagee S, Shenhar G. Transcriptional regulation of meiosis in budding yeast. *Int Rev Cytol* (2003) **224**:111–171. doi:10.1016/s0074-7696(05)24004-4

70. Primig M, Williams RM, Winzeler EA, Tevzadze GG, Conway AR, Hwang SY, Davis RW, Esposito RE. The core meiotic transcriptome in budding yeasts. *Nat Genet* (2000) **26**:415–423. doi:10.1038/82539

71. Enyenihi AH, Saunders WS. Large-scale functional genomic analysis of sporulation and meiosis in Saccharomyces cerevisiae. *Genetics* (2003) **163**:47–54.

72. Aramayo R, Metzenberg RL. Gene replacements at the his-3 locus of Neurospora crassa. *Fungal Genet Newsl* (1996) **43**:9–13.

73. Feng B, Haas H, Marzluf GA. ASD4, a new GATA factor of Neurospora crassa, displays sequence-specific DNA binding and functions in ascus and ascospore development. *Biochemistry* (2000) **39**:11065–11073. doi:10.1021/bi000886j

74. Li D, Bobrowicz P, Wilkinson HH, Ebbole DJ. A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in Neurospora crassa. *Genetics* (2005) **170**:1091–104. doi:10.1534/genetics.104.036772

75. McCluskey K, Wiest AE, Grigoriev IV, Lipzen A, Martin J, Schackwitz W, Baker SE. Rediscovery by Whole Genome Sequencing: Classical Mutations and Genome Polymorphisms in Neurospora crassa. *G3 Bethesda Md* (2011) **1**:303–316. doi:10.1534/g3.111.000307

76. Krystofova S, Borkovich KA. The predicted G-protein-coupled receptor GPR-1 is required for female sexual development in the multicellular fungus Neurospora crassa. *Eukaryot Cell* (2006) **5**:1503–16. doi:10.1128/EC.00124-06

77. Bayram O, Krappmann S, Seiler S, Vogt N, Braus GH. Neurospora crassa ve-1 affects asexual conidiation. *Fungal Genet Biol FG B* (2008) **45**:127–138. doi:10.1016/j.fgb.2007.06.001

78. Brody S, Tatum EL. The primary biochemical effect of a morphological mutation in Neurospora crassa. *Proc Natl Acad Sci U S A* (1966) **56**:1290–1297. doi:10.1073/pnas.56.4.1290

79. Kolmark HG. A sequence of colonial mutants with practical application for mutation studies in Neurospora crassa. *Hereditas* (1969) **63**:48–67.

80. Schmoll M, Dattenböck C, Carreras-Villaseñor N, Mendoza-Mendoza A, Tisch D, Alemán MI, Baker SE, Brown C, Cervantes-Badillo MG, Cetz-Chel J, et al. The Genomes of Three Uneven Siblings: Footprints of the Lifestyles of Three Trichoderma Species. *Microbiol Mol Biol Rev MMBR* (2016) **80**:205–327. doi:10.1128/MMBR.00040-15

81. Wortman JR, Gilsenan JM, Joardar V, Deegan J, Clutterbuck J, Andersen MR, Archer D, Bencina M, Braus G, Coutinho P, et al. The 2008 update of the Aspergillus nidulans genome annotation: a community effort. *Fungal Genet Biol FG B* (2009) **46 Suppl 1**:S2-13. doi:10.1016/j.fgb.2008.12.003

82. Chang P-K, Ehrlich KC. Genome-wide analysis of the $Zn(II)_2Cys_6$ zinc cluster-encoding gene family in Aspergillus flavus. *Appl Microbiol Biotechnol* (2013) **97**:4289–4300. doi:10.1007/s00253-013-4865-2

83. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S, et al. The genome sequence of the filamentous fungus Neurospora crassa. *Nature* (2003) **422**:859–868. doi:10.1038/nature01554

84. Selker EU. Premeiotic instability of repeated sequences in Neurospora crassa. *Annu Rev Genet* (1990) **24**:579–613. doi:10.1146/annurev.ge.24.120190.003051

Figure legends

Figure 2.1: Relative distribution of N. crassa transcription factor genes into major classes.

Each “slice” of the pie represents the fraction of mutants with the indicated domain. The number of mutants with each domain is indicated. The MISC (Miscellaneous) group includes the 20 domain classes with four or fewer members (see Table 1).

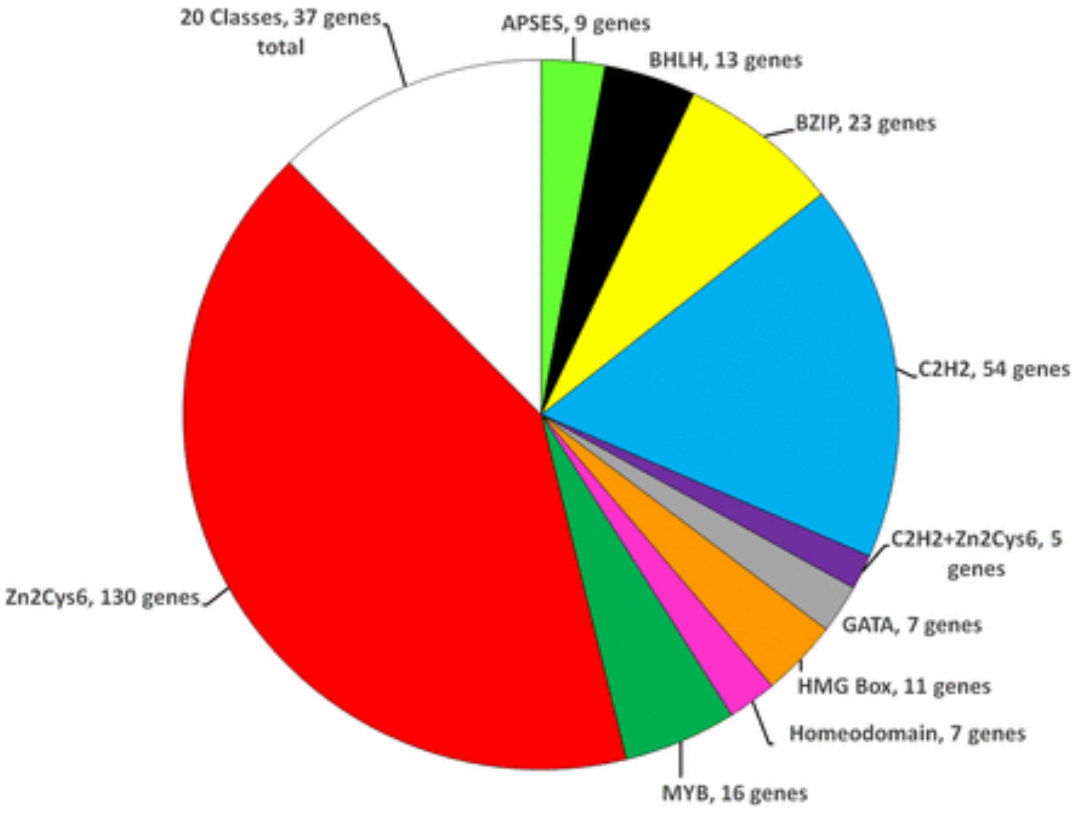

20 Classes, 37 genes total
APSES, 9 genes
BHLH, 13 genes
BZIP, 23 genes
C2H2, 54 genes
C2H2+Zn2Cys6, 5 genes
GATA, 7 genes
HMG Box, 11 genes
Homeodomain, 7 genes
MYB, 16 genes
Zn2Cys6, 130 genes

Figure 2.2: Venn diagram summary of mutants with growth and developmental phenotypes.

The total number of mutants with the indicated phenotype or combination of phenotypes is shown in each lobe of the Venn diagram.

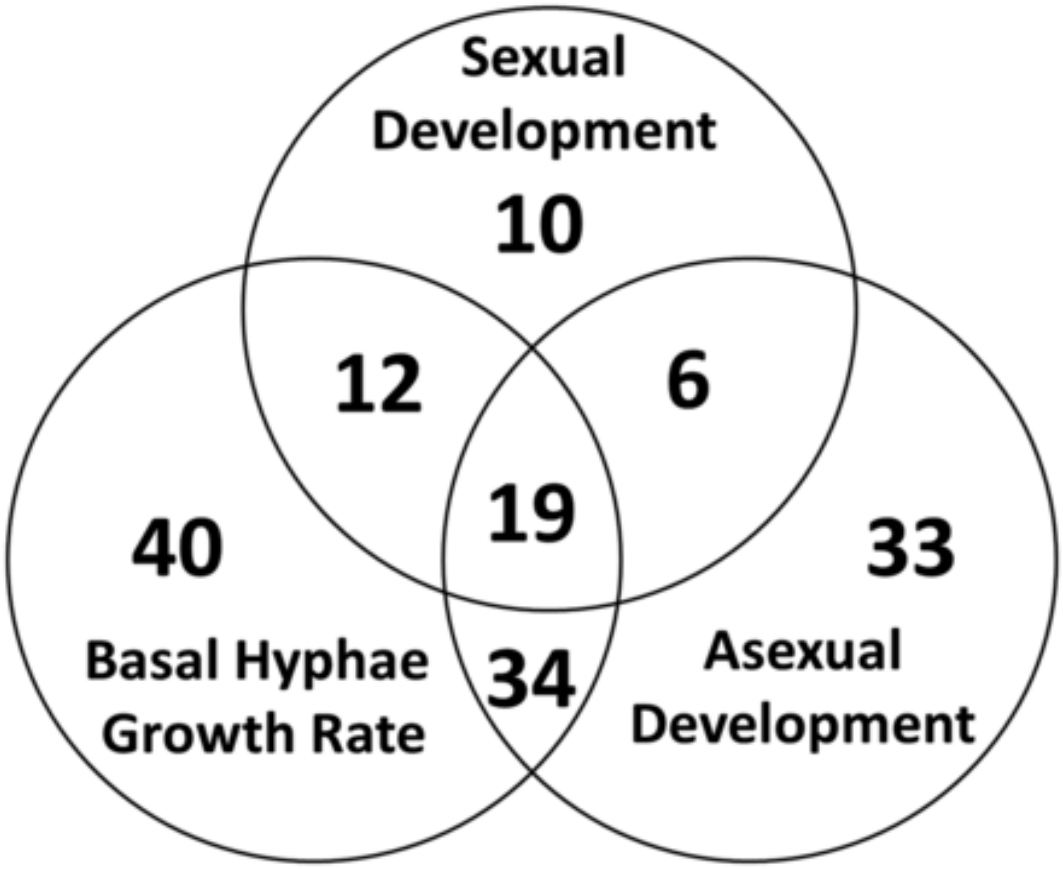
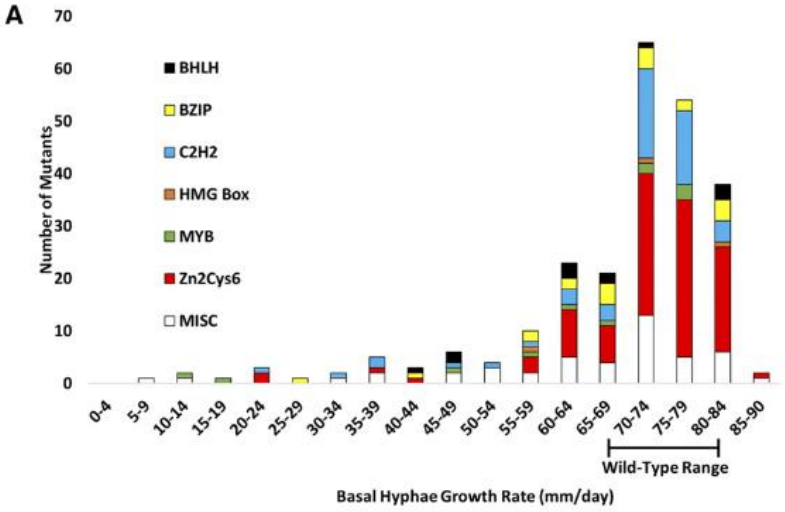
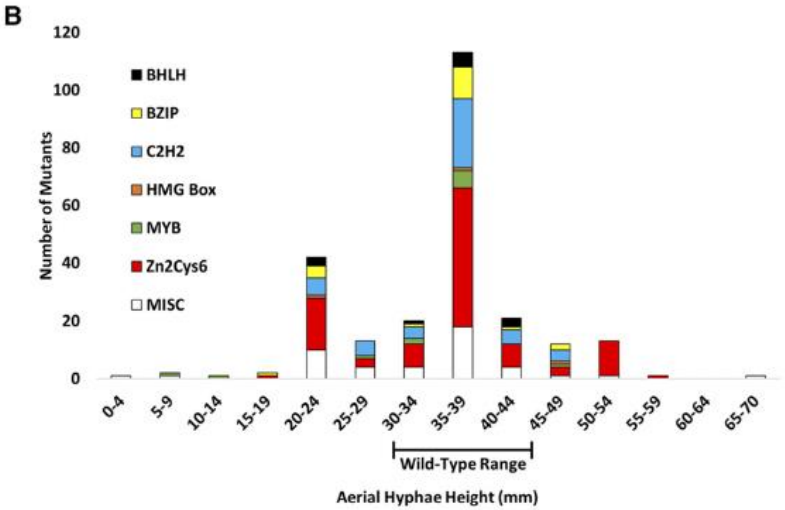

Sexual
Development
10
12
6
19
40
33
Basal Hyphae
Growth Rate
34
Asexual
Development

Figure 2.3: Basal hyphae growth rate and aerial hyphae height phenotypes for mutants in different transcription factor classes.

(A) Basal hyphae growth rate. Race tubes containing VM agar medium were inoculated with transcription factor mutants and incubated in the dark at 25°. The growth front was marked after overnight growth (t = 0), and then marked twice/day over the course of 2–3 d. Growth rate was determined using linear regression analysis (see Materials and Methods for details). Mutants were grouped in bins, as shown. The range of measurements for wild type is indicated on the x-axis. (B) Aerial hyphae height. Standing liquid VM tube cultures were inoculated with mutants and incubated statically for 3 d in the dark at 25°, after which the height of aerial hyphae was measured. Values were obtained using at least six replicates and are presented as described in (A). (C) Comparison between aerial hyphae height and basal hyphae growth rate for all mutants. The data from (A) and (B) were plotted. Mutants that fall within the range of wild type values are enclosed by the dashed-line rectangle.


A
Number of Mutants
BHLH
BZIP
C2H2
HMG Box
MYB
Zn2Cys6
MISC
0-4
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
85-90
Wild-Type Range
Basal Hyphae Growth Rate (mm/day)


B
Number of Mutants
BHLH
BZIP
C2H2
HMG Box
MYB
Zn2Cys6
MISC
0-4
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-70
Wild-Type Range
Aerial Hyphae Height (mm)

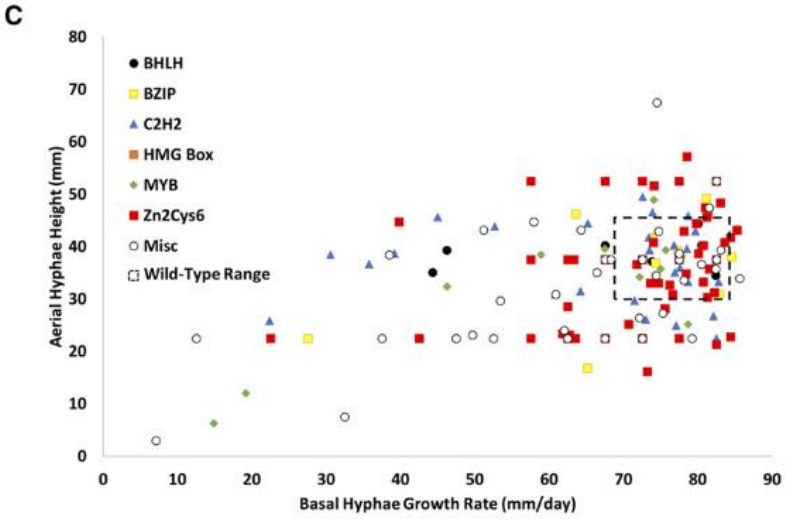

C
Aerial Hyphae Height (mm)
BHLH
BZIP
C2H2
HMG Box
MYB
Zn2Cys6
Misc
Wild-Type Range
Basal Hyphae Growth Rate (mm/day)

Figure 2.4: Transcription factor mutants with defects in sexual development.
(A) Summary of sexual cycle phenotypes. The total number of mutants with the indicated sexual cycle phenotype is shown in each lobe of the Venn diagram. Mutants with reduced number or abnormal phenotypes (red font) or complete block (black font) are scored according to the stage of their earliest defect. (B) Examples of mutants with different sexual cycle phenotypes. All strains were cultured on synthetic crossing medium plates in constant light at room temperature for 7 d to facilitate development of protoperithecia (top panels). Cultures were then fertilized using macroconidia from a wild-type strain of opposite mating type. Plates were then incubated under the same conditions for a further 7 d to allow production of fertilized perithecia and beak development (bottom panels). Images of protoperithecia were captured using an Olympus SZX9 stereomicroscope with a C-4040 digital camera, while perithecia were photographed using a S8APO stereomicroscope with a DFC280 digital camera. White arrows indicate unfertilized protoperithecia (top panels), while black arrows show protoperithecia (fmf-1) or perithecia (wild type and vsd-5 mutant) 7 d after fertilization (bottom panels). The beak at the tip of a perithecium can be seen as the darkened circular area above the black arrow in wild type (bottom left panel), while vsd-5 mutant perithecia lack this structure (bottom right panel). (C) Clustering of mRNA expression data for *N. crassa* transcription factors during a time course of sexual development. Left side of figure: RNAseq data were obtained from Wang et al. (2014). Expression data for 43 of the 47 transcription factor genes with a sexual cycle phenotype were contained in the data set. Clustering analysis and heatmap generation were performed as described in the Materials and Methods. Red

shading denotes greater levels of expression, while blue indicates lower expression. The numbers along the left side of the figure indicate groupings (1–6) based on similar patterns of expression during sexual development. The table on the right side of the figure is a phenotype summary for the mutants lacking each transcription factor. The open circles denote that the indicated structure is abnormal or that a reduced number is formed, while closed circles show that the indicated structure is not formed. The absence of a circle indicates there was no defect observed. PP, Protoperithecia; P, Perithecia; A, Ascospores.

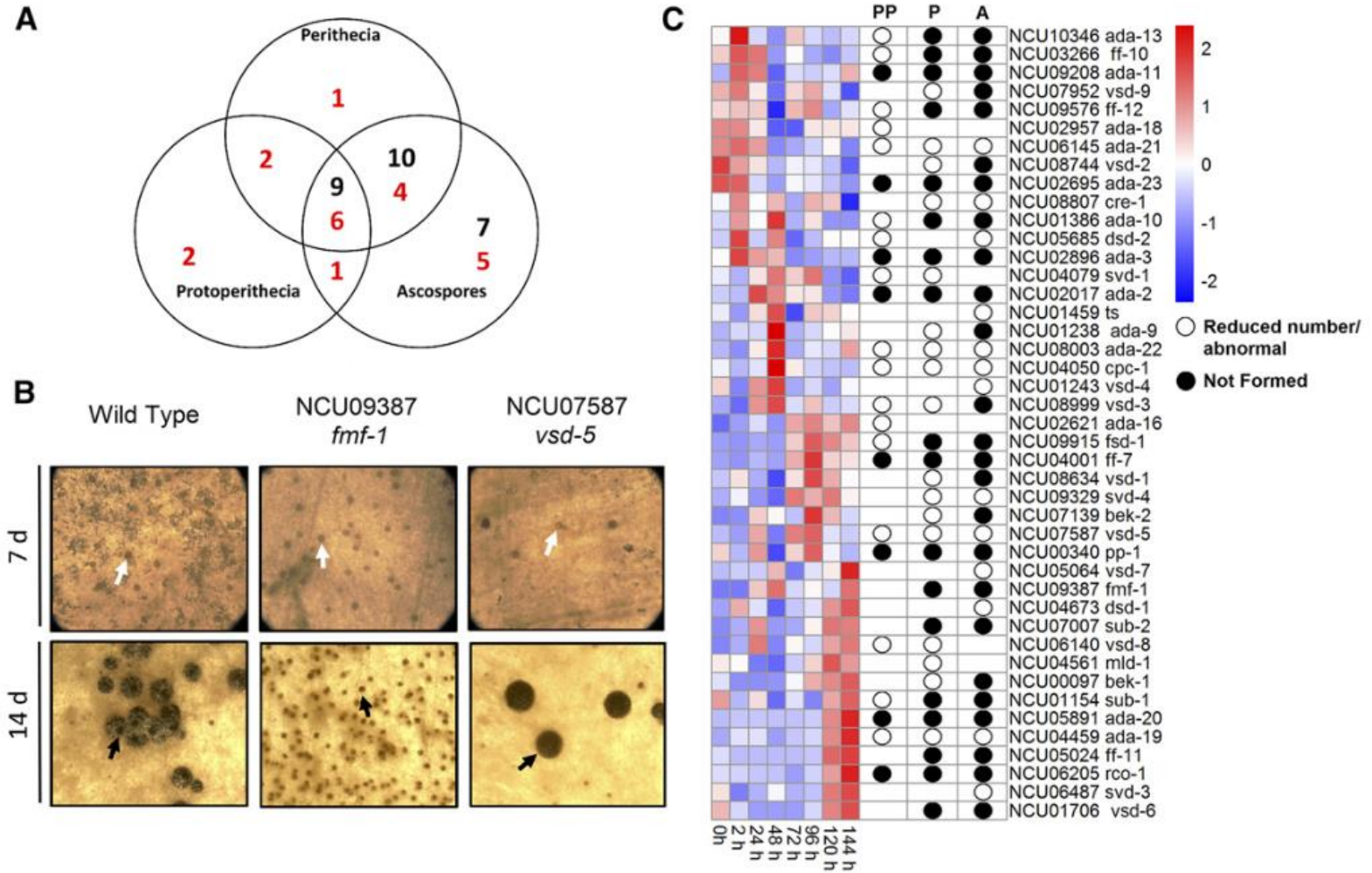

A
Perithecia
Protoperithecia
Ascospores
1
2
10
9
4
6
7
2
5
1
B
Wild Type
NCU09387
fmf-1
NCU07587
vsd-5
7 d
14 d
C
PP
P
A
NCU10346 ada-13
NCU03266 ff-10
NCU09208 ada-11
NCU07952 vsd-9
NCU09576 ff-12
NCU02957 ada-18
NCU06145 ada-21
NCU08744 vsd-2
NCU02695 ada-23
NCU08807 cre-1
NCU01386 ada-10
NCU05685 dsd-2
NCU02896 ada-3
NCU04079 svd-1
NCU02017 ada-2
NCU01459 ts
NCU01238 ada-9
NCU08003 ada-22
NCU04050 cpc-1
NCU01243 vsd-4
NCU08999 vsd-3
NCU02621 ada-16
NCU09915 fsd-1
NCU04001 ff-7
NCU08634 vsd-1
NCU09329 svd-4
NCU07139 bek-2
NCU07587 vsd-5
NCU00340 pp-1
NCU05064 vsd-7
NCU09387 fmf-1
NCU04673 dsd-1
NCU07007 sub-2
NCU06140 vsd-8
NCU04561 mld-1
NCU00097 bek-1
NCU01154 sub-1
NCU05891 ada-20
NCU04459 ada-19
NCU05024 ff-11
NCU06205 rco-1
NCU06487 svd-3
NCU01706 vsd-6
0h
2 h
24 h
48 h
72 h
96 h
120 h
144 h
2
1
0
-1
-2
Reduced number/ abnormal
Not Formed

Figure 2.5. Sexual development time course heat maps for MYB, C2H2 and Zn2Cys6 transcription factor classes. The clustering and heatmap analysis presented in Figure 2.4 was performed on all mutants in each of the displayed transcription factor classes, whether or not they had a sexual development phenotype. Boxed groups of genes include one gene with a non-blocking sexual development phenotype and another gene with no phenotype or a nonblocking defect that is similarly expressed during sexual development. The values for each gene are normalized to the RPKM at the time point with lowest expression for that gene.

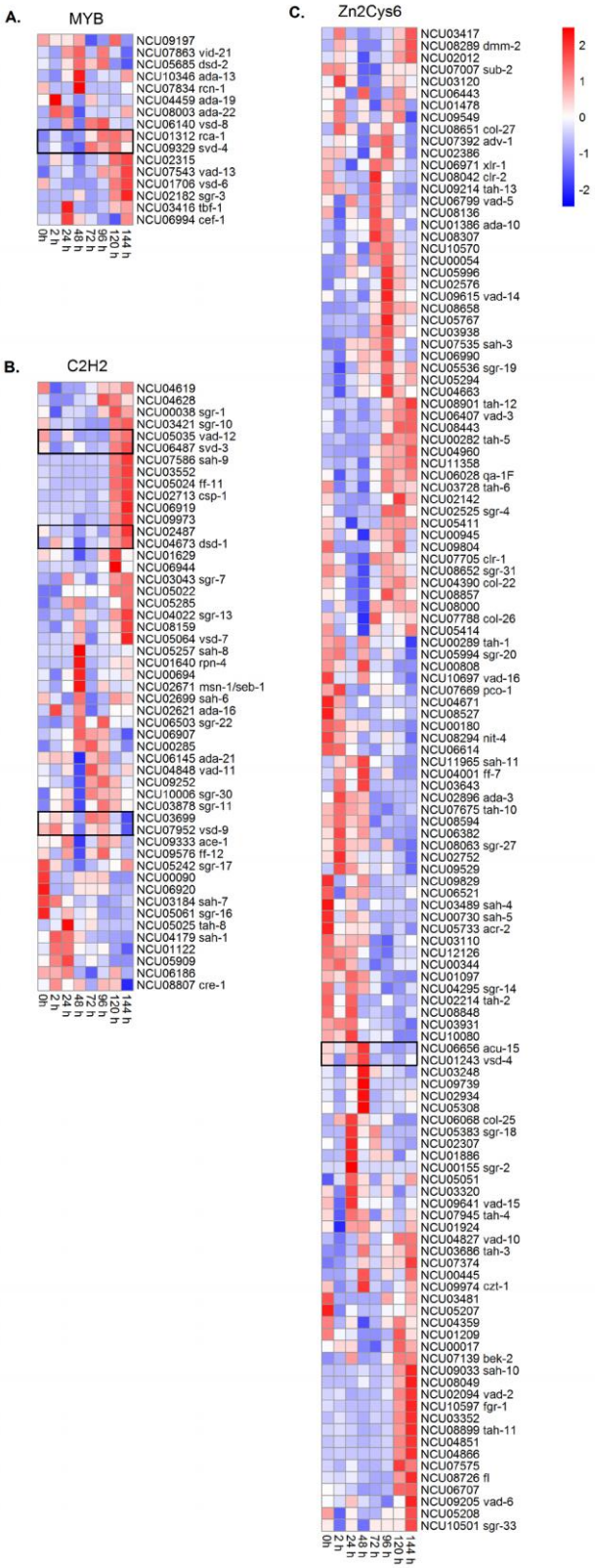

A.
MYB
NCU09197
NCU07863 vid-21
NCU05685 dsd-2
NCU10346 ada-13
NCU07834 rcn-1
NCU04459 ada-19
NCU08003 ada-22
NCU06140 vsd-8
NCU01312 rca-1
NCU09329 svd-4
NCU02315
NCU07543 vad-13
NCU01706 vsd-6
NCU02182 sgr-3
NCU03416 tbf-1
NCU06994 cef-1
0h
2 h
24 h
48 h
72 h
96 h
120 h
144 h
B.
C2H2
NCU04619
NCU04628
NCU00038 sgr-1
NCU03421 sgr-10
NCU05035 vad-12
NCU06487 svd-3
NCU07586 sah-9
NCU03552
NCU05024 ff-11
NCU02713 csp-1
NCU06919
NCU09973
NCU02487
NCU04673 dsd-1
NCU01629
NCU06944
NCU03043 sgr-7
NCU05022
NCU05285
NCU04022 sgr-13
NCU08159
NCU05064 vsd-7
NCU05257 sah-8
NCU01640 rpn-4
NCU00694
NCU02671 msn-1/seb-1
NCU02699 sah-6
NCU02621 ada-16
NCU06503 sgr-22
NCU06907
NCU00285
NCU06145 ada-21
NCU04848 vad-11
NCU09252
NCU10006 sgr-30
NCU03878 sgr-11
NCU03699
NCU07952 vsd-9
NCU09333 ace-1
NCU09578 ff-12
NCU05242 sgr-17
NCU00090
NCU06920
NCU03184 sah-7
NCU05061 sgr-16
NCU05025 tah-8
NCU04179 sah-1
NCU01122
NCU05909
NCU06186
NCU08807 cre-1
0h
2 h
24 h
48 h
72 h
96 h
120 h
144 h
C.
Zn2Cys6
2
1
0
-1
-2
NCU03417
NCU08289 dmm-2
NCU02012
NCU07007 sub-2
NCU03120
NCU06443
NCU01478
NCU09549
NCU08651 col-27
NCU07392 adv-1
NCU02386
NCU06971 xlr-1
NCU08042 clr-2
NCU09214 tah-13
NCU06799 vad-5
NCU08136
NCU01386 ada-10
NCU08307
NCU10570
NCU00054
NCU05996
NCU02576
NCU09615 vad-14
NCU08658
NCU05767
NCU03938
NCU07535 sah-3
NCU06990
NCU05536 sgr-19
NCU05294
NCU04683
NCU08901 tah-12
NCU06407 vad-3
NCU08443
NCU00282 tah-5
NCU04960
NCU11358
NCU06028 qa-1F
NCU03728 tah-6
NCU02142
NCU02525 sgr-4
NCU05411
NCU00945
NCU09804
NCU07705 clr-1
NCU08652 sgr-31
NCU04390 col-22
NCU08857
NCU08000
NCU07788 col-26
NCU05414
NCU00289 tah-1
NCU05994 sgr-20
NCU00808
NCU10697 vad-16
NCU07669 pco-1
NCU04671
NCU08527
NCU00180
NCU08294 nit-4
NCU06614
NCU11965 sah-11
NCU04001 ff-7
NCU03643
NCU02896 ada-3
NCU07675 tah-10
NCU08594
NCU06382
NCU08063 sgr-27
NCU02752
NCU09529
NCU09829
NCU06521
NCU03489 sah-4
NCU00730 sah-5
NCU05733 acr-2
NCU03110
NCU12126
NCU00344
NCU01097
NCU04295 sgr-14
NCU02214 tah-2
NCU08848
NCU03931
NCU10080
NCU06656 acu-15
NCU01243 vsd-4
NCU03248
NCU09739
NCU02934
NCU05308
NCU06068 col-25
NCU05383 sgr-18
NCU02307
NCU01886
NCU00155 sgr-2
NCU05051
NCU03320
NCU09641 vad-15
NCU07945 tah-4
NCU01924
NCU04827 vad-10
NCU03686 tah-3
NCU07374
NCU00445
NCU09974 czt-1
NCU03481
NCU05207
NCU04359
NCU01209
NCU00017
NCU07139 bek-2
NCU09033 sah-10
NCU08049
NCU02094 vad-2
NCU10597 fgr-1
NCU03352
NCU08899 tah-11
NCU04851
NCU04866
NCU07575
NCU08726 fl
NCU06707
NCU09205 vad-6
NCU05208
NCU10501 sgr-33
0h
2 h
24 h
48 h
72 h
96 h
120 h
144 h

Figure 2.6. Asexual/sexual sporulation phenotypes and sexual development expression data analysis for transcription factors regulated by the White Collar Complex (WCC).

A. Asexual and sexual development phenotype summary. Asexual sporulation (aerial hyphae height and/or conidia production) and sexual development (protoperithecia, perithecia and ascospore production) defects for available mutants lacking 18 transcription factors that are regulated by the WCC (see (Smith et al. 2010)). *Data supporting an asexual sporulation defect for ve-1 mutant taken from (Bayram et al. 2008).

B. Sexual development time course heat map. Available RNAseq data for 17/21 transcription factor genes regulated by the WCC were subjected to clustering and heatmap analysis as described in Figure 2.4C. NCU0775/clr-1 and NCU08000 (boxed) lack sexual cycle phenotypes, but are coordinately expressed during sexual development.

A.

Gene Number	Gene Name	Class	Asexual Development	Sexual Development
NCU01871		BHLH	*N/A*	*N/A*
NCU04731	*sah-2*	BHLH	AH	
NCU02713	*csp-1*	C2H2	AH	
NCU03184	*sah-7*	C2H2	AH	
NCU08159		C2H2	*N/A*	*N/A*
NCU08807	*cre-1*	C2H2	C	P,A
NCU06095	*csp-2*	CP2	AH	
NCU01154	*sub-1*	GATA		PP,P,A
NCU07728	*sre*	GATA	AH	
NCU09068	*nit-2*	GATA	AH	
NCU00097	*bek-1*	Homeodomain		P,A
NCU05964	*ve-1*	Velvet	*	
NCU01243	*vsd-4*	Zn2Cys6		A
NCU02094	*vad-2*	Zn2Cys6	AH	
NCU04295	*sgr-14*	Zn2Cys6		
NCU05994	*sgr-20*	Zn2Cys6		
NCU07392	*adv-1*	Zn2Cys6	AH,C	PP,P,A
NCU07705	*clr-1*	Zn2Cys6		
NCU08000		Zn2Cys6		
NCU09615	*vad-14*	Zn2Cys6	AH	
NCU09829		Zn2Cys6		

B.

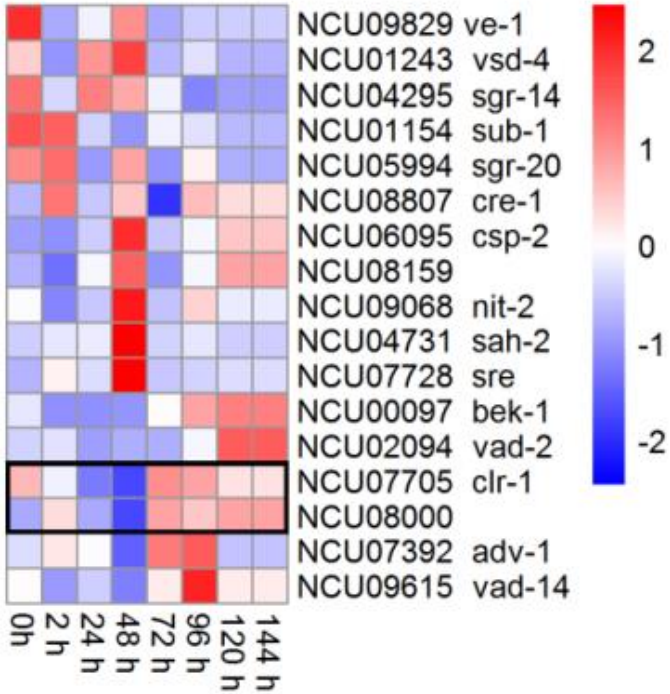


Table 2.1: Phenotype summary

TF Class	Total	Genes with complete data	Genes with complete data (%)	Genes with data with phenotype	Genes with data with phenotype (%)	Growth	Asexual	Sexual	Growth + Asexual	Growth + Sexual	Asexual + Sexual	All
APSES	9	7	78	3	43	1			1	1		
ARID/BRIGHT	2	2	100	2	100	1						1
AT hook	2	0	0	0								
BHLH	13	12	92	10	83	5		1	2	1		1
BZIP	23	20	87	13	65	3	1		5	3		1
C2H2	54	48	89	30	63	9	7	2	4	2	2	4
CAAT box/CBF	4	3	75	2	67	1						1
Copper fist	2	0	0	0								
CP2	1	1	100	1	100				1			
Far1	3	1	33	1	100		1					
Forkhead	3	3	100	2	67	1				1		
GATA	7	6	86	6	100		2		3	1		
HMG box	11	3	27	2	67						1	1
Homeodomain	7	7	100	5	71	1	1	2	1			
HSF	3	2	67	2	100				2			
MADS-box	2	0	0									
MATalpha1	1	0	0									
MBF	1	0	0									
MYB	16	12	75	11	92	2	1	1	1	2	1	3
NDT80	3	3	100	3	100	1		1	1			
NFX	1	0	0									
RFX	1	1	100	0	0							
Velvet	1	0	0									
WD40	3	3	100	3	100	1						2
Zn2Cys6	130	101	78	53	52	14	19	2	13	1	1	3
ARID + APSES	1	1	100	1	100							1
ARID + MYB	1	1	100	1	100						1	
C2H2 + APSES	1	1	100	0								
C2H2 + STE-like	1	1	100	1	100							1
C2H2 + Zn2Cys6	5	3	60	2	67		1	1				
Total	312	242	77.56	154	63.64	40	33	10	34	12	6	19

Chapter 3

Clustering analysis of large-scale phenotypic data in the model filamentous fungus *Neurospora crassa*

Abstract

With 9,730 protein-coding genes and a nearly complete gene knockout strain collection, *Neurospora crassa* is a major model organism for filamentous fungi. Phenotypic data for 10 different growth or developmental attributes have been obtained for nearly 1300 mutants in *N. crassa*. Of the 1168 mutants with complete phenotypic data (12% of genome), 386 (33%) are in the normal range, while 782 (67%) possess at least one phenotype. With the exception of unclassified functions, the distribution of functional categories for genes in the mutant dataset mirrors that of the entire *N. crassa* genome. In contrast, the majority of genes do not possess a yeast ortholog, suggesting that our analysis will reveal functions for genes that are not conserved in *Saccharomyces cerevisiae*. In order to leverage the phenotypic data to identify pathways, we tested the ability of several statistical clustering approaches to group mutants based on their growth and developmental phenotypes. Our analysis showed that a weighted Partitioning Around Medoids (PAM) approach with 40 clusters provided the most biologically relevant grouping of mutants. We found that genes encoding metabolic, transmembrane and protein phosphorylation-related genes are concentrated in subsets of clusters. Results from K-Means clustering of publicly available transcriptomic datasets for wild type *N. crassa* showed that most phenotypic clusters contain multiple expression profiles, suggesting that co-expression is not generally observed for genes with shared phenotypes. Analysis of yeast orthologs of genes that co-clustered in MAPK signaling cascades revealed potential networks of interacting proteins in *N. crassa*. Taken together, our results demonstrate that clustering analysis of phenotypes is a promising tool for generating new hypotheses regarding involvement of genes in cellular pathways in *N. crassa*.

Introduction

Neurospora crassa is a model organism for studies of cellular biology and genetics in filamentous fungi (1, 2). Filamentous fungi can be pathogens of plants and animals, but also form beneficial endophytic associations with plants (reviewed in 3-5). Many filamentous fungi are crucial players during carbon cycling in the environment and serve as commercial sources of food, drink, biofuels and other products (6-8). The contributions of *N. crassa* to many areas of cell and molecular biology include the one-gene, one-polypeptide hypothesis, genetic recombination, gene silencing by small RNAs, epigenetic phenomena, photobiology, circadian rhythms, cell signaling, plant cell wall decomposition, and self-nonself interactions during vegetative growth and sexual development (reviewed in 1 and 2).

N. crassa has a rich history of forward genetics, with more than 1000 loci identified (http://www.fgsc.net/2000compendium/2000compend.html) (9). *N. crassa* was the first filamentous fungus with a complete genome sequence (10). The ~ 40 Mb genome contains ~ 10,000 protein-coding genes distributed among seven linkage groups (10). One goal of the Neurospora Genome project was to delete all of the genes in the *N. crassa* genome using reverse genetics (11, 12). The project produced knockout mutant strains for nearly 9000 genes and these mutants are currently available at the Fungal Genetics Stock Center (13). In all knockout mutants, a gene open reading frame has been replaced with an hph selectable marker (conferring resistance to the antibiotic hygromycin) (11, 14). Another goal of the Neurospora Genome Project was to perform phenotypic characterization of these knockout mutants. An undergraduate research

program at the University of California, Los Angeles, pioneered methods for phenotyping mutants, with ~ 1000 mutants analyzed (15). In our laboratory, we have used the phenotypic methods developed by the UCLA project to analyze additional mutants, focusing on those lacking serine/threonine protein kinases, serine/threonine and tyrosine protein phosphatases, G protein coupled receptors and transcription factors (16-19). The above projects have generated phenotypic data for nearly 1300 *N. crassa* knockout mutants. However, these data have not been otherwise analyzed or interpreted using hierarchical/partitional statistical methods. Such approaches have been used in some fungal species, where phenomics was used to predict relationships between genes (20-22). All of these studies utilized either only quantitative data or converted categorical phenotypes to a numeric scale followed by generation of a distance matrix and application of Pearson's correlation (23) to perform further analysis and interpretation. However, there are no published reports of analysis of gene deletion mutants in fungi where phenotypes were clustered without conversion of categorical data to an arbitrary numerical value, nor studies that include purely categorical data in clustering analyses.

In this study, we curated phenotypic data for 10 categorical or quantitative/continuous traits for 1168 *N. crassa* knockout mutants obtained from the above projects. We then clustered the data without conversion to arbitrary values using **P**artitioning **A**round **M**edoids (PAM) (24). Further, transcriptomics data from three publicly available datasets were clustered using K-means and the resulting expression profiles compared to the phenotypic clusters to determine whether gene expression correlated with phenotype. Our results reveal previously unknown relationships between

phenotypes and genes encoding proteins with particular domains and between phenotypes and gene expression trends. These results have led to new hypotheses regarding cellular pathways that can be the subject of follow-up studies.

Materials and Methods

Data sources and curation. Knockout mutants were produced during the Neurospora Genome Project (https://geiselmed.dartmouth.edu/dunlaploros/genome/) in the Dunlap or Borkovich laboratories (12) and deposited at the Fungal Genetics Stock Center (13). Most phenotypic data for the mutants were obtained by undergraduate students in summer research programs or during courses at the University of California, Los Angeles, the University of California, Riverside (UCR), Texas A&M University and the University of Manchester, UK. Phenotypic data for kinase, phosphatase, GPCR and transcription factor mutants have been previously published (16-19). Data were initially deposited at the Broad Institute-MIT (https://www.broadinstitute.org). After downloading and curation, the data were then migrated to FungiDB (fungidb.org) (74, 75). All mutants with complete data for the 10 chosen traits (see below) were included in our analysis.

The methods used for phenotypic analysis were as previously described (11, 15-19) and will be briefly summarized here. Similar to recent publications, we have omitted measurements of pigmentation and aerial hyphae height on yeast extract-containing medium from our analysis (16, 17). Knockout mutants were either obtained from the Fungal Genetics Stock Center (FGSC; Kansas State University, Manhattan, KS; http://www.fgsc.net) or produced in the Borkovich laboratory using methods

described in (11). Near-isogenic wild-type strains FGSC4200 and/or FGSC2489 (obtained from the FGSC) were used as controls.

Hyphal growth rate. The apical extension rate of basal hyphae was measured using glass or disposable race tubes (76, 77) containing Vogel's minimal agar medium (VM (78). Tubes were inoculated at one end and were incubated at 25 °C under ambient light conditions (15) or in the dark (11, 18-19). Before marking the growth front, tubes were grown overnight to eliminate effects on growth rate due to the age of the culture or germination defects. The total growth in mm was measured for each time point and a plot of mm vs. time used to determine growth rate. A minimum of four replicates with growth rates with R squared values greater than 0.95 were used to obtain the average growth rate for each strain. Binned data from (11) and/or the Broad Database were averaged to allow comparison to actual growth rate measurements obtained for some mutants. The wild-type growth rate range was 75–85 mm/day.

Asexual development. The height of aerial hyphae was measured in liquid standing cultures containing VM. Tubes were incubated statically (typically in the dark) at 25 °C for 3–4 days. Total height (in mm) was recorded. A minimum of four replicates was analyzed for each strain. The average value in mm was reported. As for growth rate, binned data were averaged. The wild-type range was 30–45 mm. For semi-quantitative analysis of conidia number and morphology, slant tubes (13x100mm) containing 3 ml of VM agar medium were inoculated with strains and grown at room temperature for 6 to

8 days under ambient light conditions (15) or for 3 days in the dark at 30 °C and 4 days in the light at room temperature (11, 18-19). Production of conidia was scored visually, for amount and morphology. A minimum of four replicates was analyzed for each strain.

Sexual Development. Synthetic crossing medium (SCM) (78) agar slant tubes were inoculated with the various strains. After 7–8 days of incubation at room temperature in constant light, cultures were scored for the number and morphology of protoperithecia by inspection using a stereomicroscope. At least four replicates were scored for each strain. The 7–8 day old SCM cultures from the protoperithecial scoring were fertilized using a suspension of wild-type conidia of the opposite mating type. Cultures were returned to the same conditions used for protoperithecial development. After seven more days (~ 2 weeks total), cultures were scored for the number and morphology of perithecia using a stereomicroscope. The 2-week-old SCM cultures from the perithecial scoring were returned to the same culture conditions. After seven more days of incubation (~ 3 weeks total), cultures were scored for the number and morphology of ascospores using a stereomicroscope. Perithecial beak morphology was also scored at this point.

Clustering approaches. To uncover new biological pathways, we grouped mutants with similar phenotypes into clusters using several algorithms for hierarchical and partitioning clustering: Pearson's Correlation Coefficient (23), K-means (27), **F**actor **A**nalysis of **M**ixed **D**ata (FAMD) (28), Ward's minimum variance (Ward's) (29) and **P**artitioning **A**round **M**edoids (PAM) (24). Initial clustering was performed with

algorithms and distance matrices that cannot utilize non-numeric data. Data was initially transformed with the semi-quantitative categorical variables (ordinal categories) being converted to a value between 0 and 1.5 based on severity of the phenotype. The scale was chosen based on increments of 0.25 with 6 categories, from not formed [0], severely reduced [0.25], reduced [0.5], slightly reduced [0.75], normal [1.0], and increased [1.5]. These values are based on the approximate quantitation applied during the scoring. A one minus Pearson's Correlation Coefficient distance matrix (23) was created using the Factoextra package in R (79). The Pearson's distance matrix was used as the input for Hierarchical Agglomerative Clustering (HAC) with complete linkage (22). K-means clustering (27) was performed using base R (https://www.r-project.org/) with the converted dataset as the input.

Further clustering was performed with algorithms and distance matrices that can handle mixed categorical and numeric data. Factorial Analysis of Mixed Data (FAMD) (28) was performed using the FactoMineR package in R (80). The non-converted dataset with categorical data left as-is (including both ordinal and categorical data) was used as the input. (30)A Gower's distance matrix metric (30) was created using the clustering package in R (29). The Gower's matrix was used as input for Ward's minimum variance clustering algorithm (27) and Partitioning Around Medoids (PAM) clustering algorithm (24). Both algorithms were run in R using the clustering package (29). All packages and information needed to run this algorithm are available or easily added to R (https://r-project.com).

In order to judge the biological relevance of the clusters, we determined the average relative standard deviation for the two continuous traits (basal hyphae growth rate and aerial hyphae height) and the average percent consensus for the categorical traits (all other phenotypes). To determine the relative standard deviation for the two continuous phenotypes we first calculated the standard deviation for each cluster and then divided that standard deviation by the cluster mean to determine the relative standard deviation. In order to create a composite relative standard deviation for the "run" (e.g., 21 total clusters and 22 total clusters are two different "runs"), we calculated the average relative standard deviation for all clusters in that run. For categorical phenotypes, we first calculated the percentage of each category/phenotype for the cluster and then identified the most prevalent category (category with the highest percent representation). For example, if a cluster contained 10 genes/mutants, with six having reduced conidial abundance, then that cluster would have a value of 60% for conidial abundance. This was repeated for all clusters in the run. We then determined an average percent consensus by calculating the average representation (%) of the most prevalent category for all clusters in the run. The relative standard deviations and the average percent consensus were then averaged across the two continuous phenotypes and the eight categorical phenotypes to arrive at two composite values for each run utilizing each clustering approach.

Analysis of specific classes of cluster genes. Functional catalogue (Funcat) analysis was performed on *N. crassa* genes using FungiFun (https://sbi.hki-jena.de/fungifun/) to test for enrichment of specific functions among the gene clusters. A list of 5781 genes that are

annotated with functional categories was obtained from the FungiFun website. Duplicates were removed, yielding a final list of 5082 genes. All genes not in the list were categorized as "Unclassified". The p-value significance level was set to 1 to capture all possible associations of genes with functional categories. The annotation type was set to "Use also indirectly annotated top categories" to simplify the functional categories to their top-level category.

N. crassa genes with orthologs in other fungi were identified using the "Transform by Orthology" tool at FungiDB (FungiDB.org). Genetic and physical interactions between orthologs and other genes/gene products in baker's yeast were identified using the "interactions" tab at the Saccharomyces Genome Database (yeastgenome.org) (81). *N. crassa* metabolic genes were obtained from (31). Genes encoding proteins with secretion signals and/or transmembrane domains were retrieved using the "Protein targeting and localization" tool at FungiDB. Predicted proteins that are phosphorylated in *N. crassa* were obtained from (37).

Targets of the two Extracellular-signal Regulated Kinase (ERK) class Mitogen-Activated Protein Kinase (MAPK) cascades were identified in publicly available phosphoproteomic or transcriptomic datasets for *N. crassa* (whenever possible) or other fungi. For the MAK-1 MAPK pathway, clusters were checked for misregulated genes using microarray data for a Δ*mak-1* knockout mutant (44). For the MAK-2 MAPK cascade, microarray data from a Δ*mak-2* mutant (41) or a strain expressing a *mak-2* inhibitable allele (*mak-2*^{Q100G}) in the presence of inhibitor (45), as well as phosphoproteomics data from two additional studies using the *mak-2*^{Q100G} strain in the presence of inhibitor (46, 47) were

utilized. Because there currently is no transcript profiling data available for mutants lacking genes in the p38 MAPK OS-2 pathway in *N. crassa*, RNA-seq or microarray datasets were analyzed for genes controlled by the os-2 orthologs *sakA* or *hog1* in *Aspergillus fumigatus* or *Cryptococcus neoformans*, respectively (48, 49).

Publicly available transcriptomics datasets for wild-type *N. crassa* were used to compare gene expression trends to phenotypes in clusters. These included two microarray datasets, corresponding to time courses of macroconidiation (54) and colony growth (55), as well as two RNA-seq datasets for time courses during the sexual cycle (52) and conidial germination (53). Expression data was scaled to values between −2 and 2 to give comparable relative expression per gene. K-means clustering (25) was then performed to produce expression profiles. Comparisons between the phenotypic clusters and the expression profiles were made to check for relationships between mRNA expression and phenotype.

Results

***N. crassa* mutant defects are distributed across broad growth and developmental phenotypes.** This dataset contains phenotypes for 379 mutants previously reported in five publications (11, 16-19), corresponding to 242 transcription factor, 36 GPCR, 24 serine-threonine-tyrosine protein phosphatase and 77 serine-threonine protein kinase gene mutants. The dataset also includes phenotypes for 789 mutants that were not previously

published. Phenotypic data for all mutants is available at the FungiDB database on the specific gene's page.

During curation, we settled on 10 traits that were easiest for students to score and were therefore the most reliable. These include hyphal growth rate on solid medium, aerial hyphae height, conidia amount and morphology, abundance and appearance of unfertilized and fertilized female reproductive structures (protoperithecia and perithecia) and sexual spores (ascospores). Since the last eight phenotypes were scored using visual screens, amount/abundance is semi-quantitative (see Methods). In all, 1286 knockout mutants had phenotypic data for at least one trait (13% of genome) and 1168 (12% of genome) had complete data for the 10 traits included in our analysis (Figure 3.1a). Of the 1168 mutants with complete data, 903 mutants (77%) possessed at least one defect. We grouped the phenotypes for the 903 mutants with at least one defect into three global phenotypic classes: growth rate, asexual development (conidiation or aerial hyphae) and sexual development (protoperithecia, perithecia and ascospores; Figure 3.1b). This yielded 1539 total global phenotypes for the 903 mutants. A total of 742 mutants (48.2%) had growth rate defects, 553 (35.9%) possessed an asexual developmental defect and 244 (15.8%) had a phenotype during sexual development. The lower overall incidence of sexual cycle phenotypes was previously noted in published results for GPCRs and transcription factors and may reflect the large number of mutants in these two classes (36 and 242 mutants, respectively) in our combined dataset.

The genes in the phenotypic dataset broadly mirror those of the *N. crassa* genome.

We next explored the similarity between our dataset and the *N. crassa* protein-coding

genome at a global scale, for representation of genes on each of the seven chromosomes and in gene functional categories. We used a chi-square test to determine whether the distribution of genes in our dataset across all chromosomes is similar to that of the entire genome. The low p-value (0.0000038) formally rejects goodness of fit of the data in Figure 3.2 a, and, therefore, there is no strong correlation between the dataset genes and the genome. This is corroborated by the observation that our dataset is overrepresented by approximately 5–10% on chromosomes 1 and 3, and underrepresented by 2–3% on chromosomes 2 and 4, while the distribution on the other three chromosomes was within 1% of the genome (Figure 3.2a).

We investigated the distribution of functions of the mutated genes in our dataset vs. all genes in the genome using Functional Catalogue (FunCat) analysis with data from FungiFun2 (Figure 3.2b). With the exception of protein synthesis and unclassified proteins, our dataset was overrepresented in each category by 2–10%, with the average overrepresentation approximately 6.2%. Taking into account all categories over- and underrepresented, the average difference from the genome is +/− 3.7%. Thus, the findings from analyzing the distribution of genes on chromosomes and their functional categories indicate that our phenotypic dataset is broadly representative of the *N. crassa* genome.

Analysis using Partitioning Around Medoids (PAM) yielded the most biologically relevant phenotypic clusters. As there are many challenges associated with clustering mixed data (Floss et al. 2018), we initially tested several algorithms for hierarchical and partitioning clustering in order to uncover new biological pathways, including Pearson's

Correlation Coefficient (23), K-means (27), **F**actor **A**nalysis of **M**ixed **D**ata (FAMD) (28), Ward's minimum variance (Ward's) (29), **P**artitioning **A**round **M**edoids (PAM) (24) and K-prototypes (30).

We created guidelines that we could use to compare each method against each other to determine the most biologically relevant clustering. We first set a lower limit of three genes per cluster with a maximum of 40 clusters. We next calculated the average relative standard deviation for basal hyphae growth rate and aerial hyphae height for each cluster and then averaged across all clusters. For the categorical traits, we determined the most prevalent category for each cluster and determined an average across all clusters. These metrics allowed us to set guidelines for these continuous variables. We set a limit of less than 15% relative standard deviation and no less than 95% average consensus for the clustering.

We first tried methods that required a conversion of categorical data into numerical values. The two methods, Pearson's Correlation Coefficient (23) and K-means (27), provided relative standard deviations for basal hyphae growth rate and aerial hyphae height above the 15% cutoff. Based on the unsatisfactory results using the converted dataset, we turned to methods that would require little or no pre-processing of our data and that would retain the categorical data. One algorithm that can handle such mixed data is FAMD (28). However, this approach quickly failed our criteria, as the run with three total clusters contained one cluster with a single gene. Additionally, K-Prototypes (27) was found to be unstable and multiple runs would not converge on similar numbers of clusters. We next utilized the Daisy function (32) and Gower's metric (33) from the

"cluster" R package to create a dissimilarity matrix (r-project.org). We then used Ward's (24) or PAM (29) algorithms to generate clusters. Both methods did not initially meet our criteria (Table 1). However, with unequal weighting of each variable we found that these two methods could meet our criteria (Table 3.1). Overall, we found that PAM weight six with 40 clusters had the best combination of lowest relative standard error and highest percent consensus (Table 3.1).

The weighted PAM 40 method effectively separated mutants into distinct clusters, without formation of different clusters with identical phenotypes (Table 3.2). The average number of genes in each cluster was 29.2, and the median size was 14, with a range from 5 (Cluster 39) to 171 (Cluster 3) genes (Figure 3.3a). Most clusters contained 6–15 genes (Figure 3.3b). We subsequently analyzed the Weighted PAM 40 clusters for the presence of different classes of genes or other attributes, as described in the following sections.

The majority of cluster genes lack *Saccharomyces cerevisiae* orthologs. We next determined the percentage of *N. crassa* genes in each of the 40 clusters that have yeast orthologs (Figure 3.4). The range was 16% (13 genes out of 81 total) in Cluster 4 to 89% in Cluster 38 (8 genes out of 9 total) (Figure 3.4). In most clusters, a majority of genes lacked a yeast ortholog and in 21 clusters less than 50% of the genes possess yeast orthologs. Of note, the clusters with the highest proportion of yeast orthologs have reduced growth rates relative to wild type (Clusters 13, 30, 34, 38, and 40; Table 3.2), while those with the fewest yeast orthologs typically have normal or variable growth rate

phenotypes (Clusters 1, 4, 16 and 36; Table 3.2). The relative lack of shared phenotypes among clusters with a low percentage of orthologous yeast genes may reflect specialized functions that evolved within filamentous fungi and since the divergence from a common ancestor at the origin of the ascomycetes.

Two of the clusters containing the highest proportion of genes with yeast orthologs are Clusters 34 (9 genes of 12 total; 75%) and 40 (15 genes of 21 total; 71%). Cluster 34 contains the three genes for the p38 MAPK pathway, while Cluster 40 includes the three genes in each of the two ERK-class MAPK pathways. These are ancient and conserved pathways with central roles in the cellular biology of many eukaryotic organisms, including baker's yeast (34). The co-clustering of mutants in the two ERK pathways was expected, as they have similar growth and morphological defects. Overall, because genes lacking yeast orthologs predominate in our dataset, our analysis should reveal functions for genes found in *N. crassa* and other filamentous fungi that are not present in *S. cerevisiae*.

Enzymes and transmembrane proteins in clusters.

Enzymes. We used the list of genes compiled in previous work (31) to identify metabolic genes in the 40 clusters. Out of the 833 identified metabolic genes in the *N. crassa* genome, 88 were present in our dataset (11%). It should be noted that mutation of many metabolic genes results in auxotrophy in *N. crassa*, and since the knockout mutants were not cultured under conditions that would supplement all auxotrophs (11), those genes are not represented in our dataset. Therefore, it is likely that these 88 genes are

either essential under other nutritional conditions not used in our experiments and/or are functionally redundant with another gene(s).

A total of six clusters (2, 8, 11, 14, 29 and 35) consisted of at least 15% metabolic genes, as compared to a global average of 6.9%. All of these clusters except for Cluster 14 have reduced growth rates relative to wild type and, with the exception of Cluster 8, do not form ascospores (Table 2). These phenotypes may indicate diverse metabolic needs during hyphal growth and formation of the meiotic products, ascospores.

.....Predicted transmembrane proteins. We next identified proteins in the *N. crassa* genome and in clusters in our dataset that possess predicted transmembrane domains and/or are predicted G protein coupled receptors (GPCRs) (Figure 3.5a). The outcome of the analysis showed that 1796 *N. crassa* genes (18.5% of the total) encode a protein with at least one transmembrane domain. Of the genes represented by the knockout mutants in this study, 228 (19.5%) encode proteins with predicted transmembrane domains. A total of 15 clusters have 20% or more genes encoding predicted proteins with transmembrane domains, with five clusters exceeding 30% (Clusters 1, 5, 11, 19 and 36; Figure 3.5a). Within the transmembrane proteins, we also identified those that were predicted GPCRs (16). Cluster 1 contained the largest percentage of GPCR mutants (17 genes of 28 total or nearly 61% of the cluster), followed by Cluster 36, with 18.2% (6 genes of 33 total) GPCR mutants (Figure 3.5a).

There was no unifying group of phenotypes across the five clusters with the highest proportion of genes encoding transmembrane proteins (Table 3.2). However, 100% of the mutants in Clusters 19 and 36 possess tall aerial hyphae and 11 of 25 genes (44%) in

Cluster 19 encode a protein with a predicted transmembrane domain (Figure 3.5a). The 11 Cluster 19 genes encode either a transporter or an enzyme implicated in oxidative stress (Figure 3.5b). ROS stress and a hyperoxidant state have previously been shown to be important signals for tissue differentiation at three different stages of aerial hyphae development (33, 34). Oxidative stress releases repression by conidial separation-1 (csp-1; a transcription factor), allowing ergosterol gene expression, a condition associated with aerial hyphae development (35). The CSP-1 and CSP-2 transcription factors are both required for the formation of double-doublets during septation of conidiophores prior to production of the mature conidia (36). However, *csp-2* (NCU06095), not *csp-1* (NCU02713), clusters with the mutants with tallest aerial hyphae, suggesting different molecular functions for these two transcription factors during asexual development. Other genes within the cluster may have regulatory roles in pathways involved with aerial hyphae development. One such gene is the predicted 5′ to 3′ exonuclease *exr-1* (NCU01643), which may negatively regulate some mRNAs that promote aerial hyphae development.

Protein phosphorylation

Distribution of phosphorylated proteins. We used an available *N. crassa* phosphoproteome dataset (37) to ascertain whether the encoded proteins in the 40 clusters are phosphorylated when *N. crassa* is grown in liquid medium with glucose as the carbon source. Out of the 1168 genes in our study, 375 (32%) encode proteins that are phosphorylated, similar to the % in the entire genome (31%). Every cluster has at least

two phosphorylated proteins and 14% or more of the proteins in each cluster are phosphoproteins (Figure 3.6a). Clusters 10, 17, 27, 29, 30, 39 and 40 have 60% or more of their predicted proteins phosphorylated (Figure 3.6a). Growth rate defects were observed for six of these clusters, five clusters have a conidiation defect and ascospores were either reduced or not formed in four clusters. Clusters 30 and 39 have more than 80% phosphorylated targets. These two clusters share reduced/severely reduced growth rates and reduced/severely reduced conidial abundance (Table 3.2).

Serine/threonine protein kinases and serine/threonine or tyrosine protein phosphatases. There are 86 genes encoding Serine/Threonine Protein Kinases (S/T Kinases) in *N. crassa*, with 77 mutants represented in our study (19). Six clusters (14, 25, 30, 34, 39 and 40) have 20% or more S/T Kinase genes, with Clusters 30, 34 and 40 exceeding 40% (Figure 3.6b). It is known that mutants lacking the mitogen-activated protein kinase (MAPK), MAPK kinase (MAPKK) and MAPKK kinase (MAPKKK) for the same signaling pathway possess similar phenotypes in *N. crassa* (38-42). Thus, we expected to observe co-clustering of mutants for each of the three MAPK pathways. We noted that Cluster 34 contains the three kinases (os-2, NCU07024; os-4, NCU03071 and os-5, NCU00587) in the p38 MAPK osmosensing pathway (38, 39). Similarly, five of the six kinases (mak-1, NCU09842; mek-1, NCU06419 mak-2, NCU02393 mek-2, NCU04612 and nrc-1, NCU06182) that comprise the two Extracellular Signal-Regulated Kinase (ERK) MAPK pathways in *N. crassa* are in Cluster 40 (40-42). The MAPKKK for the cell integrity ERK pathway (mik-1, NCU02234) is in another cluster (Cluster 18), as the mutant has slightly taller aerial hyphae than the other two mutants in the same

pathway. Of note, Clusters 34 and 40 are among those with the highest proportion of kinase mutants (Figure 3.6b), suggesting possible regulatory interactions between MAPKs and the other kinases in the cluster.

The *N. crassa* genome contains 30 genes encoding serine-threonine or tyrosine protein phosphatases, and 24 of these are available knockout mutants and were included in our analysis (18). Protein phosphatases were not uniformly distributed throughout the clusters. There were numerous clusters (26 total) that lacked protein phosphatase genes (Figure 3.6b). In contrast, 40% (4 genes of 10 total) of the genes in Cluster 35 were protein phosphatases, and this cluster had reduced growth rate and reduced numbers of protoperithecia and perithecia.

Identification of cluster genes regulated by MAPKs in *N. crassa* and related fungi.

As mentioned above, we noted that with one exception, all three genes for each of the three MAPK cascades clustered together. Considering the importance of these three evolutionarily conserved MAPK pathways in fungi (43), we mined for potential targets in the clusters using a variety of publicly available datasets.

In order to identify genes regulated by the MAK-1 ERK MAPK, we first analyzed results from microarray analysis of the Δ*mak-1* mutant in *N. crassa* (44). Of the 424 genes down-regulated in the Δ*mak-1* mutant, 57 were included in our study (Figure 3.7). Cluster 40, which contains the three MAK-1 MAPK pathway kinases, has a relatively low percentage of genes that are targets (< 5%; Figure 3.7). In contrast, there were three clusters with > 15% MAK-1 target genes: Clusters 13 (2 genes of 12 total), 33 (1 gene of 6 total) and 39 (1 gene of 5 total). These three clusters share reduced/severely reduced

growth rates and varying degrees of sexual cycle defects (Table 4.2). These phenotypes are less severe examples of the phenotypes possessed by Δ*mak-1* mutants.

To determine putative targets of the ERK MAPK, MAK-2, we took advantage of two transcriptomic datasets, one for a Δ*mak-2* mutant (41) and the second for a strain expressing a *mak-2* inhibitable allele (*mak-2*^{Q100G}) grown in the presence of inhibitor (45). We also queried two phosphoproteomics datasets, both obtained after treatment of the *mak-2*^{Q100G} strain with inhibitor (46, 47). There were 39 genes in common between our phenotypic dataset and these four combined datasets (Figure 3.7). Clusters 8 and 3 contained more than 15% of the target genes. The only phenotype in common for these two clusters is reduced growth rate (Table 2). Cluster 40, containing the three kinases of the MAK-2 pathway, had 10% of its members as predicted targets.

There is currently no publicly available transcriptomic data for a Δ*os-2* or other p38 MAPK pathway mutant in *N. crassa*. Therefore, we turned to gene expression studies of the orthologous Δ*os-2* mutant in two filamentous fungi: microarray analysis of a Δ*hog1* mutant in Cryptococcus neoformans (48) and an RNA-seq experiment using an *Aspergillus fumigatus* Δ*sakA* mutant (49). Of the 401 genes regulated by *hog1* in *C. neoformans,* 296 had orthologs in *N. crassa* and 37 were represented in our phenotype dataset (Figure 3.7). The A. fumigatus RNA-seq analysis produced more hits, with 654 genes regulated by *sakA*, 574 of which have an ortholog in *N. crassa* and 96 that were included in our study (Figure 3.7). Analysis of these two combined datasets revealed that Cluster 19 had the highest percentage of p38 MAPK targets (7 genes of 25 total; 28%). This cluster also has the tallest aerial hyphae of any cluster. Other clusters with a

relatively high proportion of target genes (> 15%) were 6 (4 genes of 23 total), 21 (8 genes of 42 total), 26 (3 genes of 15 total) and 37 (2 genes of 13 total). These clusters all have reduced or severely reduced growth rates and reduced aerial hyphae height.

Osmotic Sensitive (OS) pathway mutants have reduced growth rates, but normal aerial hyphae height and do not produce female sexual structures. The lack of concentration of possible OS-2 pathway targets in Cluster 34 (containing the three OS-2 pathway kinases) supports a mechanism in which the diverse functions of the p38 MAPK pathway are carried out by different groups of genes. The observation of slower growth rate in mutants lacking p38 MAPK target genes is consistent *N. crassa os-2* pathway mutant defects (Table 2). It is of interest that of the clusters with a significant percentage of targets, only Cluster 37 has a sexual cycle phenotype, reduced ascospore production.

We also interrogated our dataset to identify genes that are targets of multiple MAPK pathways. A total of 188 genes encoded targets of at least one MAPK pathway. Of these, 18 are targets of two different MAPK pathways and one is a target of three pathways (Figure 3.7). The triple target is in Cluster 3, NCU03753, clock-controlled gene-1 (*ccg-1* (50, 51);. Cluster 3, which has wild type characteristics, also has the largest number of single MAPK targets (26 genes; Figure 3.7). The clusters with the greatest number of genes that are targets of two different MAPK pathways are 7 and 15 with three targets each (Figure 3.7). These last two clusters have similar phenotypes, with reduced growth rates (Table 3.2).

Transcriptional regulation

Distribution of transcription factors across phenotype clusters. *N. crassa* has 314 genes encoding transcription factors (17), with 242 (77% of the total) represented in our dataset. Transcription factors represent ~ 20.7% of the genes in our dataset and the average percentage of genes in a cluster that are transcription factors is 19.9%. Cluster 10 had the highest percentage of transcription factors, with 75% (Figure3.8) and mutants in this cluster have a severe reduction in hyphal growth rate (Table 3.2). Interestingly, none of the Cluster 10 genes are targets of the MAK-1, MAK-2 or OS-2 MAPKs (Figure 3.7).

Co-expression of genes in clusters. We utilized publicly available RNA-seq and microarray data sets for wild-type *N. crassa* to determine whether genes in the same phenotypic cluster are co-expressed. The two RNA-seq datasets are time courses during sexual development (52) or conidial germination (53), while the two microarray datasets are time courses during conidiation (54) or colony development (55). Expression data is available for ~ 99% of the cluster genes in the two RNA-seq datasets, but only 46.3% (54) and 49.7% (55) of the cluster genes were represented in the two microarray datasets. We did not analyze phenotypic clusters in which less than three genes had expression data.

Initial visual inspection of expression trends in each cluster using line plots showed few examples of co-expression. However, in clusters with more than five genes, it becomes difficult to visually determine expression patterns. Therefore, we attempted to fit linear, polynomial and sinusoidal models to the expression data for each phenotypic cluster. However, no model fit the data beyond an r-squared value of 0.2, supporting more than one expression pattern per cluster.

We next turned to K-means clustering to separate genes into discrete expression profiles for each mRNA time course and then compared the genes in these profiles to those in the 40 phenotypic clusters (See Methods). Using Within sum of squares (56), Gap statistic (57) and Davies-Bouldin index (58) as measures of cluster quality, we determined that 6–8 expression profiles provided sufficient quality clusters, depending on the dataset. We then determined the number of expression profiles that were present in each phenotypic cluster. We focused on phenotypic clusters in which one expression profile could be assigned to 40% or more of the genes in that cluster.

The genes for the sexual development (52), conidial germination (53), conidiation (54) and colony development time courses (55) were divided into seven, eight, seven and six different expression profiles, respectively (Figure 3.9a-d). The average number of expression profiles per cluster ranged from 4.1 to 6.1, revealing high diversity in expression profiles per cluster (Figure 3.10). When we focused on those clusters with a phenotype related to their respective datasets, the average number of profiles per cluster only decreased slightly, to a range of 3.9 to 5.9. In addition, there were no expression profiles that were exclusively correlated with a specific phenotype.

While the average number of expression profiles per cluster was high, there were clusters where more than 40% of genes shared the same expression profile. We investigated whether any of these dominant expression profiles were present in more than 50% of the phenotypic clusters. For sexual development and conidiation, no one expression profile was represented in more than 50% of clusters with a dominant expression profile (Figure 3.10). However, for conidial germination and colony

development there were expression profiles representing 77 and 86%, respectively, of the dominant expression profiles in the clusters (Figure 3.10). The two dominant expression profiles were expression profile 3 during conidial germination (peak expression early) and expression profile 6 during colony development (peak expression at the hyphal tip) (Figure 3.9a,c). Whereas there was little correlation between any expression profile and phenotype for sexual development and conidiation, there was some correlation between defects in growth rate and elevated expression at the hyphal tip or early during conidial germination, as those phenotypes and expression profiles most often grouped together.

..... Cluster genes co-expressed across multiple datasets Interestingly, three phenotypic clusters (14, 30 and 33) contain several genes that are consistently co-regulated across multiple datasets. In Cluster 14, the transcription factors *bek-1* (NCU00097) (59) and *bek-2* (NCU07139), the guanosine diphosphatase *gda-1* (NCU03713), hypothetical protein NCU06390 and the S/T kinase *stk-53* (NCU09064) share the same expression profile during conidial germination and sexual development. In Cluster 30, the transcription factors *vel* (NCU00406) and *ada-19* (NCU04459) are co-expressed during conidiation and conidial germination. In Cluster 33, four genes are co-expressed during sexual development and conidial germination: two signaling genes (the phosphatase *tng* (NCU0436) and kinase *div-4* (NCU04426), a chromatin remodeling factor *crf4–1* (NCU03875) and a gene implicated in mRNA stability (NCU07874). With regards to MAPK signaling, the p38 MAPKK *os-5* and MAPKKK *os-4* are co-expressed during colony development and conidiation. *mak-1* and *mak-2* are co-expressed in every time course and *os-2* is co-expressed with *mak-1* and *mak-2* during colony development.

Taken together, the results from our analysis of transcriptional data indicate that in general, no single expression profile correlates with a given phenotype. These results support earlier observations that elevated expression of a gene during a developmental time course does not necessarily correlate with an observable phenotype for the corresponding mutant in *N. crassa* (60, 61).

.

Pathway prediction utilizing yeast ortholog interaction data

In order to determine possible targets and interactors in the three MAPK pathways, we identified genetic and physical interactions between all yeast orthologs of genes in the two clusters. Of the genes that co-clustered with the two ERK MAPK pathways, 11 of 16 had a yeast ortholog. All but one yeast ortholog has evidence of either genetic and/or physical interactions with the other genes in the cluster. Based on known genetic and physical interactions in yeast, we predict that there may be two pathways associated with the two ERK MAPK signaling cascades (Figure 3.11). One possible pathway runs from MAK-2 to NGF-1 through a negative genetic interaction. NGF-1 then has various genetic and physical interactions with CAMK-1, RCO-1 and ADA-20 (Figure 3.11). The other pathway flows from MAK-1 to PRK-2 and MDK-2 through negative genetic interactions (Figure 3.11). There are five cluster genes that do not have a yeast ortholog, including two serine/threonine kinases (*stk-31* and *stk-36*), an endothiapepsin (*apr-10*), a non-anchored cell wall protein (*ncw-3*) and a protein with a tropomyosin domain (*ro-11*). These genes may represent additional pathways and/or interactors downstream of the ERK MAPKs that are not found in *S. cerevisiae*.

Nine of the 12 genes in Cluster 34, which contains the genes in the p38 MAPK cascade, have a yeast ortholog. There is substantial evidence for interaction between the yeast orthologs in this cluster. The data also suggests a large amount of crosstalk between the MAPK pathway and three genes that co-clustered (Figure 3.12); orthologs of *hda-2, div-59,* and *stk-47* each exhibit a genetic relationship with at least two out of three of the genes in the p38 MAPK cascade (Figure 3.12). Interestingly, there was one gene with a yeast homolog (*amyc*) that does not interact with any orthologs of Cluster 24 genes in yeast. There are three genes with no yeast ortholog in the cluster; a serine/threonine kinase (*stk-46*), a transcription factor (*ff-7*) and a mago nashi protein (*mrs-4*). These last four genes may be unique targets of the p38 MAPK pathway in *N. crassa*.

Discussion

In this study, we took advantage of a large dataset of phenotypic data for 10 traits obtained from 1168 knock out mutants in *N. crassa*. The distribution of genes across linkage groups and functional categories was representative of the entire genome. We found that the majority of genes lack an orthologous gene in yeast and, therefore, many relationships and genes explored in this study cannot be studied in S. cerevisiae. With this broadly representative dataset, we tested several algorithms and two distance measures to determine the optimal way to cluster our phenotypic data. We developed a novel approach to measure cluster quality, using relative standard deviation for each quantitative trait and the average percent consensus (the average percent of the most

prevalent category per categorical trait), in order to determine biological relevance of the clusters.

Previous studies have focused on using a single hierarchical clustering method for analyzing phenotypic data. For example, in *S. cerevisiae*, data was collected for 4281 gene deletion mutants subjected to 51 different environmental stressors (20). All phenotypes were quantitative and were clustered using two-way unsupervised, un-centered hierarchical agglomerative clustering, with a distance matrix generated using Pearson's correlation (20). In *Fusarium graminearum*, 17 different phenotypes were collected for 657 transcription factor knockout mutants (21). Continuous and categorical phenotypes were converted to a number scale and correlations between phenotypes were calculated using Pearson's correlation coefficient. In *Cryptococcus neoformans*, a study analyzed 30 different phenotypes for 129 kinase mutants (22). Categorical and continuous phenotypes were converted to a numeric value. A distance matrix was generated with Pearson's correlation minus one and a hierarchical agglomerative clustering algorithm was applied to generate clusters (22).

Our initial attempts to cluster relied on conversion of ordinal traits to a numerical scale. However, only four of the eight categorical phenotypes in our dataset were amenable to this conversion. Using the converted dataset, both a hierarchical method (Hierarchical Agglomerative Clustering (22)) and a partitional approach (K-means) (25) were tested for their ability to cluster genes, with unsatisfactory results in terms of large standard deviations. Therefore, we turned to distance matrices and algorithms that can utilize categorical data. Of these, the algorithm that performed the best according to our

criteria was Partitioning Around Medoids (PAM). For optimal clustering, we weighted the data in the distance matrix before applying PAM. To our knowledge, PAM has not previously been used to cluster phenotypic data in fungi.

We identified two clusters in which more than 40% of the genes in the cluster encode proteins with predicted transmembrane domains. One of the clusters, Cluster 1, is comprised mostly of G-protein coupled receptors, which have been previously shown to regulate aerial hyphae height (16). In the other cluster (Cluster 19), a majority of genes are involved in oxidative stress regulation or ion transport. ROS stress and a hyperoxidant state are important signals during aerial hyphae development (33, 34). Additionally, turgor pressure has been shown to be vital for hyphal extension (62). Indeed, the same forces must be applied to grow vertically as well, with a combination of turgor pressure and secretion of hydrophobins (*eas*, NCU08457) (63). Thus, Cluster 19 genes highlight an interesting correlation between membrane-associated proteins and aerial hyphae height in *N. crassa*.

We investigated whether gene expression was correlated with some or all of the phenotypes we observed. There was some correlation between expression at hyphal tips and defects in hyphal growth rate. However, results using most of the gene expression datasets did not reveal a strong correlation between an expression profile and a phenotype. The lack of correlation is not particularly surprising, as it has been observed that mutation of several genes that are highly expressed during conidiation did not lead to a phenotype during conidiation (60, 61). However, the lack of correlation could also be explained by differences in experimental conditions between our phenotypic analysis and

those used to isolate RNA for transcriptional profiling. There are a limited number of time-course expression datasets that are publicly available for *N. crassa*, and there was no dataset with conditions or developmental processes that perfectly matched those in our study.

Mitogen-activated protein kinase (MAPK) pathways are conserved in eukaryotes, from unicellular organisms such as yeast, to metazoans, including humans (64, 65). In animals, MAPKs regulate gene expression, mitosis, movement, metabolism, and programmed cell death (64). In plants, MAPK cascades control cell division, growth and development, and regulate resistance to pathogens, insect herbivores, and abiotic stress (66). In fungi, including saprobes and pathogens, MAPKs have been found to be involved in control of morphogenesis, development, virulence, cell wall biogenesis and stress responses (43, 67). In *N. crassa*, mutation of any of the three genes for each of the three MAPK cascades leads to severe defects in growth rate, asexual development, sexual development and failure to form conidial anastomosis tubes (19, 68). In addition, the MAK-1 and MAK-2 pathways are both involved in cell wall integrity, with the MAK-1 pathway more important for this process (42). Mutants lacking any of the three genes in the OS-2 MAPK pathway have near-normal growth rates, but possess defects in conidiation and sexual development (19).

Analysis of transcriptional and proteomic targets of the three MAPK signaling pathways revealed that targets are distributed throughout the clusters, supporting the idea that the functions of these pathways are carried out by different groups of genes. We also determined that several MAPK targets are actually genes/proteins in a different MAPK

cascade. This is not surprising, as MAPK pathways are not necessarily linear and independent, and there is evidence showing significant amounts of crosstalk in mammalian cell lines (69, 70). In *S. cerevisiae*, crosstalk has been shown between the osmosensing (HOG1) and the cell wall integrity MAPK pathways during heat stress and mating in specific mutant backgrounds (71, 72). Crosstalk has also been reported in *Aspergillus fumigatus* between the cell wall integrity and spore development pathways (73). Interestingly, we found that across all four RNA expression datasets, two of the three terminal MAPKs in *N. crassa* (*mak-1* and *mak-2*) were always co-expressed, implying some amount of co-regulation. To our knowledge, no other study has reported transcriptional co-expression of these genes.

We took advantage of the extensive interaction data available for *S. cerevisiae* to build predicted pathways for the genes that co-clustered with the ERK and p38 MAPKs. We discovered that all genes with a yeast homolog in the cluster containing the two ERK MAPK cascades showed an interaction with at least one other gene in the cluster. For the genes that co-clustered with the p38 MAPK cascade, all but one interacted with at least one other gene in the cluster. Interestingly, there was linear flow through the ERK MAPK signaling cascades, with the MAPKKK and MAPKKs only interacting with the MAPKK or MAPK from the same or the second pathway. The rest of the interacting genes formed a network that emanated from one of the terminal MAPKS (MAK-1 or MAK-2) (Figure 3.). In contrast to the interactions observed in the ERK MAPK cluster, there were three genes (*hda-2*, *div-59* and *stk-47*) that interacted with at least two of the three core kinase genes in the p38 MAPK cascade and also exhibited many interactions with each

other (Figure 3.12). Further work is needed to confirm these models, including probing physical interactions through the yeast two-hybrid assay or pull-down approaches, as well as tests of genetic epistasis.

References

1. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, et al. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev. 2004;68(1):1–108.

2. Selker EU. Neurospora. Curr Biol. 2011;21(4):R139–40.

3. Dean R, Van Kan JA, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, et al. The top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol. 2012;13(4):414–30.

4. Richardson M, Bowyer P, Sabino R. The human lung and Aspergillus: You are what you breathe in? Med Mycol. 2019;57(Supplement_2):S145–S54.

5. Weiss M, Waller F, Zuccaro A, Selosse MA. Sebacinales - one thousand and one interactions with land plants. New Phytol. 2016;211(1):20–40.

6. Glass NL, Schmoll M, Cate JH, Coradetti S. Plant cell wall deconstruction by ascomycete fungi. Annu Rev Microbiol. 2013;67:477–98.

7. Cairns TC, Zheng X, Zheng P, Sun J, Meyer V. Moulding the mould: understanding and reprogramming filamentous fungal growth and morphogenesis for next generation cell factories. Biotechnol Biofuels. 2019;12:77.

8. Park HS, Jun SC, Han KH, Hong SB, Yu JH. Diversity, application, and synthetic biology of industrially important Aspergillus Fungi. Adv Appl Microbiol. 2017;100:161–202.

9. Perkins DD, Radford A, Sachs MS. The Neurospora Compendium. San Diego: Chromosomal Loci. Academic Press. 2001. p. 325.

10. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus Neurospora crassa. Nature. 2003;422(6934):859–68.

11. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, et al. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A. 2006; 103(27):10352–7.

12. Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, et al. Enabling a community to dissect an organism: overview of the Neurospora functional genomics project. Adv Genet. 2007;57:49–96.

13. McCluskey K, Wiest A, Plamann M. The fungal genetics stock center: a repository for 50 years of fungal genetics research. J Biosci. 2010;35(1):119–26.

14. Collopy PD, Colot HV, Park G, Ringelberg C, Crew CM, Borkovich KA, et al. High-throughput construction of gene deletion cassettes for generation of Neurospora crassa knockout strains. Methods Mol Biol. 2010;638:33–40.

15. Turner GE. Phenotypic analysis of Neurospora crassa gene deletion strains. Methods Mol Biol. 2011;722:191–8.

16. Cabrera IE, Pacentine IV, Lim A, Guerrero N, Krystofova S, Li L, et al. Global Analysis of Predicted G Protein-Coupled Receptor Genes in the Filamentous Fungus, Neurospora crassa. G3 (Bethesda). 2015;5(12):2729–43.

17. Carrillo AJ, Schacht P, Cabrera IE, Blahut J, Prudhomme L, Dietrich S, et al. Functional Profiling of Transcription Factor Genes in Neurospora crassa. G3 (Bethesda). 2017;7(9):2945–56.

18. Ghosh A, Servin JA, Park G, Borkovich KA. Global analysis of serine/threonine and tyrosine protein phosphatase catalytic subunit genes in Neurospora crassa reveals interplay between phosphatases and the p38 mitogenactivated protein kinase. G3 (Bethesda). 2014;4(2):349–65.

19. Park G, Servin JA, Turner GE, Altamirano L, Colot HV, Collopy P, et al. Global analysis of serine-threonine protein kinase genes in Neurospora crassa. Eukaryot Cell. 2011;10(11):1553–64.

20. Brown JA, Sherlock G, Myers CL, Burrows NM, Deng C, Wu HI, et al. Global analysis of gene function in yeast by quantitative phenotypic profiling. Mol Syst Biol. 2006;2:2006 0001.

21. Son H, Seo YS, Min K, Park AR, Lee J, Jin JM, et al. A phenome-based functional analysis of transcription factors in the cereal head blight fungus, Fusarium graminearum. PLoS Pathog. 2011;7(10):e1002310.

22. Lee KT, So YS, Yang DH, Jung KW, Choi J, Lee DG, et al. Systematic functional analysis of kinases in the fungal pathogen Cryptococcus neoformans. Nat Commun. 2016;7:12766.

23. Freedman D, Pisani R, Purves R. Statistics. 4. Pisani R, Purves R, editors: WW Norton & Company, New York; 2007.

24. Schubert E, Rousseeuw PJ. Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms. In: Amato G, Gennaro C, Oria V, Radovanović M. (eds) Similarity Search and Applications. SISAP 2019. Lecture Notes in Computer Science. Cham: Springer. 2019;11807:171–87.

25. Hartigan JA, Wong MA. Algorithm AS 136: a K-means clustering algorithm. Appl Stat. 1979;28:100–8.

26. Audigier V, Husson F, Josse J. A principal component method to impute missing values for mixed data. Adv Data Anal Classif. 2016;10:5.

27. Szekely G, Rizzo M. Hierarchical clustering via joint between-within distances: extending Ward's minimum variance method. J Classif. 2005;22:151.

28. Huang H. Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Min Knowl Disc. 1998;2:283–304.

29. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0; 2019.

30. Gower JC. A general coefficient of similarity and some of its properties. Biometrics. 1971;27:857–71.

31. Rispail N, Soanes DM, Ant C, Czajkowski R, Grunler A, Huguet R, et al. Comparative genomics of MAP kinase and calcium-calcineurin signalling components in plant and human pathogenic fungi. Fungal Genet Biol. 2009;46(4):287–98.

32. Dreyfuss JM, Zucker JD, Hood HM, Ocasio LR, Sachs MS, Galagan JE. Reconstruction and validation of a genome-scale metabolic model for the filamentous fungus Neurospora crassa using FARM. PLoS Comput Biol. 2013; 9(7):e1003126.

33. Hansberg W, Aguirre J. Hyperoxidant states cause microbial cell differentiation by cell isolation from dioxygen. J Theor Biol. 1990;142(2): 201–21.

34. Peraza L, Hansberg W. Neurospora crassa catalases, singlet oxygen and cell differentiation. Biol Chem. 2002;383(3–4):569–75.

35. Sancar G, Sancar C, Brugger B, Ha N, Sachsenheimer T, Gin E, et al. A global circadian repressor controls antiphasic expression of metabolic genes in Neurospora. Mol Cell. 2011;44(5):687–97.

36. Springer ML, Yanofsky C. A morphological and genetic analysis of conidiophore development in Neurospora crassa. Genes Dev. 1989;3(4):559–71.

37. Horta MAC, Thieme N, Gao Y, Burnum-Johnson KE, Nicora CD, Gritsenko MA, Lipton MS, Mohanraj K, de Assis LJ, Lin L, Tian C, Braus GH, Borkovich KA, Schmoll M, Larrondo LF, Samal A, Goldman GH, Benz JP. Broad Substrate-Specific Phosphorylation Events Are Associated With the Initial Stage of Plant Cell Wall Recognition in Neurospora crassa. Front Microbiol. 2019;10:2317,1-21.

38. Zhang Y, Lamm R, Pillonel C, Lam S, Xu JR. Osmoregulation and fungicide resistance: the Neurospora crassa os-2 gene encodes a HOG1 mitogenactivated protein kinase homologue. Appl Environ Microbiol. 2002;68(2): 532–8.

39. Park G, Pan S, Borkovich KA. Mitogen-activated protein kinase cascade required for regulation of development and secondary metabolism in Neurospora crassa. Eukaryot Cell. 2008;7(12):2113–22.

40. Pandey A, Roca MG, Read ND, Glass NL. Role of a mitogen-activated protein kinase pathway during conidial germination and hyphal fusion in Neurospora crassa. Eukaryot Cell. 2004;3(2):348–58.

41. Li D, Bobrowicz P, Wilkinson HH, Ebbole DJ. A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in Neurospora crassa. Genetics. 2005;170(3):1091–104.

42. Maerz S, Ziv C, Vogt N, Helmstaedt K, Cohen N, Gorovits R, et al. The nuclear Dbf2-related kinase COT1 and the mitogen-activated protein kinases MAK1 and MAK2 genetically interact to regulate filamentous growth, hyphal fusion and sexual development in Neurospora crassa. Genetics. 2008;179(3): 1313–25.

43. Zhao X, Mehrabi R, Xu JR. Mitogen-activated protein kinase pathways and fungal pathogenesis. Eukaryot Cell. 2007;6(10):1701–14.

44. Bennett LD, Beremand P, Thomas TL, Bell-Pedersen D. Circadian activation of the mitogen-activated protein kinase MAK-1 facilitates rhythms in clockcontrolled genes in Neurospora crassa. Eukaryot Cell. 2013;12(1):59–69.

45. Leeder AC, Jonkers W, Li J, Glass NL. Early colony establishment in Neurospora crassa requires a MAP kinase regulatory network. Genetics. 2013; 195(3):883–98. Carrillo et al. BMC Genomics (2020) 21:755 Page 21 of 22

46. Jonkers W, Leeder AC, Ansong C, Wang Y, Yang F, Starr TL, et al. HAM-5 functions as a MAP kinase scaffold during cell fusion in Neurospora crassa. PLoS Genet. 2014;10(11):e1004783.

47. Jonkers W, Fischer MS, Do HP, Starr TL, Glass NL. Chemotropism and cell fusion in Neurospora crassa relies on the formation of distinct protein complexes by HAM-5 and a novel protein HAM-14. Genetics. 2016;203(1): 319–34.

48. Ko YJ, Yu YM, Kim GB, Lee GW, Maeng PJ, Kim S, et al. Remodeling of global transcription patterns of Cryptococcus neoformans genes mediated by the stress-activated HOG signaling pathways. Eukaryot Cell. 2009;8(8):1197–217.

49. Pereira Silva L, Alves de Castro P, Dos Reis TF, Paziani MH, Von Zeska Kress MR, Riaño-Pachón DM, Hagiwara D, Ries LN, Brown NA, Goldman GH. Genome-wide transcriptome analysis of Aspergillus fumigatus exposed to osmotic stress reveals regulators of osmotic and cell wall stresses that are SakAHOG1 and MpkC dependent. Cell Microbiol. 2017;19:e12681,1-21.

50. Bell-Pedersen D, Dunlap JC, Loros JJ. The Neurospora circadian clockcontrolled gene, ccg-2, is allelic to eas and encodes a fungal hydrophobin required for formation of the conidial rodlet layer. Genes Dev. 1992;6(12A): 2382–94.

51. Lauter FR, Russo VE, Yanofsky C. Developmental and light regulation of eas, the structural gene for the rodlet protein of Neurospora. Genes Dev. 1992; 6(12A):2373–81.

52. Wang Z, Lopez-Giraldez F, Lehr N, Farre M, Common R, Trail F, et al. Global gene expression and focused knockout analysis reveals genes associated with fungal fruiting body development in Neurospora crassa. Eukaryot Cell. 2014;13(1):154–69.

53. Wang Z, Miguel-Rojas C, Lopez-Giraldez F, Yarden O, Trail F, Townsend JP. Metabolism and Development during Conidial Germination in Response to a Carbon-Nitrogen-Rich Synthetic or a Natural Source of Nutrition in Neurospora crassa. mBio. 2019;10(2):e00192-19.

54. Greenwald CJ, Kasuga T, Glass NL, Shaw BD, Ebbole DJ, Wilkinson HH. Temporal and spatial regulation of gene expression during asexual development of Neurospora crassa. Genetics. 2010;186(4):1217–30.

55. Kasuga T, Glass NL. Dissecting colony development of Neurospora crassa using mRNA profiling and comparative genomics approaches. Eukaryot Cell. 2008;7(9):1549–64.

56. Lloyd SP. Least squares quantization in PCM. IEEE Trans Inf Theory. 1982;28: 129–36.

57. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. Royal Stat Soc. 2002;63:411–23.

58. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell. 1979;1(2):224–7.

59. Krystofova S, Borkovich KA. The predicted G-protein-coupled receptor GPR-1 is required for female sexual development in the multicellular fungus Neurospora crassa. Eukaryot Cell. 2006;5(9):1503–16.

60. Berlin V, Yanofsky C. Isolation and characterization of genes differentially expressed during conidiation of Neurospora crassa. Mol Cell Biol. 1985;5(4): 849–55.

61. Madi L, Ebbole DJ, White BT, Yanofsky C. Mutants of Neurospora crassa that alter gene expression and conidia development. Proc Natl Acad Sci U S A. 1994;91(13):6226–30.

62. Lew RR. How does a hypha grow? The biophysics of pressurized growth in fungi. Nat Rev Microbiol. 2011;9(7):509–18.

63. Money NP. Fungal cell biology and development. In: S.C. Watkinson LB, N. Money, editor. The Fungi. Waltham: Academic Press-Elsevier Ltd.; 2016. p. 37–66.

64. Cargnello M, Roux PP. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. Microbiol Mol Biol Rev. 2011; 75(1):50–83.

65. Johnson GL, Lapadat R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. Science. 2002;298(5600): 1911–2.

66. Liu Y, He C. A review of redox signaling and the control of MAP kinase pathway in plants. Redox Biol. 2017;11:192–204.

67. Hamel LP, Nicole MC, Duplessis S, Ellis BE. Mitogen-activated protein kinase signaling in plant-interacting fungi: distinct messages from conserved messengers. Plant Cell. 2012;24(4):1327–51.

68. Fu C, Iyer P, Herkal A, Abdullah J, Stout A, Free SJ. Identification and characterization of genes required for cell-to-cell fusion in Neurospora crassa. Eukaryot Cell. 2011;10(8):1100–9.

69. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 2002;12(1):9–18.

70. Fey D, Croucher DR, Kolch W, Kholodenko BN. Crosstalk and signaling switches in mitogen-activated protein kinase cascades. Front Physiol. 2012;3: 355.

71. Dunayevich P, Baltanas R, Clemente JA, Couto A, Sapochnik D, Vasen G, et al. Heat-stress triggers MAPK crosstalk to turn on the hyperosmotic response pathway. Sci Rep. 2018;8(1):15168.

72. O'Rourke SM, Herskowitz I. The Hog1 MAPK prevents cross talk between the HOG and pheromone response MAPK pathways in Saccharomyces cerevisiae. Genes Dev. 1998;12(18):2874–86.

73. Manfiolli AO, Siqueira FS, Dos Reis TF, Van Dijck P, Schrevens S, Hoefgen S, et al. Mitogen-Activated Protein Kinase Cross-Talk Interaction Modulates the Production of Melanins in Aspergillus fumigatus. mBio. 2019;10(2):e00215-19.

74. Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, et al. FungiDB: an integrated functional genomics database for fungi. Nucleic Acids Res. 2012; 40(Database issue):D675–81.

75. Basenko EY, Pulman JA, Shanmugasundram A, Harb OS, Crouch K, Starns D, et al. FungiDB: An Integrated Bioinformatic Resource for Fungi and Oomycetes. J Fungi (Basel). 2018;4(1):39,1-28.

76. Dharmananda S, Feldman JF. Spatial distribution of circadian clock phase in aging cultures of Neurospora crassa. Plant Physiol. 1979;63(6):1049–54.

77. White B, Woodward D. A simple method for making disposable race tubes. Fungal Genet Newslett. 1995;42:79.

78. Davis RH, de Serres FJ. Genetic and microbiological research techniques for Neurospora crassa. Methods Enzymol. 1970;71A:79–143.

79. Kassambara A, Mundt F. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.5; 2017.

80. Le S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008;25:1–18.

81. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, et al. Saccharomyces genome database: the genomics resource of budding yeast. Nucleic Acids Res. 2012;40(Database issue):D700–5.

Figure legends

Figure 3.1: Summary of results from phenotypic analysis of 1286 N. crassa gene replacement mutants.

A. Flow chart for phenotypic analysis. The number and features of mutants analyzed at each step of the process are noted. The percentage of "No phenotype" and "With phenotype(s)" mutants refers to strains with complete data. **B.** Major phenotypic classes of mutants. The distribution of mutants with phenotypes in at least one of the three major categories (growth rate, asexual development and sexual development) is shown in the lobes of the Venn diagram. The intersection of lobes indicates mutants with defects in more than one major category

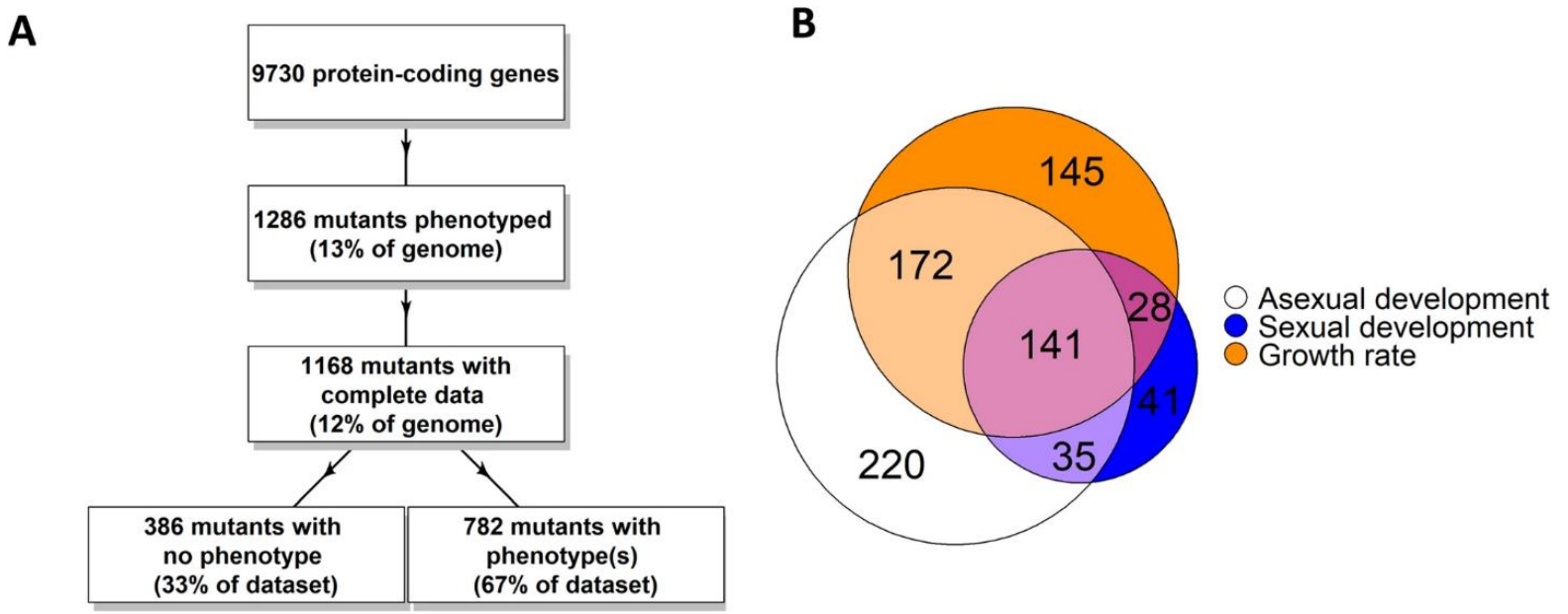

A
9730 protein-coding genes
1286 mutants phenotyped
(13% of genome)
1168 mutants with
complete data
(12% of genome)
386 mutants with
no phenotype
(33% of dataset)
782 mutants with
phenotype(s)
(67% of dataset)
B
145
172
28
141
41
35
220
Asexual development
Sexual development
Growth rate

Figure 3.2: Distribution of genes by linkage group and functional catalogue assignments.

A. Genes on different linkage groups. The distribution of all *N. crassa* genes on different linkage groups and the genes in each linkage group deleted in knockout mutants with complete phenotypic data are shown. **B.** Genes by functional catalogue (FunCat) assignment. The distribution of FunCat assignments for all genes in the *N. crassa* genome and the genes deleted in knockout mutants with complete phenotypic data is shown. Functional category data are available at https://elbe.hki-jena.de/fungifun/fungifun

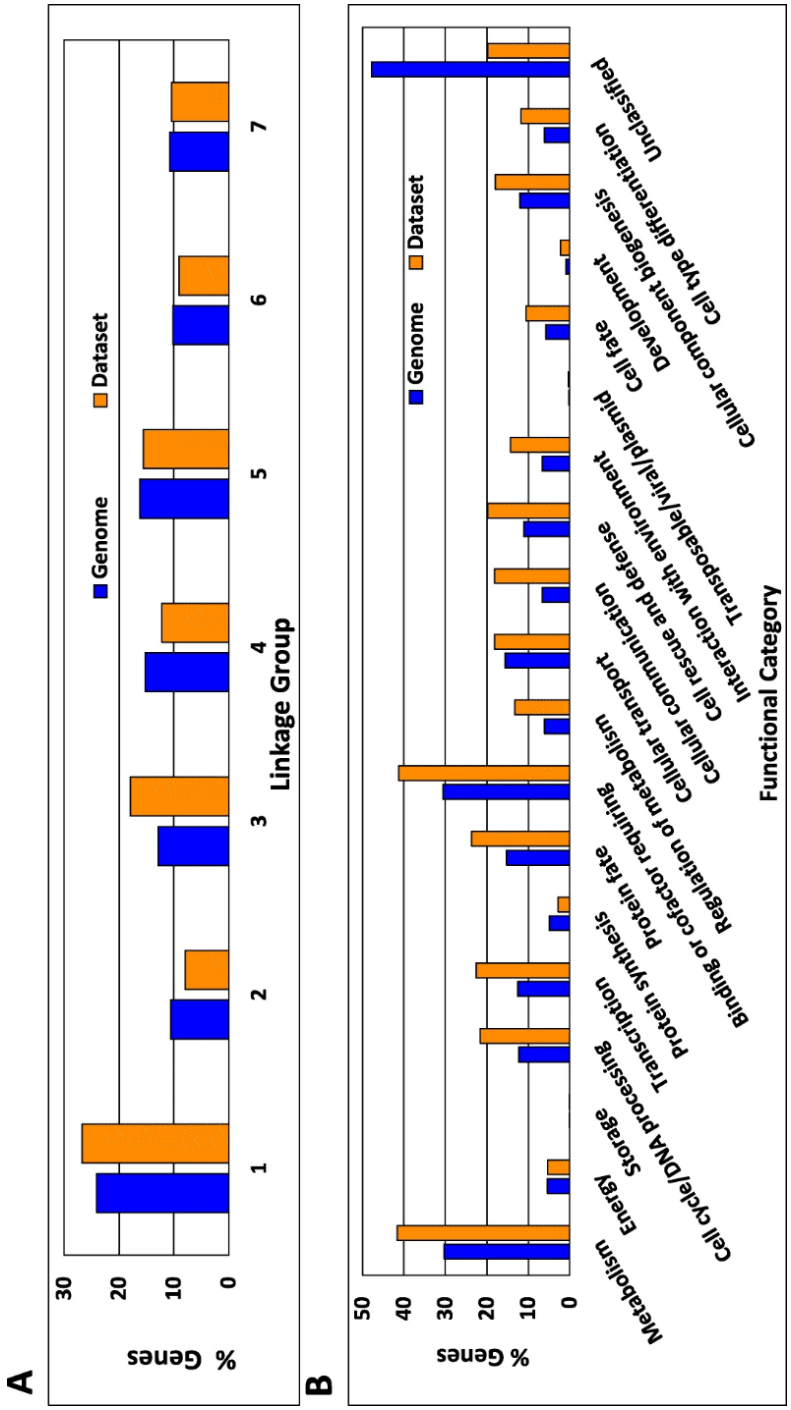

A
% Genes
30
20
10
0
Genome
Dataset
1
2
3
4
5
6
7
Linkage Group
B
% Genes
50
40
30
20
10
0
Genome
Dataset
Metabolism
Energy
Storage
Cell cycle/DNA processing
Transcription
Protein synthesis
Protein fate
Binding or cofactor requiring
Regulation of metabolism
Cellular transport
Cell rescue and defense
Interaction with environment
Transposable/viral/plasmid
Cell fate
Development biogenesis
Cell type differentiation
Unclassified
Functional Category

Figure 3.3: Grouping of mutants in clusters.

A. Number of genes per cluster. The number of mutants in each cluster is shown. The clustering algorithm was a weighted PAM. See Methods for details. The numbers above the bars indicate the number of genes per cluster. **B.** Number of clusters in different size ranges. Clusters were sorted into bins according to the number of genes in the cluster. The number of clusters corresponding to each size range (bin) is shown. The numbers above the bars indicates the number of clusters in each size range.

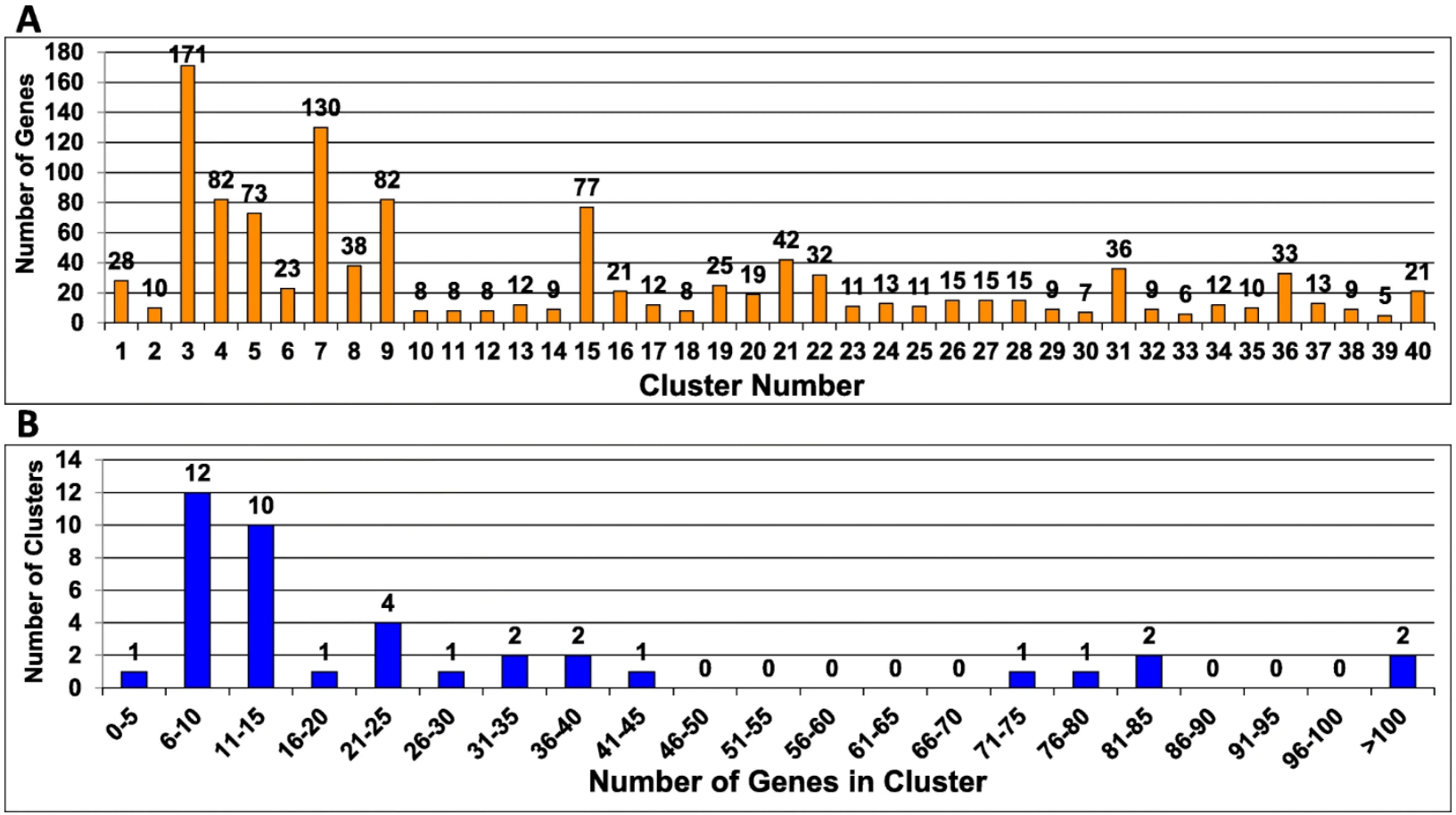

A
Number of Genes
180
160
140
120
100
80
60
40
20
0
28 10 171 82 73 23 130 38 82 8 8 8 12 9 77 21 12 8 25 19 42 32 11 13 11 15 15 15 9 7 36 9 6 12 10 33 13 9 5 21
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
Cluster Number
B
Number of Clusters
14
12
10
8
6
4
2
0
1 12 10 1 4 1 2 2 1 0 0 0 0 0 1 1 2 0 0 0 2
0-5 6-10 11-15 16-20 21-25 26-30 31-35 36-40 41-45 46-50 51-55 56-60 61-65 66-70 71-75 76-80 81-85 86-90 91-95 96-100 >100
Number of Genes in Cluster

Figure 3.4: Placement of genes with yeast orthologs in clusters.

The genes deleted in the mutants were analyzed for *S. cerevisiae* orthologs using the "Transform by Orthology" tool at the FungiDB database. The percentage of genes with orthologs in each cluster (blue bars), and in the entire dataset (orange bar), are shown. The numbers above the bars indicate the percent of genes in a cluster with a yeast ortholog.

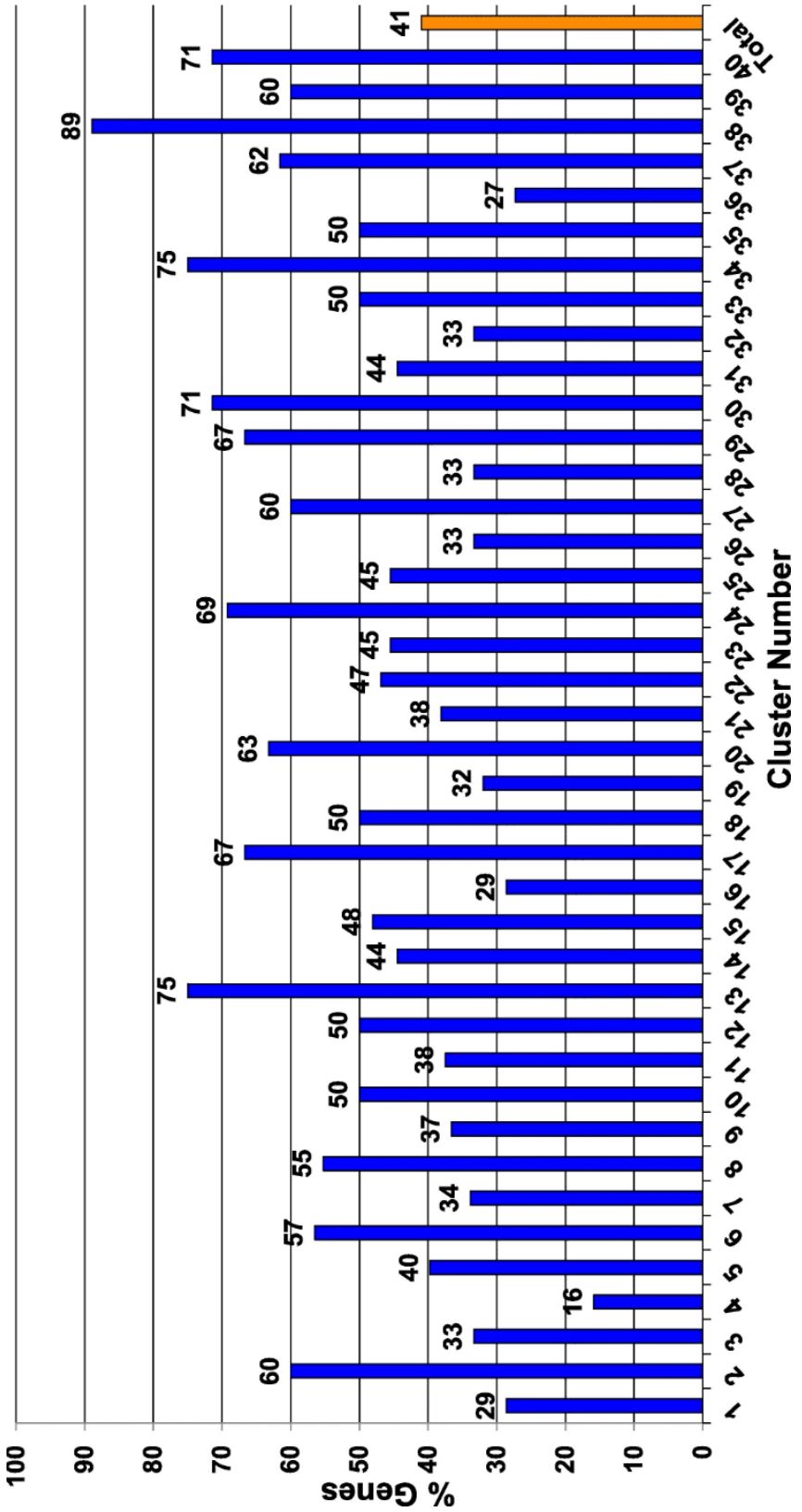

% Genes
Cluster Number
100
90
80
70
60
50
40
30
20
10
0
29
60
33
16
40
57
34
55
37
50
38
50
75
44
48
29
67
50
32
63
38
47
45
69
45
25
33
60
33
67
71
44
33
50
75
50
27
62
89
60
71
41
Total
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

Figure 3.5: Genes encoding proteins with predicted transmembrane domain(s) and/or G protein coupled receptors (GPCRs).

A. Transmembrane domains and GPCRs. The percentage of genes in each cluster encoding a protein with at least one predicted transmembrane domain is shown. The orange portion of each bar denotes the percent of genes in each cluster encoding transmembrane proteins that are also predicted G protein coupled receptors (GPCRs). Genes were retrieved using the "Protein targeting and localization - #TM Domains" tool at FungiDB. The numbers on the orange portion of the bars indicate the percent of proteins in a cluster that are GPCRs. The numbers on the blue portion of the bars indicates the percent of proteins in a cluster that have a predicted transmembrane domain.

B. Functional categories of proteins with transmembrane domains in Cluster 19. Each "slice" of the pie represents the percentage of genes encoding proteins with transmembrane domains in the indicated functional category. Functional category data are available at https://elbe.hki-jena.de/fungifun/fungifun.php

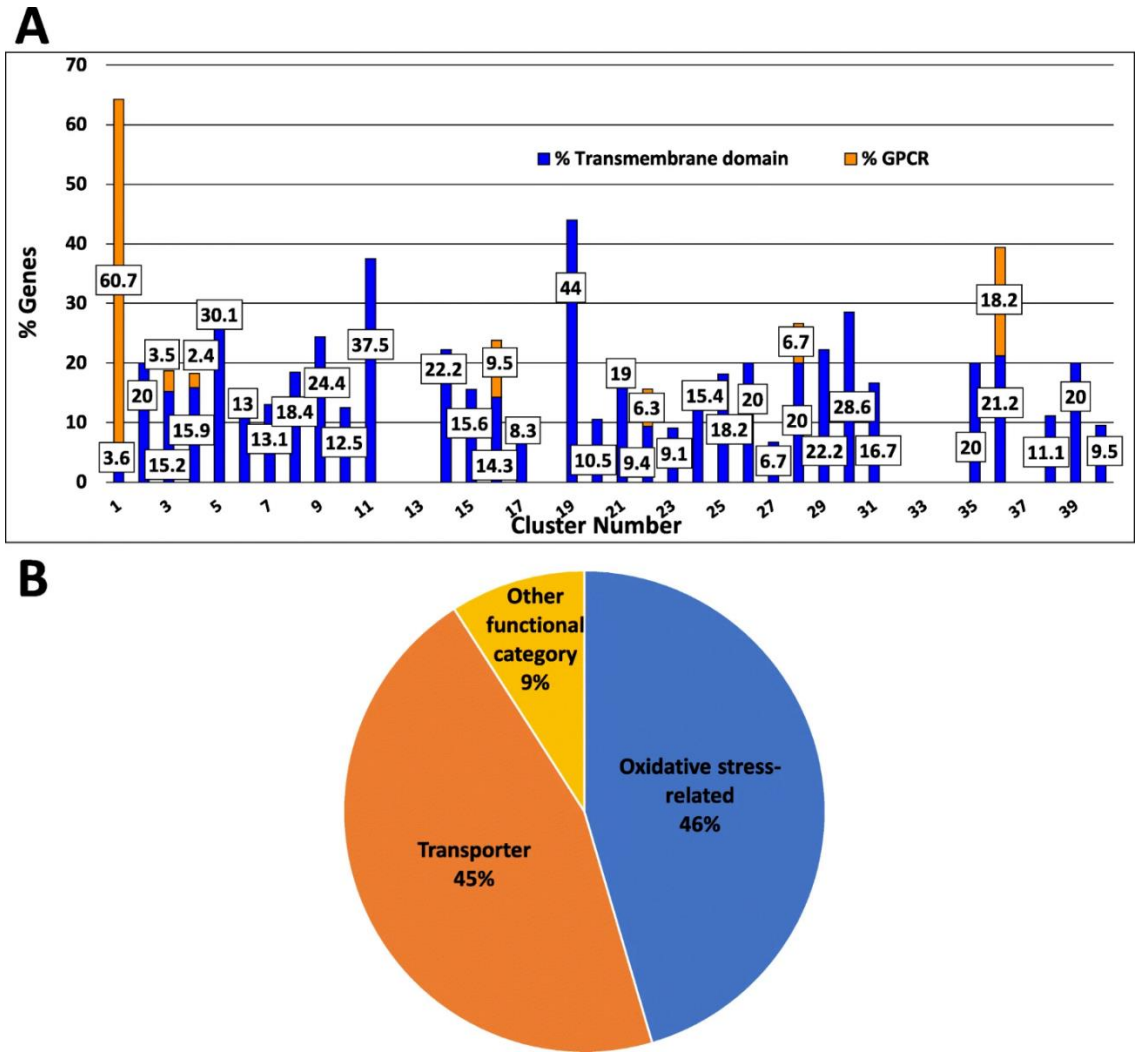

A
% Genes
Cluster Number
% Transmembrane domain
% GPCR
B
Other functional category 9%
Oxidative stress-related 46%
Transporter 45%

Figure 3.6: Protein phosphorylation-related genes. a. Assortment of phosphorylated proteins into clusters.

Proteins that are phosphorylated during growth in submerged cultures with glucose as the carbon source were obtained from Ref. (37). The number of phosphorylated proteins in each cluster (blue bars; actual number on top of bars), along with the % representation in each cluster (orange bars; % noted on top of bars) is shown. The number above the blue bars represents the number of phosphorylated proteins in each cluster. The number above the orange bars indicates the percent of phosphorylated protein that make up each cluster. b. Distribution of serine-threonine protein kinases and serine-threonine or tyrosine protein phosphatases. The percentage of genes in each cluster encoding a serine-threonine protein kinase (STKinase) or serine-threonine or tyrosine protein kinase (Protein Pase) is shown. The numbers above the blue bars indicates the percent of genes in each cluster that are serine/threonine protein kinases. The numbers above the orange bars represents the percent of genes in each cluster that are protein phosphatases.

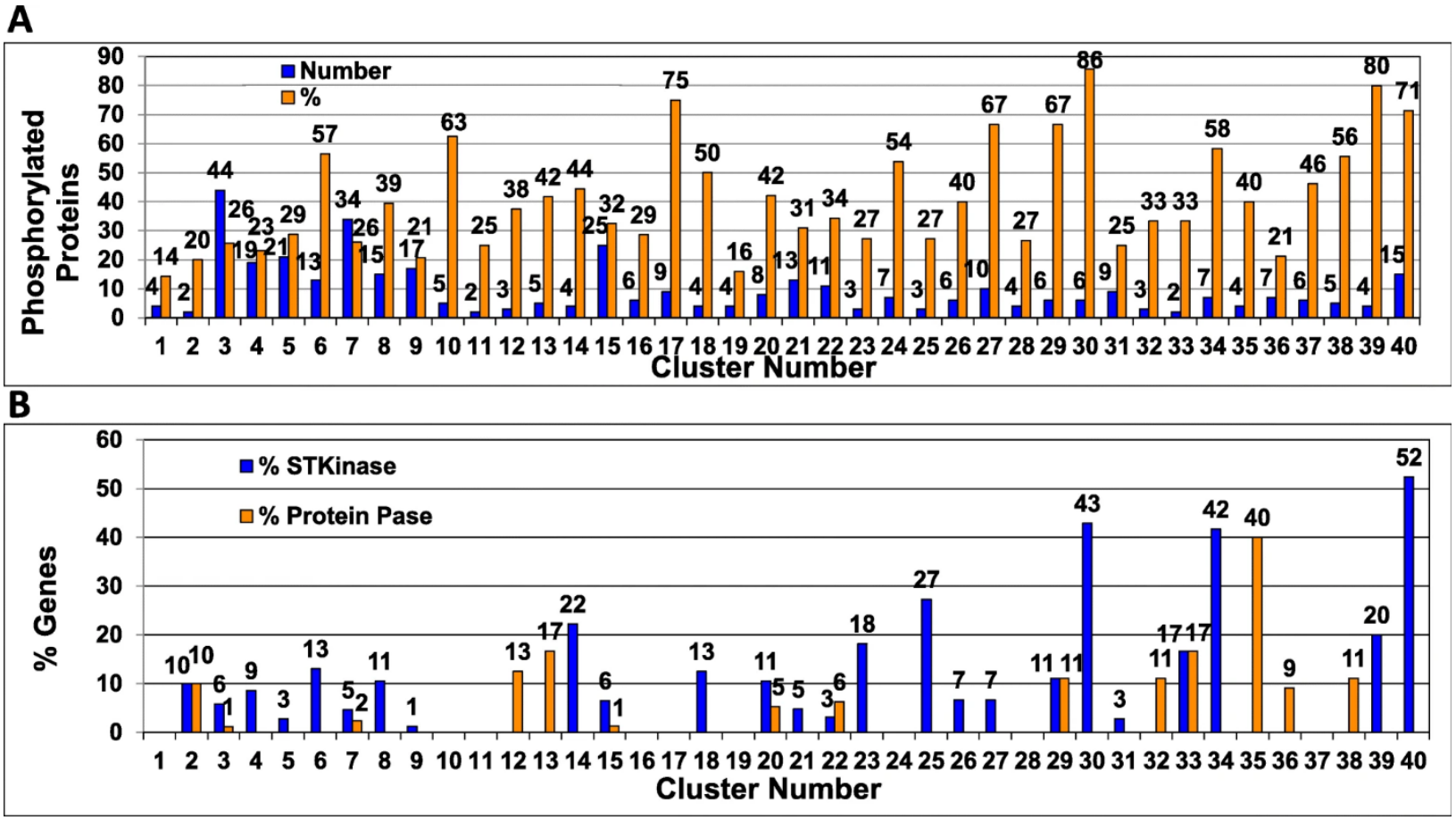

A
Number
%
Phosphorylated Proteins
Cluster Number
B
% STKinase
% Protein Pase
% Genes
Cluster Number

Figure 3.7: Cluster genes dependent on the three MAPKs: MAK-1, MAK-2 and OS-2.

Data for identification of genes with down-regulated expression in a *N. crassa* Δ*mak-1* mutant were obtained from Ref. (44). For mining cluster genes encoding transcriptional or phosphorylation targets of MAK-2 in *N. crassa*, targets were compiled from four datasets, including microarray analysis of genes down-regulated in a Δ*mak-2* mutant (41) or downregulated in a strain expressing the *mak-2*^{Q100G} strain in the presence of inhibitor (45), and MAK-2 phosphorylation targets identified during phosphoproteomics studies with the mak-2Q100G strain with inhibitor [46, 47]. Cluster genes encoding *N. crassa* orthologs of transcriptional targets of the p38 MAPK (*os-2*) in Cryptococcus neoformans and Aspergillus fumigatus were obtained using data from Ref. (48, 49). *N. crassa* orthologs in the combined dataset were retrieved using the "Transform by Orthology" tool at FungiDB. The numbers above the blue, orange and green bars indicate the percent of genes in each cluster that are either regulated by MAK-1, MAK-2 and the p38 MAPK (OS-2), respectively.

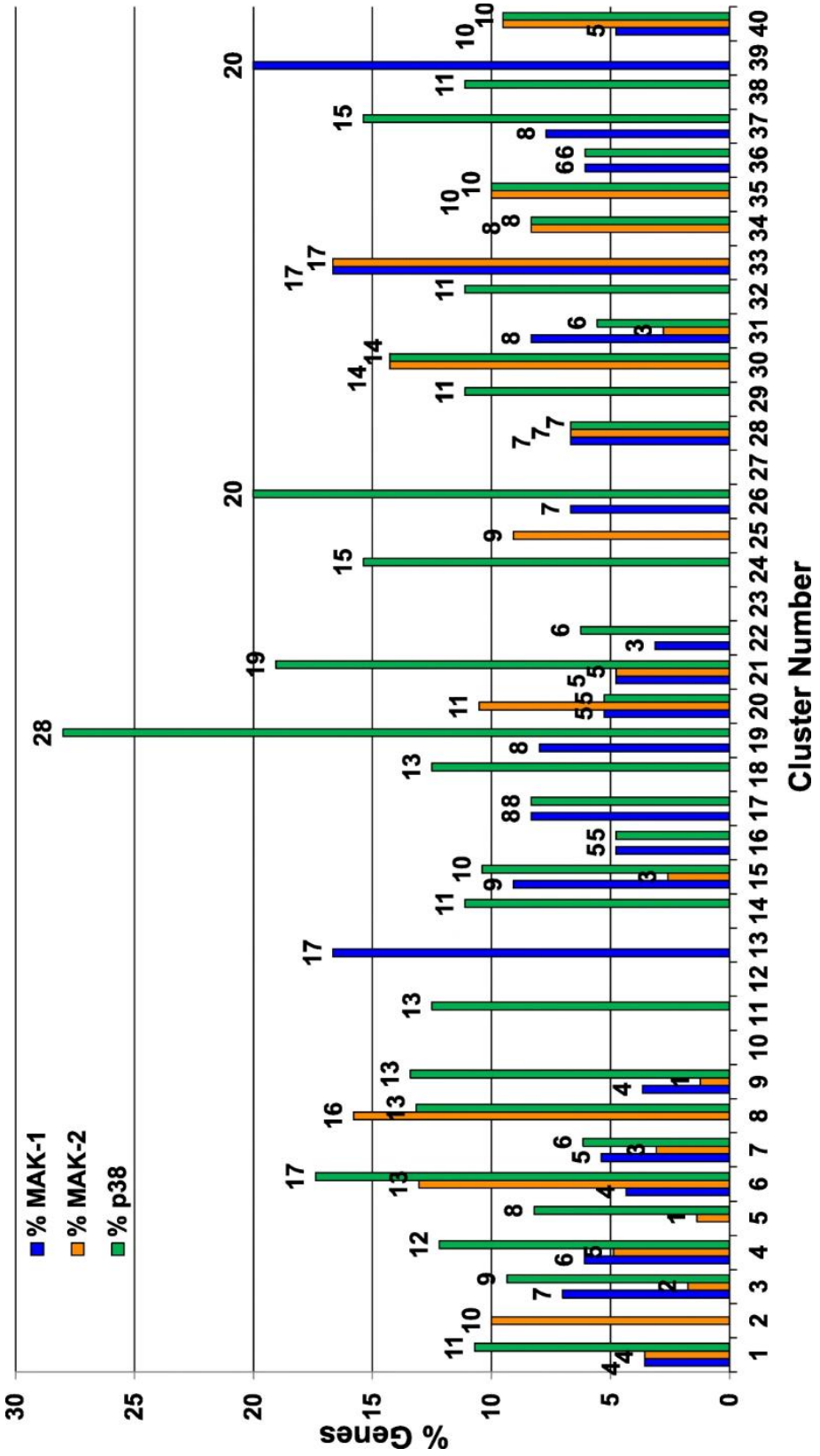

% MAK-1
% MAK-2
% p38
% Genes
Cluster Number

Figure 3.8: Distribution of transcription factors in clusters.

The number of genes encoding a transcription factor in each cluster, along with the percent representation in each cluster is shown (actual number on top of bars).

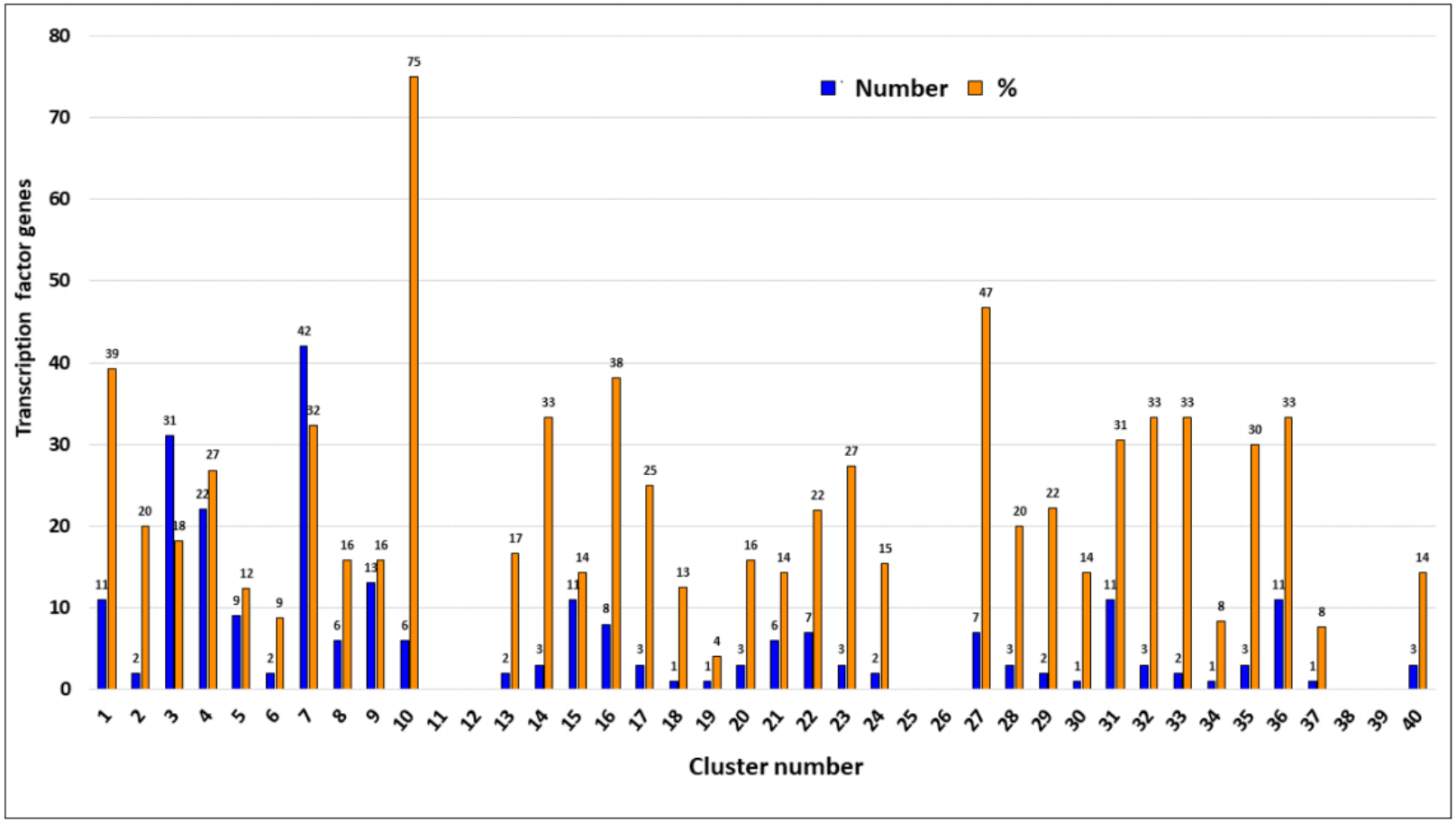

Number
%
Transcription factor genes
Cluster number
0
10
20
30
40
50
60
70
80
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
11 39
2 20
31 18
22 27
9 12
2 9
42 32
6 16
13 16
6 75
2 17
3 33
11 14
8 38
3 25
1 13
1 4
3 16
6 14
7 22
3 27
2 15
7 47
3 20
2 22
1 14
11 31
3 33
2 33
1 8
3 30
11 33
1 8
3 14

Figure 3.9: K-means clustering of transcriptomic data from wild type for genes in the phenotype dataset.

The expression values for each gene in the phenotype dataset were standardized to have a mean of 0 and a standard deviation of 2. Genes were partitioned into the indicated number of expression profiles using K-means clustering. **A.** RNA-seq data from a sexual cycle time course. RNA-seq data for eight time points were obtained from Ref. (52). Genes were partitioned into seven expression profiles. **B.** RNA-seq data from a time course of conidial germination. RNA-seq data for four time points were obtained from Ref. (53). Genes were partitioned into eight expression clusters. **C.** Microarray data from a time course of conidiation. Microarray data for 10 time points were obtained from Ref. (54). Genes were partitioned into seven expression clusters. **D.** Microarray data during a colony development time course. Microarray data for six time points were obtained from Ref. (55). Genes were partitioned into six expression clusters.

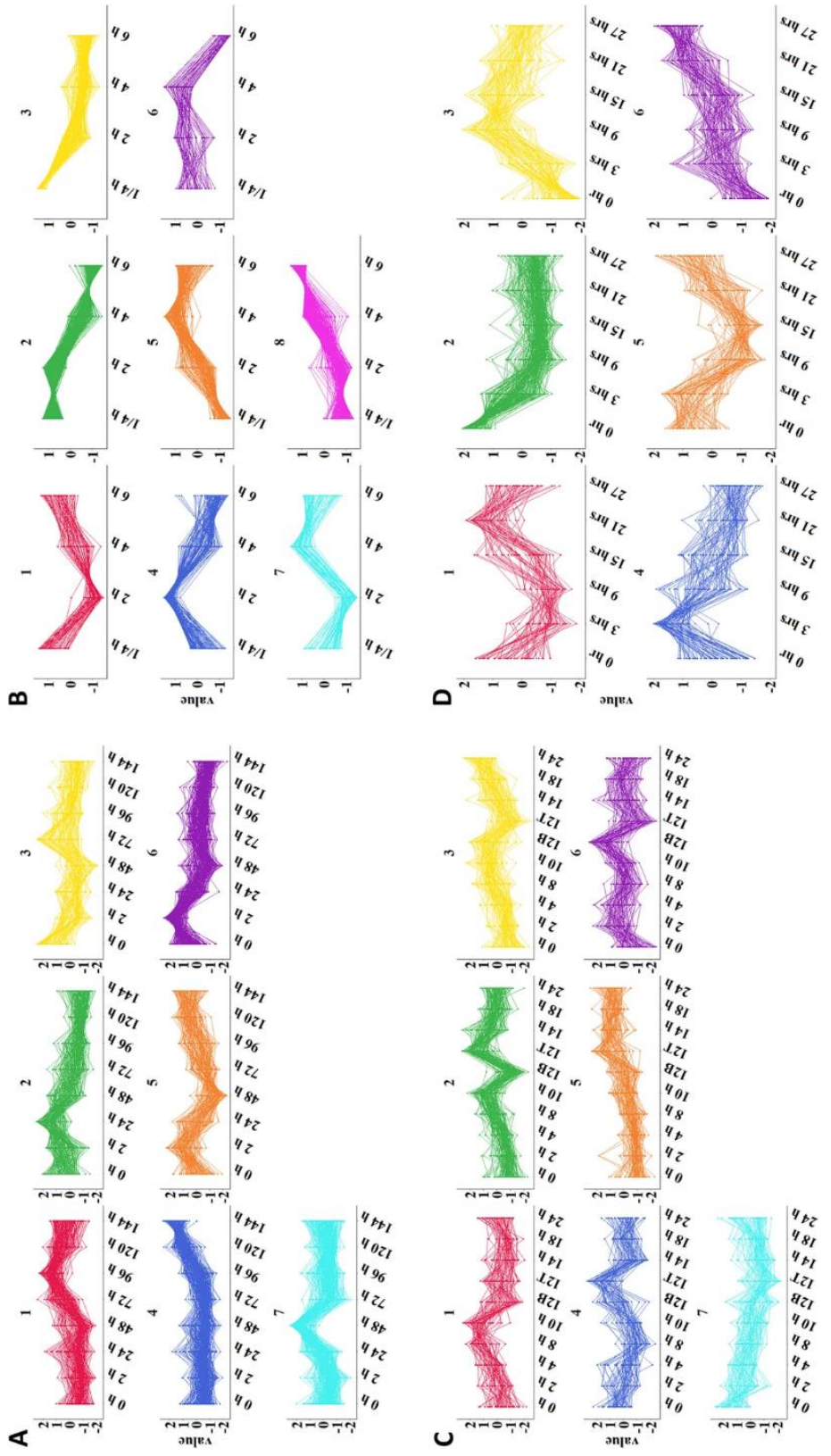

A
B
C
D
value
1
2
3
4
5
6
7
8
9
0 h
2 h
4 h
6 h
24 h
48 h
72 h
96 h
120 h
144 h
1/4 h
0 hr
3 hrs
9 hrs
15 hrs
21 hrs
27 hrs
8 h
10 h
12B
12T
14 h
18 h

Figure 3.10: Distribution of expression profiles in each cluster.

The expression value of each gene was standardized to have a mean of 0 and a standard deviation of 2 across the samples Genes were partitioned into clusters using K-means clustering. Each colored portion of each bar denotes the percent of genes in that cluster that belong to the corresponding expression profile. **A. Distribution of expression profiles during sexual development.** RNAseq data were obtained from [52]. Genes were partitioned into seven expression clusters **B. Distribution of expression profiles during conidial germination.** RNAseq data were obtained from [53]. Genes were partitioned into eight expression clusters. **C. Distribution of expression profiles during conidiation.** Microarray data were obtained from [54]. Genes were partitioned into seven expression clusters. **D. Distribution of expression profiles during colony development.** Microarray data were obtained from [55]. Genes were partitioned into six expression clusters.

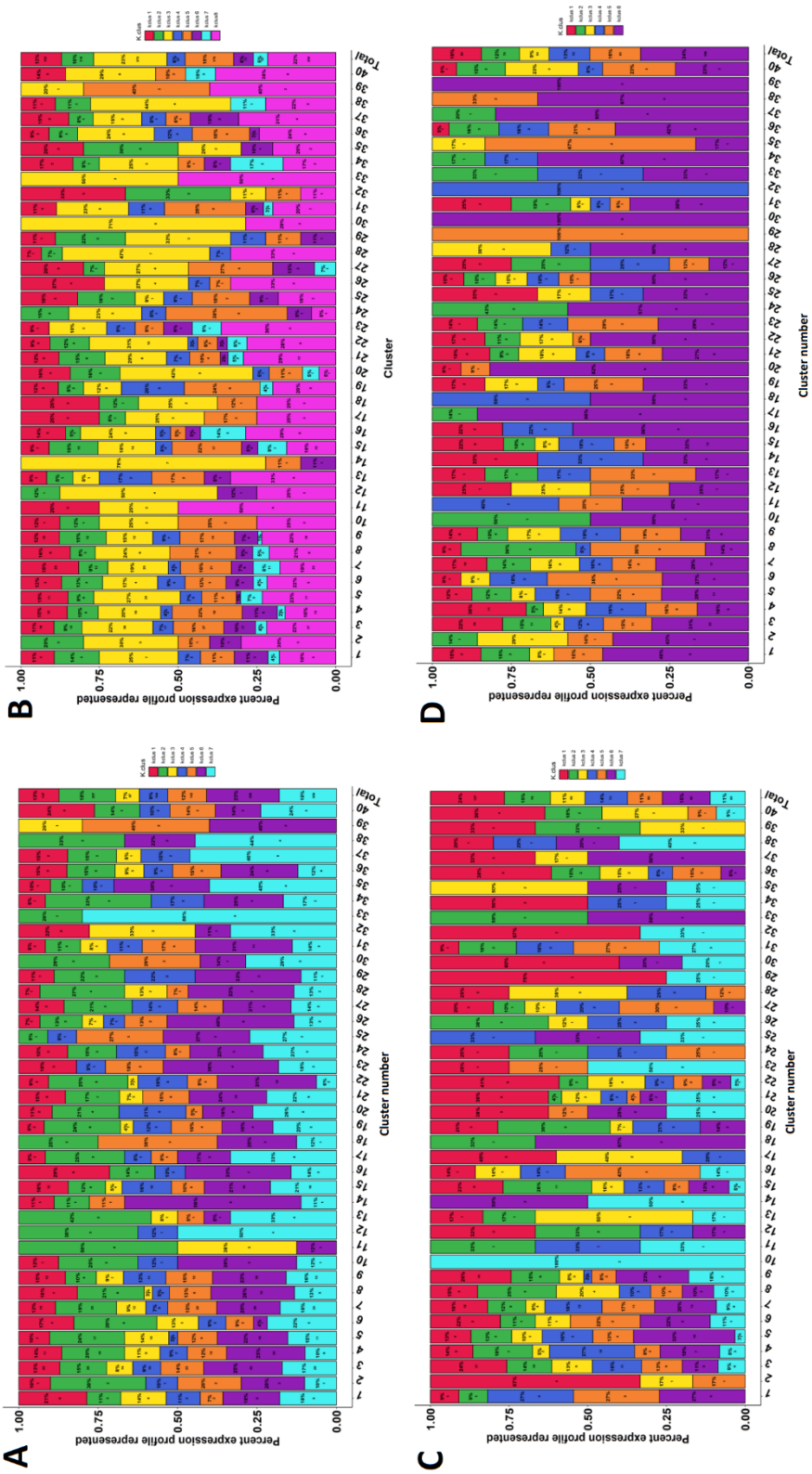

A
B
C
D
Percent expression profile represented
1.00
0.75
0.50
0.25
0.00
Cluster number
Cluster
Total

Figure 3.11: Model showing predicted pathway(s) for genes that co-cluster with the ERK MAPK genes in Cluster 40.

The MAK-1 and MAK-2 ERK MAPK cascade genes are represented by green and red shading, respectively. *S. cerevisiae* interaction data for genes in Cluster 40 were obtained from www.yeastgenome.org. Physical interactions, genetic interactions, or physical and genetic interactions between the indicated genes/gene products in yeast are shown by yellow, blue, or green lines, respectively. Genes with known interactions in *N. crassa* are connected by black lines. Genes that interact with the MAK-1 cascade kinases are shaded in yellow, while those that interact with components of the MAK-2 pathway are indicated with blue shading. Genes with no ortholog in yeast and no known interaction in *N. crassa* are represented by the white rectangles.

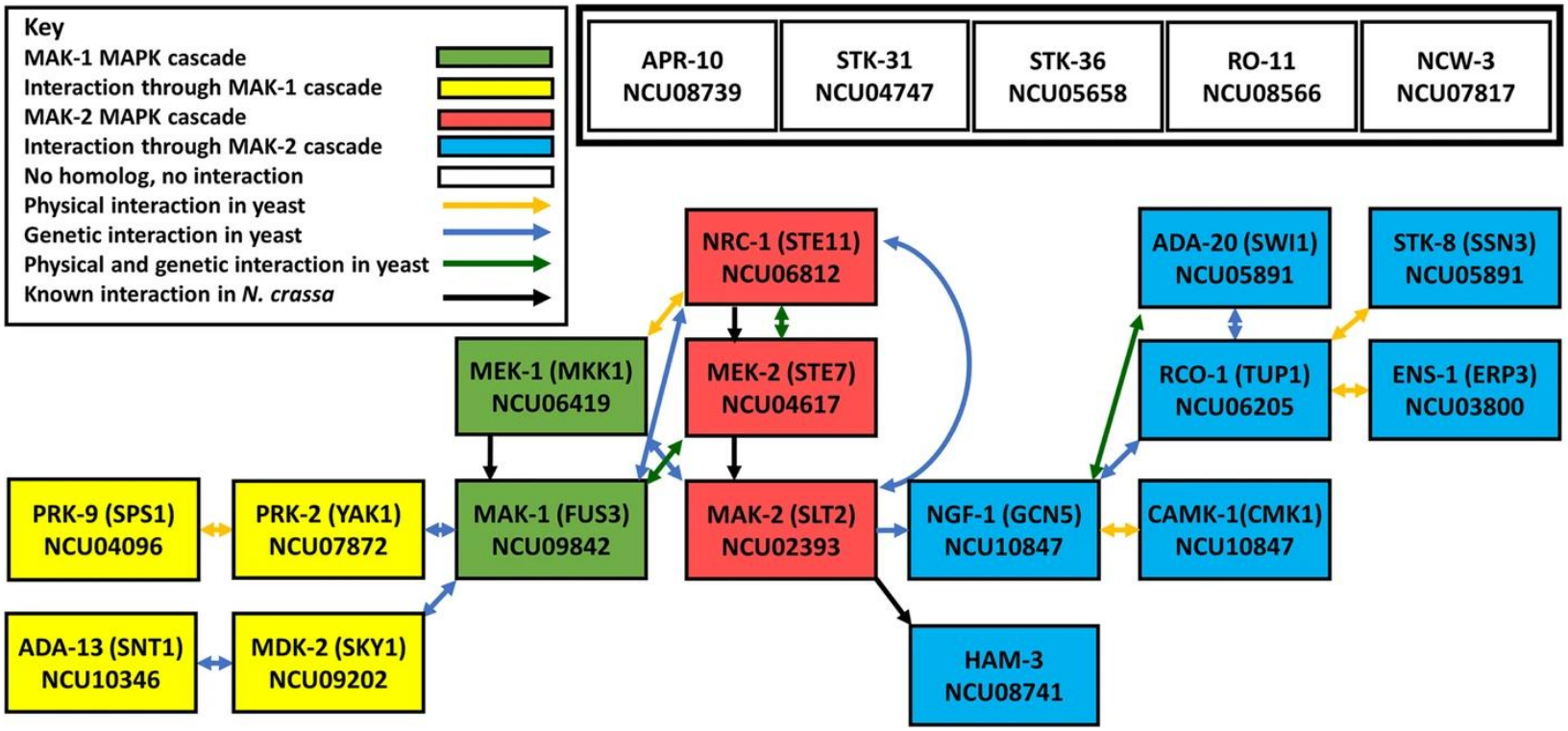

Key
MAK-1 MAPK cascade
Interaction through MAK-1 cascade
MAK-2 MAPK cascade
Interaction through MAK-2 cascade
No homolog, no interaction
Physical interaction in yeast
Genetic interaction in yeast
Physical and genetic interaction in yeast
Known interaction in N. crassa
APR-10 NCU08739
STK-31 NCU04747
STK-36 NCU05658
RO-11 NCU08566
NCW-3 NCU07817
NRC-1 (STE11) NCU06812
ADA-20 (SWI1) NCU05891
STK-8 (SSN3) NCU05891
MEK-1 (MKK1) NCU06419
MEK-2 (STE7) NCU04617
RCO-1 (TUP1) NCU06205
ENS-1 (ERP3) NCU03800
PRK-9 (SPS1) NCU04096
PRK-2 (YAK1) NCU07872
MAK-1 (FUS3) NCU09842
MAK-2 (SLT2) NCU02393
NGF-1 (GCN5) NCU10847
CAMK-1(CMK1) NCU10847
ADA-13 (SNT1) NCU10346
MDK-2 (SKY1) NCU09202
HAM-3 NCU08741

Figure 3.12: Model for predicted pathway(s) including genes that co-cluster with the p38 MAPK genes in Cluster 34.

The OS-2 MAPK cascade genes are represented by red rectangles. *S. cerevisiae* interaction data for genes in Cluster 34 were obtained from www.yeastgenome.org. Physical interactions, genetic interactions, or physical and genetic interactions between the indicated genes/gene products in yeast are indicated by yellow, blue, or green lines, respectively. Genes with known interactions in *N. crassa* are connected by black lines. Genes that interact with the OS-2 cascade kinases are represented by blue rectangles. Genes with no known interaction in *N. crassa* or *S. cerevisiae* are represented by the white rectangles.

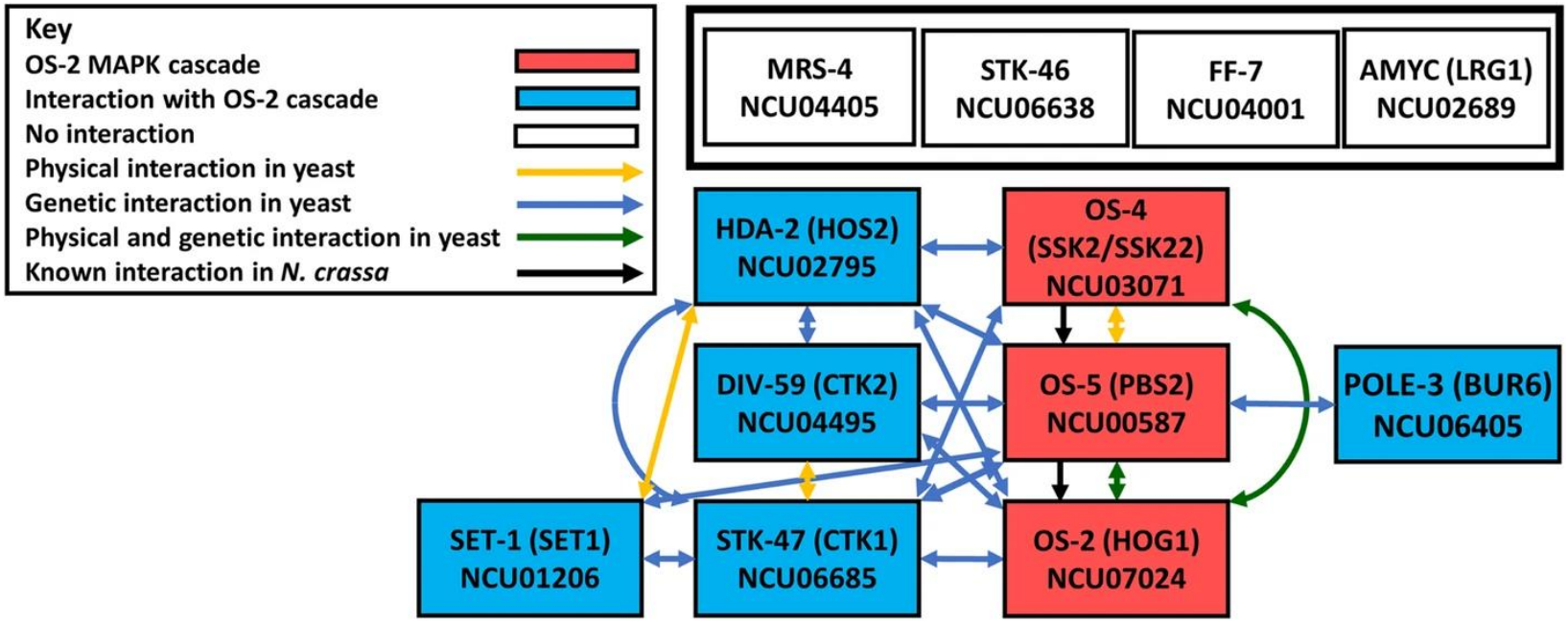

Key
OS-2 MAPK cascade
Interaction with OS-2 cascade
No interaction
Physical interaction in yeast
Genetic interaction in yeast
Physical and genetic interaction in yeast
Known interaction in N. crassa
MRS-4 NCU04405
STK-46 NCU06638
FF-7 NCU04001
AMYC (LRG1) NCU02689
HDA-2 (HOS2) NCU02795
OS-4 (SSK2/SSK22) NCU03071
DIV-59 (CTK2) NCU04495
OS-5 (PBS2) NCU00587
POLE-3 (BUR6) NCU06405
SET-1 (SET1) NCU01206
STK-47 (CTK1) NCU06685
OS-2 (HOG1) NCU07024

Table 3.1: Comparison of results using iterations of Partitioning around Medoids (PAM)

Weight[c]	Number of clusters	Average % consensus[a]									Average relative standard deviation (RST)[b]		
		C #	C morph.	PP #	PP morph.	P #	P morph.	A #	A morph.	Average % consensus overall	Basal hyphae growth rate	Aerial hyphae height	Average RST overall
0	18	91.72	94.20	93.74	94.23	95.27	93.34	96.35	98.32	95.06	26.76	34.12	13.20
0	19	92.16	94.50	94.05	94.51	95.45	93.66	96.48	98.39	95.29	25.33	32.54	12.76
0	20	92.50	94.78	94.35	94.79	95.68	93.92	96.66	98.47	95.52	24.62	30.43	12.36
0	21	93.02	95.19	94.61	95.05	95.88	94.21	96.81	97.03	95.54	28.30	31.30	13.09
1	31	95.92	96.57	93.46	94.99	95.70	95.39	97.32	98.73	96.06	19.80	24.13	11.02
1	32	95.22	96.66	94.40	95.98	95.84	95.53	97.40	98.80	96.37	20.18	24.55	10.474
1	33	95.67	97.05	94.57	96.11	95.97	95.67	97.47	97.95	96.40	22.57	25.24	10.92
2	32	96.23	94.89	94.05	95.04	94.22	95.51	98.76	98.82	95.90	17.52	26.50	10.88
2	33	96.35	95.05	94.20	95.19	94.31	95.64	98.79	98.86	96.01	16.91	25.83	10.67
2	34	95.45	95.24	94.99	96.82	94.47	95.77	98.82	98.89	96.43	17.55	26.30	10.22
2	35	95.15	94.62	95.00	96.91	93.95	94.98	98.03	98.09	95.94	18.98	25.30	11.19
3	29	96.20	96.09	93.25	94.20	92.63	95.26	95.53	97.62	94.94	13.77	27.67	10.82
3	30	96.37	96.22	93.47	94.39	92.88	95.3	95.67	97.70	95.10	13.43	26.61	10.60
3	31	97.45	97.64	91.71	94.52	94.39	94.5	95.14	96.63	94.87	15.60	27.08	10.98
3	32	97.45	97.29	92.26	94.54	93.62	93.38	95.46	96.91	94.78	15.45	27.69	11.41
3	33	96.92	96.53	92.41	94.65	93.28	92.81	95.29	96.69	94.52	17.58	26.11	12.14
4	32	90.89	94.99	93.99	95.11	94.39	94.23	94.57	97.20	94.93	14.23	18.81	11.42
4	33	90.50	94.98	93.99	94.80	94.13	93.80	94.40	97.37	94.78	14.21	19.08	11.96
4	34	90.78	94.78	94.10	94.83	94.76	94.13	94.97	97.46	95.01	14.10	20.18	11.75
4	35	90.82	94.93	94.27	94.98	94.91	94.23	95.11	97.53	95.14	13.78	19.68	11.57
5	36	85.84	95.66	94.14	95.40	95.38	93.87	94.33	97.42	95.17	14.09	13.44	11.26
5	37	85.28	95.45	94.18	95.41	95.51	94.04	94.48	97.49	95.22	13.99	12.45	11.22
5	38	85.18	95.55	94.88	95.63	94.53	94.15	94.03	97.55	95.19	13.90	12.05	11.20
5	39	84.56	95.62	95.24	96.70	94.67	94.30	94.18	97.61	95.48	13.93	12.46	11.00
5	40	84.36	95.37	96.40	96.26	94.37	94.01	94.33	97.67	95.49	13.80	12.71	10.96
6	39	89.07	95.73	94.33	94.58	93.57	93.70	93.89	97.39	94.03	14.11	11.64	11.03
6	40	89.64	95.83	94.22	94.71	93.73	93.86	94.04	97.45	94.19	13.94	11.86	10.98

[a]Average percent consensus - For each cluster the category that was most prevalent was determined and represented as a percent. Then each cluster's most prevalent category was used to create a trait average

[b]Average relative standard deviation - For continuous variables, the standard deviation for each cluster was calculated and then divided by the cluster mean to determine the relative standard deviation [c]Weight - Each phenotype was assigned a specific weight relative to the other traits when creating the Gower's distance matrix. The phenotypes are in the order as shown in Table 1 No weight = (1,1,1,1,1,1,1,1,1,1); Weight 1 = (1,1,1,1,1,1,1,1,2,2); Weight 2 = (1,0.5,1,0.5,1,0.5,1,0.5,2,2); Weight 3 = (2,2,1,1,1,1,1,1,6,2); Weight 4 = (0.5,0.5,1,1,1,1,1,1,6,5); Weight 5 = (1,1,1,1,1,1,1,1,6,4); Weight 6 = (1,1,1,1,1,1,1,1,6,6).

Column labels are as follows: conidial number (C #), conidial morphology (C morph.), protoperithecial number (PP #), protoperitheicial morphology (PP morph.), perithecial number (P #), perithecial morphology (P morph.), ascospore number (A #), ascospore morphology (A moprh.).

Table 3.2: Phenotype summary for 40 PAM clusters

Cluster #	# of genes in cluster	% Mutants with same phenotype[a]									
		Basal hyphae growth rate	Aerial hyphae height	C #	C morph.	PP #	PP morph.	P #	P morph.	A #	A morph.
1	28	75% N-A	57% N-H	93% N	100% N	100% N	100% N	100% N	100% N	100% N	100% N
2	10	90% R	80% R	100% N	100% N	100% NF	100% NF	100% NF	100% NF	100% NF	100% NF
3	171	94% N-L	95% N-A	99% N	100% N	100% N	100% N	100% N	99% N	100% N	100% N
4	82	88% N-A	85% N-A	98% N	100% N	100% N	100% N	100% N	100% N	100% N	100% N
5	73	100% SL-R	93% R	100% N	100% N	99% N	100% N	100% N	100% N	100% N	100% N
6	23	100% R	96% R	100% N	100% N	96% N	100% N	100% N	100% N	100% N	100% N
7	130	100% SL-R	96% N-A	99% N	100% N	100% N	100% N	100% N	99% N	99% N	100% N
8	38	100% R	100% R	97% N	97% N	97% N	100% N	97% N	100% N	100% N	100% N
9	82	80% N-L	95% R	100% N	100% N	100% N	99% N	99% N	100% N	95% N	100% N
10	8	62% SR	75% N-A	75% N	100% N	88% N	100 % N	62% N	88% N	75% N	100% N
11	8	50% SL-R	88% R	100% N	100% N	100% N	75% N	88% NF	88% NF	100% NF	100% NF
12	8	88% R	88% R	100% R	88% N	62% N	50% N	88% NF	88% NF	100% NF	100% NF
13	12	50% R	58% R	75% R	92% N	92% R	100% N	92% R	92% N	83% R	83% N
14	9	V	89% N-A	89% N	100% N	89% N	100% N	67% N	78% N	100% NF	100% NF
15	77	100% SL-R	96% N-A	100% N	100% N	100% N	100% N	97% N	100% N	99% N	100% N
16	21	Varied	67% N-A	95% N	95% N	86% N	90% N	100% NF	100% NF	100% NF	100% NF
17	12	100% SR	92% R	92% N	83% N	92% N	100% N	100% N	92% N	92% N	92% N
18	8	100% SR	88% R	75% N	88% N	88% NF	88% NF	100% NF	100% NF	100% NF	100% NF
19	25	52% R	100% I	96% N	100% N	100% N	100% N	100% N	100% N	96% N	100% N
20	19	74% SR	100% R	84% N	95% N	95% N	95% N	100% N	100% N	89% N	100% N
21	42	100% SL-R	100% R	100% N	100% N	100% N	98% N	100% N	100% N	100% N	100% N
22	32	100% R	84% N-A	97% N	100% N	97% N	100% N	100% NF	100% NF	100% NF	100% NF
23	11	82% R	91% R	91% N	100% N	100% NF	100% NF	100% NF	100% NF	100% NF	100% NF
24	13	85% R	85% R	92% N	92% N	92% N	63% N	100% NF	100% NF	100% NF	100% NF
25	11	91% R	100% SR	55% R	100% N	100% NF	100% NF	100% NF	100% NF	100% NF	100% NF
26	15	67% SL-R	100% SR	100% N	100% N	100% N	100% N	100% N	100% N	97% N	100% N
27	15	Varied	60% N-A	100% N	100% N	80% N	100% N	87% N	87% N	93% R	100% N
28	15	67% R	100% I	80% N	100% N	93% N	93% N	100% N	93% N	87% N	100% N
29	9	67% SL-R	78% R	89% N	100% N	89% N	89% N	78% R	89% N	100% NF	100% NF
30	7	100% SR	100% SR	Varied	71% N	100% N	100% N	57% N	86% N	V	V
31	36	100% R	92% N-A	97% N	100% N	100% N	97% N	94% N	100% N	94% N	100% N
32	9	56% SL-R	89% R	100% R	100% N	89% N	100% N	100% N	100% N	100% N	100% N
33	6	50% R, 50% SR	67% N-A	67% N	67% N	83% R	100% N	83% R	67% Abn B	83% NF	83% NF
34	12	67% R	83% N-A	83% N	75% N	92% NF	92% NF	100% NF	100% NF	100% NF	100% NF
35	10	60% SL-R	50% N-A	100% N	100% N	100% R	90% N	90% N	100% N	60% N	100% N
36	33	V	100% I	88% N	100% N	100% N	100% N	100% N	94% N	100% N	100% N
37	13	69% R	85% R	92% N	100% N	92% N	92% N	85% N	77% N	100% R	100% N
38	9	67% R	100% SR	89% N	100% N	89% N	89% N	89% N	100% N	89% N	100% N
39	5	80% R	60% R	80% N	100% N	100% N	100% N	100% N	Varied	100% NF	100% NF
40	21	100% SR	100% SR	76% R	90% N	90% NF	90% NF	100% NF	100% NF	100% NF	100% NF

[a] The value in the cell is the percentage of mutants with the indicated majority phenotype. Varied(V) indicates that no one phenotype was present in more than 50% of the mutants. Normal range for growth rate is 75–85 mm/day. 75–77.5 = Normal Low(N-L); 77.6–82.5 = Normal Average(N-A); 82.6–85 = Normal High(N-H); 65–74.9 = Slightly Reduced(SL-R); 40–64.9 = Reduced(R); < 40 = Severely Reduced(SR); > 85 = Increased(I). Normal range for aerial hyphae height is 30-45 mm. 35.1–39.9 = Normal Average(N-A); 30–35 = Normal Low(N-L); 40–45 = Normal High(N-H); 25–29 = Slightly Reduced(SL-R); 15–24.9 = Reduced(R); < 15 = Severely Reduced(SR); > 45 = Increased(I). Sexual cycle phenotypes are as follows: Normal(N), Reduced(R), Not Formed(NF), variable(V), and Abnormal beaks(Abn B).

Chapter 4

Genetic Relationships Between the RACK1 Homolog cpc-2 and Heterotrimeric G Protein Subunit Genes in Neurospora Crassa

Contributions to this chapter

I contributed to this study by helping develop and test the CPC-2 polyclonal antibody. I performed the cell fractionation assay that is figure 4.2. I performed the western blot with the CPC-2 antibody in figure 3. I captured all images in figure 4.8. I verified the genotypes of Δ*gna-1*, Δ*gna-2,* Δ*gna-3,* Δ*gnb-1,* and Δ*cpc-2*.

Abstract

Receptor for Activated C Kinase-1 (RACK1) is a multifunctional eukaryotic scaffolding protein with a seven WD repeat structure. Among their many cellular roles, RACK1 homologs have been shown to serve as alternative Gβ subunits during heterotrimeric G protein signaling in many systems. We investigated genetic interactions between the RACK1 homolog *cpc-2*, the previously characterized Gβ subunit *gnb-1* and other G protein signaling components in the multicellular filamentous fungus Neurospora crassa. Results from cell fractionation studies and from fluorescent microscopy of a strain expressing a CPC-2-GFP fusion protein revealed that CPC-2 is a cytoplasmic protein. Genetic epistasis experiments between *cpc-2*, the three Gα genes (*gna-1*, *gna-2* and *gna-3*) and *gnb-1* demonstrated that *cpc-2* is epistatic to *gna-2* with regards to basal hyphae growth rate and aerial hyphae height, while deletion of *cpc-2* mitigates the increased macroconidiation on solid medium observed in Δ*gnb-1* mutants. Δ*cpc-2* mutants inappropriately produce conidiophores during growth in submerged culture and mutational activation of *gna-3* alleviates this defect. Δ*cpc-2* mutants are female-sterile and fertility could not be restored by mutational activation of any of the three Gα genes. With the exception of macroconidiation on solid medium, double mutants lacking *cpc-2* and *gnb-1* exhibited more severe defects for all phenotypic traits, supporting a largely synergistic relationship between GNB-1 and CPC-2 in *N. crassa*.

Introduction

Heterotrimeric G protein signaling cascades consist of seven-helix transmembrane G Protein Coupled Receptors (GPCRs) and the three G protein subunits—Gα, Gβ and Gγ (1–3). In the inactive state, the Gαβγ heterotrimer is associated with the GPCR. Ligand stimulation causes exchange of GDP for GTP on the Gα, leading to dissociation of Gα-GTP from the Gβγ heterodimer. The Gα-GTP and the Gβγ dimer can then regulate downstream effectors, leading to changes in cellular physiology (3). The Gα-GTP has native GTPase activity that causes release of the inorganic phosphate from the GTP. The Gα-GDP then reassociates with the Gβ subunit and GPCR, leading to signal termination and completion of the cycle.

Neurospora crassa is a multicellular ascomycete fungus that has emerged as a model system to study G protein signaling, and comparisons with *N. crassa* have driven discoveries in pathogenic fungi and higher eukaryotes (4, 5). In *N. crassa*, there are 43 predicted GPCRs, three Gα subunits (GNA-1, GNA-2 and GNA-3), one characterized Gβ subunit (GNB-1) and one Gγ subunit (GNG-1) (6, 7). Major processes such as hyphal growth, macroconidiation, conidial germination, mating, nutrient sensing and temperature and oxidative stress resistance are regulated by G protein signaling pathways in *N. crassa* (8–14).

Receptor for Activated C Kinase-1 (RACK1) is a major scaffolding protein in many eukaryotic systems. Similar to G protein β subunits, RACK1 has a seven WD repeat structure, and is one of the best-studied proteins in the WD-repeat family (15). Initially identified as a protein that binds to the active conformation of protein kinase C

(PKC) βII, RACK1 is now known to be multifunctional (16, 17). For example, RACK1 allows cross talk between the PKC and Mitogen Activated Kinase (MAP) pathways by acting as a scaffold for the Jun N-terminal Kinase (JNK) upon stimulation, leading to PKC-mediated phosphorylation and activation of JNK (18). It has been observed that RACK1 binds to the Gβγ dimer in HEK293 cells and also regulates a subset of its functions, including promoting its dislocation from the cytosol to the membrane (19). Additionally, RACK1 is known to associate with the 40S subunit of the ribosome, near the mRNA exit channel (20). Due to its conformation when bound to the ribosome, RACK1 is believed to serve as an adaptor, bringing together proteins at the ribosome during translation (reviewed in (15)).

Homologs of RACK1 have been implicated as alternative Gβ subunits in the fungal kingdom, through direct interaction with Gα subunits (21, 22). In *Saccharomyces cerevisiae*, Asc1p functions as a Guanine nucleotide Dissociation Inhibitor (GDI) for the Gα Gpa2, and is involved in regulating glucose responsiveness through its binding to adenylyl cyclase (Cyr1) (23). gib2, an essential gene in *Crytpococcus neoformans*, encodes a protein that binds to the Gα Gpa1 and two Gγ subunits, Gpg1 and Gpg2. It also associates with Smg1, a downstream target of cAMP signaling, and to the protein kinase C homolog Pkc1 (24). In *Magnaporthe oryzae*, the RACK1 ortholog MoMip11 interacts with the Gα protein MoMagA and the Regulator of G protein Signaling (RGS) protein MoRgs7 to regulate pathogenicity (25, 26).

Additional RACK1 orthologs have been shown to regulate various aspects of growth and development in several fungal systems, but without demonstration of a

physical interaction with heterotrimeric Gα proteins. *S. pombe* Cpc2 plays a role in cell cycle regulation and stress responses through ribosomal association (27) and translational control of the stress response transcriptional factor Atf1 (27). RACK1 orthologs from *Aspergillus nidulans* and *Aspergillus fumigatus* have been demonstrated to regulate sexual differentiation and asexual growth and development, respectively (28, 29). In Ustilago maydis, Rak1 is essential for the transcription of rop1, which is a direct positive regulator of the pheromone response factor (*prf1*), making it essential for mating (30). Strains lacking RAK1 also have attenuated filamentation and virulence, and abnormal cell morphology (30).

The *N. crassa* RACK1 homolog CPC-2 was the first reported RACK1 protein in fungi, initially identified as a component of the general amino acid regulation network (31). In *N. crassa*, starvation for a single amino acid leads to an overall derepression of all amino acid biosynthetic genes at the level of transcription (32). Loss of the *cpc-2* gene blocks derepression of amino acid biosynthetic genes during amino acid limiting conditions (31). Under non-starved conditions, loss of the *cpc-2* gene decreases growth by 50% (33). During the sexual cycle, the Δ*cpc-2* mutant lacks protoperithecia, and is female-sterile (33). Other components of this cross pathway control network are *cpc-1*, homologous to GCN4 (34), and *cpc-3*, the *N. crassa* equivalent of GCN2 (35). Analysis of Δ*cpc-2* Δ*cpc-3* and Δ*cpc-2* Δ*cpc-1* double mutants showed that they possessed Δ*cpc-2* phenotypes, such as reduced growth and female sterility. These findings suggested that *cpc-2* has broader functions operating outside of amino acid control (35).

To-date, no one has explored a possible function for CPC-2 in G protein signaling in *N. crassa*. In this study, we use strains carrying single and double gene deletions or expressing constitutively activated Gα alleles to analyze genetic epistasis between components of the G protein pathway and the *cpc-2* gene. We produce a polyclonal antibody against CPC-2 and use western analysis to determine protein levels in the mutants lacking the other G protein subunits. Our results reveal that *N. crassa* mutants lacking both predicted Gβ subunits are viable, but possess major defects in growth and development. We also provide evidence for G protein dependent and independent functions for CPC-2 in *N. crassa*.

Materials and Methods

Strains and media. *N. crassa* strains were either obtained from the Fungal Genetics Stock Center (FGSC; Kansas State University, Manhattan, KS) (36) or created during this work (Table 1). Strains that are not deposited in the FGSC collection are available upon request. Strains were cultured in Vogel's minimal medium (VM) (37) to propagate vegetative hyphae or asexual spores (macroconidia; conidia). Synthetic Crossing Medium (SCM) plates containing 1% agar were used to induce development of female sexual reproductive structures (38). Sorbose-containing medium (FGS) was used to facilitate colony formation on plates (39). Media was supplemented with 100 μg/ml of histidine, 10 μg/ml pantothenate, 200 μg/ml hygromycin (Calbiochem, San Diego, CA), 200 μg/ml nourseothricin (Werner BioAgents, Germany) or 400 μg/ml phosphinothricin (purified

from Finale, Farnam Companies, Inc., Phoenix, AZ), where indicated. Conidia were propagated in VM agar flasks as described previously (39). Liquid cultures were brought to a concentration of 1x106 conidia/ml and incubated with shaking at 200 RPM at 30˚C for 16 hr. *Escherichia coli* strain DH5α was used to maintain all plasmids.

Phylogenetic analysis. Protein sequences orthologous to *N. crassa* CPC-2 (NCU05810) and GNB-1 (NCU00440) from 18 fungal species chosen to represent a diversity of fungi (40) were obtained from the FungiDB database (fungidb.org) (41). Sequences for the Gβ and RACK1 proteins from the plant *Arabidopsis thaliana* were retrieved from the National Center for Biotechnology Information (NCBI). The "One-Click Workflow" tool at NGPhylogeny.fr (42) was implemented for the phylogenetic analysis. This pipeline uses FASTA files to generate a multiple alignment using MAFFT (Multiple Alignment using Fast Fourier Transform) (43). Alignments were inspected and proteins from the 18 species resulted in good alignments for both the Gβ and RACK1. The MAFFT alignments were curated using BMGE (Block Mapping and Gathering with Entropy) (44) and FastME (Fast Minimum Evolution) (45) was used to produce the tree file. FastME uses distance algorithms to infer phylogenies. The final trees were drawn using tools at the Interactive Tree of Life (iTOL; itol.embl.de) (46). Species and gene accession numbers are in the legend to Fig 1.

***N. crassa* strain construction.** The Δ*cpc-2*::*hphR* knockout mutant was deposited at the FGSC as a heterokaryon (FGSC13695). Homokaryotic mutants were obtained from the

heterokaryon after a sexual cross to wild type strain 74-OR23-1VA and plating ascospores on medium containing hygromycin. Progeny were checked using diagnostic PCR (47) with *cpc-2* (Primer 1 or Primer 2) and *hph* (Primer 13 or Primer 14) primers (Table 4.2), and then spot-tested on phosphinothricin to check for the presence of the mus-51 mutation, which is marked with bar (48, 49). Double mutants Δ*cpc-2*, Δ*gna-1*; Δ*cpc-2*, Δ*gna-2* and Δ*cpc-2*, Δ*gna-3* were made using genetic crosses between single mutants (39) (Table 1). In cases where both single mutants in the cross were female-sterile (Δ*cpc-2* and Δ*gna-1*), the strain used as the female was a heterokaryon with the am1 helper strain (50). The presence of the mutations in the progeny was verified by diagnostic PCR using pairs of gene-specific and *hph* cassette-specific primers (Primers 1–14 in Table 4.2).

Repeated attempts to generate a Δ*cpc-2* Δ*gnb-1* double mutant through a sexual cross were unsuccessful. Therefore, the double mutant was created by electroporation of the Δcpc2#6 strain using a knockout cassette for *gnb-1*, marked with nourseothricin resistance (nat^R) (51). The Δ*gnb-1* knockout cassette was created using yeast recombinational cloning in vector pRS426 (52), with methods previously described (53). Primers used to amplify fragments for the construct are listed in Table 2. Primer pairs 15–16 and 17–18 were used to amplify the 1 kb 5’ and 3’ flanks of *gnb-1*, respectively, from genomic DNA. Primers 19 and 20 were used to amplify the nourseothricin resistance marker from plasmid pD-NAT-1 (51). The three purified PCR products plus pRS426 digested with *Xho*I and *Eco*RI were transformed into yeast strain FY834 (54). Transformants were selected on FGS plates containing nourseothricin and then checked

for the presence of the Δ*gnb-1*::*nat*^{R} mutation using diagnostic PCR with genes specific primers (Primers 21 and 22; Table 2). Positive strains were then purified to homokaryons using serial streaking of macroconidia (47) and checked again using diagnostic PCR.

Vectors containing predicted GTPase-deficient, constitutively activating mutations *gna-1*^{Q204L} (pSVK51), *gna-2*^{Q205L} (pSVK52), and *gna-3*^{Q208L} (pSVK53) were previously made using site-directed mutagenesis (55). Electroporation of *N. crassa* with 1–2 μg of pSVK51, pSVK52 or pSVK53 was as previously described (56), using the Δcpc2his3A strain as the recipient, with selection on FGS plates without histidine. Genomic DNA was extracted from transformants and subjected to Southern analysis for *gna-1*, *gna-2* and *gna-3* as described (55). Transformants determined to have a single integration event of the transforming DNA at the *his-3* locus were purified to homokaryons using microconidiation (57) or serial streaking of macroconidia (47) on FGS plates lacking histidine. Genomic DNA was extracted from these strains and analyzed using diagnostic PCR to confirm genotypes.

A vector was produced to allow expression of a GFP-tagged version of *cpc-2* in *trans* to the wild-type copy. The vector backbone (pRS426PVG) (58) was assembled in plasmid pRS426 using yeast recombinational cloning (53). The fragments included a region 1kb 5’ of the *pan-2* ORF, the *ccg-1* promoter amplified from pMF272 (59), a multiple cloning sequence, a 5xGlycine linker, a V5-tag, GFP sequence amplified from pMF272 (59), the *bar* gene, amplified from vector pTJK1 (60) and a 1 kb fragment 3’ of the *pan-2* ORF (58). All fragments were amplified using Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). The *pan-2* flanking sequences allow

targeting to, and deletion of, the *pan-2* ORF, resulting in pantothenate auxotrophy. The final expression construct for CPC-2 was produced by insertion of the *cpc-2* ORF (amplified using Primers 27 and 28; Table 4.2) into vector pRS426PVG (linearized using *Pac*I) using yeast recombinational cloning (53). Vector pRS426PVG-CPC2 was transformed using electroporation into *N. crassa* strain 51-IV-4 (Table 4.1) (58). Transformants were selected on medium containing phosphinothricin and pantothenate (58) and screened for the presence of the inserted DNA at the *pan-2* locus using PCR. Positive strains were crossed to wild-type strain 74-OR23-1VA, and ascospores were plated on medium containing phosphinothricin and pantothenate. Progeny were screened for panthothenate auxotrophy by spot testing and for the presence of the integrated DNA using diagnostic PCR. Strain CPC2-GFP-9-10 was selected for further study (Table 4.1).

A *cpc-2* complemented strain was obtained by crossing the transformants expressing GFP-tagged *cpc-2* described above to Δ*cpc-2* mutant strain Δcpc2#11 (Table 4.1). Ascospores were plated on FGS plates containing hygromycin and pantothenate to select strains carrying the Δ*cpc-2* mutation. Progeny were spot-tested on medium containing phosphinothricin and panthothenate, followed by diagnostic PCR, to determine those that also carried the *cpc-2* GFP trans gene construct at the *pan-2* locus. Positive strains were tested for the presence of the CPC-2 GFP fusion protein using western analysis with CPC-2 antiserum as described below. Strain CPC-2-GFP-13.2 was selected for further analysis (Table 4.1).

Purification of a CPC-2 fusion protein for production of a polyclonal antiserum in rabbits. CPC-2 was expressed as an in-frame, N-terminal Maltose Binding Protein

(MBP) fusion protein in *E. coli* and then purified and used as an antigen for antibody generation in rabbits. The *cpc-2* ORF was cloned as an *Eco*RI-*Pst*I fragment in *E. coli* vector pMAL-c2X (New England Biolabs). The MBP-CPC-2 fusion protein was expressed in *E. coli* strain K12 ER2508 (New England Biolabs) with induction using 300 μM IPTG (isopropyl β-D-1-thiogalactopyranoside; Sigma) and the fusion protein purified using an amylose resin according to the manufacturer's recommendations. A polyclonal antiserum specific for the MBP-CPC-2 protein was raised in rabbits by Cocalico Biologicals, Inc. (Stevens, PA, USA).

Western analysis to confirm genotypes and check protein levels in mutants. Western analysis was used to check strains for expression of CPC-2 and the G protein subunits GNA-1, GNA-2, GNA-3 and GNB-1. For confirming genotypes, submerged cultures were grown, frozen in liquid nitrogen and then pulverized in 2-ml tubes with metal beads using a TissueLyser (Qiagen Retsch GbmH, Hannover, Germany) as previously described (47). Subsequently, 500–800 μl of extraction buffer (10mM TrisCl pH 7.5, 0.5 mM EDTA, 0.1% Fungal Protease Inhibitor Cocktail (FPIC), 1 mM PMSF and 1mM DTT) was added to the tube, the solution was mixed and then centrifuged at 5000 x g for 10 min at 4˚C. Protein concentration was determined using the Bradford Protein Reagent Concentrate (Bio-Rad, Hercules, CA). Approximately 50 μg of supernatant protein (whole cell extract) was loaded onto a 10% SDS-PAGE gel and then transferred to a nitrocellulose membrane (GE Water and Process Technologies) (61). For checking G protein levels in the Δ*cpc-2* mutant, cultures were grown and the protein fraction

enriched in plasma membranes was isolated as previously described (62). For determining CPC-2 protein amount in the G protein mutants, whole cell extracts were isolated as previously described (62). The protein concentration in the preparations was determined using the Bradford Protein Reagent Concentrate. Aliquots containing equal amounts of protein were subjected to SDS-PAGE, and a western blot was prepared as described above for confirming genotypes of G protein subunit mutants.

Western blot membranes were reacted with the CPC-2 antibody at a dilution of 1:1000 or antiserum raised against GNA-1, GNA-2, GNA-3 or GNB-1 at dilutions of 1:2000 (55, 56, 63, 64). Blots were then incubated with a goat anti-rabbit antibody horseradish peroxidase conjugate (Bio-Rad; 1:10,000 dilution). Chemiluminescent detection was performed as previously described (61) using the Super Signal West Pico Plus kit (Thermo Fisher, Rockford, IL). Western blots presented in figures are representative of three biological replicates.

Phenotypic analysis. Quantitative assays for aerial hyphae height and growth rates of basal hyphae and qualitative analysis of female fertility were performed as described previously (6, 65). Twelve biological replicates were obtained for aerial hyphae height and four were used for basal hyphae growth rate calculations. Investigation of hyphal morphology and conidiation in submerged cultures and conidial germination on solid medium were conducted as described previously (55, 66) and the results shown are representative of 2–3 biological replicates. Because the Δ*cpc-2 gna3*^{Q208L} strain

KAB3210 does not produce appreciable macroconidia, 200 microliters of packed aerial hyphae were used to inoculate submerged cultures for this strain. For quantifying macroconidia, strains were inoculated in 13x100mm glass slant tubes containing 3 ml of VM agar medium and incubated for 4 days in the dark at 30˚C and 3 days in light at room temperature. Macroconidia were collected from tubes by adding 2 ml water, mixing vigorously using a vortex mixer and filtering through Handiwipes into a 15 ml conical tube using a small funnel. This step was repeated twice, once using a wooden stick to dislodge residual macroconidia from the glass tube prior to vortexing and filtering. Macroconidia were pelleted by centrifugation and the water aspirated. Water was added to an appropriate volume and the absorbance read at 600nm using a spectrophotometer. The readings for different strains were all normalized to the same volume (1 ml) to yield a macroconidial concentration expressed as OD600/ml. Eight biological replicates were obtained.

GraphPad Prism 6.0 (GraphPad Software Inc., La Jolla, CA) was used to analyze quantitative traits (hyphal growth rate, aerial hyphae height and conidia abundance). Grubb's Q test was utilized to detect and eliminate outliers and then the Ordinary One-Way ANOVA test was used for detecting statistical significance. The p-value cutoff was set to 0.05, confidence intervals were 95% and pair-wise comparisons were made. Graphs were created using Microsoft Excel (Microsoft, Redmond, WA).

CPC-2 localization experiments. Two approaches were undertaken to determine the intracellular localization of the CPC-2 protein: cell fractionation studies using centrifugation with a wild-type strain and live-cell microscopic imaging of a strain that produces GFP-tagged CPC-2. Cell fractionation of a whole cell extract of strain 74-OR23-IVA (Table 4.1) was performed as described (55). Fractions containing whole cell extract, cytosol, and the particulate fraction (membranous organelles and large macromolecular structures) were isolated. The volumes of the cytosol and particulate fractions were adjusted to the same total volume as the original whole cell extract to allow comparison. The protein concentration of the whole cell extract was determined using the Bradford Protein Reagent Concentrate (Bio-Rad). Aliquots containing a volume identical to that containing 50 μg of protein from the whole cell extract were subjected to SDS-PAGE and gels were blotted onto nitrocellulose membranes. Antibody to arginase/AGA (cytosolic marker) (67) was used at a dilution of 1:10,000 and the plasma membrane ATPase/PMA-1 (plasma membrane marker; gift from Kenneth Allen and Clifford Slayman) (68) was used at a dilution of 1:3000. Westerns shown in figures are representative of four biological replicates. Fluorescence microscopy of the CPC-2-GFP-9-10 strain was conducted essentially as described (66). The germinating conidia were visualized using differential interference microscopy on an Olympus IX71 inverted microscope (Olympus America) with a 60X oil immersion objective. For visualization of GFP fluorescence, the GFP laser was used for excitation at 400 nm. Images were captured using a QIClickTM digital CCD camera (QImaging Surrey, British Columbia, Canada).

Results

***N. crassa* CPC-2 is homologous to predicted RACK1 proteins from other fungi.** *N. crassa* CPC-2 is 316 amino acids in length and was previously reported to have 70% percent identity with RACK1 proteins (33). *N. crassa* CPC-2 and GNB-1 each possess seven WD-40 repeats and share 39% similarity and 24% identity at the protein level. In order to investigate the relationships between CPC-2, GNB-1 and RACK1 and Gβ subunit proteins from other fungi, we subjected orthologous sequences from 18 fungal species to multiple sequence alignment and tree rendering (See Materials and methods for details). Gβ and RACK1 orthologs from the plant *Arabidopsis thaliana* were included as outgroups for the analysis. The fungal species include representatives from the Ascomycota (nine species), Basidiomycota (four species), Chytridiomycota (two species) and Mucoromycota (three species) (40). Two of the species from the Mucoromycotina possessed multiple orthologs of both GNB-1 and CPC-2, and all proteins were included in our analysis.

The results for the Gβ group showed that the proteins from *N. crassa* and the other Ascomycete filamentous fungi (*Sordaria macrospora*, *Fusarium graminearum*, *Botrytis cinerea*, *Magnaporthe oryzae* and *Aspergillus nidulans*) cluster together and are more closely related to proteins from Basidiomycetes (*Ustilago maydis, Cryptococcus neoformans, Sporisorium reilianum* and *Puccinia graminis f. sp. tritici*), Chytridiomycetes (*Batrachochytrium dendrobatidis* and *Spizellomyces punctatus*) and Mucoromycetes (*Phycomyces blakesleeanus* and *Mucor circinelloides f. lusitanicus*) than to the three Ascomycete yeasts (*Saccharomyces cerevisiae*, *Candida albicans* and

Schizosaccharomyces pombe) (Figure 4.1A). These relationships are in keeping with our previous observations that a heterotrimeric Gα subunit from *N. crassa* (GNA-3) is more closely related to proteins from filamentous Ascomycetes and Basidiomycetes than to those from *S. cerevisiae* or *S. pombe* (69). We also noted that each of the GNB-1 orthologs from *M. circinelloides f. lusitanicus* cluster with 1–2 orthologs from *P. blakesleeanus*, consistent with an ancient duplication event in an ancestor of these two species and later divergence (Figure 4.1A). Evidence supporting genome duplication in these species has been previously published (70).

In contrast to the Gβ orthologs, the RACK1 proteins distribute into two major clades, with one corresponding to all of the Ascomycetes (including *N. crassa*) and the other containing the Basdiomycetes, Mucoromycetes and Chytridiomycetes (Figure 4.1B). In the case of the two Mucoromycete species, the RACK1 proteins from each species have the other protein from the same species are their closest neighbor on the tree (Figure 4.1B). This suggests a more recent gene duplication event for the RACK1 orthologs that occurred after divergence of these two species.

CPC-2 is a cytoplasmic protein. We utilized two independent methods to assess subcellular localization of CPC-2. First, differential centrifugation was performed on protein extracts from wild type and the fractions subjected to western analysis using antibodies to marker proteins and CPC-2. Since there was no antibody for CPC-2 available prior to our study, we first expressed and purified an MBP-CPC2 fusion protein

from *E. coli* and used the protein to produce polyclonal antisera in rabbits (see Materials and methods for details). Tests of the serum showed that it recognized a protein of the predicted molecular mass of CPC-2 (~35 kDa) in whole cell extracts from wild type.

For the differential centrifugation approach, we generated whole cell extracts, and samples enriched for cytosol and the particulate fraction (membranous organelles and large macromolecular assemblies). Western analysis was performed using antibodies directed against arginase/AGA (cytosolic marker) (67) and the plasma membrane ATPase/PMA-1 (plasma membrane marker) (68), with the results showing good separation of the fractions (Figure 4.2A). Some contamination of the cytosolic fraction with plasma membranes (but not vice-versa) is evident from the presence of trace amounts of PMA-1 in the cytosol and the absence of the AGA from the particulate fraction. Western analysis using the CPC-2 antibody demonstrated that the great majority of CPC-2 was localized to the cytoplasm, with a small amount in the particulate fraction.

As an alternative method, we implemented fluorescence microscopy to determine the subcellular localization of CPC-2 in a strain expressing a GFP-tagged version of the protein (Figure 4.2B). The CPC-2-GFP signal was localized in the cytoplasm and excluded from the nucleus (as represented by DAPI staining) in both macroconidia and 6 h germlings (Figure 4.2B). Thus, both subcellular fractionation and live-cell imaging approaches support a cytoplasmic localization for CPC-2 in *N. crassa*.

Creation of mutants lacking *cpc-2* and G protein subunit genes and analysis of G protein levels in Δ*cpc-2* strains. We have previously demonstrated that components of the G protein signaling pathway are crucial for hyphal growth and asexual and sexual development of *N. crassa* (8, 11, 61, 71–73). In order to explore a possible role for cpc-2 as a heterotrimeric Gβ gene in *N. crassa,* we created strains that could be used for genetic epistasis analysis. We previously employed a similar approach for analysis of genetic relationships between the Gβ gnb-1 and the three Gα subunit genes (55). We first purified Δ*cpc-2* homokaryotic knockout mutants from a transformant created during the Neurospora Genome Project (53, 74) (see Materials and methods). We constructed complemented strains carrying the Δ*cpc-2* mutation and a *pan-2* targeted, GFP tagged version of the cpc-2+ gene *in trans* (see Materials and methods and Table 4.1). The complemented strains exhibited significant complementation of several phenotypes, including hyphal growth rate (Figure 4.3) and partial complementation of aerial hyphae height (Figure 4.3). We used sexual crosses or transformation to generate deletion mutants lacking *cpc-2* alone or in combination with mutations in the three Gα genes or the Gβ, *gnb-1*. We also constructed Δ*cpc-2* strains expressing GTPase-deficient, constitutively activated Gα alleles (*gna-1*^{Q204L}, *gna-2*^{Q205L} or *gna-3*^{Q208L}; see Materials and methods and Table 4.1).

We have previously shown that, depending on the growth conditions, loss of the Gβ subunit gnb-1 leads to lower levels of one or all three Gα proteins in *N. crassa* (55, 61, 64). The exact mechanism underlying this regulation is unknown, but appears to be post-transcriptional, as Gα mRNA levels are normal in Δ*gnb-1* mutants (61, 64).

Therefore, prior to initiating genetic epistasis experiments with *cpc-2*, we utilized western blot analysis with protein-specific antisera to check levels of G protein subunits in the Δ*cpc-2* mutant (Figure 4.4). The results demonstrate that in contrast to *gnb-1*, loss of *cpc-2* does not greatly influence levels of the three Gα proteins or GNB-1 (Figure 4.4; compare wild type and Δ*cpc-2* lanes). However, Gα protein levels are still reduced when *gnb-1* is mutated in the Δ*cpc-2* background (Figure 4.4). We also consistently noted an increased level of GNA-3 in the Δ*cpc-2* Δ*gnb-1* double mutant vs. the Δ*gnb-1* single mutant, suggesting that loss of *cpc-2* partially reverses the effect of the Δ*gnb-1* mutation. The observation that the Δ*cpc-2* single mutant has normal levels of G protein subunits greatly streamlines interpretation of genetic epistasis experiments using *cpc-2*.

We next wanted to determine whether loss of any of the G protein subunits affects CPC-2 protein levels. Because CPC-2 is a cytoplasmic protein (Figure 4.2), we used protein from whole cell extracts for western analysis using the CPC-2 antiserum (Figure 4.4). The results demonstrated that CPC-2 protein levels were relatively normal in the G protein single mutants (Fig 3). Thus, similar to the situation with GNB-1 levels in the Δ*cpc-2* strain, CPC-2 levels are not affected by loss of *gnb-1*; the two predicted Gβ proteins are independent of one another in this regard. Our findings suggest that if CPC-2 does operate as a Gβ subunit, it does not share all functions with GNB-1 in *N. crassa*.

***cpc-2* is epistatic to *gna-2* during regulation of basal hyphal growth rate.** *N. crassa* grows by elongation, branching and fusion of hyphae, eventually forming a network structure called the mycelium (rev. in (75). From this mycelium, aerial hyphae grow upward and spore-forming structures (macroconidiophores) are elaborated from their tips.

Formation of cross-walls and constriction of macroconidiophores leads to formation of the mature multinucleated asexual spores, macroconidia. Macroconidia are disseminated in nature by wind currents, enabling the fungus to colonize new areas. When in the presence of water and suitable nutrients, macroconidia germinate to form a hyphal tube, which then begins the growth program described above (75).

We began our genetic epistasis analysis by investigating the set of mutants for defects in basal hyphae extension rate, using macroconidia to inoculate race tubes (see Materials and methods). The results from genetic epistasis analysis were interpreted as reported previously (55): If the phenotype of the Δ*cpc-2*, ΔGα double mutant resembles the phenotype of the ΔGα mutant, and if the mutationally activated Gα allele bypasses the phenotype of Δ*cpc-2*, then the Gα gene is epistatic to (implied downstream) to *cpc-2*. If the opposite is true, then *cpc-2* is epistatic to the Gα gene. If contradicting results are seen, this is interpreted as the two genes being partially or completely independent in regulation of the phenotype being assessed.

All of the single gene mutants had a basal hyphal growth rate phenotype (Fig 4.5). In Δ*cpc-2* mutants, the growth rate was 61% of wild type (Figure 4.5). The findings from ANOVA of the characterized strains revealed several relationships. First, *cpc-2* may operate downstream of *gna-2*. Both mutants are significantly different than wild type, the double mutant grows slower than Δ*gna-2*, but slightly faster than Δ*cpc-2*, and mutational activation of *gna-2* (*gna-2*^{Q205L} allele) does not lead to an increase in growth rate in the Δ*cpc-2* background. Second, the Δ*gna-1* and Δ*cpc-2* knockout mutations are synergistic with regards to reduction in growth rate, and mutational activation of *gna-1* does not

rescue the Δ*cpc-2* phenotype; in fact, the growth rate is further reduced in the Δ*cpc-2* *gna-1*^{Q204L} strain. The same relationships hold between *gna-3* and *cpc-2*. These results suggest that *cpc-2* regulates growth rate using a different pathway than *gna-1* or *gna-3*. Finally, Δ*cpc-2* and Δ*gnb-1* are also synergistic, with the double mutant having a significantly slower growth rate than either single mutant (Figure 4.5). This suggests that these two Gβ-like genes have some independent functions during regulation of hyphal growth in *N. crassa*.

We have previously demonstrated that strains lacking either of the G protein subunit genes *gna-1* and *gna-3*, but not *gna-2* or *gnb-1*, have a defect in germination of macroconidia, an essential step prior to hyphal growth and formation of a colony (66). Therefore, we explored this phenotype in Δ*cpc-2* mutants, using wild type as a control (Figure 4.6). Similar to Δ*gnb-1* mutants, strains lacking *cpc-2* are normal with respect to germination of macroconidia (Figure 4.6). Thus, overall colony size of Δ*cpc-2* mutants is compromised by slower extension of basal hyphae, and not by a defect in germination of macroconidia.

***cpc-2* is epistatic to *gna-2* with regards to aerial hyphae height and Δ*cpc-2* mitigates the increased macroconidia production of Δ*gnb-1* mutants on solid medium.** We next explored epistatic relationships between *cpc-2* and the other genes for two quantitative traits relevant to macroconidiation: aerial hyphae height and macroconidia abundance. Similar to the case for basal hyphae growth rate, all of the single gene deletion mutants had an aerial hyphae height defect (Figure 4.7A). For Δ*cpc-2*, aerial hyphae heights were 69% of wild type (Figure 4.7A).

ANOVA of the strain set produced results similar to those noted for basal hyphae, above. The aerial hyphae height of the Δ*cpc-2* and Δ*cpc-2* Δ*gna-2* double mutants is significantly less than that of the Δ*gna-2* single mutant and the Δ*cpc-2 gna-2*^{Q205L} strain is similar to the Δ*cpc-2* single mutant (Figure 4.7A). This result is consistent with *cpc-2* functioning downstream of *gna-2* to control aerial hyphae height. In contrast, *gna-1* and *cpc-2* appear to be independent; the double mutant has shorter aerial hyphae than either single mutant and introduction of *gna-1*^{Q204L} does not rescue the aerial hyphae defect of Δ*cpc-2* (Figure 4.7A). Δg*na-3* mutants are shorter than Δ*cpc-2* and the double mutant is similar to Δ*gna-3* (Figure 4.7A). However, the finding that aerial hyphae height is not rescued by the *gna-3*^{Q208L} allele in the Δ*cpc-2* background supports independent regulation by these two subunits. Δ*cpc-2* and Δ*gnb-1* mutants have similar aerial hyphae height and the double mutant is shorter (Figure 4.7A). As observed for regulation of basal hyphal growth, this finding supports independent signaling by *cpc-2* and *gnb-1* during control of aerial hyphae height in *N. crassa*.

Quantitative analysis of macroconidia production in agar slants did not reveal a phenotype for Δ*cpc-2* mutants (Figure 4.7B). In fact, of the single mutants analyzed, only Δ*gnb-1* possessed a phenotype (greater conidia production; Figure 4.7B) and the phenotype of the Δ*cpc-2* Δ*gnb-1* double mutant was similar to that of Δ*cpc-2* (like wild type). This suggests that loss of *cpc-2* mitigates the overproduction of conidia observed in the Δ*gnb-1* mutant, and that *cpc-2* is epistatic to *gnb-1*. For the Gα subunit double mutants, Δ*cpc-2* Δ*gna-3* produces fewer conidia than either single mutant and differentiation of macroconidia is nearly halted in the Δ*cpc-2 gna-3*^{Q208L} strain (Figure

4.7B). These results support independence of *cpc-2* and *gna-3* during regulation of macroconidiation. A similar situation exists for *gna-2*, as the Δ*cpc-2* Δ*gna-2* double mutant and the Δ*cpc-2 gna-2*^{Q205L} strain produce fewer conidia than either single mutant (Figure 4.7B). With *gna-1*, the double mutant is similar to the single mutants, but the Δ*cpc-2 gna-1*^{Q204L} mutant produces less macroconidia, consistent with independence (Figure 4.7B). The results from analysis of strains carrying the three mutationally activated Gα alleles suggest that all three Gα proteins inhibit macroconidiation when locked in the GTP-bound form.

Δ*cpc-2* mutants produce macroconidia in submerged cultures. Wild-type *N. crassa* strains do not differentiate macroconidia while growing in shaken submerged culture unless subjected to heat shock, nitrogen or carbon starvation (76–80). We have previously demonstrated that loss of the G protein subunits *gna-3, gnb-1* and *gng-1* leads to macroconidiation in submerged culture under all conditions (61, 64, 69), while Δ*gna-1* mutants only form macroconidia at high inoculation cell density (>=3x106 /ml) in liquid culture (81). Based on the precedent that the Gβ gene *gnb-1* is a negative regulator of macroconidiation in submerged cultures, we analyzed our group of strains for phenotypes at a low inoculation density (1x106 /ml). Similar to previous findings, wild type and Δ*gna-2* mutants do not produce macroconidiophores in submerged culture, while single mutants lacking *gna-3, gnb-1* and *gng-1* all produce abundant macroconidiophores (Figure 4.8). Rare macroconidiophores could also be observed in the Δ*gna-1* strain. We also noted that Δ*cpc-2* knockout mutants produce macroconidiophores in submerged culture (Figure 4.8). Double mutants Δ*cpc-2* Δ*gna-1*, Δ*cpc-2* Δ*gna-2*, Δ*cpc-2* Δ*gna-3* and

Δ*cpc-2* Δ*gnb-1* all produce conidia in submerged culture. In all four cases, loss of *cpc-2* either leads to submerged conidiation or intensifies the conidiation phenotype of the G protein subunit mutants and the Δ*gna-3* Δ*cpc-2* double mutant cultures are mostly conidia (Figure 4.8). Interestingly, introduction of mutationally activated *gna-3* corrects the submerged conidiation phenotype of Δ*cpc-2*, while the corresponding activated alleles of *gna-1* or *gna-2* do not (Figure 4.8). This result suggests that GNA-3 may operate downstream of CPC-2, but also has a CPC-2 independent function in controlling submerged conidiation.

Constitutive activation of Gα subunits does not restore female fertility to the Δcpc-2 mutant. *N. crassa* is a heterothallic organism, meaning that a given strain has one of two different mating type genes present at a single genomic locus (idiomorphs; *mat* A or *mat* a) (82, 83). Upon nitrogen limitation, *N. crassa* forms protoperithecia (the female reproductive structures) (75, 84). In the presence of a male cell (usually conidia) of the opposite mating type, pheromone detection results in chemotropic growth of specialized hyphae called trichogynes from the protoperithecium. The fruiting body, or perithecium, is then formed and contains asci, each with eight haploid spores (ascospores). Upon maturation, ascospores are ejected from the tips (beaks) of perithecia, in the direction of light. Under laboratory conditions, protoperithecial development can be induced using Synthetic Crossing Medium (SCM), and progeny are obtained from sexual crosses approximately 2–3 weeks post-fertilization (75, 84).

Our previous work showed that the mutationally activated *gna-1*^{Q204L}, *gna-2*^{Q205L} and *gna3*^{Q208L} alleles were not able to restore fertility to the Δ*gnb-1* mutant. In fact,

introduction of *gna-3*^{Q208L} resulted in complete inhibition of protoperithecial development, a phenotype that was more severe than that of the Δ*gnb-1* mutant (55). Δ*gnb-1* ΔGα double mutant strains resemble the Δ*gnb-1* mutant, in that they form protoperithecia, but no perithecia after fertilization (55).

Muller et al. (33) previously reported that *cpc-2* point mutants do not produce protoperithecia and are thus female-sterile. In contrast, our results with the Δ*cpc-2* knockout mutant indicate some protoperithecia are present, as the cultures produce rare perithecia after fertilization that are mostly submerged in the agar (Figure 4.9). This phenotype is distinct from that of Δ*gna-2* and Δ*gna-3* strains that produce perithecia similar to wild type and from Δ*gna-1* and Δ*gnb-1* mutants that do not form perithecia after fertilization (Fig 4.9).

Inspection of double mutants revealed that Δ*cpc-2* Δ*gna-1* strains do not produce visible protoperithecia, perithecia or ascospores (Figure 4.9), a more severe phenotype than either single mutant. In contrast, Δ*cpc-2* Δ*gna-2* mutants resemble Δ*cpc-2* single mutants. Δ*cpc-2* Δ*gna-3* strains exhibit a variable phenotype, with either small, submerged perithecia or no visible perithecia (Figure 4.9), and no ascospores. Mutational activation of either *gna-1* or *gna-3* in the Δ*cpc-2* background leads to no visible protoperithecia, perithecia or ascospores, while activation of *gna-2* results in the Δ*cpc-2* phenotype (Figure 4.9). These results are consistent with synergy between *cpc-2* and *gna-1* and *gna-3*. The phenotype of Δ*cpc-2 gna-3*^{Q208L} and Δ*gnb-1 gna3*^{Q208L} strains are similar (55), suggesting a common mode of action for *gna-3*^{Q208L} and/or interaction between GNA-3 and the two candidate Gβ proteins. In contrast, the different results

observed after introduction of *gna-1*^{Q204L} into the two mutants hints at a different role for CPC-2 vs. GNB-1 during regulation of female fertility.

As noted previously, the Δ*gnb-1* strain forms small, aberrant protoperithecia, but no perithecia, upon fertilization (61) (Figure 4.9). In contrast, similar to Δ*gnb-1* single mutants, Δ*cpc-2* Δ*gnb-1* double mutants do not produce perithecia (Figure 4.9). This result suggests that *gnb-1* is epistatic to *cpc-2* during sexual development.

Discussion

The *N. crassa cpc-2* gene is not essential and the encoded protein is similar to other RACK1 homologs in fungi. Genetic epistasis between *cpc-2* and components of the G protein pathway was performed using double deletion mutants and strains containing Gα activated alleles (see model in Figure 4.10). The results revealed genetic relationships between *cpc-2* and *gna-2* during growth of basal and aerial hyphae, *cpc-2* and *gna-3* during growth in submerged cultures and *cpc-2* and *gnb-1* in regulation of sexual development. In the cases of basal and aerial hyphae growth, the epistatic relationships suggest that CPC-2 operates downstream of the Gα protein, implying a tethering function for the Gα in regulation of CPC-2. However, the GNA-3 Gα acts downstream of CPC-2 during submerged culture conidiation, suggesting that the RACK1 protein is holding GNA-3 inactive. CPC-2 appears to operate upstream of the Gβ GNB-1 during sexual development and to act in an antagonistic function during production of macroconidia in agar cultures.

Our investigation of epistasis between the three Gα genes and *gnb-1* and *cpc-2* revealed some interesting parallels. As mentioned above, the results from the current study suggest that *gna-3* is at least partially epistatic to *cpc-2* during control of appropriate conidiation in submerged cultures. This is similar to the earlier relationship observed for *gnb-1* and *gna-3* for this same phenotype, with *gna-3* epistatic to *gnb-1* (55). The other two Gα subunits are independent of both *cpc-2* (this study) and *gnb-1* (55) during the regulation of this trait. This indicates that *cpc-2*, like *gnb-1*, is a negative regulator of conidiation in submerged culture, and that only activation of *gna-3* offers a genetic bypass mechanism to restore normal hyphal growth. Our previous results from epistasis studies of aerial hyphae height demonstrated that *gnb-1* is epistatic to both *gna-2* and *gna-3* and independent of *gna-1* (55). Together with the current study, the findings are consistent with a model in which the Gβ gene lies downstream of the Gα gene(s) and that *gna-1* is independent of both *gnb-1* and *cpc-2* during aerial hyphae elongation. However, any conclusions based on Δ*cpc-2* Δ*gnb-1* double mutants need to be tempered, as loss of *gnb-1* leads to decreased levels of the three Gα proteins in all genetic backgrounds tested.

It is intriguing that the Δ*cpc-2* Δ*gnb-1* double mutants have higher levels of *GNA-3* protein than Δ*gnb-1* single mutants (but still less than in wild type; Fig 3A). This finding suggests that loss of *cpc-2* partially mitigates the effects of the Δ*gnb-1* mutation. In a canonical model for G protein signaling, GNB-1 would function as a GDI for GNA-3 and loss of GNB-1 might lead to misfolding and/or proteolysis of GNA-3. Mutation of *cpc-2* partially counteracts this effect, suggesting that CPC-2 participates in the pathway

leading to decreased levels of GNA-3 protein. Furthermore, the finding that Gα single mutants have slower basal hyphae growth rates than wild type and that loss of *gnb-1* leads to lower levels of Gα proteins supports a possible tethering function for GNB-1 during hyphal growth. Loss of one Gα protein could free more GNB-1 to bind the other Gα subunits, potentially inhibiting them from serving as positive regulators of basal hyphae growth rate. Along these lines, it has been demonstrated in *S. cerevisiae* that levels of the Gα protein Gpa1p are regulated by ubiquitin-mediated proteolysis, and it has been proposed that this is a mechanism used to modulate levels of the active, free Gβγ dimer during mating (85, 86).

Analysis of the sexual cycle demonstrated that the Δ*cpc-2* forms rare protoperithecia and perithecia and is therefore female sterile. In contrast, mutants lacking other components of the cross pathway control network (*cpc-1* and *cpc-3*) have normal sexual cycles (87, 88). This indicates that the sexual cycle defect of Δ*cpc-2* mutants is not solely due to a defect in the response to amino acid limitation. However, there is a possibility that the two processes may be linked. It has been reported in *A. nidulans* that amino acid limitation arrests sexual development (28). Furthermore, loss of the RACK1/*cpc-2* homolog *cpcB* or overexpression of the *cpc-1* homolog *cpcA* also block sexual development, supporting a link between the sexual cycle program and the network that regulates amino acid biosynthesis (28).

Attempts to detect an interaction between CPC-2 and other G protein subunits in *N. crassa* using the yeast two-hybrid assay were unsuccessful. Presumably due to the large number of binding partners, the difficulty in solubilizing peripheral membrane

proteins such as Gα subunits, and protein folding concerns with heterologously expressed proteins, we were also unable to achieve co-immunoprecipitation between CPC-2 and GNB-1 or any of the three Gα proteins using cell extracts or proteins expressed and purified from *E. coli.* A similar result has been reported for the RACK1 homolog RAK1 in U. maydis (30). Knowledge of the interactions between RACK1 and G protein subunits is important for full understanding of the biology of G protein signaling. Therefore, experiments such as these and others that investigate the detailed mechanistic wiring that connects CPC-2 to heterotrimeric G protein signaling will be the focus of future work.

References

1. Bock A, Kostenis E, Trankle C, Lohse MJ, Mohr K. Pilot the pulse: controlling the multiplicity of receptor dynamics. Trends Pharmacol Sci. 2014; 35(12):630–8.

2. Chini B, Parenti M, Poyner DR, Wheatley M. G-protein-coupled receptors: from structural insights to functional mechanisms. Biochem Soc Trans. 2013; 41(1):135–6.

3. Tesmer JJ. The quest to understand heterotrimeric G protein signaling. Nat Struct Mol Biol. 2010; 17(6):650–2. https://doi.org/10.1038/nsmb0610-650 PMID: 20520658.

4. Selker EU. Neurospora. Curr Biol. 2011; 21(4):R139–40. Epub 2011/02/22.

5. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, et al. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev. 2004; 68(1):1–108.

6. Cabrera IE, Pacentine IV, Lim A, Guerrero N, Krystofova S, Li L, et al. Global Analysis of Predicted G Protein-Coupled Receptor Genes in the Filamentous Fungus, Neurospora crassa. G3 (Bethesda). 2015; 5(12):2729–43.

7. Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 2007; 61:423–52.

8. Wright SJ, Inchausti R, Eaton CJ, Krystofova S, Borkovich KA. RIC8 is a guanine-nucleotide exchange factor for Galpha subunits that regulates growth and development in Neurospora crassa. Genetics. 2011; 189(1):165–76.

9. Li L, Borkovich KA. GPR-4 Is a Predicted G-Protein-Coupled Receptor Required for Carbon SourceDependent Asexual Growth and Development in Neurospora crassa. Eukaryot Cell. 2006; 5(8):1287– 300.

10. Kim H, Wright SJ, Park G, Ouyang S, Krystofova S, Borkovich KA. Roles for receptors, pheromones, G proteins, and mating type genes during sexual reproduction in Neurospora crassa. Genetics. 2012; 190(4):1389–404.

11. Krystofova S, Borkovich KA. The predicted G-protein-coupled receptor GPR-1 is required for female sexual development in the multicellular fungus Neurospora crassa. Eukaryotic Cell. 2006; 5(9):1503– 16.

12. Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 2007; 61:423–52.

13. Turner GE, Borkovich KA. Identification of a G protein alpha subunit from Neurospora crassa that is a member of the Gi family. J Biol Chem. 1993; 268(20):14805–11. Epub 1993/07/15.

14. Kays AM, Rowley PS, Baasiri RA, Borkovich KA. Regulation of conidiation and adenylyl cyclase levels by the G-alpha protein GNA-3 in Neurospora crassa. Mol Cell Biol. 2000; 20(20):7693–705.

15. Adams DR, Ron D, Kiely PA. RACK1, A multifaceted scaffolding protein: Structure and function. Cell Commun Signal. 2011; 9:22.

16. Adams DR, Ron D, Kiely PA. RACK1, A multifaceted scaffolding protein: Structure and function. Cell Commun Signal. 2011; 9. Artn 22

17. Ron D, Chen CH, Caldwell J, Jamieson L, Orr E, Mochly-Rosen D. Cloning of an intracellular receptor for protein kinase C: a homolog of the beta subunit of G proteins. Proc Natl Acad Sci U S A. 1994; 91 (3):839–43.

18. Lopez-Bergami P, Habelhah H, Bhoumik A, Zhang W, Wang LH, Ronai Z. RACK1 mediates activation of JNK by protein kinase C [corrected]. Mol Cell. 2005; 19(3):309–20.

19. Chen S, Dell EJ, Lin F, Sai J, Hamm HE. RACK1 regulates specific functions of Gbetagamma. J Biol Chem. 2004; 279(17):17861–8.

20. Sengupta J, Nilsson J, Gursky R, Spahn CM, Nissen P, Frank J. Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. Nat Struct Mol Biol. 2004; 11(10):957–62.

21. Zeller CE, Parnell SC, Dohlman HG. The RACK1 ortholog Asc1 functions as a G-protein beta subunit coupled to glucose responsiveness in yeast. J Biol Chem. 2007; 282(34):25168–76. Epub 2007/06/27.

22. Palmer DA, Thompson JK, Li L, Prat A, Wang P. Gib2, a novel Gbeta-like/RACK1 homolog, functions as a Gbeta subunit in cAMP signaling and is essential in Cryptococcus neoformans. J Biol Chem. 2006; 281(43):32596–605. Epub 2006/09/05.

23. Zeller CE, Parnell SC, Dohlman HG. The RACK1 ortholog Asc1 functions as a G-protein beta subunit coupled to glucose responsiveness in yeast. J Biol Chem. 2007; 282(34):25168–76.

24. Palmer DA, Thompson JK, Li L, Prat A, Wang P. Gib2, a novel Gbeta-like/RACK1 homolog, functions as a Gbeta subunit in cAMP signaling and is essential in Cryptococcus neoformans. J Biol Chem. 2006; 281(43):32596–605.

25. Li G, Zhang X, Tian H, Choi YE, Tao WA, Xu JR. MST50 is involved in multiple MAP kinase signaling pathways in Magnaporthe oryzae. Environ Microbiol. 2017; 19(5):1959–74.

26. Yin Z, Zhang X, Wang J, Yang L, Feng W, Chen C, et al. MoMip11, a MoRgs7-interacting protein, functions as a scaffolding protein to regulate cAMP signaling and pathogenicity in the rice blast fungus Magnaporthe oryzae. Environ Microbiol. 2018; 20(9):3168–85.

27. Nunez A, Franco A, Madrid M, Soto T, Vicente J, Gacto M, et al. Role for RACK1 orthologue Cpc2 in the modulation of stress response in fission yeast. Mol Biol Cell. 2009; 20(18):3996–4009.

28. Hoffmann B, Wanke C, Lapaglia SK, Braus GH. c-Jun and RACK1 homologues regulate a control point for sexual development in Aspergillus nidulans. Mol Microbiol. 2000; 37(1):28–41.

29. Cai ZD, Chai YF, Zhang CY, Qiao WR, Sang H, Lu L. The Gbeta-like protein CpcB is required for hyphal growth, conidiophore morphology and pathogenicity in Aspergillus fumigatus. Fungal Genet Biol. 2015; 81:120–31.

30. Wang L, Berndt P, Xia X, Kahnt J, Kahmann R. A seven-WD40 protein related to human RACK1 regulates mating and virulence in Ustilago maydis. Mol Microbiol. 2011; 81(6):1484–98.

31. Kruger D, Koch J, Barthelmess IB. cpc-2, a new locus involved in general control of amino acid synthetic enzymes in Neurospora crassa. Curr Genet. 1990; 18(3):211–5.

32. Barthelmess IB, Kolanus J. The range of amino acids whose limitation activates general amino-acid control in Neurospora crassa. Genet Res. 1990; 55(1):7–12. Epub 1990/02/01.

33. Muller F, Kruger D, Sattlegger E, Hoffmann B, Ballario P, Kanaan M, et al. The cpc-2 gene of Neurospora crassa encodes a protein entirely composed of WD-repeat segments that is involved in general amino acid control and female fertility. Mol Gen Genet. 1995; 248(2):162–73..

34. Paluh JL, Orbach MJ, Legerton TL, Yanofsky C. The cross-pathway control gene of Neurospora crassa, cpc-1, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene v-junencoded protein. Proc Natl Acad Sci U S A. 1988; 85(11):3728–32.

35. Sattlegger E, Hinnebusch AG, Barthelmess IB. cpc-3, the Neurospora crassa homologue of yeast GCN2, encodes a polypeptide with juxtaposed eIF2alpha kinase and histidyl-tRNA synthetase-related domains required for general amino acid control. J Biol Chem. 1998; 273(32):20404–16.

36. McCluskey K, Wiest A, Plamann M. The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. J Biosci. 2010; 35(1):119–26.

37. Vogel HJ. Distribution of lysine pathways among fungi: Evolutionary implications. Am Nat. 1964; 98:435–46.

38. Westergaard M, Mitchell HK. Neurospora V. A synthetic medium favoring sexual reproduction. Amer J Bot. 1947; 34:573–7.

39. Davis RH, deSerres FJ. Genetic and microbiological research techniques for Neurospora crassa. Methods Enzymol. 1970; 71A:79–143.

40. Spatafora JW, Aime MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M. The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. Microbiol Spectr. 2017; 5(5).

41. Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, et al. FungiDB: an integrated functional genomics database for fungi. Nucleic Acids Res. 2012; 40(Database issue):D675–81.

42. Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, et al. NGPhylogeny. fr: new generation phylogenetic services for non-specialists. Nucleic Acids Res. 2019; 47(W1):W260– W5.

43. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4):772–80.

44. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010; 10:210.

45. Lefort V, Desper R, Gascuel O. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. Mol Biol Evol. 2015; 32(10):2798–800.

46. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019; 47(W1):W256–W9.

47. Ghosh A, Servin JA, Park G, Borkovich KA. Global analysis of serine/threonine and tyrosine protein phosphatase catalytic subunit genes in Neurospora crassa reveals

interplay between phosphatases and the p38 mitogen-activated protein kinase. G3 (Bethesda). 2014; 4(2):349–65.

48. Avalos J, Geever RF, Case ME. Bialaphos resistance as a dominant selectable marker in Neurospora crassa. Curr Genet. 1989; 16(5–6):369–72.

49. Pall M. The use of Ignite (basta; glufosinate; phosphinothricin) to select transformants of bar-containing plasmids in Neurospora crassa. Fungal Genet Newsl. 1993; 40:57.

50. Perkins D. Advantages of using the inactive-mating-type am1 strain as a helper component in heterokaryons. Neurospora Newslett. 1985; 31:41–2.

51. Kuck U, Hoff B. Application of the nourseothricin acetyltransferase gene (nat1) as dominant marker for the transformation of filamentous fungi. Fungal Genet Newsl. 2006; 53:9–11.

52. Christianson TW, Sikorski RS, Dante M, Shero JH, Hieter P. Multifunctional yeast high-copy-number shuttle vectors. Gene. 1992; 110(1):119–22. https://doi.org/10.1016/0378-1119(92)90454-w

53. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, et al. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A. 2006; 103(27):10352–7. Epub 2006/06/28.

54. Winston F, Dollard C, Ricupero-Hovasse SL. Construction of a set of convenient Saccharomyces cerevisiae strains that are isogenic to S288C. Yeast. 1995; 11(1):53–5.

55. Won S, Michkov AV, Krystofova S, Garud AV, Borkovich KA. Genetic and physical interactions between Galpha subunits and components of the Gbetagamma dimer of heterotrimeric G proteins in Neurospora crassa. Eukaryot Cell. 2012; 11(10):1239–48.

56. Ivey FD, Hodge PN, Turner GE, Borkovich KA. The G alpha i homologue gna-1 controls multiple differentiation pathways in Neurospora crassa. Mol Biol Cell. 1996; 7(8):1283–97.

57. Ebbole DJ, Sachs MS. A rapid and simple method for isolation of Neurospora crassa homokaryons using microconidia. Fungal Genet Newsl. 1990; 37:17–8.

58. Ouyang S, Beecher CN, Wang K, Larive CK, Borkovich KA. Metabolic Impacts of Using Nitrogen and Copper-Regulated Promoters to Regulate Gene Expression in Neurospora crassa. G3 (Bethesda). 2015; 5(9):1899–908.

59. Freitag M, Hickey PC, Raju NB, Selker EU, Read ND. GFP as a tool to analyze the organization, dynamics and function of nuclei and microtubules in Neurospora crassa. Fungal Genet Biol. 2004; 41(10):897–910. Epub 2004/09/03.

60. Park G, Servin JA, Turner GE, Altamirano L, Colot HV, Collopy P, et al. Global analysis of serine-threonine protein kinase genes in Neurospora crassa. Eukaryotic cell. 2011; 10(11):1553–64. Epub 2011/10/ 04.

61. Krystofova S, Borkovich KA. The heterotrimeric G-protein subunits GNG-1 and GNB-1 form a Gbetagamma dimer required for normal female fertility, asexual development, and galpha protein levels in Neurospora crassa. Eukaryot Cell. 2005; 4(2):365–78.

62. Turner GE, Borkovich KA. Identification of a G protein alpha subunit from Neurospora crassa that is a member of the Gi family. J Biol Chem. 1993; 268(20):14805–11.

63. Baasiri RA, Lu X, Rowley PS, Turner GE, Borkovich KA. Overlapping functions for two G protein alpha subunits in Neurospora crassa. Genetics. 1997; 147(1):137–45. Epub 1997/09/01.

64. Yang Q, Poole SI, Borkovich KA. A G-protein beta subunit required for sexual and vegetative development and maintenance of normal G alpha protein levels in Neurospora crassa. Eukaryot Cell. 2002; 1 (3):378–90. Epub 2002/11/29.

65. Turner GE. Phenotypic analysis of Neurospora crassa gene deletion strains. Methods Mol Biol. 2011; 722:191–8.

66. Eaton CJ, Cabrera IE, Servin JA, Wright SJ, Cox MP, Borkovich KA. The guanine nucleotide exchange factor RIC8 regulates conidial germination through Galpha proteins in Neurospora crassa. PLoS One. 2012; 7(10):e48026.

67. Borkovich KA, Weiss RL. Purification and characterization of arginase from Neurospora crassa. J Biol Chem. 1987; 262(15):7081–6.

68. Hager KM, Mandala SM, Davenport JW, Speicher DW, Benz EJ Jr., Slayman CW. Amino acid sequence of the plasma membrane ATPase of Neurospora crassa: deduction from genomic and cDNA sequences. Proc Natl Acad Sci U S A. 1986; 83(20):7693–7.

69. Kays AM, Rowley PS, Baasiri RA, Borkovich KA. Regulation of conidiation and adenylyl cyclase levels by the Galpha protein GNA-3 in Neurospora crassa. Mol Cell Biol. 2000; 20(20):7693–705. Epub 2000/ 09/26.

70. Corrochano LM, Kuo A, Marcet-Houben M, Polaino S, Salamov A, Villalobos-Escobedo JM, et al. Expansion of Signal Transduction Pathways in Fungi by Extensive Genome Duplication. Curr Biol. 2016; 26(12):1577–84.

71. Kim H, Borkovich KA. A pheromone receptor gene, pre-1, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in Neurospora crassa. Mol Microbiol. 2004; 52(6):1781–98.

72. Kays AM, Borkovich KA. Severe impairment of growth and differentiation in a Neurospora crassa mutant lacking all heterotrimeric G alpha proteins. Genetics. 2004; 166(3):1229–40.

73. Li L, Borkovich KA. GPR-4 is a predicted G-protein-coupled receptor required for carbon source-dependent asexual growth and development in Neurospora crassa. Eukaryot Cell. 2006; 5(8):1287–300.

74. Park G, Colot HV, Collopy PD, Krystofova S, Crew C, Ringelberg C, et al. High-throughput production of gene replacement mutants in Neurospora crassa. Methods Mol Biol. 2011; 722:179–89.

75. Springer ML. Genetic control of fungal differentiation: the three sporulation pathways of Neurospora crassa. Bioessays. 1993; 15(6):365–74.

76. Cortat M, Turian G. Conidiation of Neurospora crassa in submerged culture without mycelial phase. Arch Mikrobiol. 1974; 95(4):305–9.

77. That TC, Turian G. Ultrastructural study of microcyclic macroconidiation in Neurospora crassa. Arch Microbiol. 1978; 116(3):279–88.

78. Plesofsky-Vig N, Light D, Brambl R. Paedogenetic Conidiation in Neurospora crassa. Exp Mycol. 1983; 7:283–6.

79. Guignard R, Grange F, Turian G. Microcycle Conidiation Induced by Partial Nitrogen Deprivation in Neurospora crassa. Can J Microbiol. 1984; 30(10):1210–5.

80. Madi L, McBride SA, Bailey LA, Ebbole DJ. rco-3, a gene involved in glucose transport and conidiation in Neurospora crassa. Genetics. 1997; 146(2):499–508.

81. Ivey FD, Kays AM, Borkovich KA. Shared and independent roles for a Galpha(i) protein and adenylyl cyclase in regulating development and stress responses in Neurospora crassa. Eukaryot Cell. 2002; 1 (4):634–42. Epub 2002/11/29.

82. Metzenberg RL, Glass NL. Mating type and mating strategies in Neurospora. Bioessays. 1990; 12 (2):53–9.

83. Randall TA, Metzenberg RL. Species-specific and mating type-specific DNA regions adjacent to mating type idiomorphs in the genus Neurospora. Genetics. 1995; 141(1):119–36.

84. Raju NB. Neurospora as a model fungus for studies in cytogenetics and sexual biology at Stanford. J Biosci. 2009; 34(1):139–59.

85. Xu BE, Kurjan J. Evidence that mating by the Saccharomyces cerevisiae gpa1Val50 mutant occurs through the default mating pathway and a suggestion of a role for ubiquitin-mediated proteolysis. Mol Biol Cell. 1997; 8(9):1649–64.

86. Schauber C, Chen L, Tongaonkar P, Vega I, Madura K. Sequence elements that contribute to the degradation of yeast G alpha. Genes Cells. 1998; 3(5):307–19. PMID: 9685182.

87. Sattlegger E, Hinnebusch AG, Barthelmess IB. cpc-3, the Neurospora crassa homologue of yeast GCN2, encodes a polypeptide with juxtaposed eIF2 alpha kinase and histidyl-tRNA synthetase-related domains required for general amino acid control. J Biol Chem. 1998; 273(32):20404–16.

88. Paluh JL, Orbach MJ, Legerton TL, Yanofsky C. The Cross-Pathway Control Gene of NeurosporaCrassa, Cpc-1, Encodes a Protein Similar to Gcn4 of Yeast and the DNA-Binding Domain of the Oncogene V-Jun-Encoded Protein. Proceedings of the National Academy of Sciences of the United States of America. 1988; 85(11):3728–32.

Figure Legends

Figure 4.1: Phylogenetic analysis of Gβ and RACK1 proteins from 10 fungal species. Amino acid sequences were obtained from FungiDB or NCBI and phylogenetic analysis conducted using the "One-Click Workflow" tool at NGPhylogeny.fr. The final trees were drawn using tools at itol.embl.de (see Materials and methods for details). **A. Gβ proteins.** Organisms and protein names/accession numbers for the Gβ orthologs are *Neurospora crassa* NcGNB-1/NCU00440; *Sordaria macrospora* SmGbeta/SMAC01876; *Fusarium graminearum* GzGPB1/FGRAMPH101G14499; *Magnaporthe oryzae* MoMgb1/MGG05201; *Aspergillus nidulans* AnSfaD/AN0081; *Ustilago maydis* UmBpp1/UMAG00703; *Cryptococcus neoformans* CnGpb1/CNAG01262; *Candida albicans* CaSte4/C204210WA; *Schizosaccharomyces pombe* SpGit5/SPBC32H8.07 and *Saccharomyces cerevisiae* ScSte4/YOR212W; *Batrachochytrium dendrobatidis*/BDEG_08231; *Botrytis cinerea*/Bcin08g01420; *Puccinia graminis f. sp. tritici*/PGTG_03727; *Sporisorium reilianum*/sr11991; *Spizellomyces punctatus*/SPPG_02467; *Phycomyces blakesleeanus*/PHYBL_104565; *Phycomyces blakesleeanus*/PHYBL_139838; *Phycomyces blakesleeanus*/PHYBL_14376; *Phycomyces blakesleeanus*/PHYBL_153895; *Phycomyces blakesleeanus*/PHYBL_79980; *Mucor circinelloides f. lusitanicus*/QYA_112430; *Mucor circinelloides f. lusitanicus*/QYA_167321; *Mucor circinelloides f. lusitanicus*/QYA_177085; *Rhizophagus irregularis*/GLOIN_2v1532112; Ar*abidopsis thaliana*/AEE86382.1 **B. RACK1 proteins.** Organisms and protein names or accession numbers for the RACK1 orthologs are *Neurospora crassa* NcCPC-2/NCU05810;

Sordaria macrospora SmRACK1/SMAC07639; *Fusarium graminearum* FgRACK1/FGRAMPH101G06721; *Magnaporthe oryzae* MoRACK1/MGG04719; *Aspergillus nidulans* AnCpcB/AN4163; *Ustilago maydis* UmRACK1/UMAG10146*; Cryptococcus neoformans* CnGib2/CNAG05465; *Candida albicans* CaAsc1/C701250WA; *Schizosaccharomyces pombe* SpCpc2/SPAC6B12.15; *Saccharomyces cerevisiae* ScAsc1/YMR116C; *Batrachochytrium dendrobatidis*/BDEG_04723; *Botrytis cinerea*/Bcin14g03010*; Sporisorium reilianum*/sr10817; *Puccinia graminis f. sp. tritici*/PGTG_14970; *Phycomyces blakesleeanus*/PHYBL_133989; *Phycomyces blakesleeanus*/PHYBL_160317; *Mucor circinelloides f. lusitanicus*/QYA_155892; *Mucor circinelloides f. lusitanicus*/QYA_156660; *Rhizophagus irregularis*/GLOIN_2v1592318; *Arabidopsis thaliana*/AT OAP14939.1.

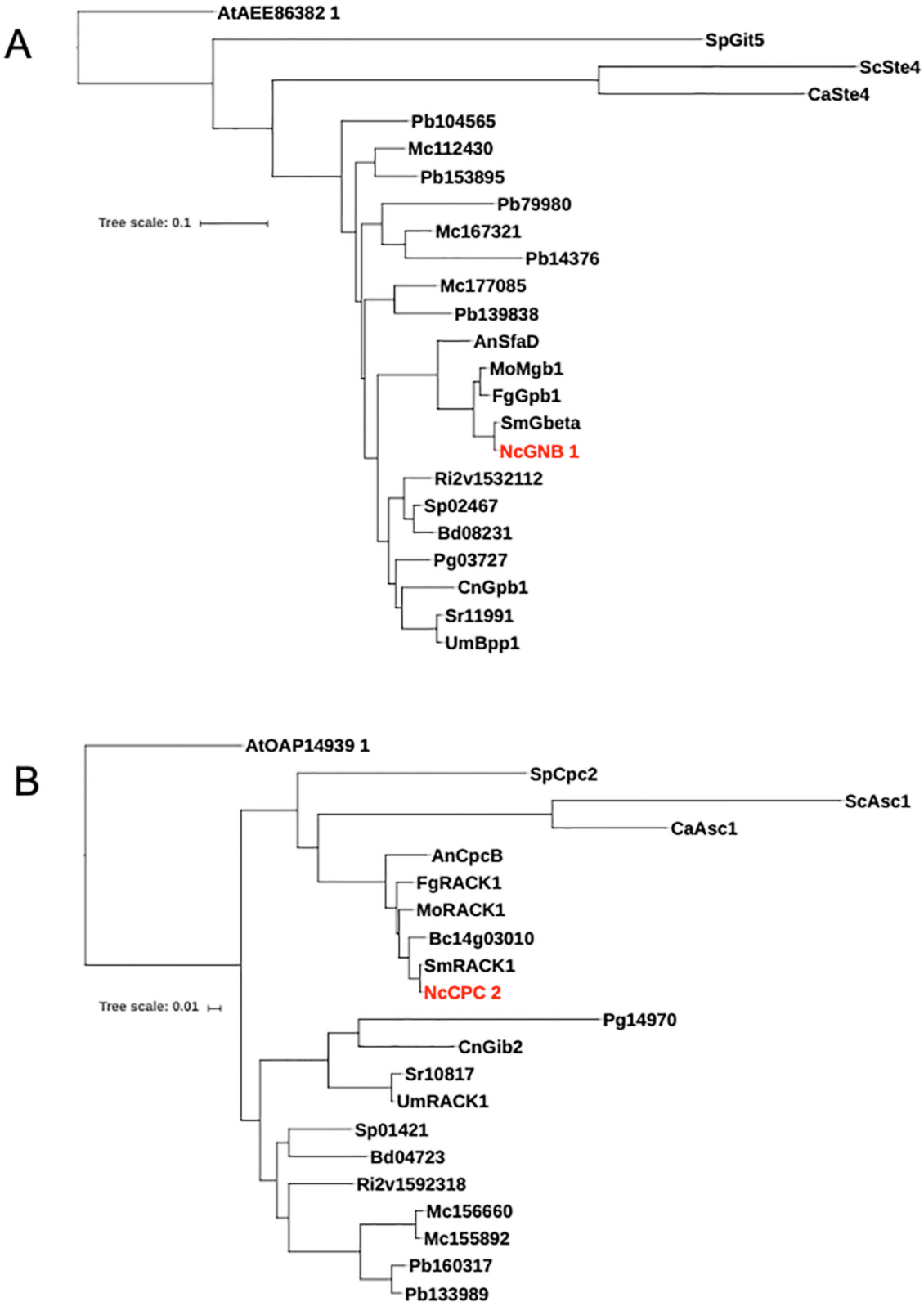
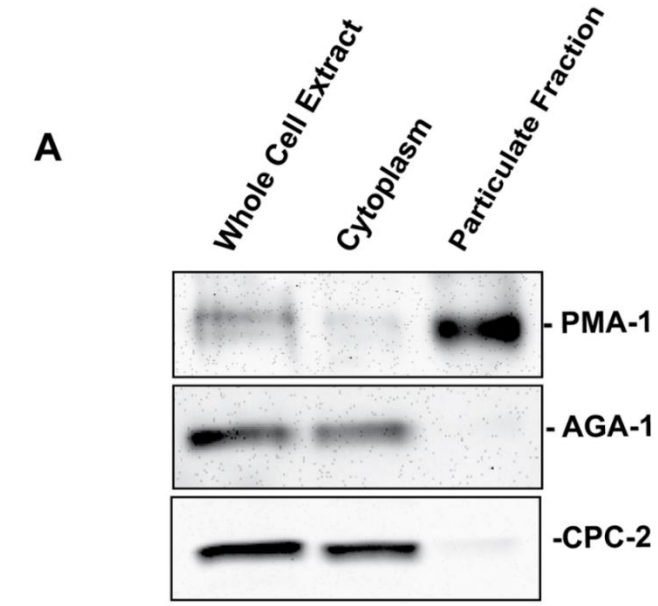

A
AtAEE86382 1
SpGit5
ScSte4
CaSte4
Pb104565
Mc112430
Pb153895
Pb79980
Tree scale: 0.1
Mc167321
Pb14376
Mc177085
Pb139838
AnSfaD
MoMgb1
FgGpb1
SmGbeta
NcGNB 1
Ri2v1532112
Sp02467
Bd08231
Pg03727
CnGpb1
Sr11991
UmBpp1
B
AtOAP14939 1
SpCpc2
ScAsc1
CaAsc1
AnCpcB
FgRACK1
MoRACK1
Bc14g03010
SmRACK1
NcCPC 2
Tree scale: 0.01
Pg14970
CnGib2
Sr10817
UmRACK1
Sp01421
Bd04723
Ri2v1592318
Mc156660
Mc155892
Pb160317
Pb133989

Figure 4.2: Subcellular localization of CPC-2. A. Fractionation of CPC-2 during differential centrifugation of cell extracts.

Cytosolic and particulate fractions were isolated from a cell extract of wild-type strain 74-OR23-1VA as described in the Materials and methods. Samples corresponding to the same volume of original cell extract were subjected to SDS-PAGE and western analysis using CPC-2, arginase (AGA; cytosol), and plasma membrane ATPase (PMA-1; plasma membrane) antibodies. The results shown are representative of four biological replicates. **B. Localization of GFP-tagged CPC-2 protein in vivo.** An aliquot containing 8 x 106 macroconidia from the CPC-2-GFP-9-10 strain was inoculated on VM agarose plates and incubated at 30°C for 0 h and 6 h. Images for the GFP channel were obtained via fluorescence microscopy and also stained with DAPI to visualize the nucleus (see Materials and methods for details). Images for GFP and DAPI were merged using ImageJ (National Institutes of Health, Bethesda, MD). Differential interference contrast (DIC) images were taken to show overall morphology of macroconidia and hyphae. Scale bar = 10 microns.


A
Whole Cell Extract
Cytoplasm
Particulate Fraction
- PMA-1
- AGA-1
-CPC-2

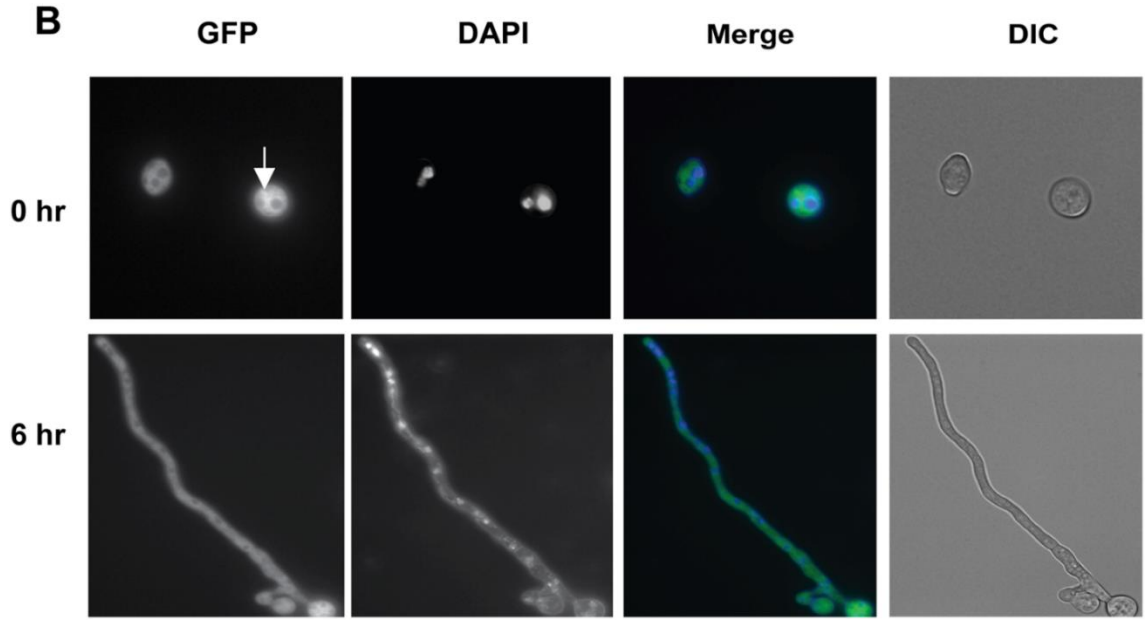

B
GFP
DAPI
Merge
DIC
0 hr
6 hr

Figure 4.3 Analysis of growth rate and asexual development in a complemented strain.

Δ*cpc-2* complemented strain CPC-2-GFP-13.2 was compared to wild type (WT matA) and Δ*cpc-2* strain Δcpc2#11 with respect to growth rate of basal hyphae (top; four replicates) and aerial hyphae height (bottom; 12 replicates) on VM medium supplemented with 10 μg/ml pantothenate. Error is indicated as the standard error of the mean. *** p value <0.001 relative to wild type.

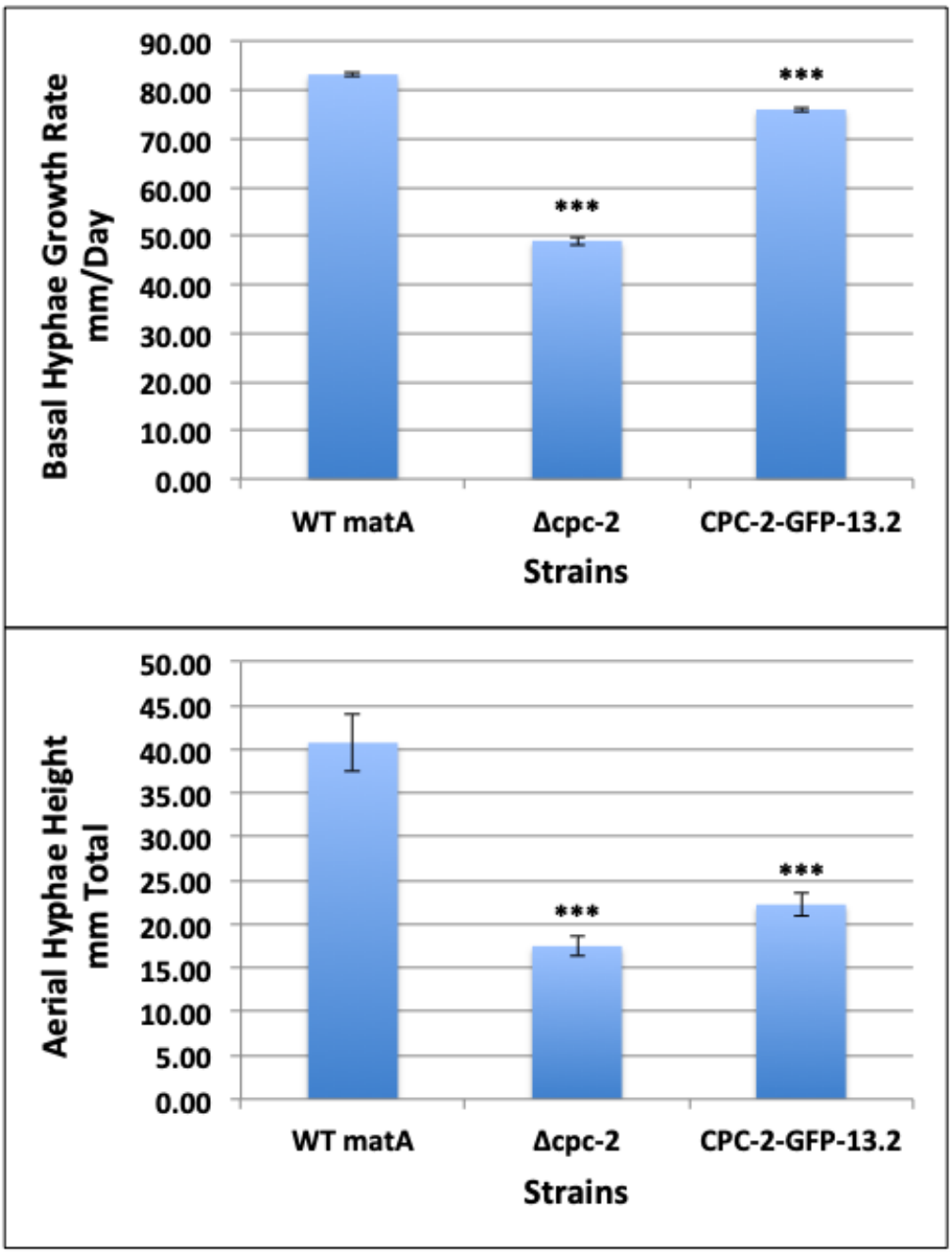

Basal Hyphae Growth Rate mm/Day
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
***
***
WT matA
Δcpc-2
CPC-2-GFP-13.2
Strains
Aerial Hyphae Height mm Total
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
***
***
WT matA
Δcpc-2
CPC-2-GFP-13.2
Strains

Figure 4.4: Levels of CPC-2 and other G protein subunits in different strain backgrounds.

For detection of GNA-1, GNA-2, GNA-3 and GNB-1, differential centrifugation was used to isolate the particulate fraction from whole cell protein extracts of the indicated strains. Samples were subjected to SDS-PAGE and western blots prepared. Blots were reacted with antiserum for GNA-1, GNA-2, GNA-3 or GNB-1. For detection of CPC-2, protein from whole cell extracts was used to prepare western blots. Blots were reacted with polyclonal antiserum raised against a MBP-CPC-2 fusion protein purified from *E. coli*. The results shown are representative of three biological replicates. The migration position of each protein is shown along the right side of the panel.

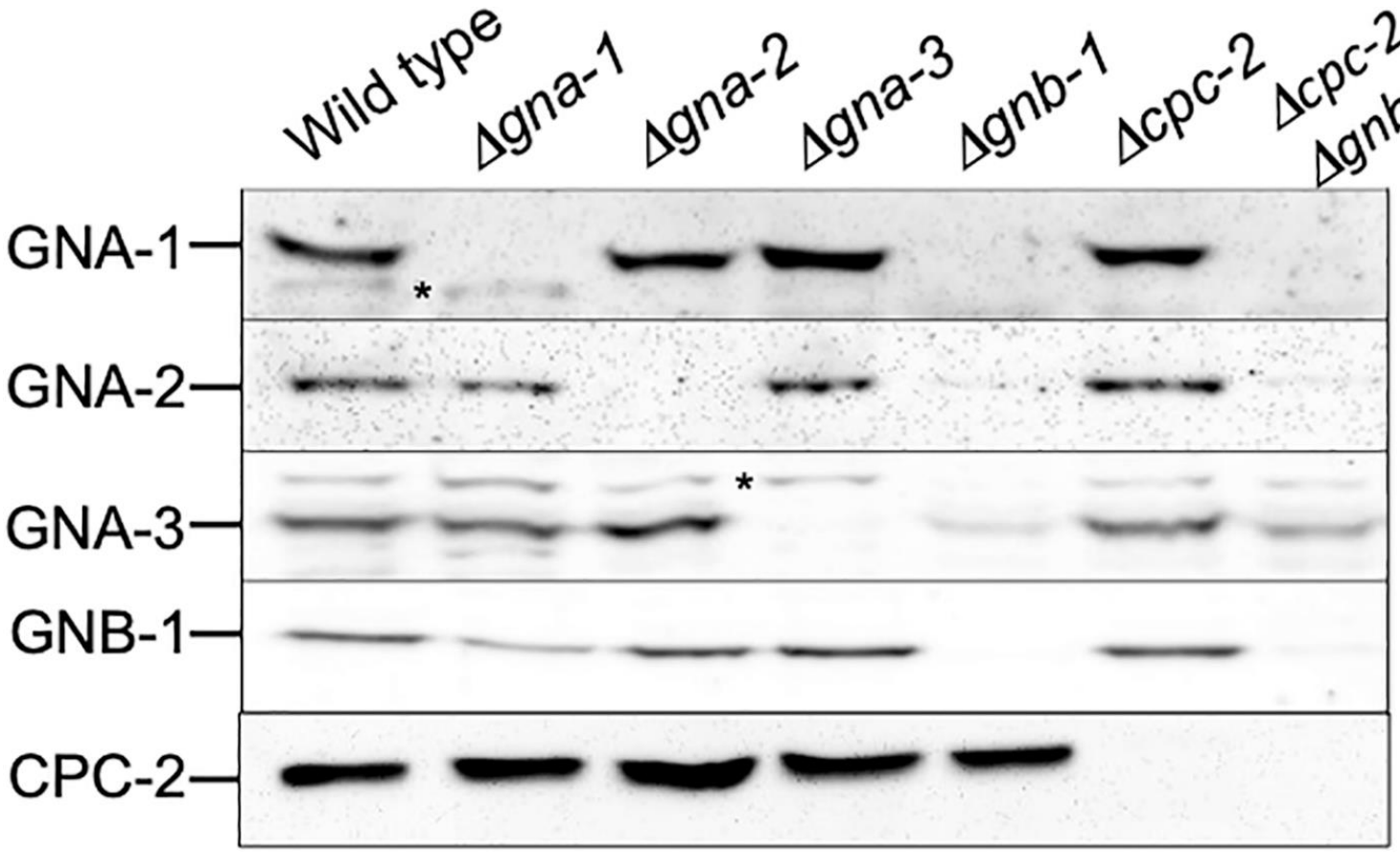

Wild type
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgnb-1
GNA-1
GNA-2
GNA-3
GNB-1
CPC-2
*
*

Figure 4.5: Growth rate of basal hyphae.

VM agar race tubes were inoculated with the indicated genotypes, incubated at 25°C and marked at various times as previously described [65]. Linear growth rates were determined with values (expressed as mm/day) taken from four biological replicates. Strains used were 74-OR23-1VA, ORS-SL6a, 3B10, Δgna2-2476, 3lc2, 42-8-3, Δcpc2#11, C2B1#2-1-1, C2G1#39, C2G2#37, C2G3#1–6, C2G1*#44, C2G2*#4 and C2G3*#1–8 (See Table 4.1 for genotypes). Error was calculated as the standard error of the mean. ANOVA was performed to identify strains that were significantly different from one another.

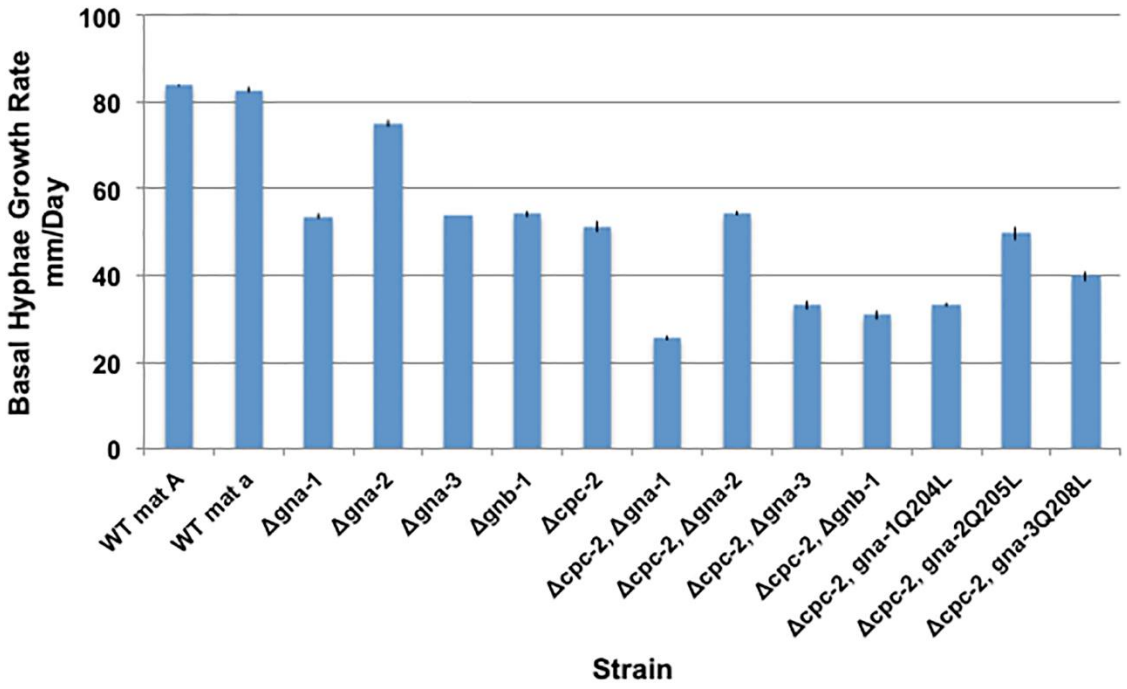

Basal Hyphae Growth Rate mm/Day
100
80
60
40
20
0
WT mat A
WT mat a
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgna-1
Δcpc-2, Δgna-2
Δcpc-2, Δgna-3
Δcpc-2, Δgnb-1
Δcpc-2, gna-1Q204L
Δcpc-2, gna-2Q205L
Δcpc-2, gna-3Q208L
Strain

Figure 4.6: Germination of macroconidia.

Macroconidia were harvested as described in [59]. An aliquot containing 8×10^6 macroconidia was spread on a VM agar plate (100mm plate containing 10 ml agar medium) and spore germination monitored microscopically at 30°C over the indicated times. DIC (differential interference contrast) micrograph images were obtained using an Olympus IX71 microscope with a QIClick digital CCD camera and analyzed using Metamorph software. Strains used were wild type, Δ*cpc-2* and Δ*gnb-1*.

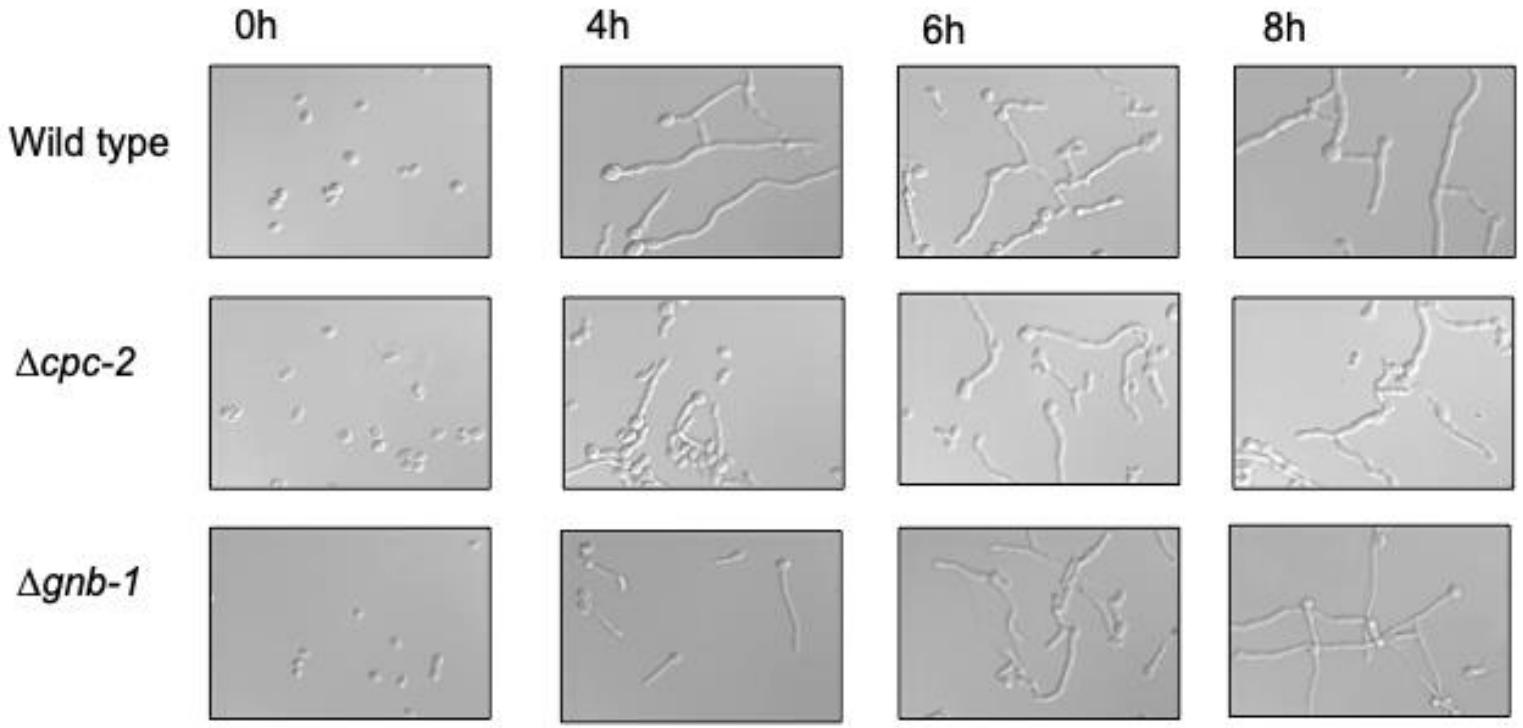
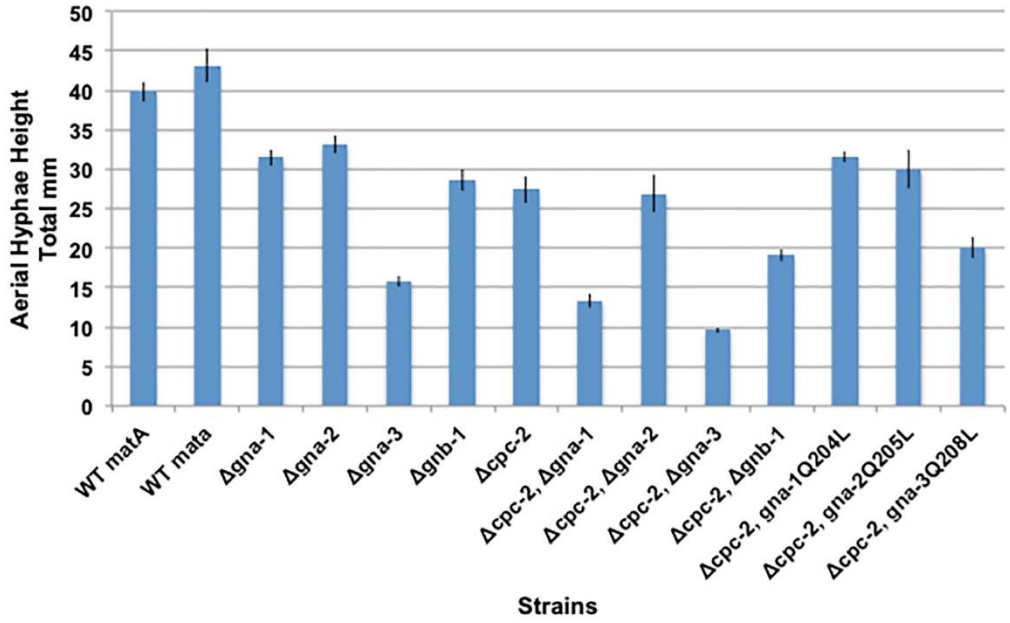

0h
4h
6h
8h
Wild type
Δcpc-2
Δgnb-1

Figure 4.7: Quantitative phenotypes during asexual development.

Strains, error calculations and ANOVA were as in Figure 4.5. **A. Aerial hyphae height**. Culture tubes containing liquid VM medium were inoculated with the indicated strains and incubated statically in the dark for three days at room temperature. The distance grown by the aerial hyphae above the medium interface was then measured. Values (mm) were taken from 12 biological replicates. **B. Macroconidia production**. Macroconidia from the indicated strains were propagated by growth in VM agar culture tubes in the dark at 30°C for four days followed by three days in light at room temperature. Macroconidia were harvested from the cultures and quantitated as described in the Materials and methods. Values represent eight biological replicates.

A


Aerial Hyphae Height
Total mm
50
45
40
35
30
25
20
15
10
5
0
WT matA
WT mata
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgna-1
Δcpc-2, Δgna-2
Δcpc-2, Δgna-3
Δcpc-2, Δgnb-1
Δcpc-2, gna-1Q204L
Δcpc-2, gna-2Q205L
Δcpc-2, gna-3Q208L
Strains

B

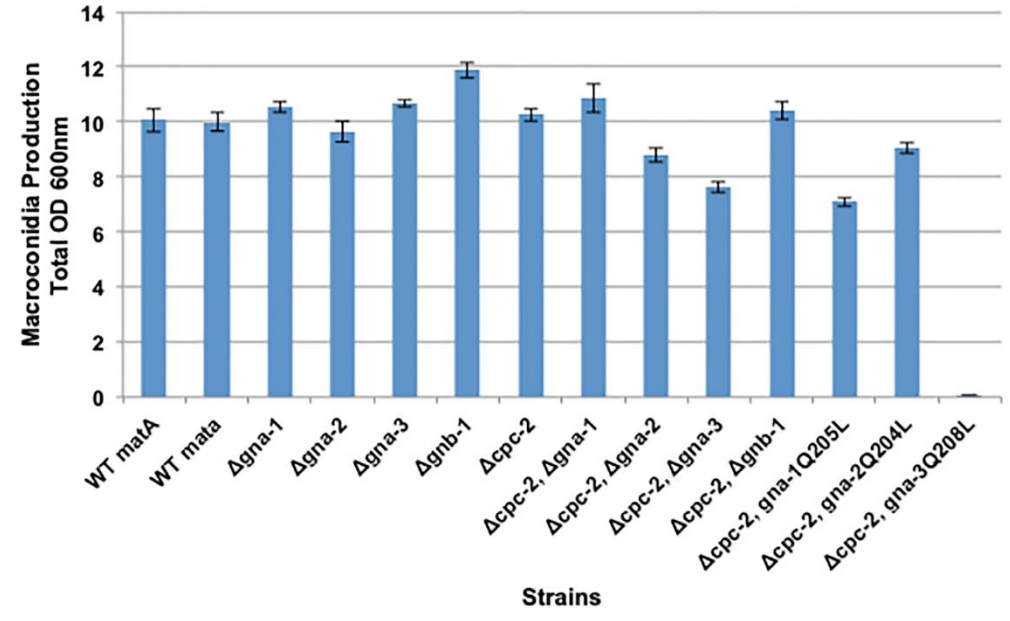

Macroconidia Production
Total OD 600nm
14
12
10
8
6
4
2
0
WT matA
WT mata
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgna-1
Δcpc-2, Δgna-2
Δcpc-2, Δgna-3
Δcpc-2, Δgnb-1
Δcpc-2, gna-1Q205L
Δcpc-2, gna-2Q204L
Δcpc-2, gna-3Q208L
Strains

Figure 4.8: Morphology in submerged culture.

Macroconidia isolated from the strains used in Figure 4.5 were inoculated at a concentration of $1x10^6$ macroconidia/ml and cultured in VM liquid medium for 16 h with shaking at 200 rpm in the dark at 30°C. In the case of strain C2G3*#1–8 (Δ*cpc-2 gna-3*^{Q208L}), a small volume of aerial hyphae was used to inoculate cultures, as this strain does not produce a significant amount of macroconidia. A sample of each culture was imaged at 40x magnification using DIC (see Materials and methods). Examples of conidiophores and/or free macroconidia are indicated by the white arrows.

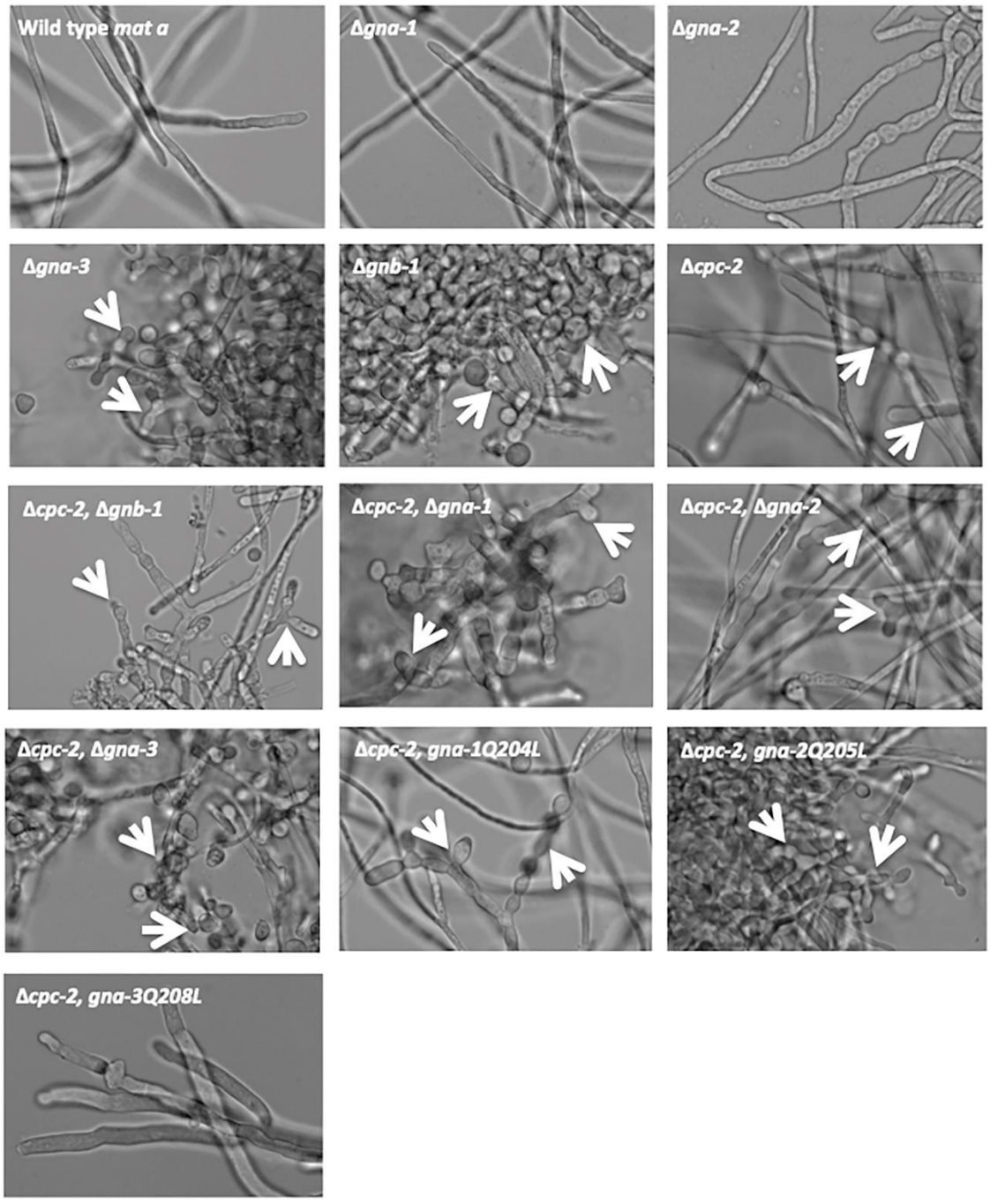

Wild type mat a
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgnb-1
Δcpc-2, Δgna-1
Δcpc-2, Δgna-2
Δcpc-2, Δgna-3
Δcpc-2, gna-1Q204L
Δcpc-2, gna-2Q205L
Δcpc-2, gna-3Q208L

Figure 4.9: Phenotypes during the sexual cycle.

Strains were ORS-SL6a, 3B10, Δgna2-2476, 3lc2, 42-8-3, Δcpc2#11, C2B1#2-1-1, C2G1#39, C2G2#37, C2G3#1–6, C2G1*#44, C2G2*#4 and C2G3*#1–8 (See Table 1 for genotypes). Macroconidia or hyphae from strains were inoculated onto SCM plates and incubated in constant light at room temperature for 7 days. At that time, half of each plate was inoculated with either macroconidia (males) of opposite mating type or water (control). Males were from wild-type strains 74-OR23-1VA (*mat A*) or ORS-SL6a (*mat a*). Incubation was continued under the same conditions for an additional 7 days. The fertilized side of each plate was then photographed using a Leica S8APO stereomicroscope with a DFC280 camera (Leica Microsystems, Buffalo Grove, IL USA). Examples of protoperithecia or submerged, aberrant perithecia are indicated by the black arrowheads, while mature perithecia are shown by the black arrows.

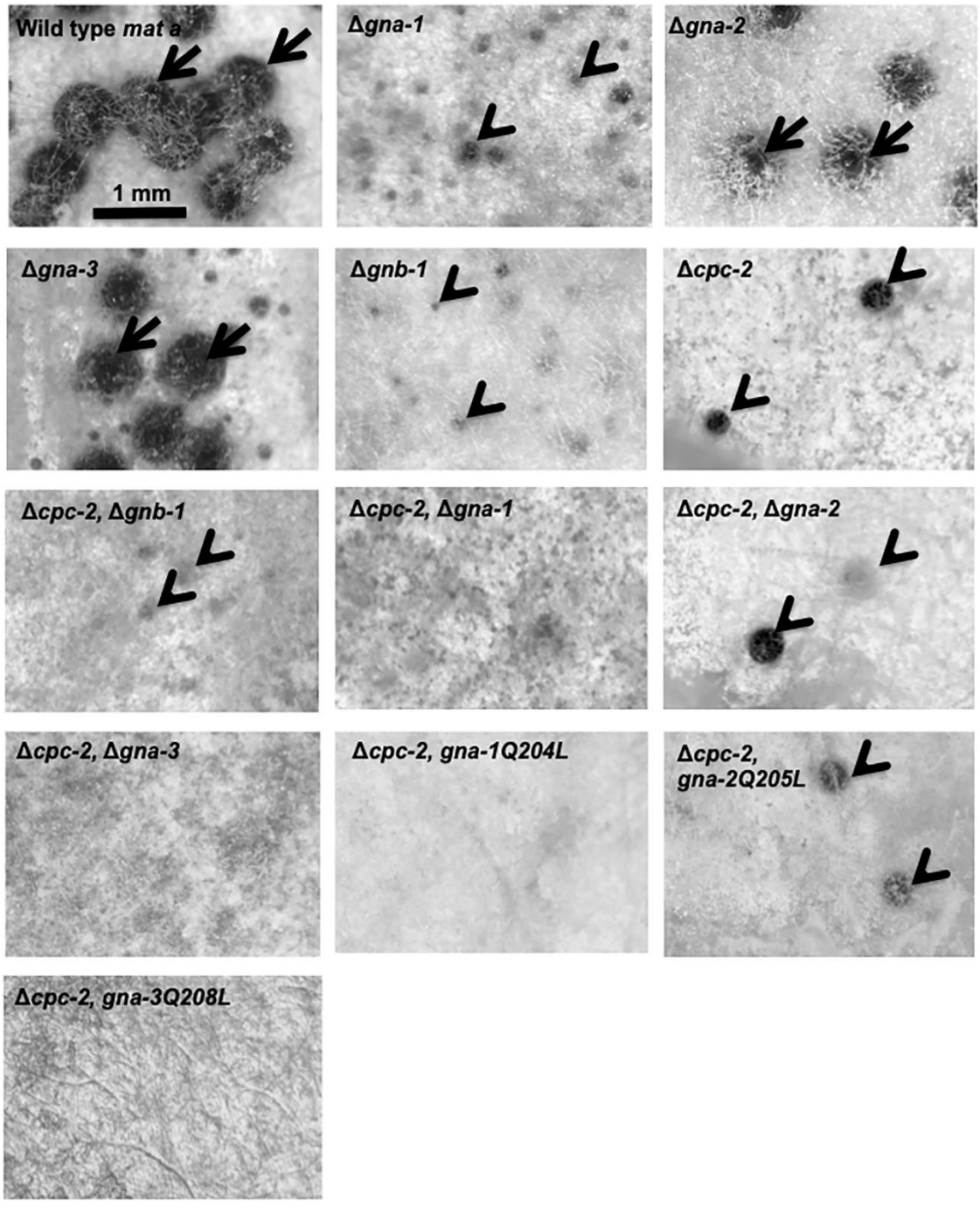

Wild type mat a
1 mm
Δgna-1
Δgna-2
Δgna-3
Δgnb-1
Δcpc-2
Δcpc-2, Δgnb-1
Δcpc-2, Δgna-1
Δcpc-2, Δgna-2
Δcpc-2, Δgna-3
Δcpc-2, gna-1Q204L
Δcpc-2,
gna-2Q205L
Δcpc-2, gna-3Q208L

Figure 4.10: Models for interactions between CPC-2 and G protein subunits in N. crassa.

The three Gα proteins are colored red, while the two predicted Gβ proteins are blue in the various panels. **A. Basal hyphae growth and aerial hyphae height**. GNA-2 operates upstream of CPC-2 to positively modulate basal hyphae growth rate and aerial hyphae height. GNA-1, GNA-3 and GNB-1 act independently of CPC-2 to regulate this trait. **B. Aerial macroconidia abundance**. GNB-1 is a negative regulator of macroconidia abundance in agar cultures and loss of *cpc-2* suppresses this effect. The asterisk indicates that GNA-3 only influences macroconidiation when mutationally activated. **C. Submerged culture conidiation**. CPC-2, GNA-1, GNA-3 and GNB-1 are negative regulators of conidiation in liquid submerged cultures. The asterisk on the GNA-1 arrow denotes the cell density dependence of the GNA-1 effect. The asterisk on the GNA-3 arrow indicates that GNA-3 may function with CPC-2 or independently, as Δ*gna-3* Δ*cpc-2* mutants have a more severe phenotype than the single mutants, but mutational activation of *gna-3* corrects the phenotype of the Δ*cpc-2* mutant. **D. Perithecial development**. GNB-1 functions downstream of CPC-2 to regulate perithecial development. The action of GNA-1 to control perithecial development is independent of CPC-2.

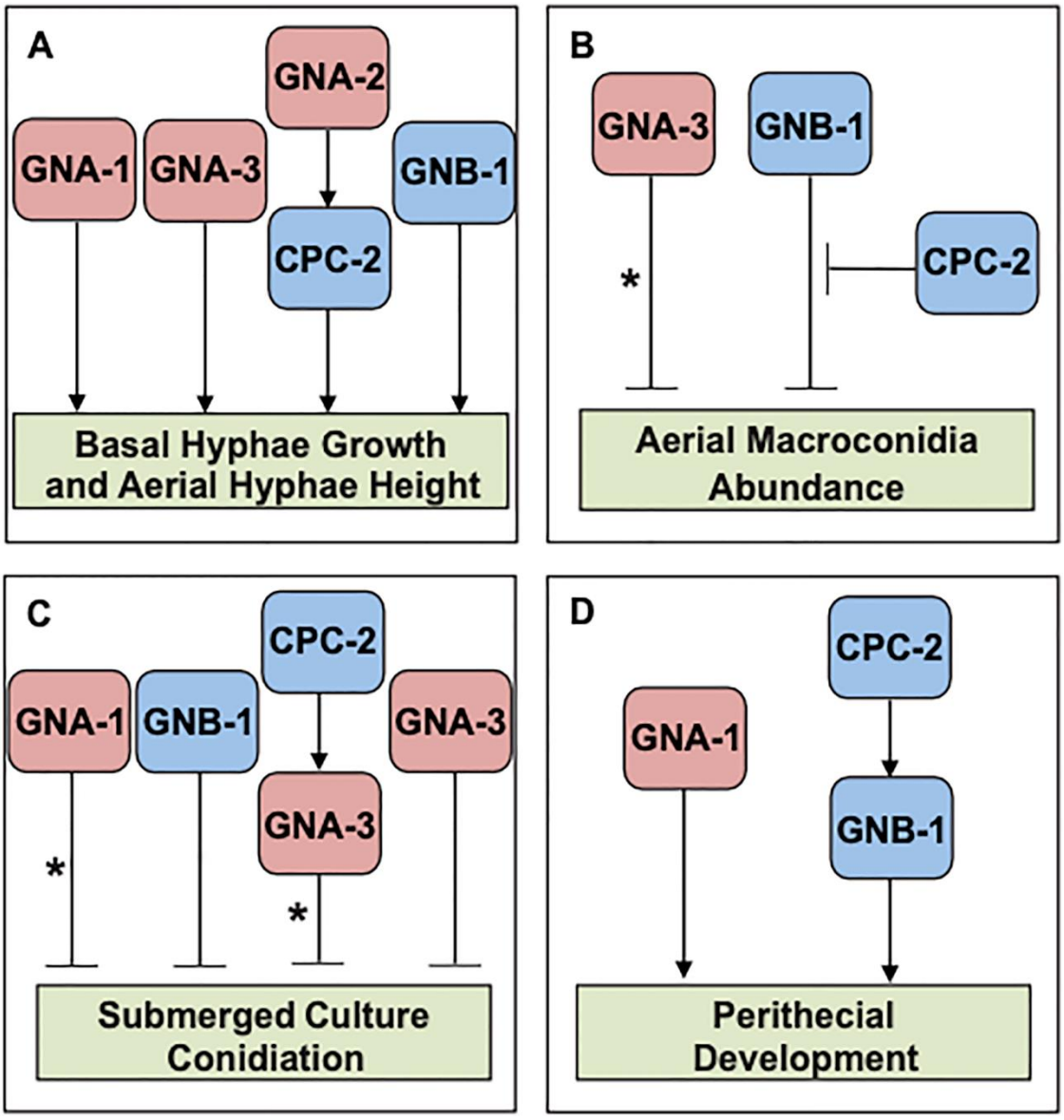

A
GNA-2
GNA-1
GNA-3
GNB-1
CPC-2
Basal Hyphae Growth
and Aerial Hyphae Height
B
GNA-3
GNB-1
CPC-2
*
Aerial Macroconidia
Abundance
C
CPC-2
GNA-1
GNB-1
GNA-3
GNA-3
*
*
Submerged Culture
Conidiation
D
CPC-2
GNA-1
GNB-1
Perithecial
Development

Table 4.1: *N. crassa* strains used in this study

Strain name	Relevant genotype	Comments	Source or Reference
74-OR23-1VA	Wild type, *mat* A		FGSC 2489
ORS-SL6a	Wild type, *mat* a		FGSC4200
74A-OR23-1A	Wild type, *mat* A		FGSC987
a^{m1}	*cyh-1, ad3B,* a^{m1}		FGSC4564
Y234M723	*his-3, mat* A		FGSC6103
his-3a#14	*his-3, mat* a		Ref. (61)
3B10	Δ*gna-1*::hph^+, *mat* a		Ref. (81)
Δgna-2-2477	Δ*gna-2*::hph^+, *mat* a		FGSC12377
Δgna-2-2476	Δ*gna-2*::hph^+, *mat* A		FGSC12376
31c2	Δ*gna-3*::hph^+, *mat* A		Ref. (14)
42-8-3	Δ*gnb-1*::hph^+, *mat* A		Ref. (64)
Δcpc2Het	Δ*cpc-2*::hph^+, Δ*mus-51::bar, mat* a (heterokaryon)		FGSC13695
Δcpc2#1	Δ*cpc-2*::hph^+, Δ*mus-51::bar, mat* a	Progeny from cross of Δcpc2Het to 74-OR23-1A	This Study
Δcpc2#6	Δ*cpc-2*::hph^+, Δ*mus-51::bar, mat* a	Progeny from cross of Δcpc2Het to 74-OR23-1A	This Study
Δcpc2#11	Δ*cpc-2*::hph^+, *mat* A	Progeny from cross of Δcpc2Het to 74-OR23-1A	This Study
Δcpc2his3A	Δ*cpc-2*::hph^+, *his-3, mat* A	Progeny from cross of Δcpc2Het to Y234M723	This Study
cpc2+a^{m1}	Δ*cpc-2*::hph^+, *his-3, mat* A + a^{m1}, *cyh-1*, *ad3B*, *mat* A (heterokaryon)	Heterokaryon of Δcpc2his3A and a^{m1}	This Study
C2G1*#44	Δ*cpc-2*::hph^+, *his-3*$^+$::*gna-1*^{Q204L}, *mat* A	Δcpc2his3A purified transformant	This Study
C2G2*#4	Δ*cpc-2*::hph^+, *his-3*$^+$::*gna-2*^{Q205L}, *mat* A	Δcpc2his3A purified transformant	This Study
C2G3*#1-8	Δ*cpc-2*::hph^+, *his-3*$^+$::*gna-3*^{Q208L}, *mat* A	Δcpc2his3A purified transformant	This Study
C2G1#39	Δ*cpc-2*::hph^+, Δ*gna-1::hph*$^+$, *mat a*	Progeny from cross of cpc2+a^{m1} to 3b10	This Study
C2G2#37	Δ*cpc-2*::hph^+, Δ*gna-2::hph*$^+$, *mat a*	Progeny from cross of cpc2+a^{m1} to Δgna-2-2477	This Study
C2G3#1-6	Δ*cpc-2*::hph^+, Δ*gna-3::hph*$^+$, *mat A*	Progeny from cross of 31c to Δcpc2#6	This Study
C2B1#2-1-1	Δ*gnb-1*::nat^+, Δ*cpc-2::hph*$^+$, *mus-51::bar*$^+$, *mat* a	Δcpc2#6 purified transformant	This Study
51-4	Δ*rid*::nat^+, Δ*mus-52::nat*$^+$, *mat* a		This Study
CPC-2-GFP-9	Δ*pan-2*::p*ccg-1*::*cpc-2*-V5-GFP::*bar*$^+$, *mat* a	51-4-1 transformant	This Study
CPC-2-GFP-9-10	Δ*pan-2*::p*ccg-1*::*cpc-2*-V5-GFP::*bar*$^+$, *mat* a	Progeny of CPC-2-GFP-9 crossed to 74-OR23-1VA	This Study
pccg-1_GFP	Δ*pan-2*::p*ccg-1*::V5-GFP::*bar*$^+$, *mat* a	Empty vector control for CPC-2-GFP-9-10	Ref (58)
CPC-2-GFP-13	Δ*pan-2*::p*ccg-1*::*cpc-2*-V5-GFP::*bar*$^+$, *mat* a	51-4-1 transformant	This Study
CPC-2-GFP-13.2	Δ*cpc-2*::hph^+, Δ*pan-2*::p*ccg-1*::*cpc-2*-V5-GFP::*bar*$^+$, *mat* a	Progeny of CPC-2-GFP-13 crossed to 74-OR23-1VA	This Study

Table 4.2: Oligonucleotides used in this study

Primer #	Primer name	Sequence 5' to 3'
1	CPC2FORDIAG	AGCAGGGCCGGGTGGAGATT
2	CPC2REVDIAG	CGAAGGTCCCACCCTAACAGCC
3	GNA1FORDIAG-1	CTTGGAGAGTGCGCGGTGGG
4	GNA1FORDIAG-2	GTCGGGTGGGCGATGGATCAA
5	GNA1REVDIAG	GTGTCGGGTGCTTTCTGCCA
6	GNA2FORDIAG-1	GCCCTGGGCACTACCGAAACG
7	GNA2FORDIAG-2	GGGCCAGAAATGGAACCTACC
8	GNA2REVDIAG	TTCCGGCCGAGTGAAACGCT
9	GNA3FORDIAG	GCGGCCTGCCCTAGCAATTCA
10	GNA3REVDIAG	GGAGTAGCGAGGTGTATGAGTGGT
11	GNB1FORDIAG	GTGCCTTCGGCCAGGCTTGT
12	GNB1REVDIAG	TTGGTTACGTATGCTGAGCAAGGG
13	HPHREV	TGCTCCTTCAATATCATCTTCTGTC
14	HPHFOR	TGTGTAGAAGTACTCGCCGATAGTG
15	GNB1NAT5'FLANK-FWD	GTAACGCCAGGGTTTTCCCAGTCACGACGGTTCCATCGGGGGTGCGGTGC
16	GNB1NAT5'FLANK-REV	CTACATGAGCATGCCCTGCCCCTGATCGCTTCTGCGAGTGGGCGGGCGGC
17	GNB1NAT3'FLANK-FWD	CTCCTTCAATATCATCTTCTGTCGAGCTCTCGCTTGTATGCATCAGGTCT
18	GNB1NAT3'FLANK-REV	GCGGATAACAATTTCACACAGGAAACAGCAAGGGCACTGGCCCACTGGAC
19	PTRPC5'GNB1	AGACCTGATGCATACAAGCGAGAGCTCGACAGAAGATGATATTGAAGGAG
20	NAT3'GNB1	GCCGCCCGCCCACTCGCAGAAGCGATCAGGGGCAGGGCATGCTCATGTAG
21	GNB1NAT-REV-DIAG	TCACGTAGACCTGATGCATA
22	NAT5' REV DIAG	CAAAAAGTGCTCCTTCAATA
23	pCCG1 FWD	CCATCATCAGCCAACAAAGC
24	GNA1ORF REV	GGGAATTCTCAAATCAAACCGCAG
25	GNA2ORF REV	GGGAATTCCTACAGGATAAGTTGT
26	GNA3ORF REV	GGGAATTCTCATAGAATACCGGAG
27	F-C2	CCACTTTCACAACCCCTCACATCAACCAAAATGGCTGAGCAACTCATCCTCAAG
28	R-C2-V5G	GTTAGGGATAGGCTTTCCGCCGCCTCCGCCAGCGCGGGACATGACACCCCAGGC
29	P-CPC-2-REV	TATGCTAGTTATGCGGCCGCTGCAGTTAAGCGCGGGACATGACACCCCAG

Chapter 5

The role of G protein signaling on metabolism in *Neurospora crassa*

Abstract

G protein signaling is important for sensing the external environment and for modulating growth. Heterotrimeric G proteins consist of Gα, Gβ, and Gγ subunits, with GDP/GTP exchange by the Gα subunit regulated by guanine nucleotide exchange factors (GEFs). GEFs include membrane-bound G protein coupled receptors and cytoplasmic proteins such as RIC8. In this study, we used a combination of RNA-seq and Liquid Chromatography-Mass Spectrometry (LC-MS) to profile the transcriptomes and metabolomes of wild type *N. crassa*, two mutants lacking Gα subunits (Δ*gna-1* and Δ*gna-3*) and the Δ*ric8* mutant. We analyzed the transcriptomes to determine the role of G proteins in transcriptional control of gene expression and compared the transcriptome data to metabolomic data to determine whether there was evidence for transcriptional and/or post-transcriptional control of metabolism. The results showed that there are large transcriptional differences between the mutants and wild type, with Δ*gna-1* mutants possessing more than 150 differentially regulated genes, and Δ*gna-3* and Δ*ric8* strains each having more than 1000 differentially regulated genes. Many of these differentially regulated genes encode metabolic enzymes, with mis-regulated genes in Δ*ric8* and Δ*gna-1* mutants being enriched for oxidoreductases. LC-MS analysis of these mutants demonstrated that there were changes in levels of primary metabolites, with Δ*gna-1*, Δ*gna-3,* and Δ*ric8* strains having 21, 27 and 68 mis-regulated metabolites out of the 120 detected. Comparison of the RNAseq and metabolomics data revealed evidence for both transcriptional and post-transcriptional control of certain metabolic proteins in the various genotypes.

Introduction

All organisms depend on cellular signaling to sense their external environment and modulate growth (1). In eukaryotes, specialized cell surface proteins called G protein coupled receptors (GPCRs) recognize external stimuli (2). GPCRs transduce the external signal (ligand) to intracellular signaling pathways via heterotrimeric G-proteins (2). The heterotrimeric G-protein is made up of three subunits: alpha (α), beta (β), and gamma (γ). The Gα subunit can be bound to GDP or GTP. When bound to GDP, the Gα is inactive and is in a complex with the Gβγ dimer and a GPCR (2,3). When bound to GTP, the Gα disassociates from the Gβγ dimer and both can regulate downstream signaling (1, 3-4).

Resistance to Cholinesterase 8 (RIC8) has been implicated as an important positive regulator of G-proteins in several species, including mammals, *Caenorhabditis elegans, Drosophila melanogaster* and *N. crassa* (5,6). RIC8 is not found in plants (5). First discovered in *C. elegans*, RIC8 was identified in screens for mutants that were resistant to inhibitors of cholinesterase and defective in synaptic vesicle priming (7). RIC8 is predicted to be comprised of 10 armadillo fold repeats, which is of interest, as armadillo repeat proteins are known to interact with multiple proteins and be involved in diverse functions (8).

In addition to its role in priming of synaptic vesicles, RIC8 is required for asymmetric cell division in *C. elegans* zygotes (9,10,11). In *D. melanogaster*, RIC8 has been found to be critical for polarity during asymmetric cell division during embryogenesis and for the stability of Gα and Gβ proteins (12). In mice, RIC8 activates Gα subunits and is essential for embryogenesis, making study of RIC8 functions difficult

in mammals (13). In another mice model study, it was shown that deletion of Ric-8A in melanocytes suppressed tumorigenesis by the uveal melanoma oncogene human $GNAQ^{Q209L}$ by depleting levels of the oncogene below the cancer-driving level (14).

RIC8 functions have also been studied in fungi. Interestingly there is no RIC8 homolog in *Saccharomyces cerevisiae* (6). In *N. crassa*, deletion of *ric8* leads to severe developmental defects, with severely reduced growth rate, limited development of sexual structures and defects in asexual structure development (6). *In vitro* guanine nucleotide exchange factor (GEF) activity assays showed that RIC8 acts as a GEF towards GNA-1 and GNA-3 in *N. crassa* (6). Loss of *ric8* leads to reduced levels of all three Gα proteins, as well as the Gβ and Gγ subunits (6). In *Aspergillus nidulans* and *Aspergillus fumigatus,* the *ric8* homolog is important for growth and asexual development (15). It has also been shown that RIC8 is required for pathogenicity in the rice blast fungus *Magnaporthe oryzae* and in the wheat head blight fungus *Fusarium graminearum* (16,17).

$G\alpha_{12}$/ $G\alpha_{13}$ proteins have been shown to have roles in regulating metabolism and energy homeostasis in mice (18). The $G\alpha_{13}$ subunit affects glucose metabolism, with loss of $G\alpha_{13}$ via a gene deletion mutant leading to systemic glucose intolerance (insulin resistance). The loss of $G\alpha_{13}$ leads to differential secretion of 67 proteins from mice liver cells, including inter-α-trypsin heavy chain 1 (ITIH1), which at increased levels is correlated with glucose intolerance and diabetes (19). $G\alpha_{12}$/ $G\alpha_{13}$ proteins have also been shown to play a role in the regulation of mitochondrial morphology, motility and respiration (20,21). Absence of the $G\alpha_{12}$ subunit increases mitochondrial motility by either actin polymerization via the Rho pathway or Rho-independent activation of JNK,

which phosphorylates kinesin-1. Morphology is affected when the $G\alpha_{12}$ is uncoupled from the RhoGEF, which causes the mitochondrial tubular network to completely fragment (20). The Sirtuin 1 (SIRT1)/ Peroxisome proliferator-activated receptor (PPARα) network regulates fatty acid beta oxidation and with the loss of specific Gα subunits, the levels of SIRT1and the transcription factor PPARα target gene transcripts are lower, leading to reduced oxygen consumption rates in mitochondria (21). Mis-regulation of mitochondrial respiration leads to liver steatosis (retention of fats) and liver metabolic dysfunction (21).

Possible effects on metabolism due to loss of a G protein subunit have also been investigated in *N. crassa*. For example, the metabolites of wild type and a Δ*gna-3* mutant have been profiled using ^{1}H Nuclear Magnetic Resonance (NMR) (22). This study showed that wild type and Δ*gna-3* have similar global metabolomes under high carbon conditions, but that the Δ*gna-3* mutant appears to possess a carbon-sensing defect under low carbon conditions (22). However, to-date, metabolomics studies have not been performed with any other mutants implicated in G protein signaling in *N. crassa*, including Δ*ric8*.

In submerged liquid cultures, wild type *N. crassa* maintains hyphal development, unless subjected to stresses such as heat or carbon/nitrogen starvation (23-27). Δ*gna-1* mutants produce hyphae in submerged liquid cultures unless inoculated at high cell densities (28). In the Δ*gna-3* mutant, inappropriate conidiation is observed in submerged cultures even when inoculated at lower cell densities (30). Further, when *gna-1* is deleted in addition to *gna-3* there is an increase in production of conidiophores in submerged

cultures (29,30). The Δ*ric8* strain produces conidiophores and conidia in submerged cultures at levels similar to a Δ*gna-1*, Δ*gna-3* double mutant (6).

In this study, RNA sequencing was used to determine gene expression profiles in wild type and three G protein signaling mutants cultured with glucose as the carbon source. We analyzed two strains that produce only hyphae in submerged liquid cultures (wild type and the Δ*gna-1* mutant) and two strains that produce hyphae and conidia (Δ*gna-3* and Δ*ric-8*), with the Δ*ric8* strain being more severe than the Δ*gna-3* mutant. Substantial mis-regulation of metabolic genes was observed in the three mutants in comparison to wild type. To understand possible downstream effects on metabolism, Liquid Chromatography-Mass Spectrometry was used to profile and compare the metabolomes of the same four strains. Levels of each detected metabolite were then compared across all samples. The results revealed considerable differences as compared to wild type in the levels of specific metabolites in the Δ*gna-1*, Δ*gna-3* and Δ*ric-8* mutants. In additional, the results provide evidence for both transcriptional and post-transcriptional regulation of metabolic protein levels by G protein signaling in *N. crassa*.

Methods and Materials

Strains and media. The *N. crassa* strains used in this study were wild-type OR23-1VA (74A, *mat* A, Fungal Genetics Stock Center #2489; Fungal Genetics Stock Center, Kansas City, MO), Δ*gna-1* strain 3B10 (28), Δ*gna-3* strain 31c2 (29), and Δ*ric8* strain ric8 1-5a (6). Conidia were propagated in Vogel's Minimal Medium (VM; (31)) agar flasks as described previously (32).

Glass Erlenmeyer flasks (125-ml total volume) used for growth of liquid cultures were treated with dichlorodimethylsilane (5% [vol/vol] in chloroform; Alfa Aesar, Ward Hill, MA) in order to prevent hyphae from adhering to the inside walls of the flasks. Each strain was grown in 50 mL of liquid VM, substituting sucrose with 100 mM Glucose as the carbon source (VM-Glucose). Liquid cultures were brought to a concentration of $1x10^6$ conidia/ml and incubated shaking at 200 RPM at 30°C for 16 hrs. Cultures were collected by vacuum filtration using filter paper (Whatman #405), transferred to 1.5 ml centrifuge tubes and stored at -80°C until extraction. Tissue was collected from five biological replicates.

Tissue collection, RNA extraction and RNAseq. Conidia collected from 5-day-old cultures were used to inoculate 50 ml of VM-Glucose at a final concentration of $1x10^6$ conidia/ml. Cultures were incubated with shaking at 200 RPM for 16 hrs at 30°C. Tissue was collected via vacuum filtration using filter paper (Whatman #405) and then ground into a fine powder with a mortar and pestle (CoorsTek, Golden, CO).

RNA was extracted from ground tissue using TRIzol (Invitrogen, Waltham, MA) and chloroform (Fisher, Hampton, NH). A sample containing 0.5 ml of ground tissue was added to 1 ml TRIzol in a 2-ml snap cap centrifuge tube. The mixture was incubated for 5 minutes at room temperature. A volume containing 200 µl of chloroform was added to the mixture, the tube was briefly vortexed and left to incubate for 3 minutes at room temperature (RT). The mixture was then centrifuged at 12,000xg for 15 minutes at 4°C for phase separation. Subsequently, 500 µl of the aqueous phase from each sample was

combined with 500 µl isopropanol and incubated at RT for 10 minutes. Samples were then centrifuged at 12,000xg for 10 minutes at 4°C. The supernatant was removed and the RNA pellets were washed with 75% ethanol and centrifuged at 7,500xg for 5 minutes at 4°C. Ethanol was aspirated using a P200 pipette tip and the RNA pellets were allowed to dry for 10 minutes at RT with the caps open. The RNA pellets were then resuspended in 50-100 µl of RNase-free water. Samples were then quantified using nanodrop 2000 (Thermo fisher, Waltham, MA).

Samples containing 20 µg of RNA were treated with 2 µl of DNaseI (New England Biolabs, Ipswich, MA) in a 30 µl reaction. Samples were incubated at 37°C for 30 minutes and then cleaned up using either Zymo RNA clean & concentrator kit (Zymo, Irvine, CA) or NEB Monarch RNA cleanup kit (New England Biolabs, Ipswich, MA). RNA was tested for quality using agarose gel electrophoresis and by running a sample on the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Three biological replicates were obtained for every strain.

RNA samples were sent to the University of California, Davis DNA Technologies and Expression Analysis Core Laboratory for Illumina library preparation and sequencing. Paired-end sequencing (2x150 bp) was performed using a NovaSeq 6000 with an S4 flow cell (Illumina, San Diego, CA). Raw reads were mapped against a generated index file via Kallisto version 0.46.1 (33) to generate raw counts and transcripts per million. Differential expression analysis was performed on raw counts using DEseq2 version 1.26.0 (34). Genes with p-values greater than 0.05 were not

considered differentially expressed. Differential genes were then filtered to exclude genes with less than 10 transcripts per million.

Metabolite extraction. Tissue was lyophilized for 2 days at -80°C (Labconco, Kansas City, MO). Dried samples were transferred to 1.5 mL screw cap tubes. Four to five ceramic beads were added to the dried tissue and the sample was bead-beat in a bead mill homogenizer (Thermo fisher, Waltham, MA) at -80°C for one minute. Approximately 8.5 to 9.8 mg of each sample was weighed and 100 µL of extraction buffer (methanol: water) was added per 1 mg of tissue. The samples were then briefly vortexed and placed in a 4°C ice-water bath and sonicated for 30 minutes. The samples were vortexed at 4°C for 30 minutes and then centrifuged at 16,000xg for 15 minutes. The supernatant was transferred to a new tube and stored at -80°C.

LC-MS. Targeted analysis of polar, primary metabolites was performed at the UC Riverside Metabolomics Core Facility as previously described (35). Briefly, analysis was performed on a TQ-XS triple quadrupole mass spectrometer (Waters) coupled to an I-class UPLC system (Waters). Separations were carried out on a ZIC-pHILIC column (2.1 x 150 mm, 5 µM) (EMD Millipore). The mobile phases were (A) water with 15 mM ammonium bicarbonate adjusted to pH 9.6 with ammonium hydroxide and (B) acetonitrile. The flow rate was 200 µL/min and the column was held at 50° C. The injection volume was 1 µL. The gradient was as follows: 0 min, 90% B; 1.5 min, 90% B; 16 min, 20% B; 18 min, 20% B; 20 min, 90% B; 28 min, 90% B.

The mass spectrometer was operated in selected reaction monitoring mode. Source and desolvation temperatures were 150° C and 500° C, respectively. Desolvation gas was set to 1,000 L/hr and cone gas to 150 L/hr. Collision gas was set to 0.15 mL/min. All gases were nitrogen, except the collision gas, which was argon. Capillary voltage was 1 kV in positive ion mode and 2 kV in negative ion mode. A quality control sample, generated by pooling equal aliquots of each sample, was analyzed every 3-5 injections to monitor system stability and performance. Samples were analyzed in random order.

Identification of metabolites. Data were processed with Skyline software (36). There were ~210 different potential molecules detected. Individual molecular peaks were manually checked for peak intensity, peak height above background, and retention time accuracy, and consistent peak boundaries were assigned for each molecule between samples. Peaks were then integrated using Skyline to analyze peak ratios. Experimental data was compared to preexisting standard LC-MS readings of metabolites to identify peaks to their corresponding small molecule.

Metabolite levels were compared between samples using relative abundance values. In R Studio, unpaired t-tests were used to calculate statistically significantly differences in abundance levels of metabolites between strains and to provide standard error for the five replicates for each strain. Significant differences were determined and labeled for different degrees of significance according to p-values of $p<0.05$, $p<0.001$, and $p<0.0001$. All mutants were compared to wild type, and relative abundance

percentages were calculated by dividing the mutant value by the wild type value for each metabolite. Principal component analysis was performed using R (37), using integrated peaks from the spectral analysis with mean centering and scaling of data.

Results

Differentially expressed genes in the three mutants. To identify mis-regulated genes in the gene deletion mutants, the transcriptomes of Δ*gna-1*, Δ*gna-3* and Δ*ric8* were determined using RNA-seq. Variable transcriptional differences were observed, with Δ*gna-1* having a relatively modest 159 differentially expressed genes, while Δ*gna-3* and Δ*ric8* mutants exhibited larger differences relative to wild type, with 1138 and 1395 differentially expressed genes, respectively (Figure 5.1).

To investigate the extent to which RIC8 might control transcription through GNA-1 and/or GNA-3, the unique and shared differentially expressed genes in the different genotypes were determined. Similar to the trend noted in the total number of differentially expressed genes in the three genotypes, it was observed that 601 genes were uniquely mis-regulated in the Δ*ric8* mutant and 322 and 37 in the Δ*gna-3* and Δ*gna-1* strains, respectively (Figure 5.1). The pair of genotypes with the smallest overlap in mis-regulated genes is Δ*gna-1* and Δ*ric8*, sharing only five differentially expressed genes (Figure 5.1). Of the five genes: one is of unknown function (NCU16741), one is involved in steroid synthesis (NCU02042), another is a glucanase (NCU01353; gh16-1) and the other two are transporters (NCU01968 and NCU08397). The two strains that had the most overlap are Δ*gna-3* and Δ*ric8*, with 699 shared differentially expressed genes

(Figure 5.1). Included in the 699 genes are 27 transcription factors (TFs), 10 of which are down-regulated and the other 17 up-regulated. The four most up-regulated TFs are NCU02142, NCU08307, *csp-1* (NCU02713) and *csp-2* (NCU06095). The first two TFs have been studied for their growth and development phenotypes with both NCU02142 and NCU08307 exhibiting no phenotypes (38). The second two TFs are involved in conidiation, which is of note, as Δ*gna-3* and Δ*ric8* mutants inappropriately produce conidia in submerged cultures (6,29). Another interesting TF (and the most down-regulated) is *vsd-8* (NCU06140), which is required for normal growth and sexual development; these are functions shared by *gna-3* and *ric8* (38).

All three mutants share 90 mis-regulated genes (Figure 5.1). Interestingly there are three GPCRs (NCU02903 [*gpr-16*], NCU03253 [*gpr-8*] and NCU06312 [*gpr-4*]) in this group. NCU06312 (*gpr-4*) which is up-regulated at least 4.6-fold over wild type in all three mutants, is of particular interest because of its importance for growth on poor carbon sources (39). Another interesting up-regulated gene is NCU05046 (*ena-1*), a Ca^{2+}-ATPase (40). There are six transcription factors (NCU00090 [*pacc-1*], NCU01312 [*rca-1*], NCU02787 [*sah-12*], NCU03552, NCU04866 [*ada-6*], NCU08726 [*fl*]) that are up-regulated in all three mutants. One TF is of unknown function (NCU03552), *ada-6* regulates all aspects of growth and development, while *sah-12* and *fl* regulate asexual development and *rca-1* has a more subtle or redundant role. (38,41).

Expression of genes involved in energy metabolism. Many metabolic reactions include ATP/ADP and/or NADH/NAD^+ as substrates or products. In the electron transport chain there are five main complexes (42). Complex I transfers electrons from NADH to ubiquinone; this is coupled to the translocation of protons across the inner mitochondrial membrane to set up the gradient to drive ATP formation (see below; 42). Complex II transfers electrons from succinate to ubiquinone, with no coupling to proton translocation. Complex III (cytochrome C) and complex IV (cytochrome C oxidase) transfer electrons from ubiquinone to oxygen producing water. Proton pumping results in a proton gradient that is used by the FoF1-ATP synthase (complex V) to generate ATP (42). There is another type of ATPase, the V-type, which instead of generating ATP uses ATP to pump protons through membranes (43). There is a third type of ATPase, P-type ATPases, which utilize ATP, but in contrast to V-type ATPases, translocate ions to maintain ion gradients (44).

Complex I in *N. crassa* is comprised of 35 nuclear-encoded proteins and 8 mitochondrially encoded proteins (45). In Complex I, 20 of the 35 (95%) protein-coding genes have transcripts that are significantly down-regulated in the Δ*gna-3* and Δ*ric8* mutants. On average, gene expression for Δ*gna-3* was 74% of wild type and gene expression in Δ*ric8* was 48% of wild type (Figure 5.2a). Complexes II, III, IV are mostly similar to wild type for all mutants, with the exception of NCU01808 (*cyc-1*) which is up-regulated 2.5-fold in the Δ*ric8* mutant (Figure 5.2b). CYC-1 or Cytochrome C is the electron carrier in complex IV and becomes a signal for apoptosis when released from the mitochondria (46).

Many differentially expressed genes have predicted metabolic functions. FunCat analysis revealed that many of the mis-regulated genes in the three mutants were implicated in metabolic pathways (Figure 5.3). Of the 38 uniquely regulated genes in the Δ*gna-1* mutant, 17 (44.7%) are predicted metabolic genes. The encoded proteins range from alcohol dehydrogenases, NADPH dehydrogenase, to siderophore transport genes. We analyzed assigned EC numbers to determine whether there was an enrichment in any enzyme class(s). Of the 17 metabolic genes in Δ*gna-1,* 14 (82.3%) have an assigned EC number (computation and/or curated). The most prevalent class is EC class 1 (oxidoreductases; 7 of 17 genes (50%). EC class 3 (hydrolases) and EC class 2 (transferases) are the next most abundant, with 4 (28.6%) and 3 (21.4%) of 17 genes, respectively. There were no other classes represented.

In the Δ*gna-3* mutant, 89 of 322 (27.6%) of the uniquely genes were identified by FunCat as metabolic genes (Figure 5.3). Of the 89 genes, EC numbers were available for 58 (65.1%). The most common class was EC class 2 (transferases) with 21 genes (36.2%), followed by EC class 1 (oxidoreductases) with 17 genes (29.3%), EC class 3 (hydrolases) with 13 genes (22.4%) and EC classes 4-6 (lyases, isomerases, and ligases) with 7 genes (12.1%).

The Δ*ric8* mutant has 186 of 601 (30.9%) uniquely differentially expressed genes that are implicated in metabolism (Figure 5.3). EC numbers were available for 140 of the 186 (75.2%) metabolic genes. The most common EC number class was 1 (oxidoreductases) with 53 genes (37.9%). This class is of note as many of these enzymes

require NAD for their reactions. The next largest class was EC class 3 (hydrolases) with 40 genes (28.6%), then EC class 2 (transferases) with 30 genes (21.4%) and EC classes 4-7 making up the last 17 genes (12.1%).

Targeted Metabolite profiling. Having observed that many of the differentially expressed genes in the three mutants were implicated in metabolism, we implemented targeted metabolite profiling using LC-MS to determine whether there was any difference in the levels of specific metabolites in the three mutants. Targeted metabolite profiling involves comparing observed spectra to a known standard. In our study, a total of 201 standards were used to detect a maximum of 201 metabolites (Table 5.1). Out of the 201 metabolites, 120 were unambiguously detected in wild type and the three mutants (Table 5.1). There were no examples of metabolites that could be detected in one strain, but not the others. Four metabolites (proline, glutamine, phenylalanine and glycerophosphocholine) were present at levels higher than the threshold error and threshold concentrations, but were distinct spectra and therefore were detectable, but not quantifiable. The detected metabolites were grouped into 10 categories: acetyl amino acids, amino acids, electron carriers, organic acids, oxidative stress-related, purines, pyrimidines, TCA cycle-related, urea cycle-related and miscellaneous (Figure 5.4).

Principal component analysis was conducted to represent metabolite differences between the different genotypes (47). We used PCA with three dimensions (accounting for 58.9% of the variance), with results showing that the three mutant strains are in groups distinct from wild type (Figure 5.5).

The metabolome of the Δ*gna-1* mutant is most similar to wild type. The Δ*gna-1* mutant had the greatest number of metabolites with relative levels similar to wild type: 99 out of 120 metabolites (83%). Of the 20 metabolites that differed, only 5 (25%) were uniquely mis-regulated in the Δ*gna-1* strain (Figure 5.6). The five metabolites fell into four categories: acetyl amino acids (2 of 5; N-acetylglutamic acid and N-acetylproline), amino acids (1; 2-aminobutyric acid), organic acids (1; gluconic acid), and urea cycle-related (1; urea). Following the same pattern as the RNA-seq, the mutant with the next highest number of non-shared metabolites was Δ*gna-3*, with 19 (Figure 5.5). The most represented category among these metabolites was amino acids, (8; 42.1%), with no other class having more than two metabolites in a category. Finally, Δ*ric8* had the most metabolites that differed from wild type, with 46 metabolites (Figure 5.6). The category with the most mis-regulated metabolites was amino acids, with 13 (28.2%), followed by purines with 7 metabolites (15.2%), and the next three categories each with three members (oxidative stress related, TCA cycle and urea cycle) for 9 total (19.6%).

In contrast to the results from the RNAseq data, the pair of mutants with the largest overlap in mis-regulated metabolites was Δ*gna-1* and Δ*ric8*, with 13 (Figure 5.6). Of these 13 metabolites, the two categories with the greatest representation are purines and amino acids, with four metabolites in each category (8 total; 61.5%). The category with the highest percent of members represented was electron carriers with 2 of 4 (50%) in lower abundance in the two mutants relative to wild type. The two electron carriers are NADP and nicotinamide mononucleotide.

The pair of mutants with the next highest number of shared mis-regulated metabolites was Δ*ric8* and Δ*gna-3,* with seven metabolites. In the group of seven, there were no enriched categories. Interestingly, of these seven shared metabolites, five are more abundant and two are lower. The two metabolites that are less abundant are NAD^+ and hydroxypyruvic acid.

Evidence for transcriptional and post-transcriptional regulation of metabolites. We next investigated whether any of the enzyme-encoding genes associated with differentially regulated metabolites exhibit evidence for regulation at the transcriptional or post-transcriptional level. In Δ*gna-1* mutants, N-acetylglutamic acid (234% of wild type) and urea (40% of wild type) are the two most mis-regulated metabolites. These metabolites are involved in arginine biosynthesis (N-acetylglutamic acid) or catabolism (urea). When comparing the precursor and product enzyme genes, it was noted that none were differentially expressed, suggesting that GNA-1 may regulate these enzymes post-transcriptionally (Table 5.2). In Δ*gna-3* strains, there was evidence for some transcriptional regulation of tryptophan biosynthesis (Table 5.2). Tryptophan was detected at elevated levels and the gene that encodes the tryptophan 2,3-dioxygenase (forms N-formylkynurenine from tryptophan) was expressed at a level 4-fold lower than wild type, suggesting that lower levels of the enzyme might lead to accumulation of tryptophan (Table 5.2). In the Δ*ric8* mutant, there is evidence for both transcriptional and post-transcriptional regulation. Aspartate was detected at increased levels in Δ*ric8* (327%). The enzyme that initiates the process of converting aspartate into homoserine is

aspartate-semialdehyde dehydrogenase (NCU00554; *hom-1*(48)). *hom-1* is down-regulated 2.2-fold in the Δ*ric8* mutant (Table 5.2). Other metabolites, such as ornithine, have evidence for being regulated at the post-transcriptional level. Ornithine levels are 28% of wild type in the Δ*ric8* strain, however both the precursor and product enzyme genes are not differentially expressed (Table 5.2).

Discussion

We began our study with the goal of identifying genes that are differentially regulated in Δ*ric8* mutants and strains lacking two Gα subunits regulated by RIC8: *gna-1* and *gna-3*. We discovered that metabolic genes made up a large portion of the significantly differentially expressed genes. Having determined that metabolism was a major target of RIC8, GNA-1 and GNA-3, we implemented LC-MS targeted metabolomics to identify mis-regulated metabolites in the three mutants. The combination of a transcriptional and metabolite "snapshots" allows for investigations into whether the metabolites are regulated transcriptionally or post-transcriptionally.

As mentioned above, two of the strains in our study produced only hyphae in submerged cultures (wild type and Δ*gna-1*), while Δ*gna-3* and Δ*ric8* strains consist of hyphae and conidia. Two previous metabolic studies have been performed in *N. crassa* to measure levels of amino acids and other compounds in hyphae and conidia, one using column chromatography and the other study using H-NMR (22,49). While our data generally agreed with the previous two studies, there were differences noted. Proline,

methionine and cysteine could not be detected in conidia using column chromatography in (49) and were not detected in any tissue using NMR in (22). In our study, we were able to detect two of these three metabolites in the hyphae of wild type/Δ*gna-1* and the hyphae/conidia of Δ*gna*-3/Δ*ric8* with LC-MS. Proline was detected, but not quantifiable due to saturation. Methionine was detected in all strains, including Δ*gna-3* and Δ*ric8*, which are predominantly composed of conidia. Cysteine was not one of the targeted metabolites in our study.

Levels of several metabolites differ between conidia and hyphae, with glutamate, glutamine and alanine all found at higher levels in conidia (49). In our study, we observed that the Δ*gna-3* mutant had increased levels of glutamate relative to wild type, which correlates with the increase in levels seen in conidia relative to hyphae in (49). Interestingly, the Δ*ric8* mutant had lower levels of glutamate, suggesting that the differences from wild type are not just due to inappropriate conidiation in submerged cultures, but due to a defect in glutamate metabolism in this mutant.

In this study, we observed down regulation of complex 1 of the electron transport chain in both Δ*gna-3* and Δ*ric8* mutants. The loss of this complex may explain the lower levels of NAD^+ that were observed. NAD^+ serves as a crucial co-enzyme for redox reactions and co-substrates for NAD^+-dependent enzymes (50). Thus, NAD^+ and its co-enzymes control a broad range of physiological processes, including redox homeostasis, genomic stability, gene expression, RNA processing, energy metabolism and circadian clock (50). When acting as a co-enzyme, NAD^+ plays important roles in energy metabolism pathways, such as glycolysis, the TCA cycle and alcohol metabolism (51).

The wide range of metabolic roles for NAD^+ may explain why so many diverse metabolites are mis-regulated in Δ*gna-3* and Δ*ric8* mutants. Another interesting observation is that the most common mis-regulated enzymes in Δ*gna-1* and Δ*ric8* mutants are in EC class 1 (oxidoreductases). Many enzymes in this class require NAD^+ or $NADP^+$ as an electron acceptor.

In the Δ*ric8* mutant the arginine biosynthesis pathway exhibits many differences from wild type including misregulation of both arginine and ornithine. In *N. crassa* the majority of the ornithine and arginine are stored in vacuoles, along with several other amino acids (52-55). The pools of ornithine and arginine are relatively inert metabolically and exchange slowly with the more dynamic cytosolic pools (53, 56, 57). Therefore, the expression of amino acid permeases may affect the levels of available arginine and ornithine and their downstream metabolites. We investigated the transcript levels of amino acid permeases and found that out of 20 detectable amino acid permeases, only one is upregulated in Δ*gna-1*, six (three up and three down-regulated) in Δ*gna-3*, and in Δ*ric8* seven are differentially regulated permeases, with four down and three up-regulated. The one permease that is differentially regulated in Δ*gna-1* (NCU10276, [*aap-11*]) is not shared in Δ*gna-3* and Δ*ric8.* In Δ*gna-3* and Δ*ric8* the three permeases that are upregulated (NCU08880, [*aap-18*]; NCU04435, [*aap-22*]; NCU03784 [*aap-4*]) are the same, however, the two mutants only share one down-regulated permease (NCU00721; *aap-24*). With several amino acid permeases having lower expression as compared to wild type, it is possible that there may be an issue with transport of arginine and ornithine

into the vacuole in the Δ*ric8* mutant. This could reduce the amount of stored arginine and ornithine, leading to reduction in abundance of those metabolites.

Loss of Gα subunits leads to changes in carbon metabolism in both mammalian systems and in *N. crassa*. We observed changes in expression of important carbon metabolism genes, such as *gpr-4* (NCU06312) and *acu-6* (NCU09873). *gpr-4* is important for growth on poor carbon sources and is upregulated on glucose in all three mutants. *acu-6* encodes phosphoenolpyruvate carboxykinase, an important enzyme in gluconeogenesis and is transcriptionally down-regulated in all three mutants. Inspection of levels of phosphoenolpyruvate revealed that amounts were lower in all three mutants as compared to wild type. Additionally, in mammals the loss of Gα subunits lead to changes in mitochondrial respiration. In the Δ*ric8* mutant, down-regulation of electron transport chain complex I was observed, suggesting effects on mitochondrial respiration. As Complex I is involved in NAD^+/NADH recycling, we analyzed levels of NAD^+ and observed that NAD^+ was reduced in the Δ*ric8* mutant. Overall, similarly to in mammals, Gα subunits may play an important role in carbon metabolism and mitochondrial respiration in *N. crassa*.

In this study we used RNAseq and targeted LC-MS analysis to determine the role of *gna-1*, *gna-3* and *ric8* on metabolism in *N. crassa*. We observed that Δ*gna-1*, Δ*gna-3* and Δ*ric8* have differentially expressed genes as compared to wild type. In Δ*gna-3* and Δric8 mutants more than half of the differentially expressed in both mutants are mis-regulated in the same way. Additionally, many of the genes that are differentially

expressed are metabolic genes. To determine if effects on transcription lead to changes in metabolite levels targeted LC-MS analysis was performed against 201 different metabolites standards. Of the 201 potential metabolites 120 were unambiguously detected. LC-MS analysis revealed that all three mutants have significant differences from wild type in metabolite levels. Unlike the RNAseq results, the two mutants with the highest shared metabolite changes are Δ*gna-1* and Δ*ric8*. Combining the both the RNAseq and metabolomics datasets revealed evidence for both transcriptional and post-transcriptional regulation of metabolism by G protein signaling in *N. crassa*.

References

1. Koelle MR. Heterotrimeric G protein signaling: Getting inside the cell. Cell. 2006;126:25–7.

2. Syrovatkina V, Alegre KO, Dey R, Huang X-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. Journal of Molecular Biology. 2016;428:3850–68.

3. Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 2007;61:423–52.

4. Chakravorty D, Assmann SM. G protein subunit phosphorylation as a regulatory mechanism in heterotrimeric G protein signaling in mammals, yeast, and plants. Biochem J. 2018;475:3331–57.

5. Wilkie TM, Kinch L. New roles for Galpha and RGS proteins: communication continues despite pulling sisters apart. Current biology: CB. 2005;15:R843-854.

6. Wright SJ, Inchausti R, Eaton CJ, Krystofova S, Borkovich KA. RIC8 is a guanine-nucleotide exchange factor for Galpha subunits that regulates growth and development in Neurospora crassa. Genetics. 2011;189:165–76.

7. Miller KG, Alfonso A, Nguyen M, Crowell JA, Johnson CD, Rand JB. A genetic selection for Caenorhabditis elegans synaptic transmission mutants. Proc Natl Acad Sci U S A. 1996;93:12593–8.

8. Coates JC. Armadillo repeat proteins: beyond the animal kingdom. Trends Cell Biol. 2003;13:463–71.

9. Miller KG, Rand JB. A role for RIC-8 (Synembryn) and GOA-1 (G(o)alpha) in regulating a subset of centrosome movements during early embryogenesis in Caenorhabditis elegans. Genetics. 2000;156:1649–60.

10. Afshar K, Willard FS, Colombo K, Johnston CA, McCudden CR, Siderovski DP, et al. RIC-8 is required for GPR-1/2-dependent Galpha function during asymmetric division of C. elegans embryos. Cell. 2004;119:219–30.

11. Reynolds NK, Schade MA, Miller KG. Convergent, RIC-8-dependent Galpha signaling pathways in the Caenorhabditis elegans synaptic signaling network. Genetics. 2005;169:651–70.

12. Hampoelz B, Hoeller O, Bowman SK, Dunican D, Knoblich JA. Drosophila Ric-8 is essential for plasma-membrane localization of heterotrimeric G proteins. Nature Cell Biology. 2005;7:1099–105.

13. Gabay M, Pinter ME, Wright FA, Chan P, Murphy AJ, Valenzuela DM, et al. Ric-8 Proteins Are Molecular Chaperones That Direct Nascent G Protein α Subunit Membrane Association. Science signaling. 2011;4:10.1126/scisignal.2002223.

14. Patel BR, Tall GG. Ric-8A gene deletion or phorbol ester suppresses tumorigenesis in a mouse model of GNAQ(Q209L)-driven melanoma. Oncogenesis. 2016;5:e236.

15. Kwon N-J, Park H-S, Jung S, Kim SC, Yu J-H. The putative guanine nucleotide exchange factor RicA mediates upstream signaling for growth and development in Aspergillus. Eukaryot Cell. 2012;11:1399–412.

16. Li Y, Yan X, Wang H, Liang S, Ma W-B, Fang M-Y, et al. MoRic8 Is a novel component of G-protein signaling during plant infection by the rice blast fungus Magnaporthe oryzae. Mol Plant Microbe Interact. 2010;23:317–31.

17. Wu J, Liu Y, Lv W, Yue X, Que Y, Yang N, et al. FgRIC8 is involved in regulating vegetative growth, conidiation, deoxynivalenol production and virulence in Fusarium graminearum. Fungal Genet Biol. 2015;83:92–102.

18. Yang YM, Kuen D-S, Chung Y, Kurose H, Kim SG. Gα12/13 signaling in metabolic diseases. Exp Mol Med. 2020;52:896–910.

19. Kim TH, Koo JH, Heo MJ, Han CY, Kim Y-I, Park S-Y, et al. Overproduction of inter-α-trypsin inhibitor heavy chain 1 after loss of Gα13 in liver exacerbates systemic insulin resistance in mice. Sci Transl Med. 2019;11:eaan4735.

20. Andreeva AV, Kutuzov MA, Voyno-Yasenetskaya TA. G alpha12 is targeted to the mitochondria and affects mitochondrial morphology and motility. FASEB J. 2008;22:2821–31.

21. Kim TH, Yang YM, Han CY, Koo JH, Oh H, Kim SS, et al. Gα12 ablation exacerbates liver steatosis and obesity by suppressing USP22/SIRT1-regulated mitochondrial respiration. J Clin Invest. 2018;128:5587–602.

22. Kim JD, Kaiser K, Larive CK, Borkovich KA. Use of 1H nuclear magnetic resonance to measure intracellular metabolite levels during growth and asexual sporulation in Neurospora crassa. Eukaryotic Cell. 2011;10:820–31.

23. Cortat M, Turian G. Conidiation of Neurospora crassa in submerged culture without mycelial phase. Arch Mikrobiol. 1974;95:305–9.

24. Guignard R, Grange F, Turian G. Microcycle Conidiation Induced by Partial Nitrogen Deprivation in Neurospora crassa. Can J Microbiol. 1984;30:1210–5.

25. That TC, Turian G. Ultrastructural study of microcyclic macroconidiation in Neurospora crassa. Arch Microbiol. 1978;116:279–88.

26. Plesofsky-Vig N, Light D, Brambl R. Paedogenetic conidiation inNeurospora crassa. Experimental Mycology. 1983;7:283–6.

27. Madi L, McBride SA, Bailey LA, Ebbole DJ. rco-3, a gene involved in glucose transport and conidiation in Neurospora crassa. Genetics. 1997;146:499–508.

28. Ivey FD, Kays AM, Borkovich KA. Shared and independent roles for a Galpha(i) protein and adenylyl cyclase in regulating development and stress responses in Neurospora crassa. Eukaryot Cell. 2002;1:634–42.

29. Kays AM, Rowley PS, Baasiri RA, Borkovich KA. Regulation of conidiation and adenylyl cyclase levels by the Galpha protein GNA-3 in Neurospora crassa. Mol Cell Biol. 2000;20:7693–705.

30. Kays AM, Borkovich KA. Severe impairment of growth and differentiation in a Neurospora crassa mutant lacking all heterotrimeric G alpha proteins. Genetics. 2004;166:1229–40.

31. Vogel HJ. Distribution of lysine pathways among fungi: Evolutionary implications. Am Nat. 1964;98:435–46.

32. Davis RH, de Serres FJ. [4] Genetic and microbiological research techniques for Neurospora crassa. In: Methods in Enzymology. Academic Press; 1970. p. 79–143. doi:10.1016/0076-6879(71)17168-6.

33. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525–7.

34. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15:550.

35. Maternal-to-zygotic transition as a potential target for niclosamide during early embryogenesis - PubMed. https://pubmed.ncbi.nlm.nih.gov/31398420/. Accessed 15 Aug 2021.

36. Skyline for Small Molecules: A Unifying Software Package for Quantitative Metabolomics. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7127945/. Accessed 15 Aug 2021.

37. R Core Team. R: A language and environment for statistical computing. https://wwwR-project.org/. 2017. https://www.R-project.org/.

38. Carrillo AJ, Schacht P, Cabrera IE, Blahut J, Prudhomme L, Dietrich S, et al. Functional Profiling of Transcription Factor Genes in Neurospora crassa. G3 (Bethesda). 2017;7:2945–56.

39. Li L, Borkovich KA. GPR-4 is a predicted G-protein-coupled receptor required for carbon source-dependent asexual growth and development in Neurospora crassa. Eukaryot Cell. 2006;5:1287–300.

40. Benito B, Garciadeblás B, Rodríguez-Navarro A. Molecular cloning of the calcium and sodium ATPases in Neurospora crassa. Molecular Microbiology. 2000;35:1079–88.

41. Shen WC, Wieser J, Adams TH, Ebbole DJ. The Neurospora rca-1 gene complements an Aspergillus flbD sporulation mutant but has no identifiable role in Neurospora sporulation. Genetics. 1998;148:1031–41.

42. Joseph-Horne T, Hollomon DW, Wood PM. Fungal respiration: a fusion of standard and alternative components. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2001;1504:179–95.

43. Forgac M. Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. Nat Rev Mol Cell Biol. 2007;8:917–29.

44. Dyla M, Basse Hansen S, Nissen P, Kjaergaard M. Structural dynamics of P-type ATPase ion pumps. Biochemical Society Transactions. 2019;47:1247–57.

45. Pereira B, Videira A, Duarte M. Novel insights into the role of Neurospora crassa NDUFAF2, an evolutionarily conserved mitochondrial complex I assembly factor. Mol Cell Biol. 2013;33:2623–34.

46. Tait SWG, Green DR. Mitochondria and cell death: outer membrane permeabilization and beyond. Nat Rev Mol Cell Biol. 2010;11:621–32.

47. Ringnér M. What is principal component analysis? Nat Biotechnol. 2008;26:303–4.

48. Jenkins MB, Garner HR. Studies of a homoserineless bradytroph of Neurospora crassa: demonstration of an altered aspartate beta-semialdehyde dehydrogenase. Biochim Biophys Acta. 1967;141:287–95.

49. Schmit JC, Brody S. Biochemical genetics of Neurospora crassa conidial germination. Bacteriol Rev. 1976;40:1–41.

50. Xie N, Zhang L, Gao W, Huang C, Huber PE, Zhou X, et al. NAD+ metabolism: pathophysiologic mechanisms and therapeutic potential. Signal Transduct Target Ther. 2020;5:227.

51. Cantó C, Menzies K, Auwerx J. NAD+ metabolism and the control of energy homeostasis - a balancing act between mitochondria and the nucleus. Cell Metab. 2015;22:31–53.

52. Weiss RL. Intracellular localization of ornithine and arginine pools in Neurospora. J Biol Chem. 1973;248:5409–13.

53. Wiemken A, Nurse P. Isolation and characterization of the amino-acid pools located within the cytoplasm and vacuoles of Candida utilis. Planta. 1973;109:293–306.

54. Nurse P, Wiemken A. Amino acid pools and metabolism during the cell division cycle of arginine-grown Candida utilis. J Bacteriol. 1974;117:1108–16.

55. Wiemken A, Dürr M. Characterization of amino acid pools in the vacuolar compartment of Saccharomyces cerevisiae. Arch Microbiol. 1974;101:45–57.

56. Subramanian KN, Weiss RL, Davis RH. Use of external, biosynthetic, and organellar arginine by Neurospora. J Bacteriol. 1973;115:284–90.

57. Karlin JN, Bowman BJ, Davis RH. Compartmental behavior of ornithine in Neurospora crassa. Journal of Biological Chemistry. 1976;251:3948–55.

Figure legends

Figure 5.2: Venn diagram for differentially expressed genes.

Only genes with more than 10 TPMs present in at least one strain and that differed significantly from wild type in the indicated mutant (with $p<0.05$) with a fold change of >2 were included. The numbers represent the number of shared differentially regulated genes that are either up or down regulated in one, both or all three mutants as compared to wild type. The white circle contains the 159 differentially regulated genes in the Δ*gna-1* mutant, the blue circle contains the 1138 differentially regulated genes in the Δ*gna-3* mutant, and the orange circle contains the 1395 differentially regulated genes in the Δ*ric8* mutant.

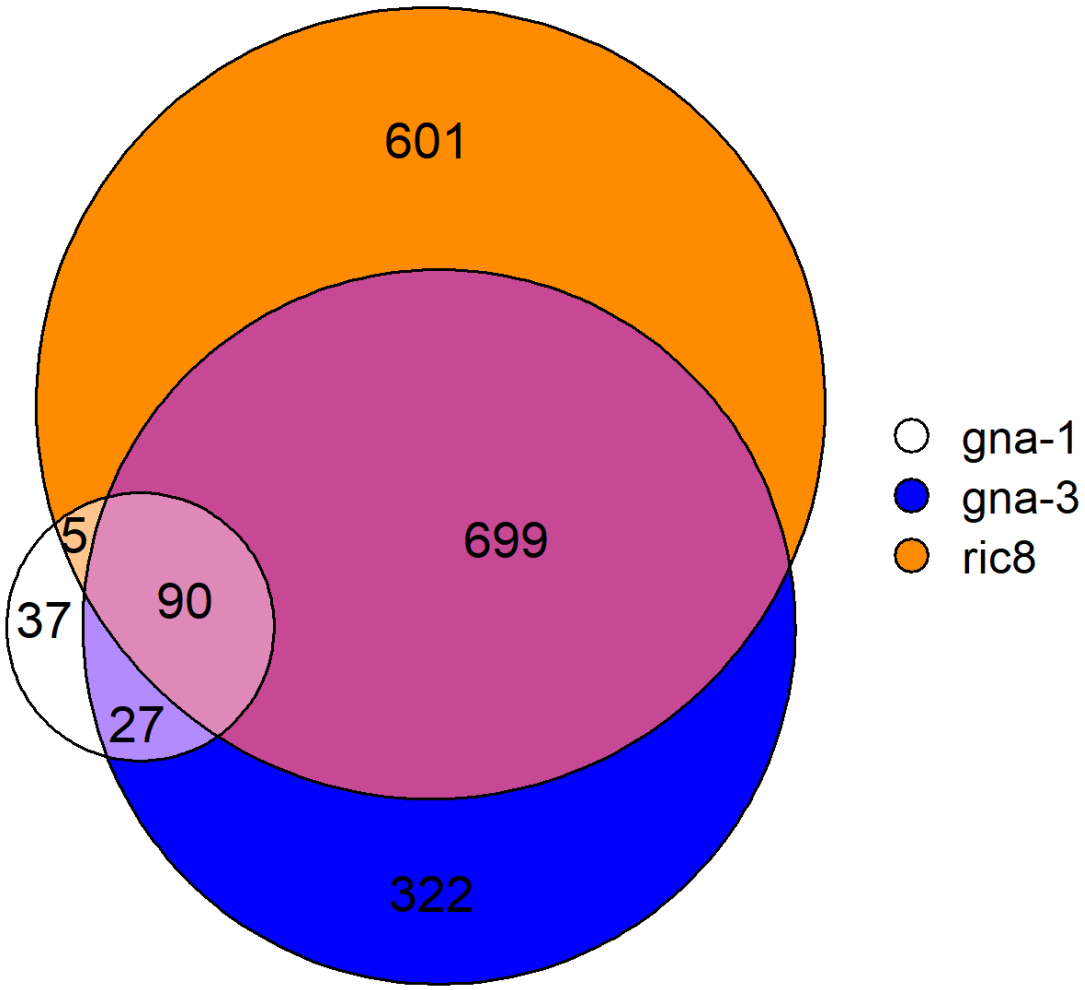

601
699
5
37
90
27
322
gna-1
gna-3
ric8

Figure 5.3. Expression levels of genes.

The relative RNA abundance for each gene transcript is expressed as a percent of wild type. Error bars indicate the standard error. **A.** The 32 nuclear genes that encode proteins that form electron transport chain complex 1. **B.** The nuclear genes that encode electron transport chain complex 3.

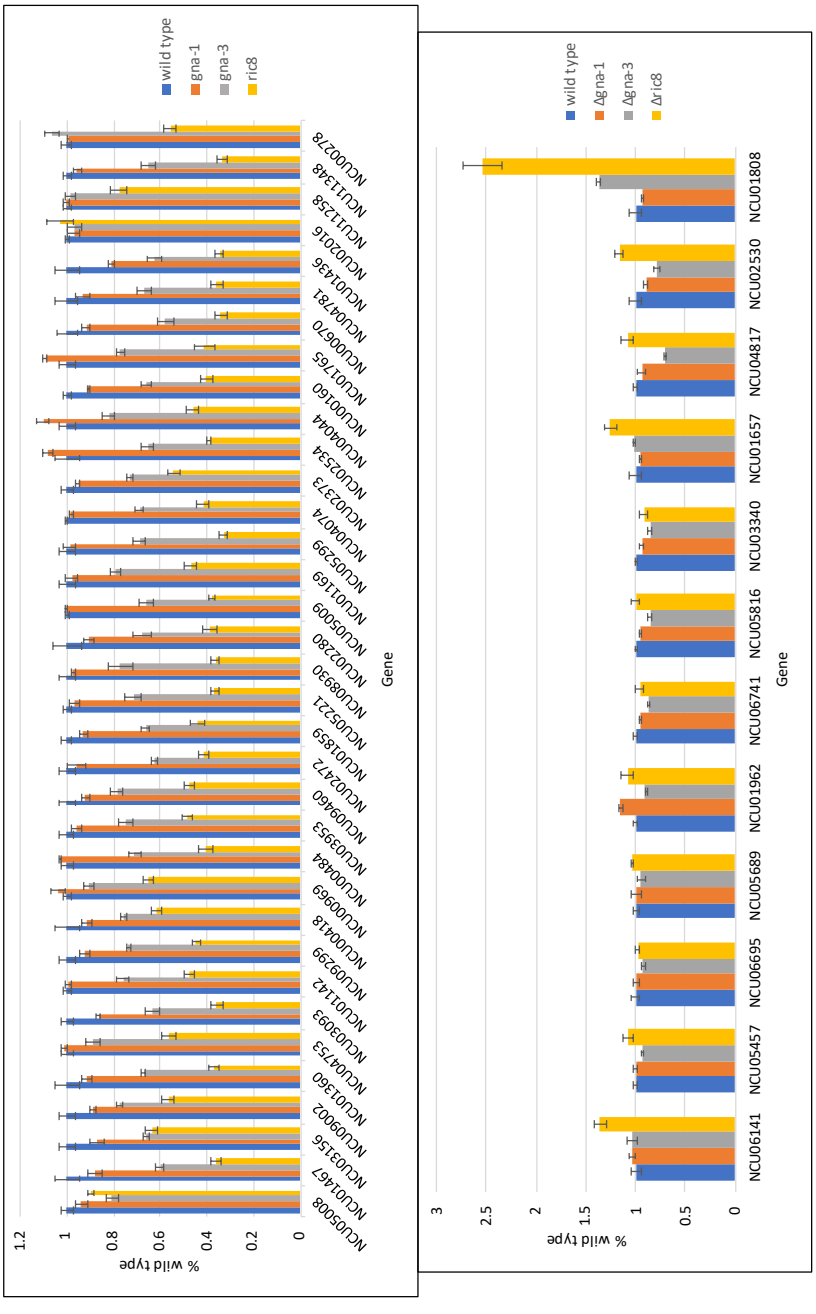

A.
wild type
gna-1
gna-3
ric8
% wild type
Gene
B.
wild type
Δgna-1
Δgna-3
Δric8
% wild type
Gene
NCU06141
NCU05457
NCU06695
NCU05689
NCU01962
NCU06741
NCU05816
NCU03340
NCU01657
NCU04817
NCU02530
NCU01808

Figure 5.4: Differentially expressed genes grouped by functional catalogue (FunCat) assignment.

The distribution of FunCat assignments for all differentially expressed genes in the three mutants. Functional category data were obtained from https://elbe.hki-jena.de/fungifun/fungifun.php **A.** The 159 differentially expressed genes in the Δ*gna-1* mutant. **B.** The 1138 differentially expressed genes in the Δ*gna-3*. **C.** The 1395 differentially expressed genes in the Δ*ric8*.

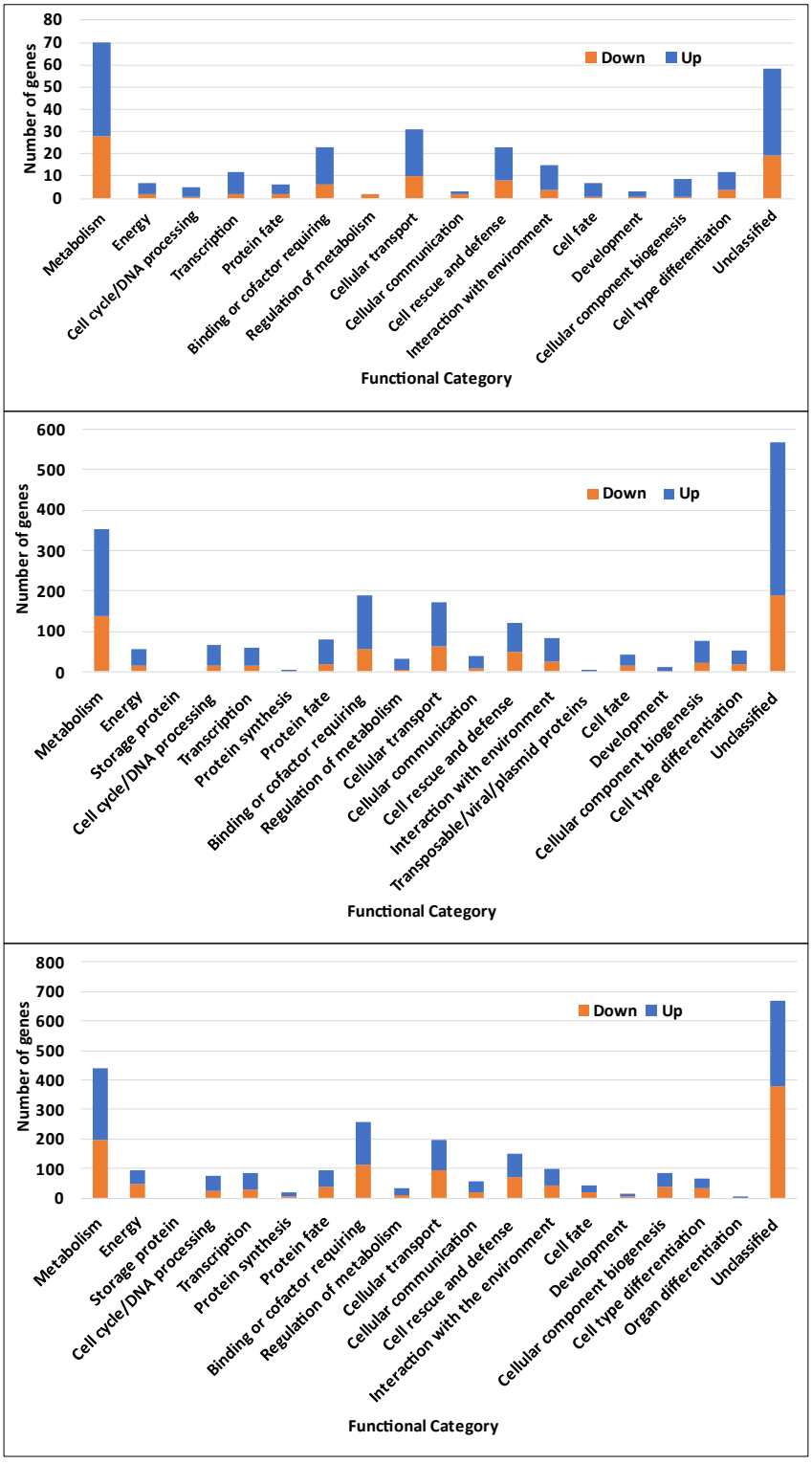

Number of genes
Down
Up
Functional Category
Metabolism
Energy
Cell cycle/DNA processing
Transcription
Protein fate
Binding or cofactor requiring
Regulation of metabolism
Cellular transport
Cellular communication
Cell rescue and defense
Interaction with environment
Cell fate
Development
Cellular component biogenesis
Cell type differentiation
Unclassified
Storage protein
Protein synthesis
Transposable/viral/plasmid proteins
Interaction with the environment
Organ differentiation

Figure 5.5: Principal component analysis (PCA) score plot of metabolite abundances.

PCA analysis of the relative abundances from LC-MS spectra measured for replicate samples. The *x* y and z axis 58.9% of the variance.

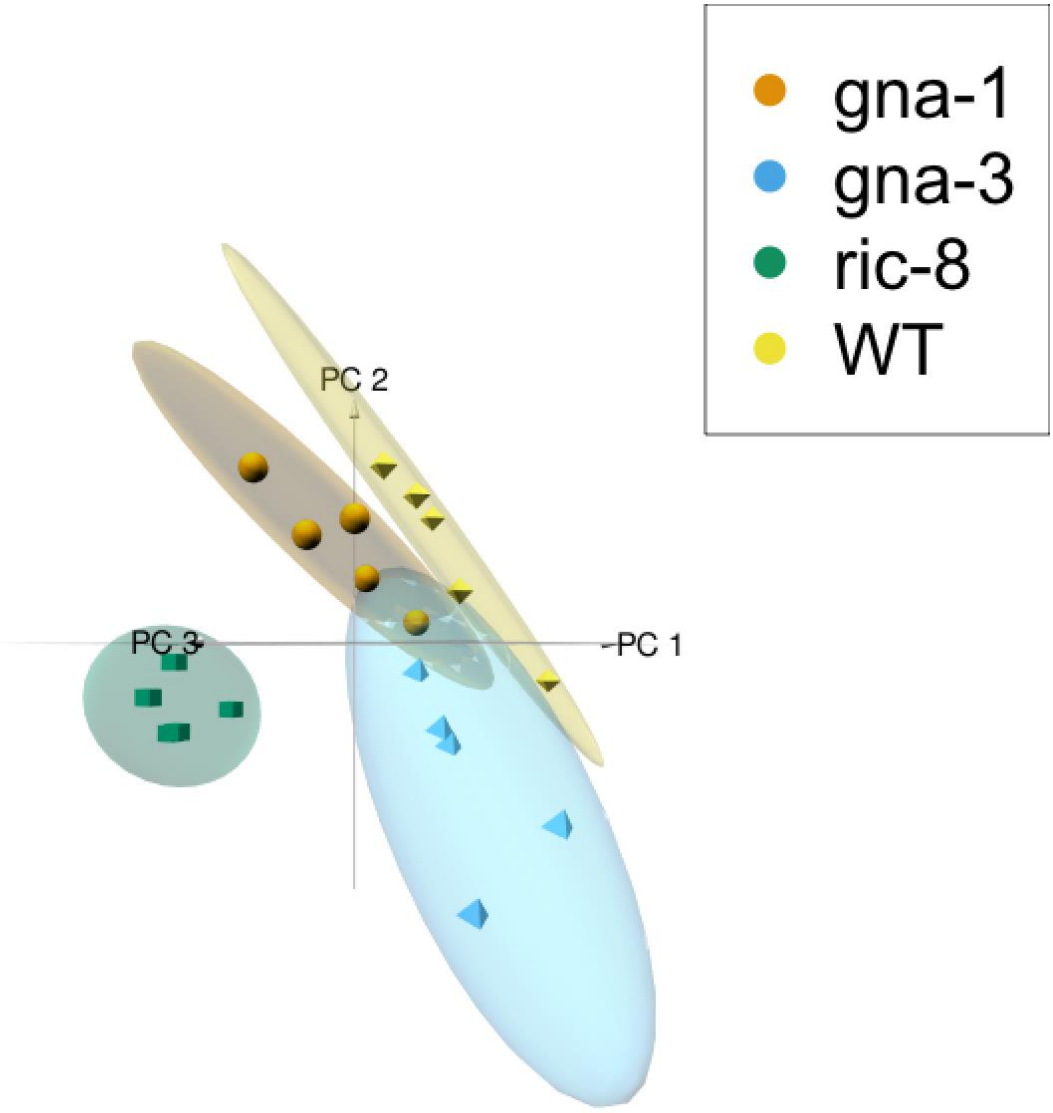

gna-1
gna-3
ric-8
WT
PC 2
PC 3
PC 1

Figure 5.6. Distribution of metabolites into categories.

Each “slice” of the pie represents a fraction of the metabolites within the indicated category. The number of metabolites in each category is indicated. The MISC (Miscellaneous) group includes 18 categories with three or fewer members.

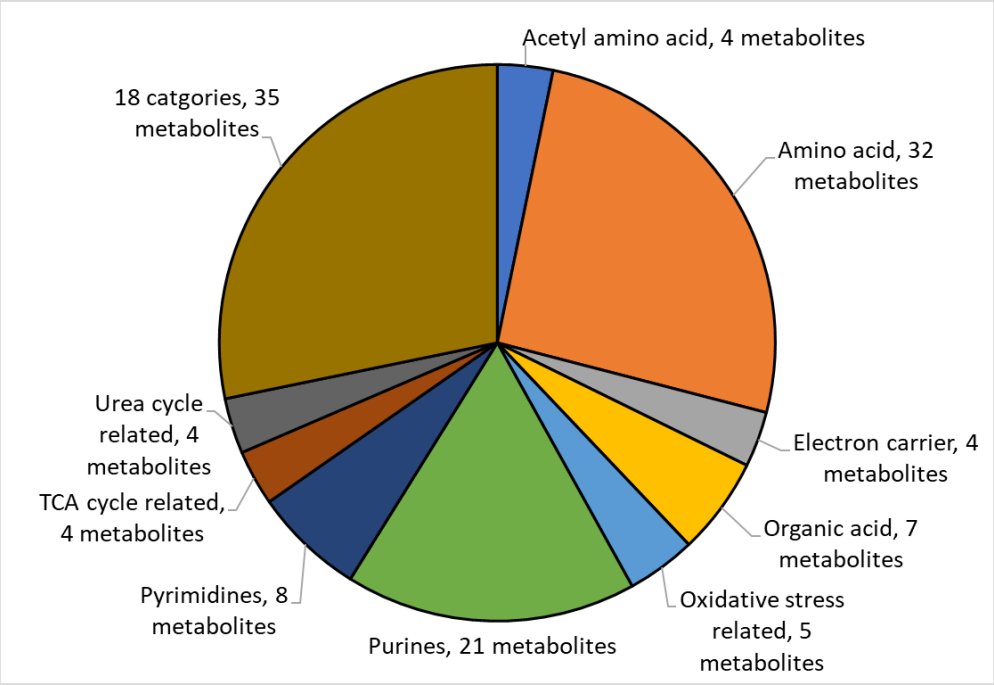

Acetyl amino acid, 4 metabolites
Amino acid, 32 metabolites
Electron carrier, 4 metabolites
Organic acid, 7 metabolites
Oxidative stress related, 5 metabolites
Purines, 21 metabolites
Pyrimidines, 8 metabolites
TCA cycle related, 4 metabolites
Urea cycle related, 4 metabolites
18 catgories, 35 metabolites

Figure 5.7: Venn diagram of regulated metabolites.

All metabolites that differed significantly from wild type at $p<0.05$ or lower were included. The numbers represent the number of shared differentially regulated metabolites that are either up or down regulated in one, both or all three strain as compared to wild type. The white circle contains the 25 differentially regulated metabolites in the Δ*gna-1* mutant, the blue circle contains the 27 differentially regulated genes in the Δ*gna-3* mutant, and the orange circle contains the 67 differentially regulated gene in the Δ*ric8* mutant.

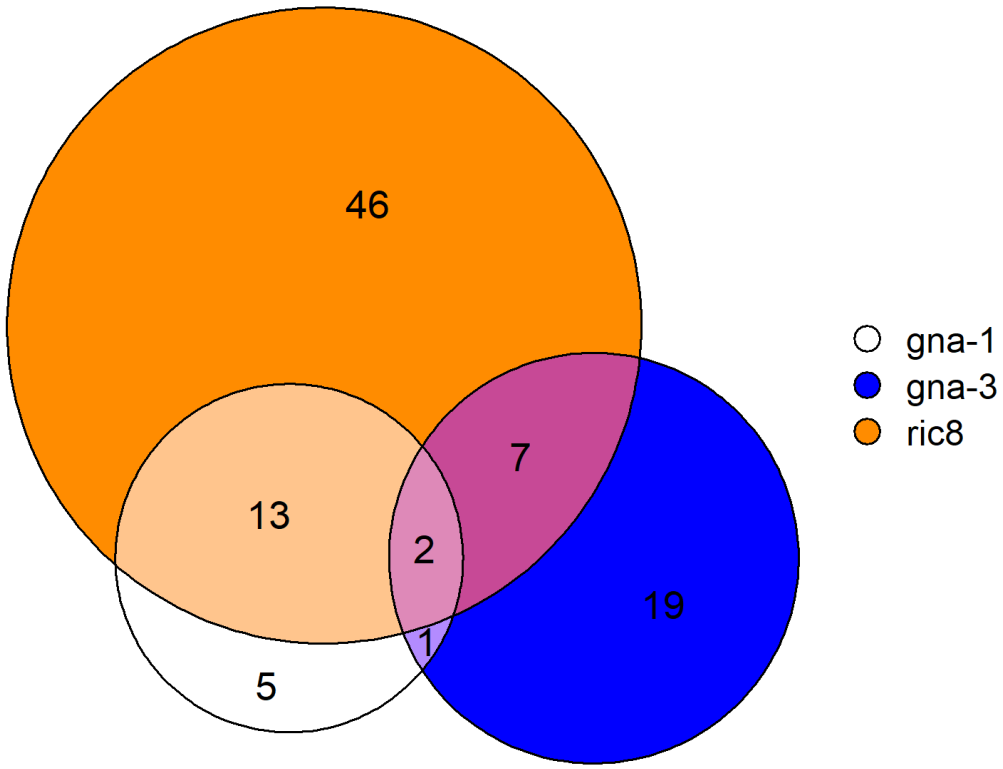

46
13
2
7
19
1
5
gna-1
gna-3
ric8

Table 5.1: List of the 201 metabolites

			Relative abundance			
Metabolite	**Category**	**Detected**	**Wild type**	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***
N-Acetylglutamic acid	Acetyl amino acid	Yes	804423174.4	1874725027	1381451446	357840230.4
N-Acetylglycine	Acetyl amino acid	Yes	6077851	5547780.2	6283351.8	5163450.2
N-Acetyllysine	Acetyl amino acid	Yes	17602755.6	13830016.6	26805270.2	9194723.2
N-Acetylproline	Acetyl amino acid	Yes	3208229.6	3957286.8	3504768.2	820115.8
Trimethylamine N-oxide	Amine oxide	Yes	816068.8	827896.6	941896	1098750.6
Malylglutamic acid	Amino acid	Yes	86391	77094.8	106340.6	52122.8
Taurine	Amino acid	Yes	4085936.2	3350510.8	2975296.2	3202095
Asparagine	Amino acid	Yes	28759554.8	27030504	42531856	19411269.6
4-Guanidinobutyric acid	Amino acid	Yes	9882528	7558653.8	9659564	3564059.8
Lysine	Amino acid	Yes	2447184090	1540166234	3744693350	903047574.4
Threonine	Amino acid	Yes	5031059021	5008972723	5900597965	4054736947
Arginine	Amino acid	Yes	12035400000	1.0177E+10	1.3327E+10	7134725427
Pipecolic acid	Amino acid	Yes	1657635536	780397596	1403411859	277348538.6
Urocanic acid	Amino acid	Yes	7932659.6	8908988	9233369.4	5695056.8
O-Succinylhomoserine	Amino acid	Yes	49931719.6	60868025.2	54584128	29378632.6
Tyrosine	Amino acid	Yes	617278921.6	634527958	901281434	603893510.4
4-Hydroxyproline	Amino acid	Yes	5242315.8	6605211	7371032.4	5185722.2
Methionine	Amino acid	Yes	1250119194	1967972762	1159612269	1071230618
5-Aminovaleric acid	Amino acid	Yes	2616463.4	2367335.8	2454350.4	2210013.4
Kynurenine	Amino acid	Yes	13190760.8	12423597.6	7169262.6	3425097.2
Aspartic acid	Amino acid	Yes	106592710.2	202180558	144099326	348678693.6
2-Aminobutyric acid	Amino acid	Yes	1565105582	1943769554	1733052380	1208856732
Glutamic acid	Amino acid	Yes	6473990630	7418981709	8016131072	4749030733
Leucine	Amino acid	Yes	2342585736	2102346511	2889521535	1996885224
Serine	Amino acid	Yes	801819670.4	851924661	879636880	889101735.2
beta-Alanine	Amino acid	Yes	20751088.6	25221569.6	51335827.4	96899233
Tryptophan	Amino acid	Yes	559313363.2	520647654	735311450	446527238.4
Cystathionine	Amino acid	Yes	1322052371	1361020410	1064071213	589062592
5-Hydroxytryptophan	Amino acid	Yes	703184.6	565267.4	768612.4	510446.2
5-Hydroxylysine	Amino acid	Yes	120755.6	179299.2	186508.4	180068.8
3-Methylhistidine	Amino acid	Yes	8985459.4	8459811.6	10149332.4	12335678
Isoleucine	Amino acid	Yes	2787517811	2391660563	3068951034	2158040243
gamma-Aminobutyric acid	Amino acid	Yes	768131513.6	288431134	230423190	550345673.6
Valine	Amino acid	Yes	9077268774	8234958259	1.053E+10	7604706675
Nicotinic acid	B vitamin	Yes	86922949.4	109026286	129659992	96473358.6
Thiamine	B vitamin	Yes	1037183.6	1762497	1593788.4	1522453.2
Thiamine pyrophosphate	B vitamin	Yes	1616591	1778911.2	2028167	1766651.6
Acetylcarnitine	Carnitine	Yes	975878408	1086483152	816644040	2025107306
Carnitine	Carnitine	Yes	700071766.4	753682067	530765968	822869945.6
3-Hydroxy-3-methylglutaryl CoA	Cholesterol synthesis	Yes	892591	665962.8	624432.4	315721.8
Choline	Choline	Yes	432136805.6	476602021	839255936	1099937874
Nicotinamide mononucleotide	Electron carrier	Yes	55211208.8	43909111.4	46620782.4	31770487.4
NADP	Electron carrier	Yes	2932658.4	1697653.8	2629792.6	1767760.6
FAD	Electron carrier	Yes	68149258.4	57792462.8	78556238	56410395.2

			Relative abundance			
Metabolite	**Category**	**Detected**	**Wild type**	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***
NAD	Electron carrier	Yes	399571827.2	265300355	164557075	194221268
Creatinine	Energy	Yes	1267209.8	1518927.2	1231988.8	1076607.6
Phosphoenolpyruvic acid	Glycolysis	Yes	1415128.6	664429.8	577628	579081
Glycerol-3-phosphate	Glycolysis	Yes	77073672.2	106352681	122699285	67647353.4
DXP	MEP pathway	Yes	1531574.2	1407418.8	1260402.8	1059072.8
MEcPP	MEP pathway	Yes	189661.6	210229.8	263831.4	255096.2
MEP	MEP pathway	Yes	381652.4	371285.2	194328.6	405182.2
Dimethylarginine	Methyl amino acid	Yes	90496210.8	82636093.2	84663672.4	83949140
Trimethyllysine	Methyl amino acid	Yes	91155176.8	132195616	81167447.8	170242969.6
S-Adenosylhomocysteine	Methylation	Yes	449918276.8	503896859	649931213	409576972.8
S-Adenosylmethionine	Methylation	Yes	127305286	166009341	231356122	125294153.6
Betaine	Methylation	Yes	3698346.8	2239783.6	3202210.2	5494687.4
Glutaric acid	Organic acid	Yes	521466	522756.2	631369.2	349648.4
Glycolic acid	Organic acid	Yes	38407.8	40879	36205.6	69762.8
Hydroxypyruvic acid	Organic acid	Yes	574416	557234.6	412516.4	193542.2
Glyceric acid	Organic acid	Yes	11789837	9539455.2	17745094.4	10001828.2
Tartaric acid	Organic acid	Yes	550818	1164373	724571	634391.2
Gluconic acid	Organic acid	Yes	22179911.8	35801170.2	35528539	25613323.4
Malonic acid	Organic acid	Yes	3532659.4	2512602.6	2690254.2	1697035.4
3-Nitrotyrosine	Oxidative stress	Yes	658985.4	1225139.4	2179309.4	4691829.2
Methionine sulfoxide	Oxidative stress	Yes	8495892.2	12491986	8127119.2	8215304.6
GSH	Oxidative stress	Yes	6135951821	5688457882	6138252800	4391524710
Ophthalmic acid	Oxidative stress	Yes	965971328	1182925958	842538659	778847840
GSSG	Oxidative stress	Yes	364845384	414797726	628193760	163535826
Glu-asp-OH	Peptide	Yes	370258	278665.8	505771.8	242901.4
4-Hydroxybenzoic acid	Phenolic	Yes	965192	922553.4	973287.8	785598.4
O-Phosphoethanolamine	Phospholipid turnover	Yes	1328813.2	753533.2	1391642.2	1133192.2
CDP-choline	Phospholipid turnover	Yes	129077839.6	109852908	114800153	174514207.2
CDP-ethanolamine	Phospholipid turnover	Yes	12049321.2	10469646	14646774.6	13163793
N-Acetylputrescine	Polyamine	Yes	29910702	31799724.2	26664312.4	39466931.4
Allantoin	Purine	Yes	2426115.4	2245541.2	3319139.4	3900691.2
ATP	Purine	Yes	17815358.4	13007334.6	14368358.6	11402566.6
dGDP	Purine	Yes	1594564.6	747578.2	1550340.4	821913.6
dGMP	Purine	Yes	996313	816241	1429007.2	1163560.4
GMP	Purine	Yes	94653966.8	45680523	114773245	55200473.4
IDP	Purine	Yes	320326.8	184143.4	279992.4	185916.4
IMP	Purine	Yes	18440255.6	13917519	20812771.4	8499128
Xanthine	Purine	Yes	10156516.8	13340781	13479628.4	24706388.8
Hypoxanthine	Purine	Yes	66717186.6	76368234.2	58984941.8	126410800
Guanosine	Purine	Yes	566344730.8	542548254	852683907	680802638.4
Adenosine	Purine	Yes	6947113808	4475693037	1.4052E+10	4537293350
ADP ribose	Purine	Yes	169009950.4	222537298	347991024	250739089.6
Deoxyguanosine	Purine	Yes	280438.8	489629.4	435579.2	1416997
dATP	Purine	Yes	125222.2	101892.4	162664.6	69585.4
Inosine	Purine	Yes	12688850.2	9974863	33127967.4	9171751.4
Deoxyadenosine	Purine	Yes	12284975	13787389.8	13071813.6	15130858.2

			Relative abundance			
Metabolite	**Category**	**Detected**	**Wild type**	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***
AMP	Purine	Yes	1817265144	1245737782	2522675048	1115987533
Guanine	Purine	Yes	183998027.2	184666163	218751677	241963917.6
ADP	Purine	Yes	169381795.6	98286174.6	174575004	82241284.2
Adenine	Purine	Yes	649385059.2	630426659	838229187	702915960
cyclic GMP	Purine	Yes	10071199	11068704.6	11268721.6	10877077.4
cyclic CMP	Pyrimidine	Yes	5282117.6	5653426.8	5698412.8	5479867.8
dCMP	Pyrimidine	Yes	315962	337195.4	410782.2	377997
Cytidine	Pyrimidine	Yes	62045298.4	65297237.8	111546770	74786831.6
UMP	Pyrimidine	Yes	74099687	44038460	128661564	32360066.2
Uridine	Pyrimidine	Yes	18245779.2	11307836.4	38206699	21224671.4
Orotic acid	Pyrimidine	Yes	10214773.4	12032009.4	7435913.4	5192563.6
UDP-glucose	Pyrimidine	Yes	1257600486	1369101117	1289922518	1015519878
dUMP	Pyrimidine	Yes	581243.6	456568.8	483636.4	507757.2
Disaccharide C12	Sugar	Yes	1609492994	1442484102	1393950638	1843140218
Hexose	Sugar	Yes	10970080	13998258.6	11605565	16911705
Adonitol	Sugar alcohol	Yes	3467781.8	2787769.6	2482418.4	1050189
myo-Inositol	Sugar alcohol	Yes	2913462	3273221.6	5963872.8	6898317.4
Sorbitol	Sugar alcohol	Yes	1342009645	1723219904	1438724326	2076658445
Hexose phosphate	Sugar phosphate	Yes	50889373	59430529.4	61582997.4	48750599.6
Pentose phosphate	Sugar phosphate	Yes	32046598.4	47123766.6	55673935.6	26654367.2
Trehalose-6-P	Sugar phosphate	Yes	2717141	5146657.8	4435833.4	5008377.4
Iso/Citric acid	TCA cycle	Yes	558448116.8	367809565	361230334	1040661866
Malic acid	TCA cycle	Yes	867270688	1153769811	1058139549	1348390970
alpha-Ketoglutaric acid	TCA cycle	Yes	91056367.6	78315643.8	77059516.8	34737375.8
Succinic acid	TCA cycle	Yes	75510886.6	83866707.6	89068558.4	185174162.4
Urea	Urea cycle	Yes	112095595.2	44426901.6	110536042	88369873.6
Citrulline	Urea cycle	Yes	3159413197	2664412826	3709102234	869167756.8
Ornithine	Urea cycle	Yes	4844321485	3758314906	5951520154	1336966118
Argininosuccinic acid	Urea cycle	Yes	835141760	983371629	1107404733	1127033408
UDP-N-acetylglucosamine		Yes	1007863283	1107408742	1042321158	897369088
Proline	Amino acid	Yes	N/A	N/A	N/A	N/A
Phenylalanine	Amino acid	Yes	N/A	N/A	N/A	N/A
Glutamine	Amino acid	Yes	N/A	N/A	N/A	N/A
Glycerophosphocholine	Choline	Yes	N/A	N/A	N/A	N/A
O-Acetylserine	Acetyl amino acid	No				
N-Acetylserine	Acetyl amino acid	No				
Trigonelline	Alkaloid	No				
N-Methyltryptamine	Alkaloid	No				
Cysteinesulfinic acid	Amino acid	No				
Tryptophanamide	Amino acid	No				
Tyramine	Amino acid	No				
3-methoxytyramine	Amino acid	No				
Homocystine	Amino acid	No				
Aminoadipic acid	Amino acid	No				
Methionine sulfoximine	Amino acid	No				
S-Carboxymethylcysteine	Amino acid	No				
Selenomethionine	Amino acid	No				
Xanthurenic acid	Amino acid	No				
2,6-Diaminopimelic acid	Amino acid	No				

			Relative abundance			
Metabolite	**Category**	**Detected**	**Wild type**	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***
Cystine	Amino acid	No				
Cysteic acid	Amino acid	No				
N-Acetylneuraminic acid	Amino sugar	No				
Glucosaminic acid	Amino sugar	No				
Hexoseamine	Amino sugar	No				
N-Acetylhexoseamine	Amino sugar	No				
Folic acid	B vitamin	No				
Thiamine monophosphate	B vitamin	No				
Glutarylcarnitine	Carnitine	No				
Acetylcholine	Choline	No				
3-dephospho-CoA	CoA	No				
L-DOPA	Dopamine pathway	No				
NADPH	Electron carrier	No				
NADH	Electron carrier	No				
Phosphocreatine	Energy	No				
Creatine	Energy	No				
Naringenin-7-O-glucoside	Flavonoid glycoside	No				
2/3-Phosphoglyceric acid	Glycolysis	No				
CDP-ME	MEP pathway	No				
N-Methylglutamate	Methyl amino acid	No				
1-Methylhistidine	Methyl amino acid	No				
Sarcosine	Methylation	No				
N,N-dimethylglycine	Methylation	No				
Methylmalonic acid	Organic acid	No				
3-Methylglutaric acid	Organic acid	No				
Adipic acid	Organic acid	No				
2-Methylglutaric acid	Organic acid	No				
Azelaic acid	Organic acid	No				
Guanidinosuccinic acid	Oxidative stress	No				
Anserine	Peptide	No				
Thyrotropin releasing hormone	Peptide	No				
Epi/catechin	Phenolic	No				
p-Coumaric acid	Phenolic	No				
3,4-Dihydroxycinnamic acid	Phenolic	No				
Syringic acid	Phenolic	No				
Vanillic acid	Phenolic	No				
Gallic acid	Phenolic	No				
Phosphocholine	Phospholipid turnover	No				
2-phosphoglycolate	Photosynthesis	No				
6-Phosphogluconic acid	PP pathway	No				
GTP	Purine	No				
dAMP	Purine	No				
GDP	Purine	No				
ITP	Purine	No				

			Relative abundance			
Metabolite	**Category**	**Detected**	**Wild type**	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***
Xanthosine	Purine	No				
ADP glucose	Purine	No				
8-OH-deoxyguanosine	Purine	No				
XMP	Purine	No				
3-Methylxanthine	Purine	No				
CTP	Pyrimidine	No				
UDP-glucuronic acid	Pyrimidine	No				
Thymidine	Pyrimidine	No				
dCDP	Pyrimidine	No				
Deoxyuridine	Pyrimidine	No				
Deoxycytidine	Pyrimidine	No				
Dihydroortic acid	Pyrimidine	No				
3-Ureidopropionic acid	Pyrimidine	No				
Tetrasaccharide C24	Sugar	No				
Trisaccharide C18	Sugar	No				
Glucosamine-6-phosphate	Sugar phosphate	No				
Glucosamine-6-sulfate	Sugar sulfate	No				
3,5-Diiodotyrosine	Thyroid hormone pathway	No				

Table 5.2: Misregulated metabolites and genes encoding possible associated enzymes

	Metabolite abundance (% wild type)			Precursor enzyme	Precursor differentially regulated			Product enzyme	Product differentially regulated		
Compound	**Δ*gna-1***	**Δ*gna-3***	**Δ*ric8***		**Δ*gna-1***	**Δ*gna-3***	**Δric8**		**Δ*gna-1***	**Δ*gna-3***	**Δric8**
N acetyl-glutamic acid	233%**	172%	44%**	Arginino-succinate lyase (NCU08162) Arg-10	No	No	No	Arginase (NCU02333) Aga-1	No	Yes	No
Urea	40%*	99%	79%	Arginase (NCU02333) *aga-1*	No	Yes	No	Urease (NCU01246) *ure-1*	No	No	No
Ornithine	78%	123%	28%*	Acetylornithine-glutamate transacetylase (NCU05622) *arg-15*	No	No	No	Ornithine carbamoyl transferase (NCU01667) *arg-12*	No	No	No
Tryptophan	93%	131%*	80%	-				Tryptophan 2,3-dioxygenase (NCU05752) *nt*	Yes	Yes	Yes
Aspartate	190%	135%	327%***	-				Aspartate-semialdehyde dehydrogenase (NCU00554) *Hom-1*	No	No	Yes

Chapter 6

Conclusions and future directions

The main objective of this thesis has been to generate and utilize phenotypic data to reveal potential gene pathways in *N. crassa* and to determine the ways in which G protein signaling regulates transcription and metabolite levels.

In Chapter 2, 312 transcription factors were annotated and 242 had phenotypes determined for 10 traits in *N. crassa*. Of the 242 mutants analyzed 64% had at least one phenotype, 21% had at least two phenotypes in growth and development and 8% exhibited phenotypes in all aspects of growth and development. Combining the phenotypic data with mRNA expression data revealed that some transcription factors are co-transcribed and yet do not individually exhibit phenotypes, suggesting redundancy. Creating and analyzing double gene deletion mutants of the co-transcribed genes may reveal redundant transcriptions factors and implicate them as acting in the same pathways.

In Chapter 3, ~1200 gene deletion mutants had phenotypic data available for 10 different traits in growth and development. To utilize this data, we clustered the mutants into groups by phenotypes using a weighted partitioning around medoids approach. We then looked for correlations between phenotypes, protein features and mRNA expression. We observed that mutants with substantially taller aerial hyphae height correlated with genes that encoded proteins with transmembrane domains. The genes that encoded proteins with transmembrane domains were predicted as being related to oxidative stress or as transporters, revealing that these genes may work together to regulate aerial hyphae height. The induction of aerial hyphae requires a hyperoxidant state and we could

measure the strains sensitivity to oxidative stress to see if these genes play a large role in the oxidative stress response.

We also found that all genes with a yeast homolog had evidence in yeast of interaction with the at least one of the other co-clustered genes revealing a potential MAPK network of genes. In the future, this work could be followed up on by exploring epistatic relationships and effects on MAPK phosphorylation in the *N. crassa* mutants in the cluster.

With this clustering method, as phenotypic data becomes available for additional mutants, the clustering analysis can be re-run and the new pathways may be revealed. Additionally, this clustering method could be applied to phenotypic data in different conditions looking at other aspects of growth and development such as stress and drug resistance.

In Chapter 4, we investigated the role of Receptor for Activated C Kinase-1 (RACK1) homolog *cpc-2* in regulation and interaction with G protein signaling. RACK1 has been shown to act as alternative Gβ protein in many systems. Genetic epistasis experiment between *cpc-2* and the G proteins (Δ*gna-1,* Δ*gna-2,* Δ*gna-3,* and Δ*gnb-1)* revealed that *cpc-2* is epistatic to Δ*gna-2* for basal hyphae growth rate, and mutational activation of *gna-3* suppressed an inappropriate conidiation defect present in Δ*cpc-2*. Follow-up studies could explore epistasis between *cpc-2* and G protein subunit genes for additional traits, such as utilization of cellulose or hemicellulose as a carbon source.

Finally, in Chapter 5, we used transcriptomics to determine the role of G proteins in transcriptional control of gene expression. We found that all mutant strains had differentially expressed genes. Many of the differentially expressed genes were categorized by FunCat analysis as metabolic genes. We employed LC-MS analysis to determine levels of primary metabolites in the strains, with all strains exhibiting mis-regulated metabolites. Comparing the RNAseq data to the metabolomics data revealed evidence for both transcriptional and post-transcriptional regulation of metabolites. It also showed mis-regulation in both energy-carrying metabolites and in energy metabolite recycling genes, suggesting a general mis-regulation of energy metabolism. The next steps would be to assay several of these energy metabolites directly using spectrophotometric or colorimetric tests to determine the pools of these metabolites in each mutant. In addition, the activities or protein levels for the various ATPases could be determined using available assays and antibodies in *N. crassa*.