UNDERSTANDING CIRCADIAN OUTPUT NETWORKS IN *NEUROSPORA CRASSA*

A Dissertation

by

RIGZIN NGODUP DEKHANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Deborah Bell-Pedersen
Committee Members,	Paul E. Hardin
	Terry L. Thomas
	Martin B. Dickman
Head of Department,	Thomas D. McKnight

May 2015

Major Subject: Biology

ABSTRACT

The *Neurospora crassa* circadian clock is based on a highly regulated molecular negative feedback loop, similar to molecular clocks in all eukaryotes. A core component of the *N. crassa* molecular clock is the White Collar complex (WCC), composed of the blue light photoreceptor WC-1 and its partner WC-2. The WCC serves as a master regulator that controls light signaling, and the precise timing of target gene expression. Up to 40% of the eukaryote genome is under the control of the clock at the level of transcript abundance, but the molecular links between the core oscillator and downstream target genes, as well as the mechanisms controlling the phase of rhythmic gene expression, are not understood. Using chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-seq), about 400 binding sites for the WCC were identified throughout the *N. crassa* genome. We found that 24 transcription factors (TFs) were significantly enriched among the direct WCC target genes. As expected for genes that are controlled by the WCC, the first-tier TFs are both clock- and light-regulated. These data led to the hypothesis that the WCC functions to control rhythms in TFs, which in turn control rhythmicity and phase of downstream target genes and processes.

To test this hypothesis, the first-tier TF ADV-1 (Arrested Development-1) was investigated in detail to characterize the downstream circadian genetic network. ADV-1 target genes were identified using ChIP- and RNA-seq, and as expected many ADV-1

downstream target genes were light-responsive and/or clock-controlled. An enrichment for ADV-1 target genes involved in cell fusion, a process that is critical for normal vegetative and sexual development in *N. crassa,* provided a rationale for the observed developmental defects in ADV-1 deletion cells, and suggested that cell fusion is clock-controlled. Importantly, this work revealed that the transduction of time-of-day information through ADV-1 to its downstream targets is more complex than anticipated. Specifically, I show that deletion of ADV-1 does not always lead to predicted changes in rhythmic gene expression and/or phase, suggesting that ADV-1 functions in combination with other first-tier TFs to control rhythmicity. In support of this idea, genome-wide binding profiles of all of the first-tier TFs uncovered complex feedback and feed forward regulation involving ADV-1. Thus, my data revealed that in order to fully understand how the clock signals phase information to downstream targets, we need to go beyond the candidate gene approach, and instead develop computational models from our TF ChIP-seq and rhythmic transcriptome data to model how time of day information is transduced in the molecular circadian output gene network. Predictions of the model can then be validated using ADV-1 deletion cells alone, or in combination with deletion of other first-tier TFs in the network, with the goal of deriving design principles that define conserved aspects of the circadian output network in all eukaryotes, and important in human health.

DEDICATION

I would like to dedicate this thesis to my father, Wangchuk Norbu Dekhang, and my mother, Dolma Choedon Dekhang.

ACKNOWLEDGEMENTS

I would like to thank my PhD advisor, Dr. Deb Bell-Pedersen, for her continued support, encouragement and guidance throughout my research. I have matured as a person and as a researcher under her tutelage. I would also like to thank my committee members, Dr. Paul Hardin, Dr. Terry Thomas, and Dr. Marty Dickman, for their support and suggestions.

I am thankful to former lab members, Drs. C. Goldsmith, L. Bennett, I. Nsa and R. De Paula and current lab members, N. Ojha, O. Ibarra, J. Fazzino, Dr. N. Karunarathna and Dr. T. Lamb, of the Bell-Pedersen lab. Special thanks to Dr. T. Lamb for her help with Chapters I and V. I would also like to thank all the collaborators and the personnel involved in my research project- Drs. M. Sachs (Texas A&M), M. Freitag (Oregon State University), J. Galagan (Boston University) and J. Dunlap (Dartmouth College). Thanks also go to my colleagues and the department faculty and staff for making my time here a good experience.

Finally, thanks to my mother and father for their immense love and support, my younger brothers, Tseten and Jangchup and sister, Tsering Samten for their support and to my wife, Tenzing Zompa for her patience and love. I would like to thank my uncle and aunt, Dorjee and Kalsang Ngamdung, and their two sons, Tsering and Tashi for always

making me feel at home, and my grandparents especially grandma Kalden and uncles, Tsultrim. Tsewang, Rabgyal and Lobsang Chama and their family for all the support.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

PROBLEM STATEMENT

The circadian clock is an evolutionary conserved time-keeping mechanism that, through the regulation of rhythmic gene expression, coordinates the physiology of an organism with the daily environmental cycle. The regulatory network that links the clock to rhythmic transcription and that controls the phase of rhythmic gene expression, is not understood. This dissertation is a body of studies using the circadian model system *N. crassa* to map the clock output regulatory network.

BIOLOGICAL CLOCKS AND CIRCADIAN RHYTHMS

The daily 24-hour light dark cycle caused by earth's rotation on its axis influences every aspect of an organism from daily molecular oscillations, physiology and metabolism to behavior. Organisms have therefore evolved an internal time-keeping mechanism, known as the circadian clock, to anticipate environmental cycles and align biological processes with the cyclic variations in light, temperature, humidity and other environmental factors so that they can occur at the most appropriate times of the day for maximal output and performance. For example, in *N. crassa*, asexual spores (conidia)

are produced during the night when the damaging effects of UV-light are minimal and humidity low, thus reducing the chance of DNA damage and increasing the spread of the hydrophobic spores (Dunlap et al., 2004). Another example is that in certain species of cyanobacteria, two necessary but incompatible biological processes, -oxygenic photosynthesis and oxygen-sensitive nitrogen fixation-, are segregated in time through the action of the clock (Berman-Frank et al., 2003). The temporal coordination between cellular rhythms and the environmental cycle confers adaptive advantages. This was demonstrated through resonance studies in cyanobacteria, assessing reproductive fitness where cyanobacteria with a circadian period matching the environmental cycle outcompeted those that are not synchronous with the imposed cycle (Ouyang et al., 1998). Similarly, there is mounting evidence to suggest that in humans, chronic exposure to shift work schedules, which causes misalignment of clocks with the day-night cycle, can increase the risk of cancer, obesity, cardiovascular disease, and other metabolic disorders (Hsu et al., 2013; Ukai-Tadenuma et al., 2008) and through clock regulation of metabolism, drugs are more effective or more toxic depending on the time of administration (Ptacek et al., 2007).

The first documented study of the internal timing mechanism was made in the 18th century by a French astronomer, Jean Jacques d'Ortus de Marian who noticed that rhythmic opening and closing of leaves of a perennial herb, *Mimosa pudica* persisted in the dark under constant conditions. Then, Augustin de Cadolle observed that the leaves opened and closed every 22-23 h, not exactly 24 h in constant conditions, strongly

indicating the presence of endogenous clocks (McClung, 2006). Over the years, circadian clocks (derived from Latin *circa diem* meaning 'about a day') have been identified across all phyla, ranging from bacteria to humans. The canonical properties of circadian rhythms were defined by the works of Pittendrigh and Aschoff (Pittendrigh, 1960): 1) circadian rhythms are endogenous and self-sustaining oscillation with a period of approximately 24 h that persist in constant conditions (constant light and temperature), 2) circadian rhythms are temperature compensated, meaning that the period of circadian rhythms remains stable over a wide range of physiological temperatures. This is unlike most biochemical reactions that accelerate with rise in temperature. This compensation also extends to nutrition, and therefore is sometimes called 'noise compensation', and 3) circadian rhythms are entrainable-the clock (circadian oscillator) that generates the rhythms can be reset or phase-shifted by environmental cues, and be synchronized to the day night cycle. Light and temperature are the two major entraining signals (Dunlap and Loros, 2004) however other inputs such as feeding (Damiola et al., 2000; Hara et al., 2001; Stokkan et al., 2001), and social interactions (Levine et al., 2002), have been shown to entrain the clock.

While the individual genes and proteins that make up the clock are not very well conserved among different organisms, pointing to their independent evolutionary origin (Tauber et al., 2004), the overall network architecture and underlying clock mechanisms are highly conserved (Bell-Pedersen et al., 2005; Young and Kay, 2001). This has led to a simplified view of the circadian clock, consisting of three key components, an input

pathway, an oscillator and an output pathway (Figure 1-1). However it is important to note they are not mutually exclusive; there are examples where an output gene can feed back into the input, and/or affect the oscillator. For example, the *vvd* (Vivid) gene is an output gene in *N. crassa* that feeds back to regulate light input into the oscillator (Heintzen et al., 2001).

At the core of every circadian clock is an oscillator or a network of coupled oscillators that keep time, even in the absence of any time cues from the environment. All known circadian oscillators are based on an autoregulatory, molecular feedback loop, and consist of positive and negative elements that control rhythmic gene expression at the levels of transcription and translation (Bell-Pedersen et al., 2005). In this loop, the positive elements activate the expression of the clock genes that encode for negative elements. As the concentration of the negative elements rise and accumulate over time, they reach sufficient levels to inhibit expression and function of the positive elements. Phosphorylation of the negative elements lead to their degradation, allowing the positive elements to start the cycle all over again (Dunlap, 1999). In addition, the feedback loop also controls rhythmic transcription of other feedback loop components, which lends robustness to the clock system, and the positive elements signal out from the clock to regulate rhythms in gene expression (Young and Kay, 2001). Input pathways sense environmental cues such as light and temperature and synchronize the molecular oscillator to the 24-h light-dark cycle, whereas the output pathways convey temporal information from the oscillator to regulate rhythmic gene expression and ultimately

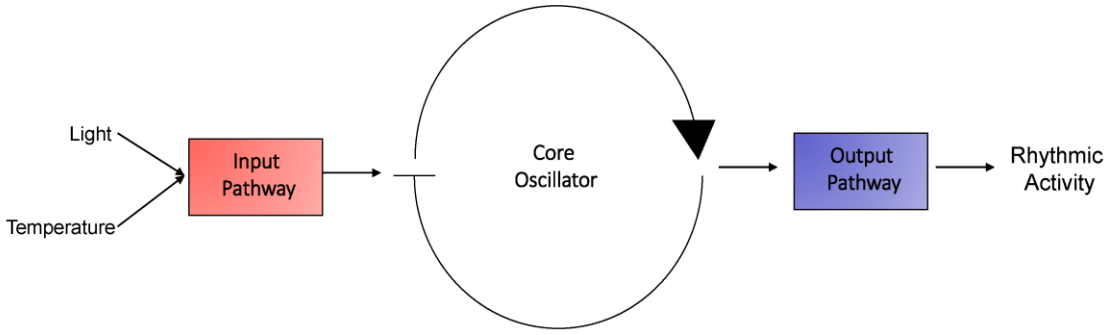


Figure 1-1. Schematic representation of the circadian clock system. Refer to the text for more details.

generate overt rhythms in development, physiology and behavior, such as the conidiation rhythm in *N. crassa* and sleep-wake cycles in mammals.

N. CRASSA AS A MODEL SYSTEM TO STUDY CIRCADIAN RHYTHMS

The filamentous fungus *N. crassa* was the pioneer model organism studied to demonstrate that fundamental life processes are under genetic control. George Beadle and Edward Tatum established the 'one gene-one enzyme hypothesis' using metabolic mutants in *N. crassa* (Beadle and Tatum, 1941). Since then *N. crassa* has proven to be an important model organism to understand eukaryotic biology, including circadian biology. The first genetic screens for mutations that affect circadian rhythmicity were done around the same time in *Drosophila melanogaster* (Konopka and Benzer, 1971) and *N. crassa* (Feldman and Hoyle, 1973). In *N. crassa*, mutants that altered the period of rhythmic conidiation ranging from 16.5 to 29 h were identified. This work ushered in the modern era of molecular studies on the circadian clock (Feldman and Hoyle, 1973). Interestingly, mutations that shortened and lengthened period, or abolished rhythmicity all mapped to the same genetic locus, called *frequency* (*frq*). Subsequent cloning of the *frq* gene established *N. crassa* as a major experimental system for the analysis of the molecular mechanisms of the circadian rhythm (McClung et al., 1989).

N. crassa continues to be a major model for understanding eukaryotic circadian clocks because of its relatively simple lifestyle, excellent genetics, fully sequenced genome

(Galagan et al., 2003), and availability of a whole genome knockout library (Colot et al., 2006). Moreover, *N. crassa* possesses a robust and easily monitored circadian rhythm in conidiation, observed as a series of dense, orange conidial bands followed by undifferentiated vegetative hyphae on a solid agar media (Sargent et al., 1966). The conidiation rhythm is typically assayed using a hollow glass tubes (called race tubes). When grown in constant conditions, the fungus shows a free-running circadian rhythm in conidia formation of about 22.5 h. The linear growth rate across the tube is relatively constant; thus, the position of the conidial bands relative to the 24-h growth marks allows the determination of period and phase of the rhythm. This simple way of analyzing the circadian rhythm and its formal properties contributed significantly to our current understanding of the eukaryotic circadian clock.

An array of genetic, molecular and biochemical tools are available for work with *N. crassa*. For molecular analysis of circadian rhythms in gene expression or protein accumulation, *N. crassa* mycelial tissues are grown in liquid cultures in DD, and harvested every 4-h over at least 2 days (Perlman et al., 1981). The use of codon-optimized firefly luciferase reporters allows for automated monitoring of mRNA or protein rhythms in living cells in real time (Gooch et al., 2008; Morgan et al., 2003). Recently, genome-wide studies coupled to high-throughput sequencing such as ChIP-seq and RNA-seq have been successfully used in *N. crassa* to gain insights into light and clock signaling pathways (Hurley et al., 2014; Sancar et al., 2011; Smith et al., 2010; Wu et al., 2014).

THE *N. CRASSA* FRQ/WCC OSCILLATOR

The *N. crassa* circadian FRQ/WCC oscillator, similar to oscillators in other model systems, is based on an autoregulatory transcriptional-translational feedback loop (TTFL). During the late subjective night in constant conditions (DD and 25°C), the positive components of the FRQ/WCC oscillator, two GATA-type transcriptional activators, White Collar-1 (WC-1) and White Collar-2 (WC-2), interact through their PAS (PER-ARNT-SIM) domains to form the White Collar Complex (WCC) (Ballario et al., 1996; Cheng et al., 2002b; Denault et al., 2001; Linden and Macino, 1997; Talora et al., 1999). The WCC binds to the distal light response elements (dLRE) or the Clock-box (C-box) elements in the *frq* promoter to activate *frq* transcription, which peaks in the early subjective morning (Froehlich et al., 2002; Froehlich et al., 2003). Subsequently, FRQ protein reaches its peak in the late subjective day, with 4-5 h delay relative to *frq* mRNA levels (Garceau et al., 1997). FRQ protein contains a nuclear localization signal, and is shuttled into the nucleus (Luo et al., 1998). In the nucleus, FRQ binds to FRH (FRQ-interacting RNA helicase), which is essential for FRQ stability (Cheng et al., 2005; Cheng et al., 2001a; Guo et al., 2010) to form the FRQ-FRH complex (FFC), and then recruits several kinases including Casein kinases 1 and 2 (CK1 and CK2) to phosphorylate the WC proteins (He and Liu, 2005b; He et al., 2005; Schafmeier et al., 2005). Phosphorylation of the WCC leads to a decrease in its DNA-binding activity, and therefore a decrease of *frq* transcription. After its synthesis, FRQ undergoes progressive

phosphorylation by several protein kinases (Baker et al., 2009; Heintzen and Liu, 2007; Tang et al., 2009), and once fully phosphorylated, it interacts with FWD-1, an F-box/WD-40- repeat containing protein and substrate recruiting subunit of an SCF-type ubiquitin ligase complex, that facilitates FRQ ubiquitination and degradation by the proteasome (He and Liu, 2005a; Liu, 2005).

The transcriptionally active unphosphorylated WCC found in the nucleus is not stable, whereas phosphorylated WCC is inactive but stable and accumulates in the cytosol (Schafmeier et al., 2008; Schafmeier et al., 2005). In the cytosol, the WCC undergoes dephosphorylation by phosphatases, PP1 and PP2A. This results in activation and translocation of the WCC into the nucleus, thereby initiating a new cycle to start the next morning (Yang et al., 2004). These events form a negative feedback loop that takes approximately 24 h to complete, and is considered to be the driving force of the circadian oscillations in *N. crassa*. *frq* mRNA and FRQ protein exhibit robust rhythms with a period close to 22 h, which phenotypically correlates with rhythmic conidiation on race tubes. Mutations in any of the clock genes can either lead to period length changes or complete loss of rhythmicity. In addition, certain mutations in *frq* and *wc-1* affect temperature compensation, and mutations in *wc-1* or *wc-2* abolish light resetting of the clock (Aronson et al., 1994a; Liu and Bell-Pedersen, 2006). WC-1 is also a part of a positive feedback loop, whereby FRQ regulates WC-1 levels post-transcriptionally, such that in the absence of FRQ, WC-1 levels are reduced. This reduction is attributed to

FRQ-promoted phosphorylation of WC-1, which stabilizes the protein in an inactive form (Lee et al., 2000).

Blue light is an important cue for entrainment of the circadian clock in *N. crassa* to the environment. All of the known light responses in *N. crassa* are blue light-dependent, and include the development of asexual spores and sexual structures, bending of perithecial beaks, biosynthesis of carotenoids and entrainment of the clock (Chen et al., 2010b). Light signals are received directly by the FRQ/WCC oscillator through the blue light photoreceptor WC-1. In response to light, the flavin adenine dinucleotide (FAD) chromophore binds to the LOV (for light-, oxygen- and voltage-sensing) domain of WC-1 and activates it by an unknown mechanism (Froehlich et al., 2002). Light-activated WCC binds to the *frq* promoter and activates *frq* transcription independent of the time of the day (Crosthwaite et al., 1997). The induced change in the levels of *frq* mRNA in response to light leads to a phase shift of the clock depending on when the light pulse is given. If a light pulse is given in the late subjective night when *frq* mRNA levels are rising, the abrupt increase in *frq* mRNA transcription will result in a phase advance. Alternatively, a light pulse in the subjective evening when *frq* levels are decreasing will result in a phase delay (Crosthwaite et al., 1995). This provides a molecular mechanism for light resetting and entrainment in *N. crassa*. Light treatment of cells induces binding of WCC to light-responsive elements (LREs) in the promoters of hundreds of light responsive genes besides *frq* to activate transcription, suggesting a common molecular basis for light responses in *N. crassa* (Smith et al., 2010).

Temperature is another environmental cue that can reset the clock. While light acts directly on the WCC, the effect of temperature seems to be mainly through translational control of FRQ. FRQ protein oscillates around a higher mean level at high temperature. As a consequence, if cells are rapidly shifted from lower (20°C) to higher temperature (28°C), regardless of the time of the day the culture is in, the existing FRQ levels will now match the lowest FRQ levels expected at the new temperature, which corresponds to those observed at dawn, thus resetting the clock to dawn. Alternatively, if the culture goes rapidly from high (28°C) to a lower temperature (20°C), the existing FRQ levels will match the highest point of expression at the lower temperature, which corresponds to those observed at dusk, and the clock will be reset to dusk (Liu et al., 1998).

Output from the FRQ/WCC oscillator drives daily rhythms in gene expression, and ultimately controls overt rhythmicity, including rhythms in metabolic processes and the development of asexual spores (Dunlap and Loros, 2006). The pervasive control by the circadian clock of essentially all known biological processes in *N. crassa* is well documented (Hurley et al., 2014; Vitalini et al., 2006). More details on major advances and future directions in connecting the components of the oscillator with the downstream output pathways will be presented in the following section.

CIRCADIAN OUTPUT PATHWAYS IN *N. CRASSA* AND OTHER ORGANISMS

Initial studies to map the circadian output pathways worked backwards from a terminal clock-controlled gene (ccg) to the clock, and identify the upstream *cis*-regulatory elements, conferring rhythmicity (Bell-Pedersen et al., 1996a). Nuclear run on experiments showed that *ccg-1* and *ccg-2* are regulated at the level of transcription and the promoter elements conferring rhythmicity have been defined (Bell-Pedersen et al., 1996a; Bell-Pedersen et al., 2001). However, this approach did not succeed in linking the known clock components in a defined pathway to control ccgs. An alternative approach, involved a genetic selection screen for mutations in the output pathway regulating *ccg-1* (Vitalini et al., 2004). This work uncovered clock regulation of conserved MAPK pathways, which in turn, control sets of genes regulated by the signaling pathway (Bennett et al., 2013; Lamb et al., 2011; Vitalini et al., 2007). This discovery provided a direct connection from the FRQ/WCC oscillator to a ccg, and suggested a hierarchy in which the clock could control signaling pathway elements and thereby put all of downstream signaling pathways under clock control.

The output pathways remain the least understood aspect of the circadian clock system, despite substantial progress made in *N. crassa* in connecting the clock to MAPK signaling (Bennett et al., 2013; Lamb et al., 2011; Vitalini et al., 2007), and in cataloging ccgs in different organisms (Duffield, 2003; Hurley et al., 2014). The challenge in

dissecting the output gene network arises due to the inherent complexities, which we are beginning to appreciate, and therefore requires a systems level approach that can integrate vast amounts of experimental data, model this data reiteratively and produce meaningful information. One ultimate goal is to define the design principles required for rhythmicity and phase determination.

In order to begin to understand the molecular basis underlying regulation of circadian output pathways, studies were carried out to identify ccgs in different organisms. With the help of microarray-based studies in the past and now deep-sequencing, we know that clock control of gene expression is not limited to few genes, but is more ubiquitous, with a significant proportion of the transcriptome in eukaryotes under clock control, in all organisms examined (Duffield, 2003; Koike et al., 2012; Menet et al., 2012; Nagel and Kay, 2012; Vitalini et al., 2006).

The first systematic screen for ccgs was done in *N. crassa* using subtractive hybridization and two morning-specific ccgs, *ccg-1* and *ccg-2* were identified (Loros et al., 1989), which were subsequently shown to be regulated at the transcriptional level, suggesting the presence of *cis*-acting elements conferring rhythms (Bell-Pedersen et al., 1996a; Bell-Pedersen et al., 2001; Loros and Dunlap, 1991). Subsequently, screens for ccgs in *N. crassa* were performed using differential hybridization of time-of-day specific cDNA libraries (Bell-Pedersen et al., 1996c), high-throughput cDNA sequencing (Zhu et al., 2001) and DNA microarrays ((Dong et al., 2008; Lewis et al., 2002; Nowrousian et

al., 2003). Together, these studies suggested that as much as 30% of the *N. crassa* transcriptome may be under clock control, and that these ccgs encode proteins that are involved in a wide range of biological processes such as DNA damage repair, development, metabolism, cell cycle and stress responses (Hurley et al., 2014; Vitalini et al., 2006).

In cyanobacterium *Synechococcus elongatus*, using promoter: luciferase fusions, it was found that nearly all genes in the genome are under the control of the clock, and that 80% of the circadian-regulated genes show peak expression during subjective dusk (Liu et al., 1995). Later studies supported the hypothesis that the global regulation of gene expression by the cyanobacterial clock is mediated through clock regulation of chromosomal compaction and supercoiling (Woelfle and Johnson, 2006). In *Arabidopsis thaliana*, only 2-6% of the genome was shown to be under clock regulation at the level of transcript abundance (Harmer et al., 2000). Surprisingly, when changes in transcriptional rate were measured by promoter-trap experiment, the influence of clock was far more prevalent with ~36% of promoters found to be under circadian regulation (Michael and McClung, 2003).

In mice, clock control of gene expression has been studied in different tissues (Akhtar et al., 2002; McCarthy et al., 2007; Panda et al., 2002; Storch et al., 2002; Ueda et al., 2002; Zvonic et al., 2006). The majority of these rhythmic mRNA outputs are tissue-specific (Panda et al., 2002; Storch et al., 2002; Ueda et al., 2002). Similar results were

also obtained from comparison between rhythmic mRNAs identified in the Drosophila head and body (Ceriani et al., 2002). The tissue-specificity in ccg expression observed in Drosophila and mouse highlights the fact that although the fundamental mechanism driving rhythmic expression is the same across all cell types, only a specific subset of ccgs that are relevant to that cell's unique function are expressed, and not others. It was shown in Drosophila that tissue-specific transcription factors (TFs) function synergistically with the core clock TFs CLOCK/CYCLE (dCLK/CYC) to achieve tissue-specificity in ccg expression (Meireles-Filho et al., 2014).

Despite the ability to monitor gene expression on a large scale, microarrays had limitations; due to a lack of sensitivity, they were unable to detect poorly expressed genes, and low amplitude rhythms in mRNA levels. Therefore, microarrays underestimated the number of ccgs. In addition, the choice of probe sets and mistakes in gene annotations for some of the probes were also cause of concerns. Recent advances in high-throughput sequencing have led to a better resolution of the extent of circadian control in different organisms. Microarrays have been supplanted by RNA-seq, which allows one to quantify all known transcripts in a genome. RNA-seq has been employed in *N. crassa* (Hurley et. al., 2014), flies (Hughes et al., 2012; Rodriguez et al., 2013) and mice (Yoshitane et al., 2014; Koike et al., 2012; Vollmers et al, 2012; Menet et al., 2012) to monitor circadian gene expression with better coverage and resolution than was previously possible with microarray. In *N. crassa*, approximately 40% of the

transcriptome was found to have rhythmic abundance based on rhythmic RNA-seq (Hurley et al., 2014).

Depending on the organism and tissue, anywhere between 10-40% of the eukaryotic genome is under clock control at the level of transcription. These studies also revealed that the temporal expression profiles exhibited by these ccgs are distributed at all times of the circadian cycle, implicating the action of a network of TFs determining phase. Consistent with this idea, TFs were significantly enriched among ccgs in the RNA-seq studies discussed above. In summary, transcriptional regulation is a defining architecture for circadian output pathways. However, we lack basic information on how the clock coordinates rhythmic gene expression to particular times of the day. Research in *N. crassa* has begun to shed light on these questions, including my studies described in this thesis.

TRANSCRIPTIONAL CONTROL OF LIGHT AND CIRCADIAN GENE EXPRESSION

Based on the core clock mechanism, our overall hypothesis that the core clock TFs WCC would function to control rhythms in key output molecules that in turn would regulate downstream processes. To test this hypothesis, we started with identifying the direct targets of the core clock component and TF WCC by performing WC-2 ChIP-seq after a light pulse (See Chapter III). Our prediction was that the first-tier TFs driven by the core

clock complex would regulate downstream TFs and so on, thereby amplifying the number of ccgs and imparting phase specificity. As predicted, we found that WCC binds to the promoters of about 200 genes, a number of which are clock- and light-regulated, and genes encoding TFs were significantly enriched as WCC targets. Our data, therefore suggested that the WCC sits at the top of the hierarchy of TFs driving rhythmic gene expression and light responses. In *N. crassa*, light and clock signaling pathways are intertwined and share WC-1, which acts both as a core clock TF and a blue light photoreceptor. Upon light exposure, the WCC binds to light response elements (LREs) in the promoters of target light-responsive genes. In order to understand transcriptional response to light in *N. crassa* at a genome-wide level, we also performed RNA-seq in wild type *N. crassa* cells in the dark and after light treatment. At least 31% of expressed genes in *N. crassa* are regulated by light. Together, our data support expression studies suggesting a hierarchical network for light signaling pathways, in which the light-activated WCC controls expression of early light induced genes including some TFs. These early light induced TFs would in turn control the expression of late light responsive genes. Consistent with this idea, the *sub-1* gene is an early light induced gene whose promoter is bound by the WCC and deletion of *sub-1* alters light responses for many late-light induced genes (Chen et al., 2009). The findings of that study are presented in greater detail as Chapter II. Our studies in Chapter II and III have resulted in a wealth of genome-associated information, which then led to studies presented in Chapter IV to understand direct mechanistic connections from the core oscillator to clock-controlled processes.

Following our study presented in Chapter III, there have been similar studies in other circadian model organisms, using ChIP-seq to identify genes directly bound by the core clock TFs, PRR5 and RVE8 in Arabidopsis (Hsu et al., 2013; Nakamichi et al., 2012), dCLK/CYC in Drosophila (Abruzzi et al., 2011) and CLOCK/BMAL1 in mice (Rey et al., 2011). All of these studies similarly found TFs as the most enriched class of direct targets, and support our initial findings in *N. crassa* (Smith et al., 2010), indicating that the rhythmic transcriptome is achieved through a hierarchical transcriptional cascade, in which expression of the first-tier TFs are driven by the core TF complex. For transcriptional activators, it is thought that peak in expression would occur at the time of peak activity of the clock component. TF targets of the first-tier TFs then drive expression of ccgs that peak at a later phase and so on. In addition, first-tier TFs can act as repressors, in which case their target ccgs will peak in expression in opposite phase to the time of peak activity of the clock component. This mode of action was exemplified in *N. crassa* for CSP-1 (conidial separation-1), a first-tier TF, which through its repressor activity causes evening-specific expression of a subset of its direct target genes (Sancar et al., 2011). In this way, it is thought that a small number of phase-specific TFs can potentially achieve any peak expression phase (Ukai-Tadenuma et al., 2008). However, TFs often control gene expression in a combinatorial fashion rather than in isolation. In support of this idea for circadian output networks, dCLK/CYC functions in coordination with tissue-specific TFs bound to target genes to control tissue-specific ccgs (Meireles-Filho et al., 2014). Furthermore, in mammals, a significant number of BMAL1 direct target genes peak much later than the time of maximal Bmal1 binding

(Rey et al., 2011), which support more complex regulatory network, similar to that in *N. crassa*.

Consistent with the idea that first-tier TFs connect the core oscillator to the output pathway, many of the first-tier TFs in *N. crassa* are rhythmically expressed (Smith et al., 2010). To further characterize the transcriptional network associated with light and clock regulation of gene expression in *N. crassa*, I focused on ADV-1 (arrested development-1), one of the 24 first-tier TFs identified as direct WCC target. ADV-1 is a zinc finger TF necessary for vegetative and sexual development in *N. crassa*. ADV-1 knockouts are defective in protoperithecia formation, which is required for the female sexual structure called the perithecia to form (Colot et al., 2006). In addition, an *adv-1* deletion strain also lacks the ability to undergo cell fusion during conidial germination, and later in the mature colony during hyphal growth (Fu et al., 2011). Consistent with *adv-1* being a direct WCC target, the gene is both light- and clock-regulated, suggesting a role for ADV-1 in connecting downstream developmental processes to circadian clock and light signaling pathways (Smith et al., 2010). To understand the role of ADV-1 in regulating light and clock output pathways I used ChIP-seq in WT cells and RNA-seq in WT versus ADV-1 cells, to identify direct and indirect target genes. Consistent with the *adv-1* KO phenotype, ADV-1 direct targets are enriched for genes involved in sexual and asexual development, cell fusion and metabolism. Our data also suggest that the circadian clock, through rhythmic activation of ADV-1, controls the timing of cell fusion. Furthermore, while light responses of ADV-1 direct target genes were altered as expected in *adv-1* KO

cells, the temporal aspects were not. These data suggested that unlike the flat hierarchical network observed for light signaling, the circadian TF network is more complex, with feed forward and feedback loops contributing to rhythmic transcription. Our preliminary data indicate that in addition to WCC, several first-tier TFs bind to the *adv-1* promoter and that ADV-1 feeds back to bind the promoters of these same TFs. These same TFs also bind and potentially co-regulate each other, and the direct targets of ADV-1. Together, these data suggest a complex regulatory network linking WCC to ADV-1 and the downstream output genes involved in development and metabolism. Deciphering these mechanisms will require an integrative approach combining computational and experimental studies that allows for an iterative process of model-driven predictions, experimental validation and refinement of the model. In Chapter 5, I will discuss the future directions to map the complex clock regulatory network using systems biology.

BEYOND TRANSCRIPTIONAL REGULATION OF CIRCADIAN GENE EXPRESSION

Genome-wide studies have also suggested that along with transcriptional regulation, there are other gene-regulatory mechanisms, including post-transcriptional, translational and post-translational control, which can influence dynamics governing mRNA rhythms and protein rhythms, including their phase. For example, by analyzing circadian RNA-seq data with sequenced reads mapped not just to exons, but also to introns (to measure

pre-mRNA expression), it was shown that 80% of the ccgs mRNAs did not undergo rhythmic *de novo* transcription, suggesting a role for circadian post-transcriptional regulation (Koike et al., 2012). Similar findings were observed in a study using a technique called Nascent-seq, which measures nascent RNA levels, and therefore provides a measure of *de novo* transcription. Here, more than 50% of the transcripts that are rhythmic in *de novo* transcription did not show rhythmic mRNA expression (Menet et al., 2012). The poly (A) tail length was shown to be under circadian control across the whole genome in mouse liver, which can influence translation and mRNA stability (Kojima et al., 2012). Most of the rhythmically polyadenylated genes showed rhythmic transcription based on rhythmic pre-mRNA expression. More interestingly, for some genes, rhythms in polyadenylation were observed despite the absence of rhythms in pre-mRNA or mRNA levels and this correlated with oscillations in protein expression. Importantly, the rhythmic change in the poly (A) tail length, not the mRNA levels, correlated with the rhythm in protein levels (Kojima et al., 2012). Furthermore, ribosome profiling, a technique used to extract actively translated RNA from polyribosome fractions, discovered that ~2% of the expressed genes are rhythmically translated independent of any circadian changes in the mRNA levels (Jouffe et al., 2013). Therefore, it is of interest to integrate these different regulatory processes that influence rhythmic genes expression with the transcriptome data. In summary, the clock-regulated output network that controls gene expression, and ultimately rhythms in physiology and behavior is very complex. Also, given the fact that circadian regulation of gene

expression takes place at different levels, it is likely that number of ccgs is grossly underestimated.

SIGNIFICANCE OF UNDERSTANDING CLOCK REGULATORY NETWORK

Given that circadian clocks in organisms, including humans, control most physiological processes, it is not surprising to find that disruption of the clock leads to impaired health and susceptibility to different pathological disorders. Understanding how the clock controls downstream processes will likely have a profound effect on human health, since several studies have demonstrated a link between disruption of the clock and disease, including cancer, asthma, metabolic syndrome, coronary heart disease, and learning disorders (Ptacek et al., 2007). However, this link is not well understood. Studies of ccgs have revealed that they include genes involved in the cell cycle, DNA repair, and metabolism in a wide range of organisms. These data provide a rationale for the diversity of circadian-related diseases. Thus, an understanding of the mechanisms by which the clock regulates ccgs and overt rhythmicity will likely help decipher molecular pathogenesis of metabolic disorders, and provide new approaches to treatments for circadian disorders. Furthermore, phase relationships are key, for example if the phase between cycles of DNA damage and repair are not aligned, it can lead to instability in genome. Similarly, misalignment can superimpose incompatible biochemical processes such as the oxidative and reductive phases of the metabolic cycle. Because circadian

oscillators, and often the output genes, are conserved, *N. crassa* provides an ideal model system for uncovering these mechanisms.

CHAPTER II

GENOME-WIDE CHARACTERIZATION OF LIGHT-REGULATED GENES IN *NEUROSPORA CRASSA*[*]

INTRODUCTION

The ability to sense and respond to light is critical for the survival of most organisms. In *N. crassa*, one of the best studied model systems for light responses, blue light controls many physiological processes including the synthesis of protective photopigments (carotenoids), asexual and sexual spore formation, the direction of sexual spore release, and entrainment and resetting of the circadian clock (Corrochano, 2007, 2011; Herrera-Estrella and Horwitz, 2007; Hsu et al., 2013; Linden and Macino, 1997; Nakamichi et al., 2012; Ptacek et al., 2007; Purschwitz et al., 2006). Underlying these light-regulated physiological processes is the transcriptional control of gene expression.

N. crassa perceives blue light through the photoreceptor and GATA zinc finger transcription factor encoded by *white collar-1* (*wc-1*, NCU02356); the chromophore flavin adenine dinucleotide (FAD) is bound by WC-1 and undergoes a transient covalent addition to the protein upon illumination (Ballario et al., 1996; Chen and Loros, 2009; Froehlich et al., 2002; He et al., 2002; Linden et al., 1997b; Sargent and Briggs, 1967).

---

[*]Reprinted with permission from "Genome-wide characterization of Light-regulated Genes in *Neurospora crassa*" by Wu C, Yang F, Smith KM, Peterson M, Dekhang R, *et al*. (2014) G3, 4 (9), 1731-1745, reproduced under a Creative Commons Attribution Unported License.

WC-1 interacts with the zinc-finger protein WC-2 (encoded by NCU00902) through its Per-Arnt-Sim (PAS) domain to form a heterodimeric transcription factor, the White Collar Complex (WCC) (Ballario et al., 1996; Cheng et al., 2002b; Crosthwaite et al., 1997; Denault et al., 2001). Upon light exposure, the WCC can bind to light-responsive elements (LREs) in the promoters of many light-responsive genes to activate their transcription (Cheng et al., 2003; Froehlich et al., 2002; He and Liu, 2005b; Olmedo et al., 2010a; Smith et al., 2010). In addition to WC-1, the blue light photoreceptor VVD, encoded by *vivid* (*vvd*, NCU03967), which is strongly light-induced under the control of the WCC, plays a key role in photoadaptation by desensitizing the WCC-mediated light response, and thereby reducing the transcription of WCC target genes (Gin et al., 2013; Heintzen et al., 2001; Hunt et al., 2010; Malzahn et al., 2010; Ptacek et al., 2007; Schwerdtfeger and Linden, 2000, 2003; Shrode et al., 2001). The *N. crassa* genome sequence revealed several additional putative photoreceptors, including a cryptochrome (*cry*, NCU00582), two phytochrome homologs (NCU04834 and NCU05790), and an opsin (*nop-1*, NCU10055) (Bieszke et al., 1999b; Galagan et al., 2003). However, the effects of deletion of these candidate photoreceptors on physiology and light-controlled gene expression are subtle (Chen et al., 2009; Froehlich et al., 2010; Froehlich et al., 2005; Olmedo et al., 2010b), consistent with a primary role for the WCC in *N. crassa* light signaling cascades (Chen et al., 2009).

To better understand gene regulation in *N. crassa* in response to light, several studies have identified light-controlled genes (Chen et al., 2009; Dong et al., 2008; Lewis et al.,

2002; Smith et al., 2010). Estimates of the number of light-responsive genes based on microarray analyses have varied widely, ranging from 3-14 % of the genome (Chen et al., 2009; Dong et al., 2008; Lewis et al., 2002), reflecting the use of different microarray platforms, strains, culture conditions and statistical cut-offs. Light-induced genes can be roughly classified as early-response (with peak mRNA levels between 15-45 min of light treatment), and late-response (with peak mRNA levels between 45-90 min of light treatment), consistent with earlier patterns detected from analyses of individual genes (Linden et al., 1997a).

More recently ChIP-seq was used to identify ~400 direct targets of light-activated WCC (Smith et al., 2010); genes encoding transcription factors (TFs) were overrepresented. This was consistent with expression studies that uncovered a hierarchical network in which the light-activated WCC directly controls expression of early light-induced genes, including some of the TF genes identified as direct WCC targets (Chen et al., 2009). Early light-induced proteins in turn control the expression of late light-induced genes. Among the early light-induced TFs, SUB-1 (*submerged protoperithecia-1*, NCU01154), are necessary for light-induction of a large set of the late light-inducible genes (Chen et al., 2009), and CSP-1 (*conidial separation-1*, NCU02713) acted at a similar place as *sub-1* in the hierarchy, just below WC-1 (Chen and Loros, 2009; Sancar et al., 2011).

To better understand the organism's response to light and to comprehensively describe the light-controlled gene regulatory network, we used RNA-seq to identify *N. crassa*

transcripts whose levels are responsive to light. We compared RNA-seq data from dark-grown cells to that from cells exposed to light for between 15 and 240 min, and identified both known and novel light-induced genes. We also discovered many transcripts whose levels decreased in response to light. This latter class of transcripts is enriched for genes whose products function in ribosome biogenesis; this response to light has not been described previously.

RESULTS

RNA seq to identify light-regulated transcripts

We grew *N. crassa* in liquid cultures in the dark for 24 hr, and then exposed the cultures to light for 15, 60, 120, or 240 min. Capped and polyadenylated mRNA was purified from harvested cells, hexanucleotide-primed cDNA was produced and sequenced, and results of two independent biological replicates are reported here. The reproducibility between biological replicates was excellent for cultures grown in the dark or exposed to light for 15, 60 and 120 min ($R^2>0.95$); there was more divergence between the replicate cultures exposed to light for 240 min ($R^2=0.75$) (Figure 2-1). For determining expression levels of individual genes, and for statistical analyses, the data from the biological replicates were pooled and analyzed with CuffDiff 2.0 (Trapnell et al., 2013). The analyses of the combined data are given in Appendix F File S1; the analyses of each independent experiment are given in Appendix F File S2 and File S3. Based on the

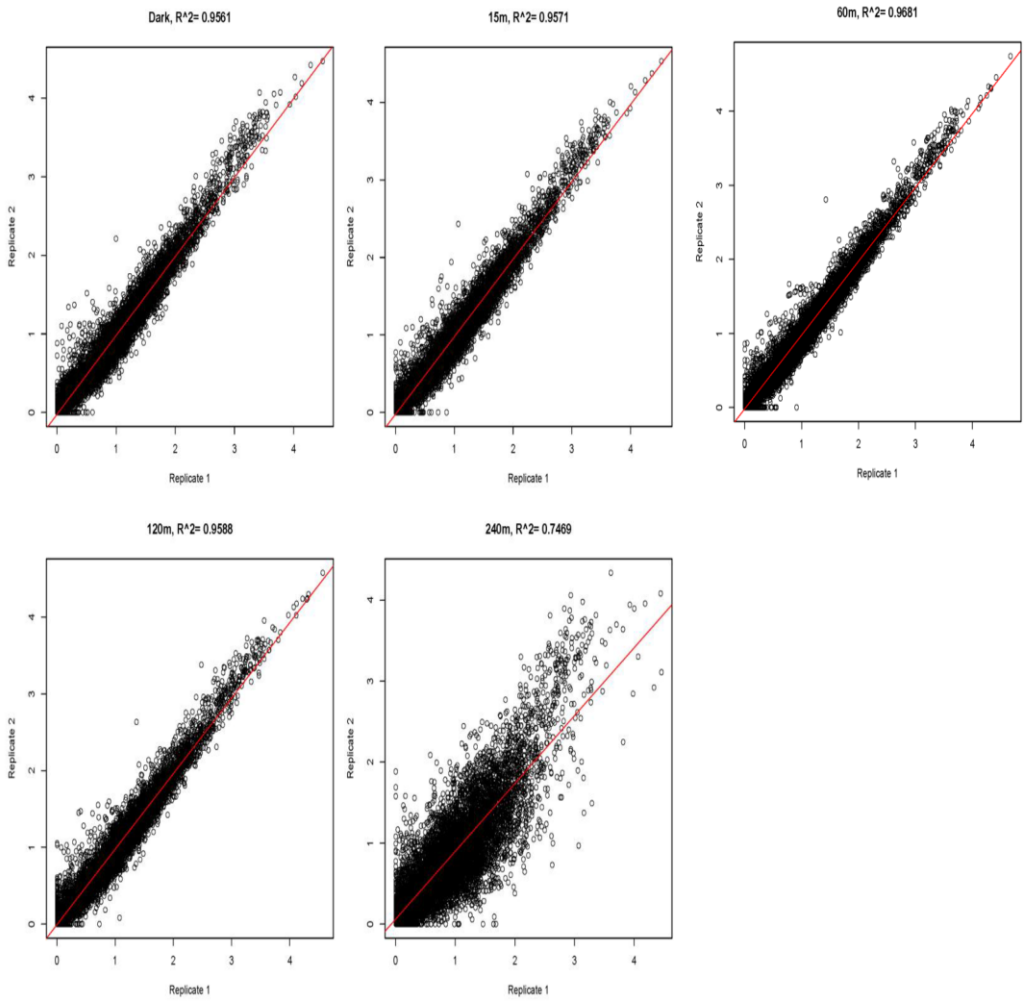


Figure 2-1. Comparison of RNA-seq replicate experiments. The FPKM for biological replicate 1 is plotted against biological replicate 2 for each gene, demonstrating strong correlation between replicate experiments at each time point. The correlation coefficient, R, is shown for each time point.

analyses of the combined data, 79% of predicted protein coding genes (7660 of 9728) were expressed with FPKM>1 (FPKM, fragments per kilobase per million reads) at one or more time points analyzed, and these were considered "expressed genes" under these growth conditions. Most expressed genes (55%) specified proteins that have known functions in *N. crassa* or other organisms or whose sequences are conserved in the fungi or other organisms. In contrast, only 16% (341 of 2068) of the poorly or non-expressed genes (FPKM<1 at all time points analyzed) specified known or conserved proteins. Importantly, among the poorly expressed genes that specified proteins with known functions were many that had roles in secondary metabolism or were associated with the sexual cycle (Appendix F File 2-S4), consistent with their low expression under the vegetative culture conditions used here.

Our observations of genes abundantly expressed in the dark (and also in the light) were generally consistent with prior work using microarrays (Kasuga et al., 2005). The most abundant transcripts in the cell were those implicated in thiamine biosynthesis (NCU06110 and NCU09345) (Cheah et al., 2007; Faou and Tropschug, 2003, 2004; McColl et al., 2003). Also present at high levels in these cultures (FPKM>400 or FPKM>1000) were transcripts encoding proteins involved in translation, including those specifying ribosomal proteins and translation factors, and those with roles in energy metabolism (Figure 2-2A and Appendix F File S5). Light has a major effect on the physiology of *N. crassa* and this is reflected in light-driven changes in the levels of many transcripts. Transcript levels changed at least twofold for at least one time point

with respect to the dark-grown sample in 24% of the predicted protein coding genes (2353 of 9728), or 31% of the expressed genes (2353/7660) (Appendix F File 2-S6). Among the transcripts whose levels increased in response to light, the major fraction has unclassified functions in the FunCat scheme (Ruepp et al., 2004); genes with roles in carbon metabolism and in stress responses are also among those that are significantly over-represented (Figure 2-2B and Appendix F File S7). Among the transcripts whose levels decreased in response to light, genes with roles in metabolism and biogenesis were significantly over-represented (Figure 2-2C and Appendix F File S8).

Cluster analysis of light-regulated transcripts

As noted above, a characteristic of the light response is that subsets of genes are regulated with different kinetics, and this suggested that cellular functions might be temporally coordinated. To set the stage for hierarchical clustering of genes by regulation, we identified a subset of transcripts (5%, 532 of 9728 genes) that demonstrated a change in level with a q-value of 0.2 or less in the combined analysis of the two independent experiments (Appendix F File 2-S9). The q value provides a measure of the false discovery rate, and 532 genes had $q \leq 0.2$, 392 genes had $q \leq 0.1$ and 300 genes had $q \leq 0.05$for at least one time point (Appendix F File 2-S9). In the aggregate this subset of 532 genes with $q \leq 0.2$ generally reflected the same functional categories as the larger set of 2352 genes with a 2-fold change in levels in response to light (compare

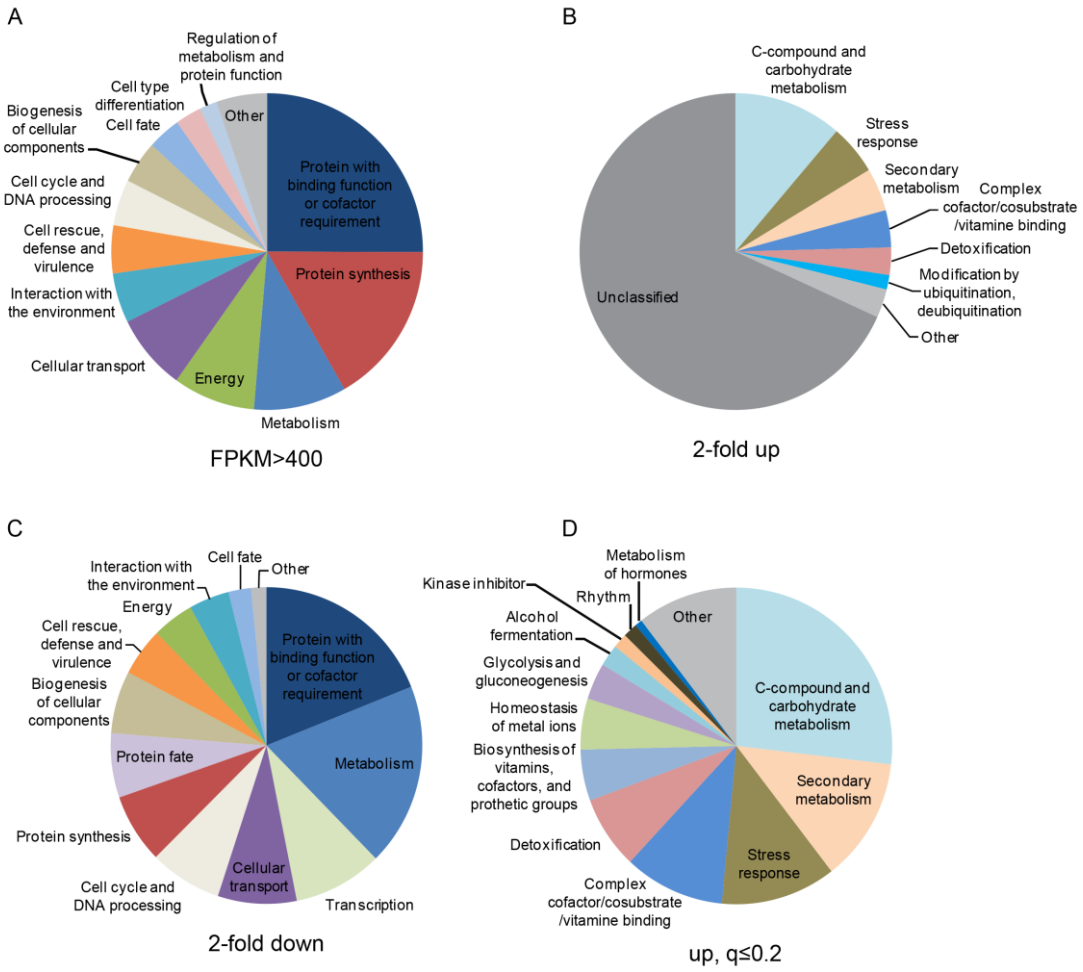


Figure 2-2. (A) FunCat Analyses of genes with FPKM>400. The complete analyses are given in Table S5. (B, C) FunCat analysis of the most up-regulated (B) and most down-regulated (C) subsets of the 2353 genes regulated 2-fold or more in response to light. The top 999 genes in each category were analyzed. The complete analyses are given in Appendix F File S7 and Appendix F File S8. (D) FunCat analyses of the up-regulated genes whose levels changed with a q value of 0.2 or less. The complete analyses are given in Appendix F File S9.

Figure 2-2B with Figure 2-2D and Cluster 5 in Figure 2-3C), except that unclassified genes are not a major enriched category among the genes whose expression level changes were significant at q≤0.2. Comparisons of genes based on gene ontology (GO) to examine enrichments in the two datasets (all genes that were 2-fold regulated by light and all genes whose light regulation met the q≤0.2 stringency requirement) showed overall similar correspondences in GO terms for the top classes (Appendix F File 2-S10), with unclassified genes missing in the q≤0.2 set.

Hierarchical clustering based on changes in expression of the 532 gene set relative to the dark sample (Figure 2-3A) revealed 310 transcripts up-regulated and, surprisingly, 222 transcripts down-regulated in response to light, the latter representing a class not heretofore described in *N. crassa*. We demarcated genes with common expression profiles into five main clusters based on analyses of overall differences in expression patterns in the tree-structure (Figure 2-3A and Figure 2-3B). The average transcript level of all genes in each cluster at each time point in the light relative to the average transcript level in the dark are given in Figure 3B. FunCat (Ruepp et al., 2004) and GO analyses of the genes in each cluster identified significantly enriched functional gene categories (Figure 2-3C and Appendix F File S11). Cluster 1 includes transcripts with rapid, high level, sustained light responses. This cluster was significantly enriched for genes involved in carotenoid biosynthesis, such as *albino-1* (*al-1*) (NCU00552) and *al-2* (NCU00585), and blue light photoresponses, including the photoreceptor genes *vvd* and

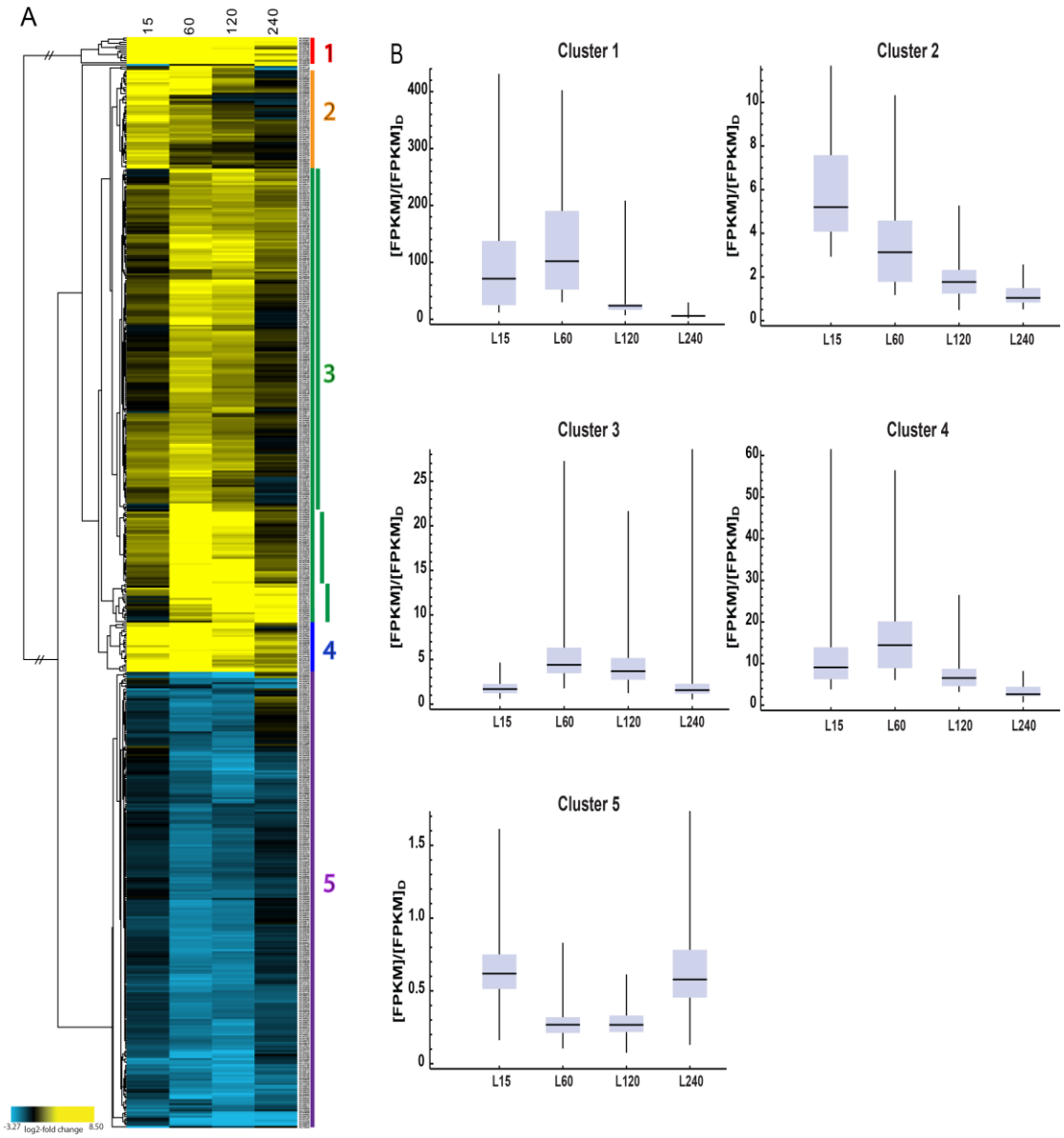


Figure 2-3. Light responses of selected transcripts. (A) Unsupervised hierarchical clustering of light-inducible responses identifies early and late light-responsive genes for genes whose transcript levels are different with q≤ 0.2 between cells grown in the dark and exposed to light for 15, 60, 120 or 240 min. Yellow: up-regulated; blue: down-regulated. Five major clusters determined by visual analysis of the tree-structure are indicated; cluster 3 was further divided into three sub-clusters based on similarities in expression-patterns within this relatively large cluster. (B) Expression-changes for transcripts in each of the five major clusters. Values for each time-point (L15, L60, L120 and L240) are normalized to expression in the dark. The horizontal black bar is the median, the box top and bottom are the 75% and 25% quantiles, and the whiskers extend to the maximum and minimum values. (C) FunCat enrichment analyses of genes in the five major clusters shown in panel A. The charts show major enriched classes; the complete data are given in Appendix F File S11.

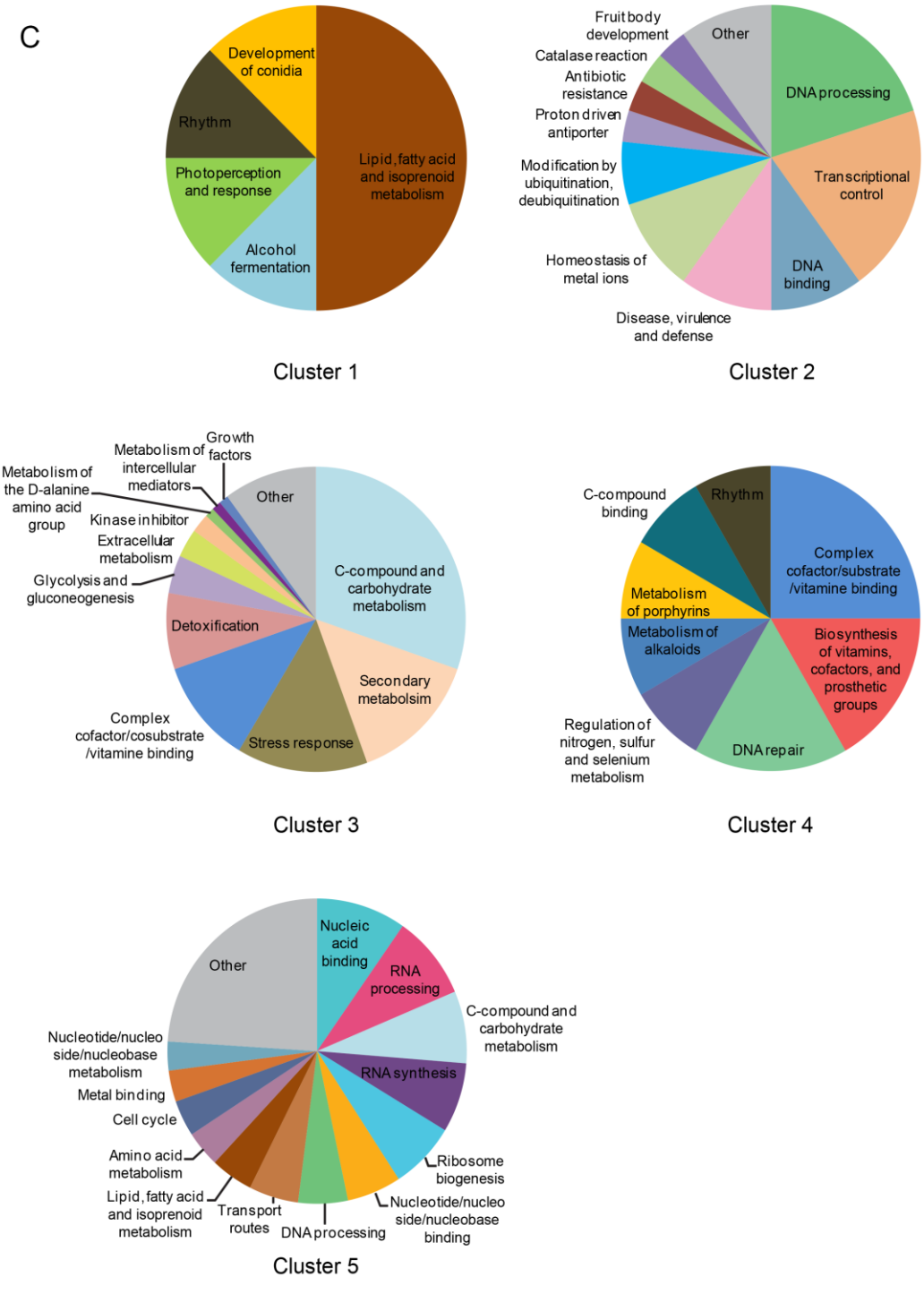


Figure 2-3. Continued.

cry, both of which attenuate WCC light responses (Heintzen et al., 2001; Olmedo et al., 2010b). The observation that *vvd* mRNA levels are induced >100 fold, while the transcripts for its activator WC-1 (Lee et al., 2003) are rising and peak later in the light, suggests that basal levels of WC-1 are sufficient to maintain photoresponses. The conidiation (con)-related genes *con-10* (NCU07325) and *con-6* (NCU08769) (Berlin and Yanofsky, 1985), originally identified by their induction during conidiation, and previously shown to be light-responsive (Corrochano et al., 1995; Lauter and Russo, 1991) are also included in this group.

Cluster 2 represents genes that were typically induced by light within 15 min, but returned to dark levels by 60 or 120 min. This response is typical of the light adaptation response of *N. crassa* that is mediated through feedback inhibition of the WCC by the photoreceptor VVD present in cluster 1 (Heintzen et al., 2001; Malzahn et al., 2010; Ptacek et al., 2007; Schwerdtfeger and Linden, 2001, 2003; Shrode et al., 2001). Cluster 2 includes WC-1 (NCU02356), and several of the TFs known to be direct targets of WCC: conidial separation-1 (*csp-1*, NCU02713), conidial separation-2, (*csp-2*, NCU06095), *vos-1* (NCU05964), *sub-1* (NCU01154), and siderophore regulation (*sre*, NCU07728) (Chen and Loros, 2009; Smith et al., 2010). Also included in this group of TFs was the *fluffy* (*fl*) gene (NCU08726) encoding a major regulator of conidiation in *N. crassa* (Rerngsamran et al., 2005; Ukai-Tadenuma et al., 2008). Although *fl* was not observed in general microarray studies to be light-regulated, its behavior in RNA-seq was similar to that observed in directed studies of *fl* expression (Belden et al., 2007;

Olmedo et al., 2010a). Also included in this cluster were genes involved in metal ion homeostasis (*mig-12*,NCU09830; *sre*, NCU07728; and *cax*, NCU07075), uncovering a previously not discerned need for the fungus to control metal ion homeostasis during exposure to light.

Cluster 3 includes genes that typically peaked in expression between 60 and 120 min in the light, the so-called late light-induced genes. Cluster 3 was visually subdivided into clusters 3a, 3b and 3c based on differences in expression patterns within the tree-structure of this cluster (Figure 2-3A and Appendix A Figure S1). Each of the subclusters showed higher levels of RNA at 120 min than at 15 min, as did the main cluster. FunCat analyses showed differences in functional enrichment categories for these subclusters (Appendix F File 2-S11). Cluster 3a is highly enriched for genes involved in metabolism and responses to oxidative stress, including the genes for the well-described detoxification enzymes glutathione-S transferase (NCU05706), glutamate decarboxylase (NCU00678), and NADH cytochrome B5 reductase (NCU03112). This cluster also includes a large number of genes encoding hypothetical proteins (75 genes), suggesting that at least some of these genes function in metabolism or cellular stress responses. Cluster 3b is enriched for genes involved in sugar metabolism. This cluster also included *catalase-1* (*cat-1*, NCU08791), which is important in hydrogen peroxide detoxification following light-triggered production of reactive oxygen species (Wang et al., 2007a, b). Interestingly, cluster 3b includes *clock-controlled gene-1* (*ccg-1*/*grg-1*, NCU03753), a gene of unknown function that is regulated by the circadian clock as well

as by light, oxidative stress and glucose starvation (Loros et al., 1989; McNally and Free, 1988). Both *cat-1* and *ccg-1* were previously shown to be regulated by the TF ATF-1 (also called ASL-1; NCU01346) which functions downstream of the osmosensing MAPK pathway in *N. crassa* (Lamb et al., 2012; Yamashita et al., 2008). The *atf-1*/*asl-1* gene displayed a >2-fold light induction, peaking at 15 min (Appendix F File S6); however, it did not make our stringent cutoff of q≤0.2. Cluster 3c is enriched for genes primarily involved in responses to environmental stress, including heat shock and cell wall integrity stress response pathways. These data are consistent with previous studies demonstrating regulation of these pathways by the WCC complex and the clock (Bennett et al., 2013; Smith et al., 2010).

Cluster 4 genes were rapidly and highly induced by light, and generally stayed induced over the course of light-treatment; this pattern of induction resembles that of Cluster 1. This cluster includes the core circadian oscillator gene *frequency* (*frq*, NCU02265), previously shown to be insulated from photoadaptation (Crosthwaite et al., 1995). Also included in this cluster are several oxidoreductases (*mig-3*, NCU04452; NCU01861; and NCU08291) important in redox reactions.

Genes in Cluster 5 are repressed by light, with repression being the strongest 60-120 min after light treatment. Although some light-repressed genes were noted in earlier work (Dong et al., 2008) a large cluster of such genes has not been previously described in detail. Because of the novelty of this observation we confirmed light-repression on a

subset of transcripts by RT-qPCR (Figure 2-4). The repression is observed when mRNA levels are normalized either to 25S rRNA levels or to levels of *cox-5*, a transcript whose levels did not respond to light. Genes implicated in ribosome biogenesis were highly enriched in this cluster, suggesting that light triggers a temporary reduction in the production of new ribosomes, which in turn would likely limit protein synthesis. As light is a stress to the organism (reflected here by the induction of numerous stress response genes in response to light), these data are consistent with previous studies demonstrating global protein synthesis repression following environmental stress (Holcik and Sonenberg, 2005; Lindquist, 1981; Shalgi et al., 2013; Spriggs et al., 2010).

Taken together, functional analyses of the clusters revealed that light generally increases cellular metabolism (Clusters 3a and 3b), and at the same time causes significant oxidative stress to the organism. To deal with this stress, protective photopigments are made (Cluster 1), antioxidants are produced (Clusters 2, 3a, 3b, 3c, and 4), genes involved in ribosome biogenesis are transiently repressed (Cluster 5), and the overarching regulatory pattern driven by the circadian system is reset to subjective morning (Cluster 4) in anticipation of a long period of continued light.

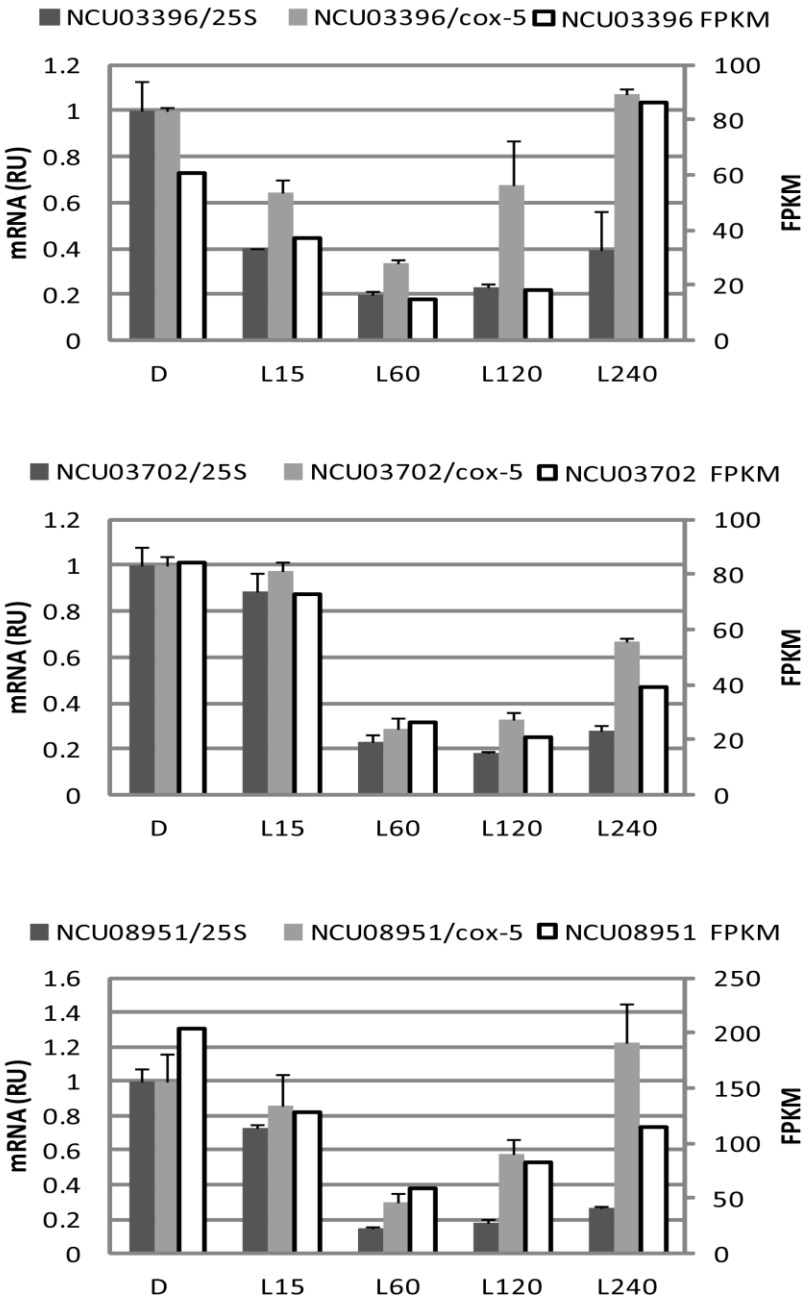


Figure 2-4. Validation by RT-qPCR of selected down-regulated transcripts in Cluster 5. FPKM at each time point (white bars) are compared to RT-qPCR values for each transcript that were normalized to either 25S rRNA (dark gray bars) or to *cox-5* (light gray bars); the *cox-5* transcript does not show a light-response in the RNA-seq data. Error bars show the standard error obtained from triplicate technical replicates of each of the two RNA preparations used for RNA-seq.

Global regulation by light

Nearly 25% of the genome showed a >2-fold change in gene expression in response to light under our growth conditions. As previous studies have implicated a hierarchical network of TFs controlling light induction (Chen et al., 2009; Smith et al., 2010), it was not surprising to find that TFs are overrepresented in several of the light-induced clusters. Overall, 58 of 252 identified TFs are regulated by light (Figure 2-5). Among these, 12 TFs were significantly regulated at a q-value ≤ 0.2 (indicated by asterisks in Figure 2-5) and are represented in clusters 2, 3a, and 5. Furthermore, of the 27 TFs identified as direct targets of the WCC (Smith et al., 2010), 14 showed a >2-fold change in response to light for at least one time point, and 7 with a q-value ≤ 0.2 (Appendix F File 2-S12). Of these 14 genes, 3 TF genes, *adv-1*, NCU07846, and NCU05994, showed decreased levels for at least one time point. Together, these data demonstrate that, with the growth conditions used here, not all steady state mRNAs from genes that are direct targets of the WCC show significant light responses, and that some direct targets of the WCC are repressed by light.

A relatively large number of transcripts showed sizeable increases in expression upon light exposure; 27 mRNAs had 16-fold or greater increases at one or more time points

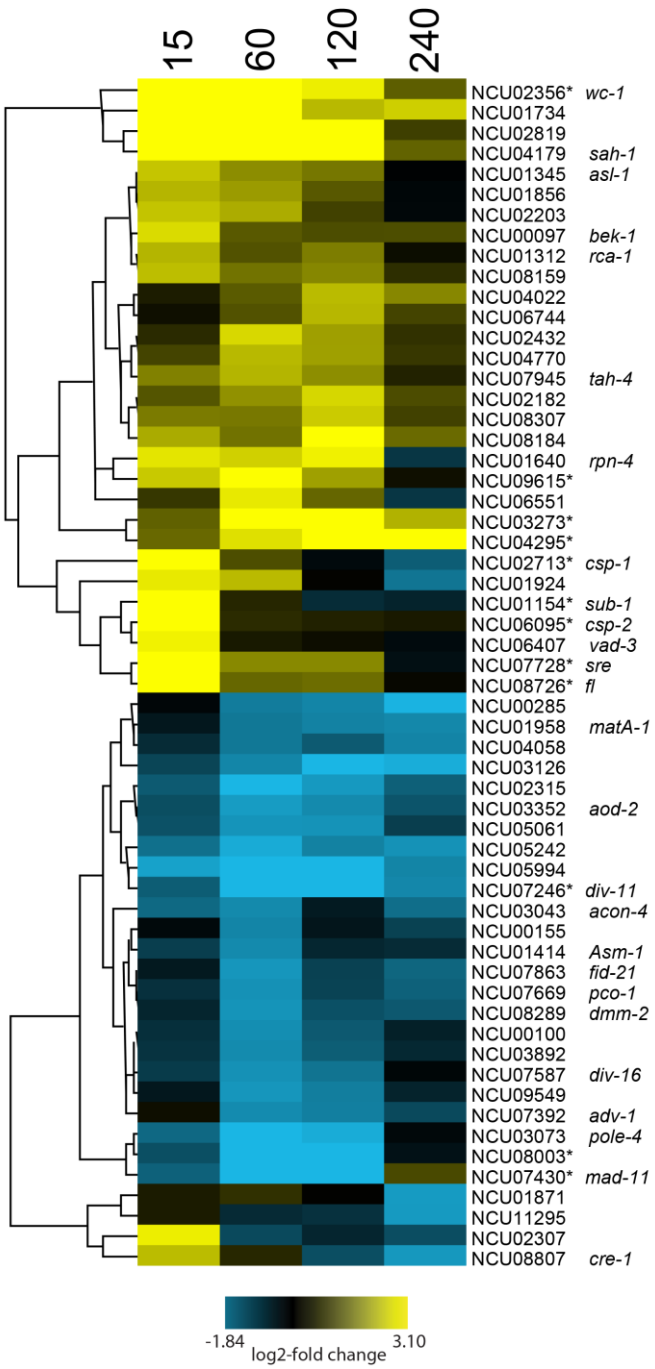


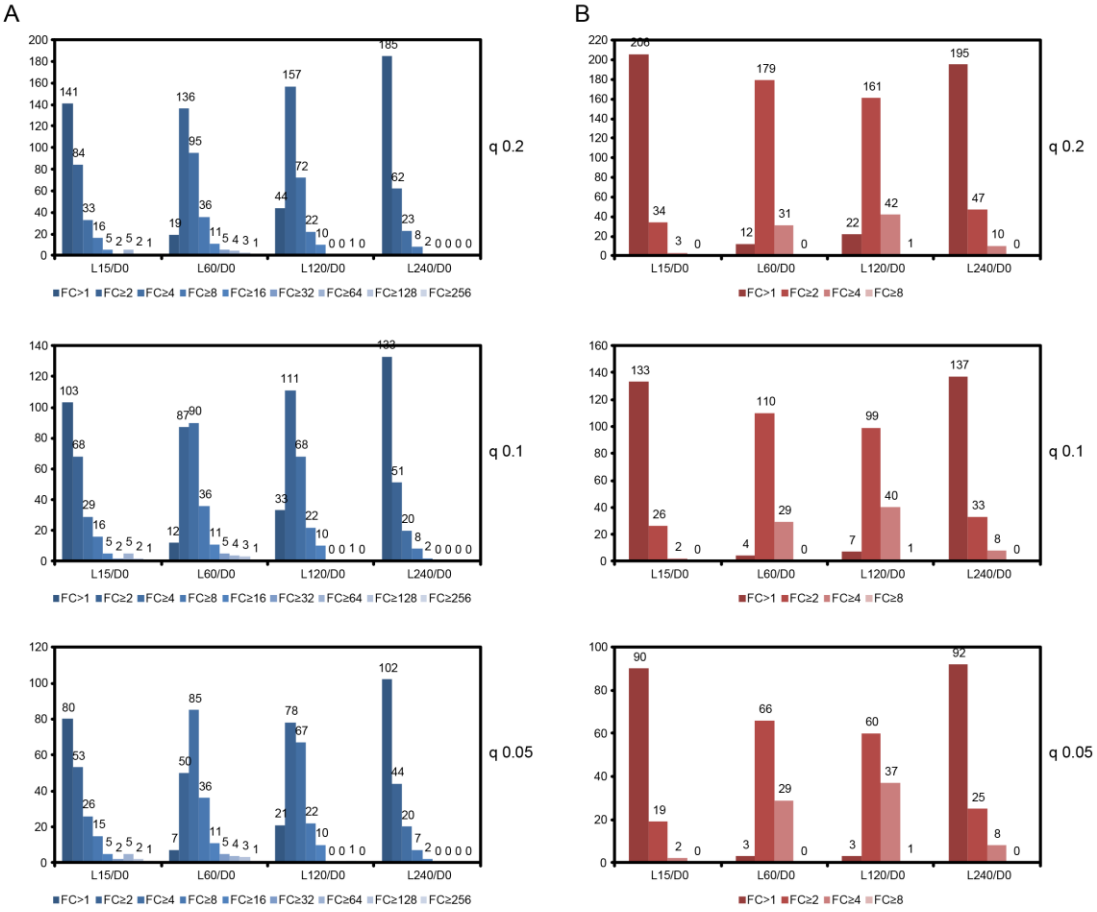
Figure 2-5. Hierarchical clustering of transcription factors whose mRNA levels are light-regulated. Factors whose mRNA levels were regulated at q≤0.2 are indicated with asterisks.

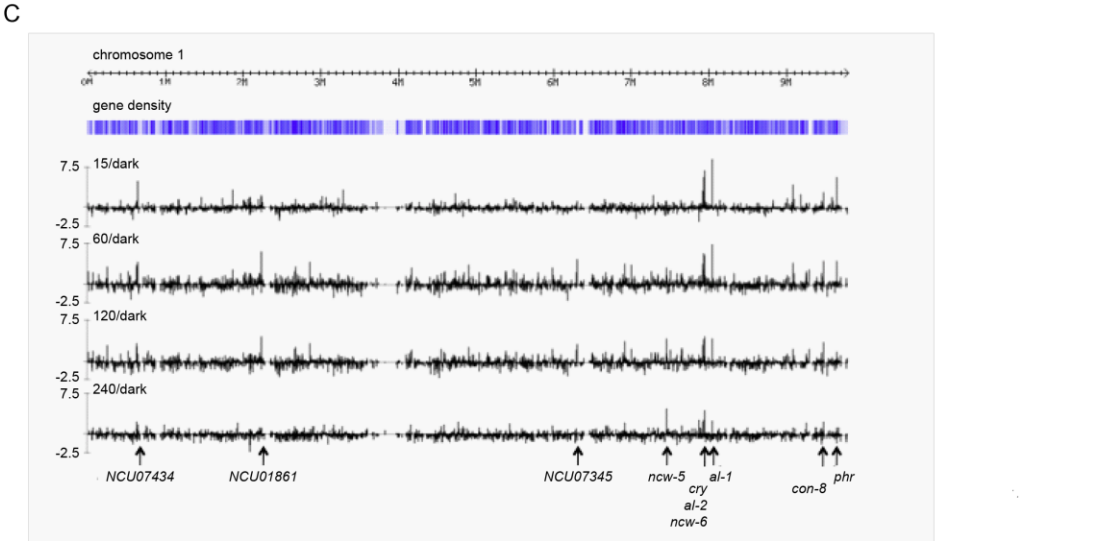
Table 2-1. Transcripts with 16-fold or more induction by light

Locus	Symbol	Name	L15/D	L60/D	L120/D	L240/D
NCU00552	*al-1*	albino-1	8.3	6.9	4.1	2.4
NCU08699	*bli-4*	bli-4 protein	7.5	7.8	4.8	3.5
NCU02190		oxysterol binding protein	7.3	5.6	2.5	1.5
NCU03967	*vvd*	vivid	6.6	5.1	4.0	3.1
NCU08770		hypothetical protein	6.4	7.6	2.3	0.7
NCU00582	*cry*	cryptochrome DASH	6.3	4.9	4.4	4.1
NCU10063		sugar isomerase	6.2	4.3	2.8	2.1
NCU08769	*con-6*	conidiation-6	6.0	8.5	7.5	2.4
NCU08626	*phr*	photoreactivation-deficient	5.3	4.0	1.9	1.2
NCU00585	*al-2*	albino-2	5.2	5.3	3.9	2.5
NCU07325	*con-10*	conidiation-10	4.8	7.1	4.3	1.5
NCU11424	*cao-2*	carotenoid oxygenase-2	4.8	4.1	2.8	2.1
NCU07434		short-chain dehydrogenase/reductase SDR	4.6	3.8	1.5	1.7
NCU07267	*bli-3*	blue light-induced-3	4.6	6.8	4.6	2.4
NCU11395		S-(hydroxymethyl)glutathione dehydrogenase	4.0	6.1	4.4	2.7
NCU06420		hypothetical protein	3.8	4.1	2.7	1.3
NCU01060		hypothetical protein	3.5	4.9	3.9	2.8
NCU09049		hypothetical protein	3.4	6.1	4.3	1.3
NCU03506		hypothetical protein	3.1	4.3	3.0	1.5
NCU05844		hypothetical protein	2.7	4.5	2.6	-0.5
NCU01861		short chain dehydrogenase/reductase family	2.2	5.6	4.5	1.1
NCU06597		hypothetical protein	2.0	5.4	3.3	0.6
NCU04823		NADP-dependent alcohol dehydrogenase C	1.6	4.5	3.5	1.0
NCU05652		hypothetical protein	1.4	4.6	3.5	0.7
NCU07345		hypothetical protein	1.2	4.3	3.2	0.4
NCU07337		hypothetical protein	1.2	4.2	4.1	2.1
NCU00716	*ncw-5*	non-anchored cell wall protein-5	-0.3	2.8	4.1	4.3

The $\log_2$ change in expression in the light versus the dark is given for each time-point in the light (15, 60, 120 and 240 min).

Figure 2-6. Genome wide view of light responses. (A) A plot of the number of light-induced genes as a function of time after lights on at each of the time points. These plots show the number of genes for a given level of expression-change (FC, fold-change) in response to light (up- or down-regulated) at $q \leq 0.2$, 0.1, or 0.05 (the primary data are in Appendix F File S9). (B) A plot of the number of light-repressed genes as a function of time after lights on with all of time points. (C) Pattern of light-regulation of genes on Linkage Group I (Chromosome 1). The $\log_2$ of the fold-change in expression in the light versus the dark is given on the Y-axis for each time-point (15, 60, 120 and 240 min). Several examples of strongly light-induced genes mentioned in the text are marked by arrows.


A
B
q 0.2
q 0.1
q 0.05
L15/D0
L60/D0
L120/D0
L240/D0
FC>1 FC≥2 FC≥4 FC≥8 FC≥16 FC≥32 FC≥64 FC≥128 FC≥256


C
chromosome 1
gene density
15/dark
60/dark
120/dark
240/dark
NCU07434
NCU01861
NCU07345
ncw-5
cry
al-2
ncw-6
al-1
con-8
phr

(Table 2-1). FunCat analysis indicated enrichment for functions similar to those identified for Cluster 1 genes (compare Appendix F File S13). Plots of the numbers of genes showing increased or decreased responses to light at different q-value thresholds are shown in Figure 2-6A and B. In general, reductions in response to light were smaller than induction in response to light, with no reduction more than 8-fold. The greatest reductions were seen 60-120 min after lights on, in contrast to induction, which peaked 15-60 min after lights on. Global analyses of the distribution of light-regulated transcripts across *N. crassa* chromosomes did not reveal obvious regions where expression patterns of genes clustered (Figure 2-6C and Appendix A Figure S2).

We examined the phenotypes of a subset of genes whose transcripts were altered 2-fold or more in response to light with a $q \leq 0.2$ with both known and unknown functions. The phenotypes of knockouts of these selected genes are shown in Figure 2-7. Many of these genes showed obvious phenotypes affecting vegetative growth or sexual development. Interestingly, the *frq* deletion strain, in an otherwise WT background, showed slower linear growth in race tubes, and a phenotype on plates that differed from the wild type. This result was surprising given growth differences between *frq*-null mutations in the *ras-1*^{bd} background (commonly used for clock studies) and control *ras-1*^{bd} strains, have not been observed or reported (Aronson et al., 1994b; Loros and Feldman, 1986). We confirmed that an independent disruption of *frq* with the Bar^{r} marker in the wild-type background also showed slower linear growth at 25°, but in the (already slower-growing) *ras-1*^{bd} strain, disruption of *frq* with the Bar^{r} marker did not further reduce the

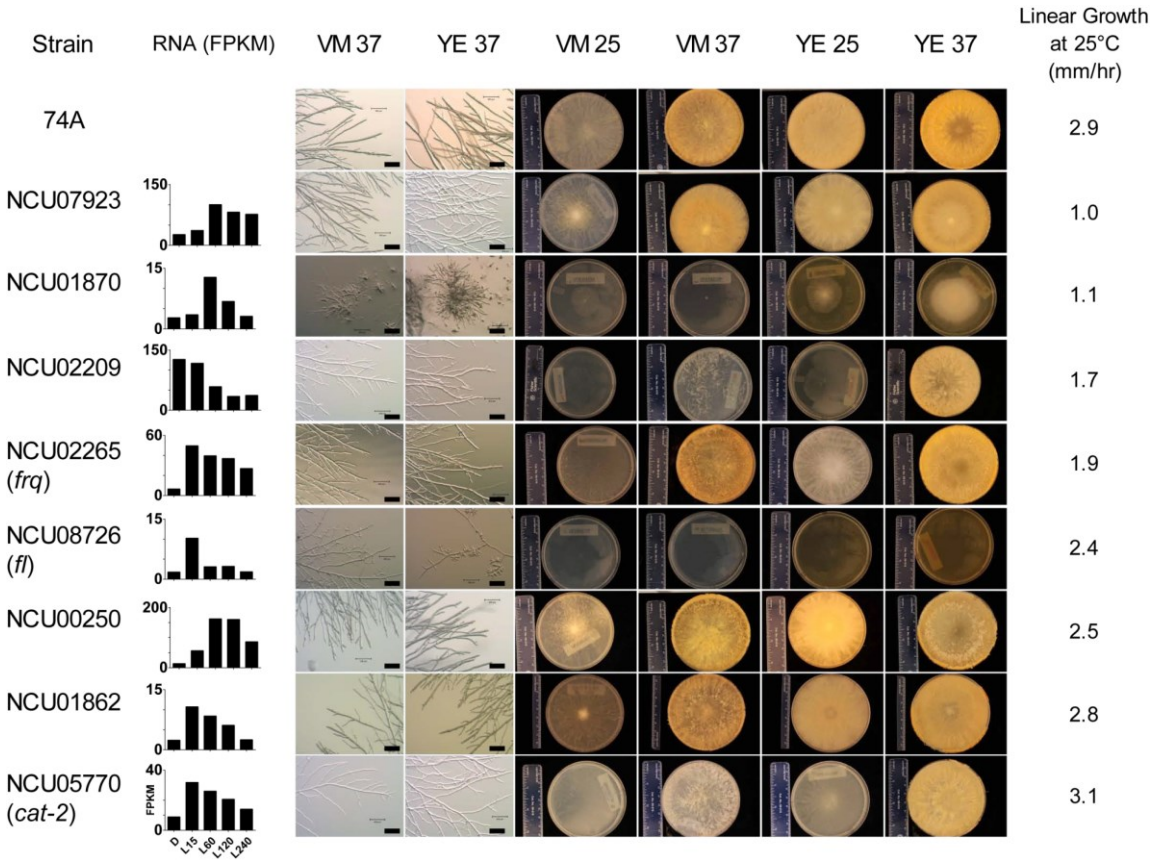


Figure 2-7. Phenotypes of selected light-regulated genes. A sampling of mutants containing disruptions of genes showing light regulation with q≤0.2 were analyzed for vegetative-growth phenotypes. The strain name indicates the gene that was disrupted (74A is the wild-type strain). The indicated gene's expression in the wild-type strain as determined by RNA-seq is shown for reference. VM: Vogel's Minimal medium; YE, VM supplemented with yeast extract; 25 and 37, 25°C and 37°C, respectively. Analyses of sexual growth phenotypes showed that each strain was female-fertile, except for NCU01870, which was female-sterile.

linear growth rate (data not shown). The *fluffy* deletion strain showed a morphological phenotype when yeast extract was included in the growth medium. While the functions of many of these gene products are hypothetical, the results of phenotyping illustrate that light affects the expression of many genes whose functions are important for the growth of the organism. Additional phenotypic analyses of strains containing deletions of these light-regulated genes should provide clues to the functions of the hypothetical genes and reveal broader connections among the functions of genes with similar expression profiles.

DISCUSSION

The goal of this work was to achieve a comprehensive understanding of how gene expression in *N. crassa* changes in response to light through the use of RNA-seq. We therefore purified capped and polyadenylated mRNA from vegetatively growing *N. crassa* mycelium grown in the dark or grown in the dark and exposed to light for 15, 60, 120 or 240 minutes. We sequenced cDNA obtained from this mRNA by using Illumina short-read methodology, and analyzed gene expression by measuring the relative abundance of mRNA for known or predicted *N. crassa* protein coding genes (based on assembly Nc10 and annotation 10.6). The abundance of transcripts for approximately 25% of all predicted *N. crassa* genes changed 2-fold or more based on these data. Transcript abundance levels for 532 genes (5% of all predicted genes) were light-regulated using a false discovery rate cutoff of $q \leq 0.2$ for the data from two independent

experiments. The increased power of the RNA-seq approach compared to previous microarray-based approaches enabled the identification of genes which were not highly expressed, but which were regulated in response to light. For example, we obtained evidence for a major class of genes predicted to have roles in rRNA processing that were down-regulated in response to light.

Estimates of the fraction of *N. crassa* genes induced by light have ranged from 3-14 % of the predicted protein-coding genes in the genome (Chen et al., 2009; Dong et al., 2008; Lewis et al., 2002). A conservative estimate based on the data obtained here (using a cutoff of $q \leq 0.2$) is 5% of the genome, while a more liberal estimate is 25% since this is the fraction of genes whose expression changed 2-fold in response to light. A major category of genes whose predicted functions are significantly enriched in response to light in the larger group of genes (2-fold regulated) are those with uncategorized or unknown functions. Among these are genes that are fungal-specific. For example, within this category of genes for which there is strong statistical support for light induction is NCU07923 (4-fold induction with $q<0.02$, Table S9), a hypothetical protein that appears strongly conserved within the ascomycetes but not outside of them; deletion of this gene has an obvious vegetative growth phenotype (Figure 2-7).

While a substantial fraction of predicted *N. crassa* protein-coding genes are regulated at the transcript level in response to light, there were no obvious large chromosomal clusters of genes that showed common regulation. This indicates that the action of light

to increase or decrease transcript levels is not generally operating on clustered genes, and further that these mechanisms are not affecting large contiguous domains of chromatin.

The functions of genes that respond early to light appear different than those that respond later. Light induces the expression of many genes associated with stress responses 60-120 min after exposure, and this can be rationalized because light can generate reactive oxygen species (Wang et al., 2007b). Consistent with the damaging effects of light, direct targets of the WCC are enriched for DNA repair enzymes (Smith et al., 2010). Thus, it is not surprising that we found genes involved in DNA repair mechanisms, and encoding light absorbing photopigments, are rapidly light-induced. The kinetics of the responses of stress-response genes are similar to those of a subset of genes with roles in C-compound and carbohydrate metabolism whose expression is also induced by light (Cluster 3, Figure 2-3C and Appendix F File S11). FunCat functional enrichment for this category of genes among late light response genes has been observed previously in microarray studies using the same growth medium (Chen et al., 2009). The expression of a large set of genes is also reduced 60-120 min after light exposure. While the mRNAs for the protein components of the ribosome are not reduced by light exposure, many of the mRNAs specifying factors involved in rRNA processing and ribosome assembly are reduced. Thus, it may be more efficient for cells to control ribosome assembly, as opposed to adjusting the levels of abundantly expressed mRNAs encoding ribosome proteins, in response to environmental signals.

Among the early light-induced genes that are associated with DNA repair are NCU08850 (*mus-18*) and NCU08626 (*phr*). Mutations in *mus-18* were originally identified because they were UV-sensitive (Ishii et al., 1991), and mutant strains are deficient in excision-repair. This light response is deeply conserved because *UVE1*, the homolog of this gene in the basidiomycete *Cryptococcus neoformans*, is also strongly light induced through the WCC (Verma and Idnurm, 2013). UV-irradiation can result in the formation of cyclobutane pyrimidine dimers, and *N. crassa phr* specifies a cyclobutane pyrimidine dimer photolyase that reduces this DNA damage through light-dependent photoreactivation (Shimura et al., 1999) .

The value of RNA-seq in discovery for a better understanding of the behavior of relatively well-characterized genes is illustrated by the results obtained here with *nop-1* (*new eukaryotic opsin 1*). *nop-1* (NCU10055) encodes a retinal binding protein that affects the expression of genes that are themselves light-regulated, and thus would be anticipated to have light-specific functions (Bieszke et al., 1999a; Bieszke et al., 2007). In previous studies, the *nop-1* mRNA level was not observed to increase in response to light (Bieszke et al., 2007; Chen et al., 2009). However, in each of our two independent experiments, and in analyses of the pooled experimental data, *nop-1* mRNA increased at 60 min and 120 min (and came down at 240 in one while remaining up at 240 in the other). This increase in *nop-1* was significant at a q-value <0.2 in each case. The RNA-seq data thus demonstrate that the expression of *nop-1* is light-regulated, and provides

the basis for further experiments to identify how its increased expression relatively late in the light-response impacts the biology of the organism.

The transcripts for *con-6* (NCU08769) and *con-10* (NCU07325) are strongly induced in response to light as shown here and elsewhere (Chen et al., 2009; Corrochano et al., 1995; Lauter and Russo, 1991; Lauter and Yanofsky, 1993; Olmedo et al., 2010b). The role of these genes in *N. crassa* has remained elusive since single mutants do not display significant phenotypes. However, a phenotype for the equivalent double mutant of the *A. nidulans* homologs *conF* and *conJ* has been described (Suzuki et al., 2013). The double knockout strain resulted in significant increases in the amount of cellular glycerol and erythritol, which delayed conidial germination and provided an increase in resistance of the cells to desiccation. As is the case for *N. crassa*, both *conF* and *conJ* are rapidly light-induced, and the single knockouts displayed no obvious phenotypes. These data suggest the likelihood that *con-6* and *con-10* have redundant functions in spore germination in response to light.

Light plays a key role in synchronizing the *N. crassa* circadian clock to local time, and therefore, it is not surprising to find that light affects the levels and activities of core clock components. We found that *wc-1* (NCU02356) is transiently light induced, consistent with previous work (Ballario et al., 1996; Linden and Macino, 1997). However, while *wc-2* (NCU00902) is reported to be weakly light induced (Ballario et al., 1996; Linden and Macino, 1997), no increase in *wc-2* transcripts was observed

following light treatment in our experiments. In agreement with these data, light induces a transient increase in WC-1 protein levels, but little or no change in the WC-2 (Schwerdtfeger and Linden, 2000; Talora et al., 1999), whose levels, unlike WC-1, are not limiting in cells (Cheng et al., 2001b; Denault et al., 2001). Light-activated WCC binds to light-responsive elements (LRE's) in the *frq* promoter, leading to subsequent activation of *frq* transcription (Froehlich et al., 2002; He and Liu, 2005b) This change in *frq* mRNA and protein levels is responsible for resetting the phase of the clock to the appropriate time of the day (Crosthwaite et al., 1995). Interestingly, none of the other clock components and modifiers of the components, including FRH (FRQ-Interacting RNA Helicase), a binding partner of FRQ necessary for negative feedback (Guo et al., 2010) and several kinases (CK1, CK2, PKA, and CAMK-1) and phosphatases (PP1, PP2A, and PP4) that modify the activities of the clock components, met our stringent criteria for light-regulation. Three genes, *camk-1* (NCU09123), *ppp-1* (NCU00043 encoding the catalytic subunit of protein phosphatase 1), and *rgb-1* (NCU09377 encoding the regulatory subunit of protein phosphatase 2A) had a greater than 2X change in expression following light treatment, but the q value was >0.2. These data suggest that circadian light responses are mediated by the absolute changes in *frq* mRNA levels through the activity of the WCC.

RNA expression analyses of 27 TF genes shown to be direct targets of the WCC has been accomplished using RT-qPCR to compare RNA levels in dark-grown cells and cells given a 15-min light pulse (Smith et al., 2010). These data are similar to the RNA-

seq data we obtained (Appendix F Table S12) with the exception of *adv-1* discussed below. First, in both studies, not all of the genes that are direct targets of the WCC were light-regulated. Second, while not noted before (Smith et al., 2010), the steady-state mRNA levels of some of the TF genes are reduced in response to light (e.g. NCU05994). These data suggest that the WCC has repressive, as well as activating functions for specific targets, or that other TFs participate in the regulation of the light-repressed genes. Identification of the direct targets of the TFs will help to resolve this question. The gene encoding the ADV-1 TF was the only case where differences were observed between the two data sets. In our experiments, the levels of *adv-1* mRNA stayed fairly constant in DD and L15, but then decreased more than 2-fold at L60 and L120. In contrast, in previous RT-PCR assays, *adv-1* mRNA levels were induced more than 2-fold following a 15 min light treatment (Smith et al., 2010). The basis for this difference is not understood. It may reflect the use of different media in the two studies, or differences in the culture conditions. Specifically, the amount of time the cultures were in the dark before the light treatment varied in the two data sets, 12 h (Smith et al., 2010) versus 24 h in our experiments, which for a rhythmic gene, such as *adv-1*, would result in time-of-day differences in the dark mRNA levels.

Together, the identification of light-responsive genes will provide the foundation for our ongoing efforts to decipher the specific roles of TFs that respond to light and the clock in the regulation of the global photo-responses and circadian rhythmicity. These studies include determining the direct and indirect targets of the light-controlled TFs, and the

interplay of TFs in orchestrating the light and circadian response. With a substantial fraction of the genome showing altered gene expression in response to light, and with different patterns of expression, we anticipate a complex light-controlled and circadian-regulated TF network.

MATERIALS AND METHODS

Strains

Wild-type strains 74-OR23-IVA (FGSC 2489; *mat A*) and ORS-SL6a (FGSC 4200; *mat a*), and single gene deletion strains (Colot et al., 2006) ΔNCU00250 (FGSC12215; *mat a*), ΔNCU01862 (FGSC19012; *mat a*), ΔNCU01870 (FGSC13270; *mat a*), ΔNCU02209 (FGSC16054; *mat a*), ΔNCU02265 (FGSC11554; *mat a*), ΔNCU05770 (FGSC11532; *mat A*), ΔNCU07923 (FGSC19046; *mat A*), ΔNCU08726 (FGSC11044; *mat a*) were obtained from the Fungal Genetics Stock Center (Kansas City, MO).

Light treatment

Macroconidia were obtained from flask cultures containing 1X Vogel's medium with 2% sucrose and 2% agar (Sachs and Yanofsky, 1991). Conidia were harvested with water, filtered through cheesecloth, and counted with a hemacytometer. For each time-point, 200-ml of Bird's Medium (Metzenberg, 2004) was inoculated to a final

concentration of 10^7 conidia/ml, and grown in the dark for 24 h at 25°C with orbital shaking (150 rpm). Some of the cultures were exposed to white-light using cool white fluorescent bulbs (1200 lux), and cells were harvested in a darkroom at time 0 (dark), 15, 60, 120 and 240 min after light exposure. The cells were harvested by centrifugation (1000g for 1 min) using an IEC clinical centrifuge and washed once with ice-cold sterile water (50 ml). The mycelia pad was cut into small pieces (approximately 100 mg/piece) with a razor. Individual pieces were placed into 15ml screw-cap tubes and snap-frozen in liquid nitrogen. Mycelia were stored frozen at –80°C.

Total RNA and poly(A) RNA isolation

Total RNA was isolated from frozen mycelia (approximately 100 mg) using 1 g of autoclaved Zirconia/Silica Beads (Biospec Products, Inc, Bartlesville OK, product number 11079105Z) and a Mini-BeadBeater-8 (Biospec Products, Inc, Bartlesville OK) with ice-cold 580 µl extraction buffer (100 mM Tris-HCl, pH 7.5; 100 mM LiCl; 20 mM DTT), 420 µL phenol, 420 µL chloroform, and 84 µL 10% SDS (Sachs and Yanofsky, 1991). Immediately after removal from -80°C storage, mycelia were homogenized in the bead beater for 1 min. After a 4 min end-over-end rotation, the homogenate was centrifuged at 12000 g at 4 °C for 1 min to separate phases. The aqueous phase was extracted with phenol/chloroform and chloroform, and precipitated in 0.3 M sodium acetate and ethanol. The pellet was washed twice with 70% ethanol that was prepared with diethylpyrocarbonate (DEPC)-treated water, briefly air-dried, and then dissolved in

100µl filter-sterilized DEPC-treated water. RNA concentration was determined using a Nanodrop spectrophotometer, and quality was assessed by denaturing gel electrophoresis in formaldehyde gels and northern analyses. Poly(A) mRNA was purified from total RNA using oligo-dT cellulose (Sachs and Yanofsky, 1991) for replicate 1, or using the Poly(A)Purist MAG kit (replicate 2) (Ambion, Grand Island, NY). Residual DNA was removed from poly(A) mRNA using Ambion Turbo DNase and intact mRNA was selectively enriched using Epicentre mRNA-Only. (Illumina, San Diego, CA) to degrade mRNA species lacking a cap. Poly(A) mRNA concentration was determined by RiboGreen fluorometric assay (Life Technologies, Grand Island, NY).

cDNA preparation

The procedure for preparing cDNA was modified from a previously described procedure as follows. To synthesize the first-strand cDNA, 1 µg poly(A) mRNA was mixed with 1.2 µl of 10 mM dNTPs, 1 µg of random hexamer, heated at 65 °C for 5 min, and then placed on ice for 3 min to remove any secondary structure. 10 µl of 5X first strand buffer (250 mM Tris-HCl, pH 8.3; 375 mM KCl; 15 mM MgCl2), 4.55 µl of 0.1 M DTT, 5.6 µl of 80% glycerol, 2.88 µl of 1 µg/µl BSA, 20 units RNasin (Promega, Madison WI), and 300 units SuperScript III reverse transcriptase (Life Technologies, Grand Island, NY) were then added to the reaction tube and the reaction was incubated at 25 °C for 5 min, 50 °C for 50 min, and 70 °C for 15 min in a thermocycler. The reaction mix was precipitated with 0.3 M sodium acetate and ethanol, and the pellet was washed twice

with 70% ethanol, briefly air-dried, and dissolved in 22 µl of DEPC-treated water. The RNA-DNA hybrid was treated with 2.5 units of *E. coli* RNaseH (New England Biolabs, Ipswich, MA) at 37 °C for 20 min. Then 10 µl of 10X second strand buffer (200 mM Tris-HCl, pH 8.93; 350 mM KCl; 50 mM $(NH4)_2SO4$; 10 mM $MgSO_4$; 10 mM $MgCl_2$; 0.5% Triton X-100), 3 µl of 10 mM dNTPs, 6 units RNaseH, and 50 units *E. coli* DNA polymerase I (New England Biolabs, Ipswich, MA) were added to the reaction tube. The reaction was incubated at 16 °C for 2.5 h and then stopped by freezing.

Sequencing library preparation

cDNA was resuspended in 300 µl TE in a 1.5 ml TPX micro tube (Diagenode, Denville, NJ) and sheared using the low energy cycle for 15 min (15 sec on/15 sec off) in a Diagenode Bioruptor, whose water bath was maintained at 4°C with a recirculating chiller. Sheared cDNA was transferred to a 1.7 ml polypropylene microcentrifuge tube and precipitated with 0.3 M sodium acetate and ethanol. The pellet was washed once with 70% ethanol, air-dried, and dissolved in 15 µl sterile water and stored at -20°C. The yield of recovered DNA was determined by PicoGreen fluorimetric assay with DNA (New England Biolabs, Ipswich, MA) as the standard using a Victor3V 1420 multilabel counter (Perkin Elmer, Waltham, MA). An aliquot of sheared DNA (1.5 µl) was examined on a 1% TAE agarose gel to verify sizes ranging between 100-800-bp. RNA-seq libraries were prepared from 1 µg of sheared DNA using an Illumina (San Diego, CA) TruSeq Sample Prep kit according to the manufacturer's instructions. Each

sample was ligated to an indexing adapter and selectively enriched through a 10-cycle amplification. The library was then separated on a 1% TAE agarose gel, and DNA ranging between 200-500 bp was excised and purified using QIAEX II gel purification kit (Qiagen, Germantown, MD). The yield of amplified DNA was determined using the PicoGreen assay. An aliquot (3 µl) was examined on a 1% TAE agarose gel to monitor amplification. The DNA was diluted in water to a final concentration of (10 nM in 30 µl) for Illumina sequencing.

High throughput sequencing

cDNA libraries were sequenced on an Illumina HiSeq 2000 for 57 cycles, and processed with Illumina pipeline RTA 1.13.48 and CASAVA v1.8.2. Trimming the 6-mer adapter generated 51-mer reads. The yield for each sample is shown in Appendix B Table S1. Raw and processed data were submitted to the NCBI GEO database with accession number GSE53534.

Data analysis

Illumina reads were mapped to N. crassa assembly 10 with TopHat v2.0.5 (options -i 30 -I 2000). FPKM and fold-changes were calculated using the cuffdiff command from version 2 of the Cufflinks suite. Version 10.6 of the *N. crassa* annotation was used, and

fragment bias (the "-b" option) and multiple-read correction (the "-u" option) options were enabled.

Quantitation and statistical tests were performed both separately for each time course, as well as for the combined set of replicates. FPKM values across the light time course were hierarchically clustered and visualized as heat maps with Cytoscape 2.8.3 and the clusterMaker plugin (Morris et al., 2011). Functional enrichments for each cluster were calculated using the FunCat database (FunCatDb) (Ruepp et al., 2004). Gene identifiers in FunCatDb were mapped to the identifiers in the current annotation using the mapping information available at the Broad Institute *N. crassa* Database (http://www.broadinstitute.org/annotation/genome/neurospora/MultiDownloads.html). Gene identifiers in the Gene Ontology were mapped using the GO Term Finder tool (http://go.princeton.edu/cgi-bin/GOTermFinder) with a *N. crassa* GO gene association file (go_for_nc12.tsv downloaded from http://www.broadinstitute.org/annotation /genome/neurospora/Downloads.html.

Quantitative RT-PCR

Eight ng of cDNA (0.125 ng for quantification of 25S cDNA) was used as qPCR template in a reaction containing: 1X Platinum® Taq PCR Buffer (200 mM Tris-HCl pH 8.4, 500mM KCl) (Life Technologies, Grand Island, NY), 2.5 mM MgCl2, 0.2 mM dNTPs, 1X ROX Reference Dye (Life Technologies, Grand Island, NY) 1X SYBR

Green I (Life Technologies, Grand Island, NY), 500 nM each primer and 5 U/µl Platinum Taq DNA Polymerase (Life Technologies, Grand Island, NY) or 5 U/µl TaKaRaTaq DNA Polymerase (Clontech, Mountain View, CA) in a 20µl reaction with filtered sterile DEPC-treated water. Amplification was as follows: 50°C for 2 min, 95°C for 10 min followed by 40 cycles at 95°C for 15s, and 60°C 1 min. 25S rRNA was used as an internal control for normalization.

Phenotype analyses

Phenotyping of *N. crassa* single gene deletion strains, including strains containing deletions of genes whose expression was affected by light, was accomplished in an undergraduate research course at Texas A&M University (BIOL452, Fungal Functional Genomics) taught by MSS, DBP, YZ and RD. Phenotyping was performed as described (Colot et al., 2006) with minor modifications. Vogel's minimal medium (VM)/1.5% sucrose/2% agar was used for vegetative growth and synthetic crossing medium (SCM)/1% sucrose/2% agar for development of female sexual structures (Davis and de Serres, 1970). VM was supplemented with 2% yeast extract (YE) where indicated. All strains were analyzed in at least triplicate for each phenotype. All strains were incubated at either 25°C or 37°C in Percival chambers with a 12-hr light/12-hr dark cycle. Images of hyphae were obtained using Olympus SZX16 microscopes with IDEA 5 cameras and IDEA SPOT software. Plate photos were obtained using Canon APS-C sensor cameras. Tests for female fertility were accomplished using 3 ml SCM in Falcon T12.5 Tissue

Flask (353108) with the arrow on the flask's cap pointed to the top of the flask to ensure sufficient airflow.

CHAPTER III

TRANSCRIPTION FACTORS IN LIGHT AND CIRCADIAN CLOCK SIGNALING NETWORKS REVEALED BY GENOMEWIDE MAPPING OF DIRECT TARGETS FOR *NEUROSPORA* WHITE COLLAR COMPLEX*

INTRODUCTION

Light perception leads to changes in gene expression that ultimately alter physiology. Even after extensive research, little is known about mechanisms that directly link photoreceptor activation to signaling pathways eliciting light responses. Typically, these regulatory modules are tightly coupled to endogenous circadian clocks (Hsu et al., 2013). In *N. crassa*, a eukaryotic model for light responses and the circadian clock, the blue-light receptor and PAS domain GATA-type transcription factor (TF) WHITE COLLAR-1 (WC-1) dimerizes with a second PAS domain GATA TF, WC-2, to form the White Collar Complex (WCC) (Ballario et al., 1996; Linden and Macino, 1997). In the FRQ/WCC oscillator, the WCC is the positive element that directly activates frequency (*frq*) transcription in the morning by binding the *frq* promoter at the proximal light response element (pLRE), required for light regulation of *frq*, and the distal LRE

---

* Reprinted with permission from "Transcription Factors in Light and Circadian Clock Signaling Networks Revealed by Genomewide Mapping of Direct Targets for *Neurospora* White Collar Complex" by Smith, K.M., Sancar, G., Dekhang, R., Sullivan, C.M., Li, S., Tag, A.G., Sancar, C., Bredeweg, E.L., Priest, H.D., McCormick, R.F., *et al*., 2010. Eukaryotic Cell 9, 1549-1556.;

(dLRE) or “clock box” (C box), which is required for clock and light regulation (Cheng et al., 2005; Froehlich et al., 2002; He and Liu, 2005b; Nakamichi et al., 2012). This regulation results in rhythmic expression of *frq* mRNA and FRQ protein. FRQ protein dimerizes and binds to the FRQ-RNA helicase (FRH). The FRQ/FRH complex (FFC) functions as the negative element in the circadian negative-feedback loop (Cheng et al., 2005; He et al., 2006). Once FRQ protein is made, it becomes progressively phosphorylated, and when fully phosphorylated, it is degraded, allowing the cycle to restart the next morning (Dunlap et al., 2007; Liu et al., 2000; Schafmeier et al., 2008).

The WCC is also required for all known blue-light responses, including resetting the circadian clock, carotenoid synthesis, asexual spore development, formation of female sexual structures, and ascospore release (Vitalini et al., 2006). In response to light, activated WCC functions as a TF and binds to LREs to regulate the expression of target genes. More than 100 light-responsive genes have been identified in *N. crassa*, primarily through transcript microarray studies. Importantly, not all of these genes contain an obvious LRE (Chen et al., 2009; Lewis et al., 2002; Linden and Macino, 1997; Ptacek et al., 2007), which implies that a genetic regulatory cascade orchestrates expression of the light-responsive gene network. Furthermore, most light-regulated genes are also clock regulated. The identification of key molecules that function in light-regulated/circadian output pathways and control rhythms in target gene expression is one major goal of circadian biology. Despite significant efforts, these molecules have been difficult to identify. While the predicted functions of the *N. crassa* light-inducible genes suggest

which processes are regulated by light, microarray studies have failed to directly link the action of the WCC to these genes. Identifying the genes directly regulated by the WCC is an essential step toward generating hierarchical or network models to describe the molecular and physiological responses to light and the circadian clock.

To identify the direct targets of the light-activated WCC, we used chromatin immunoprecipitation (ChIP), followed by high-throughput sequencing of bound DNA on an Illumina genome analyzer (ChIP-sequencing [ChIP-seq]) (Johnson et al., 2007; Mikkelsen et al., 2007). We found that the WCC binds to hundreds of sites, predominantly upstream of genes. WCC targets fell into various functional categories, but genes encoding TFs were overrepresented. We tested all TF genes and numerous other genes with WCC binding sites in their promoters for light induction and found that, as expected, most, but not all, responded to light. I show that one of the genes encoding a TF, *adv-1*, is a clock-controlled gene that is necessary for circadian rhythms in development. Our data suggest a "flat" hierarchical network in which 20% of all annotated *N. crassa* TFs are regulated during the early light response by the WCC, the key TF factor of the circadian clock.

RESULTS

WCC ChIP

To identify direct targets of the WCC, we performed ChIP-seq with anti-WC-2 antibody on cultures subjected to an 8-min light pulse. We verified that the ChIP-seq library was enriched for known targets of the WCC (Appendix C Figure S-1). Of 4,866,015 32-nt-long ELAND-processed (Johnson et al., 2007) sequences, 92% were mapped to assembly 7 of the *N. crassa* genome by SOAP (Li et al., 2008b). Sequence reads are available through the NCBI sequence read archive (SRA010801.1).

The WCC binding sites were compared to ChIP-seq data for H3K4me2 (K. M. Smith, C. M. Sullivan, K. R. Pomraning, and M. Freitag, unpublished data) because the presence of this epigenetic modification is correlated with transcriptional activity. To establish a cutoff value for statistically significant WCC binding sites, we used the CASHX mapping algorithm (Fahlgren et al., 2009) and sliding-window read counts to calculate the mean and standard deviation of reads per 500 bp for the entire genome. A count of 83 reads per 500-bp window was significantly above the mean level of background signal ($P<0.001$ and $z=3.09$).

In previous studies, only the *frq*, *vvd*, *al-3*, *fl*, and *sub-1* promoters had been identified as direct targets for the WCC (Chen et al., 2009; Froehlich et al., 2002; He and Liu, 2005b; Olmedo et al., 2010a). By ChIP-seq, we identified 400 significant regions of WCC enrichment, with 200 falling in known or predicted promoters of at least one gene. Here, we focus on two groups: (i) the most significant peaks located in promoters and (ii) peaks in promoters of TF genes (Table 3-1). We also compiled a complete list of regions with significant reads, including genes involved in the circadian clock, chromatin function, kinases, phosphatases, cell cycle, DNA replication, DNA repair, and metabolism (Appendix G File S1). Target genes were summarized according to their functional categories. While almost half of all target genes encode unclassified proteins, we noticed an enrichment for genes with functions in the cell cycle, transcription, protein binding, response to the environment, and cellular components when we compared the best 109 targets (z score>5; $P<2.9e^{-7}$) to all 584 targets (z score>3.09; P<0.001) (Figure 3-1A; Appendix C Figure S3).

The Weeder (Pavesi et al., 2004) and SCOPE (Carlson et al., 2007) algorithms were used to derive a consensus binding site for WCC based on 1-kb regions centered on the WCC ChIP-seq peak at 29 genes that were confirmed to be light induced (Figure 3-1B). Both algorithms identified a common consensus binding site, GATCGA (with variability in the first and last bases), which extends the most recently published consensus (GATC) that was derived from studies with expression arrays (Chen et al., 2009). Every 1-kb

Table 3-1. Regions of WCC enrichment after 8 min of light induction

Contig[a]	Start nt	No. of reads	Z score[b]	Gene(s)[c]	Array induction[d]	qPCR[e]
Largest peaks[f]						
7.10	277601	738	45.66	NCU02265 *frq*	2.35 (10)	13
7.12	422401	696	42.93	NCU03967 *vvd*	4.4 (15)	130
7.2	298501	657	40.40	NCU00582 *cry*	4.8 (15)	269
7.57	14601	415	24.68	NCU08699 *bli-4*	4.9 (15)	ND
7.9	872451	387	22.87	NCU03071 *os-4*/NCU03072	1.6 (10)/2.3 (15)	4.2/7.1
7.10	272251	378	22.28	NCU02264/NCU02265 *frqAS*	2.3 (15)/*frqAS*[g]	2.6/13.4
7.48	193601	321	18.58	NCU10063	6.5 (10)	ND
7.12	623601	298	17.09	NCU04021	NF	2.7
7.22	197301	262	14.75	NCU05594	8 (15)	ND
7.7	782201	253	14.17	NCU02800/NCU02801	NF/3.8 (15)	2.4/ND
7.45	138451	235	13.00	NCU07541	2.8 (30)	ND
7.2	196651	230	12.67	NCU00552 *al-1*	14.3 (15)	ND
7.12	426151	226	12.41	NCU03968	1.8 (30)	ND
7.21	6451	210	11.37	NCU11300/NCU06017	NF	4.7/1.7
7.2	305951	200	10.72	NCU00584/NCU00585 *al-2*	3.7 (15)/5.1 (45)	ND
Peaks close to transcription factor genes						
7.13	558401	340	19.81	NCU04179 sah-1 TF	3.4 (30)	4
7.10	644051	245	13.65	NCU02356 *wc-1* TF	3.1 (10)	4.7
7.13	125751	236	13.06	NCU04295 TF	1.4 (10)	1.2
7.81	66601	203	10.92	NCU09615 TF	NF	7.1
7.59	47101	182	9.56	NCU09068 *nit-2* TF	NF	NI
7.4	336501	171	8.84	NCU05964 TF	NF	2.2
7.88	31301	170	8.78	NCU09829/NCU09830 TF	NF/11.6 (15)	7.2
7.15	93601	162	8.27	NCU01242/NCU01243 TF	NF/1.8 (10)	2.4/1.3
7.66	66151	158	8.00	NCU01871 TF/NCU01873	NF	NI/1.4
7.54	127201	150	7.48	NCU08480 *hsf-2* TF	NF	1.8
7.6	191451	124	5.79	NCU02713 *csp-1* TF	4.4 (10)	16.4
7.60	44551	122	5.66	NCU08807 *cre-1* TF/NCU08806 *rhp-55*	NF	1.3/2.1
7.38	129051	115	5.20	NCU07392 *adv-1* TF	NF	3.8
7.1	389051	115	5.20	NCU00097 *bek-1*TF	NF	5.5
7.47	109401	107	4.69	NCU07705 TF	NF	1.2
7.15	391951	96	3.97	NCU01154 *sub-1* TF	1.8 (15)	8.1
7.28	101201	92	3.71	NCU06534 TF/NCU06536	NF	NI/NI
7.16	202151	91	3.61	NCU04731 *sah-2* TF	NF	1.2
7.47	25201	90	3.58	NCU07728 *sre* TF	NF	3.1
7.4	173001	89	3.52	NCU07846 TF	NF	2
7.1	1039001	87	3.39	NCU00275 TF/NCU00276 *mip-1*	NF/1.7 (30)	2.5
7.3	949151	86	3.32	NCU02094 *vad-2* TF/NCU02095	NF	2.3/NI
7.4	200401	86	3.32	NCU05994 TF	NF	NI
7.9	175701	85	3.26	NCU03273 TF/NCU03271	NF	NI/NI
7.9	498501	85	3.26	NCU03184 TF	NF	1.7
7.51	99701	84	3.19	NCU08000 TF	NF	NI
7.52	161651	83	3.13	NCU08159 TF	NF	1.9
7.21	260701	83	3.13	NCU06095 TF	1.8 (10)	6.2

[a] Assembly 7 of the *N. crassa* genome (http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html).

[b] *z* score calculated as described in Materials and Methods.

[c] If binding sites were located between two divergently transcribed genes, both locus names are listed.

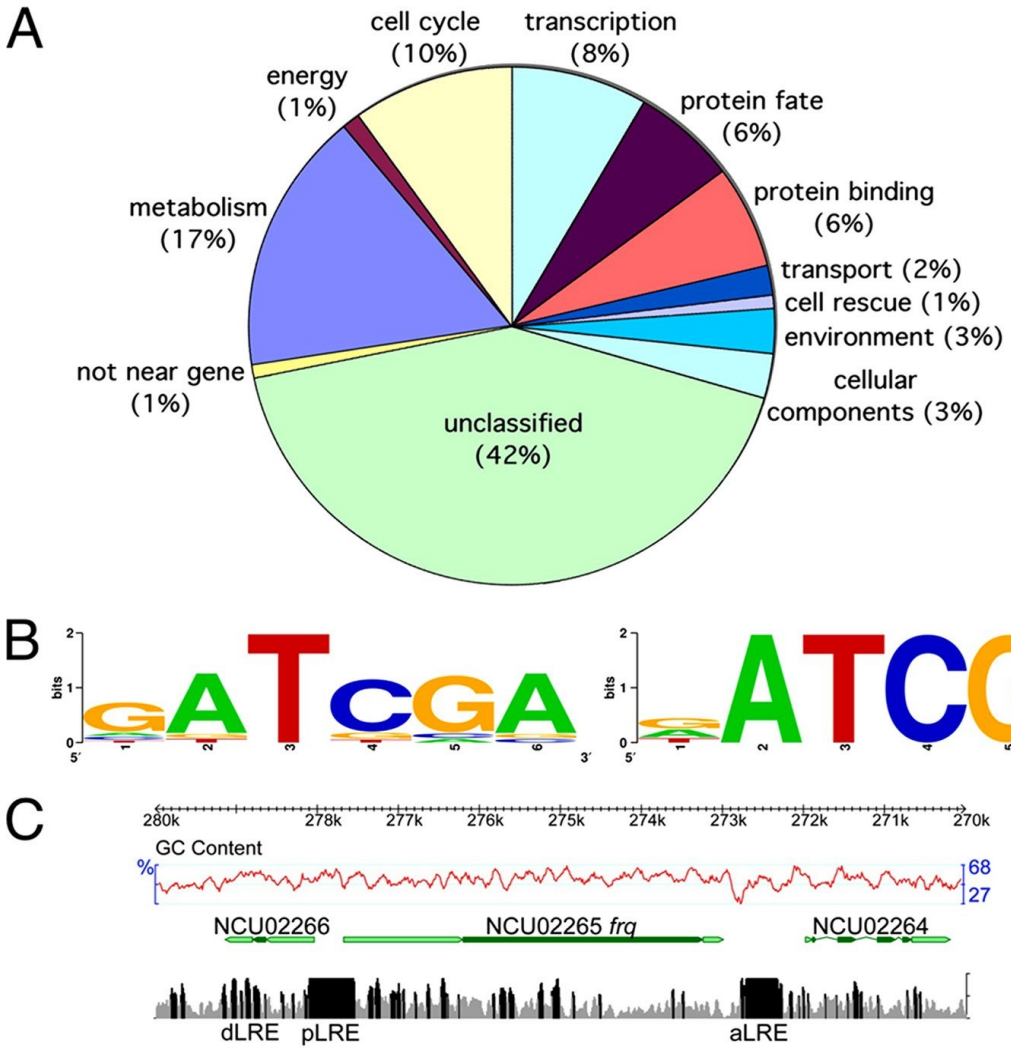
[d] Level of induction (*x*-fold) calculated from raw data (Chen et al., 2009) at the time in minutes shown in parentheses. If two genes were near WC-2 binding sites, the fold induction levels are separated by a slash. NF, not found in array experiments.

[e] Expression level changes determined in this study by qPCR. If genes were known to be light induced from earlier studies and reference 6, qPCR was usually not done. NI, not induced; ND, not determined because there was previous evidence for light induction.

[f] Only regions with 200 or more reads/500-bp sliding window are listed (all peaks are shown in Apenndix G File S1.

[g] The antisense *frq* transcript was not tested in the array experiments.

Figure 3-1. Genomewide analysis of WCC binding. (A) Genes with WCC binding sites fall into diverse functional categories (indicated by the percentage of genes with binding sites within a class). Only the most significant 109 WCC binding sites (*z* score>5; $P<2.9e^{-7}$) were included here, but a comparison between functional categories of the best targets and all targets is shown in Fig.S3 in the supplemental material. (B) Analyses of WCC binding sites found by ChIP-seq with Weeder (Pavesi et al., 2004) (left) and SCOPE (Carlson et al., 2007) (right) revealed a consensus binding site for the WCC. The relative height of each nucleotide (shown in 5'-to-3' direction on the x axis) indicates the degree of sequence conservation, with a maximal score of 2 (as indicated on the y axis). (C) ChIP-seq verified known light response elements (dLRE and pLRE) and revealed a novel binding site (aLRE) for the WCC in the *frq* region. SOAP-mapped sequence reads from the WC-2 ChIP library were plotted as a histogram of sequence coverage per base along a 10-kb fragment containing the *frq* locus (the negative strand of contig 7.10; nt 270,000 to 280,000) in a customized genome browser to integrate ChIP-seq results with gene annotation data http://gb.fungalgenomes.org/gb/gbrowse/neurospora_crassa_OR74A_7/). The y axis has a maximum value of 20 reads per base, and regions with 10 reads are shown in black, with others in gray. Gene annotations are based on known or predicted cDNAs (UTRs are shown in light green, coding regions in dark green, and introns as thin lines). (D) WCC binding results in light induction of *frq*, antisense *frq*, and NCU02264 transcripts and is dependent on WC-2. Quantitative RT-PCR was performed on RNA isolated from control (black bars) and Δ*wc-2* (white bars) strains grown in the dark (D) or following a 15-min light induction (L). The values shown are the averages of two replicates, normalized to the level of actin expression (y axis). The error bars indicate standard deviations.


A
cell cycle (10%)
transcription (8%)
energy (1%)
protein fate (6%)
protein binding (6%)
metabolism (17%)
transport (2%)
cell rescue (1%)
environment (3%)
not near gene (1%)
cellular components (3%)
unclassified (42%)
B
bits
C
280k
278k
277k
276k
275k
274k
273k
272k
271k
270k
GC Content
%
68
27
NCU02266
NCU02265 frq
NCU02264
dLRE
pLRE
aLRE
D
NCU02265 frq
NCU02265 as frq
NCU02264
D L D L

fragment used to generate the consensus contained at least one, but often two or more, copies of this motif interrupted by a variable number of nucleotides.

For a known WCC target gene, *frq*, we observed peaks in two regions of the *frq* promoter that corresponded exactly to the previously identified pLRE and dLRE sites (Froehlich et al., 2002; He and Liu, 2005b) (Figure 3-1C). Enrichment at the frq dLRE (or C box), which deviates from the consensus WCC binding motif, was less pronounced than at the pLRE. WCC binding at the pLRE correlated with increased *frq* transcription in response to light (Figure 3-1D, black bars). This induction was absent in a Δ*wc-2* strain (Figure 3-1D, white bars). We found an additional peak downstream of the *frq* coding sequence, which we named the "antisense LRE" (aLRE) because we predicted it would control light-induced expression of the antisense *frq* transcript (Crosthwaite et al., 1995).

The aLRE also falls in the promoter of NCU02264, encoding a predicted protein with a prefoldin chaperone domain. Binding of WCC to the aLRE is associated with induction of both the antisense *frq* transcript and the divergently transcribed NCU02264 (Figure 3-1D). Approximately 15% of all WCC binding sites in promoters occur between two divergently transcribed genes, and most often, both genes are regulated by one shared binding site (Table 3-1). We measured the light induction of transcription by quantitative PCR (qPCR) and compared our results to previous results from microarray experiments (Chen et al., 2009) performed under similar conditions (Table 3-1). This list included

known clock and light-regulated genes and several novel genes. A replicate WC-2 ChIP was used to validate the ChIP-seq results by using duplex ChIP-PCRs for numerous regions. In all cases, enrichment shown by ChIP-PCR validated the WC-2 ChIP-seq results. Of the other known WCC binding sites, both *vvd* and *sub-1*, listed in Table 3-1, have highly significant WCC binding sites in their promoters. Both were also light induced (Table 3-1; Appendix C Figure S2A for *sub-1*). The *vvd* and *al-3* promoter fragments were used in ChIP-PCRs to check the quality of the ChIP libraries prior to sequencing (Appendix C Figure S1), and both were enriched in the library. Even though *al-3* promoter sequence was present in the library, the number of reads sequenced from this region fell below our cutoff for significance. The *fl* promoter was recently shown to be a target of WCC (Olmedo et al., 2010a) , but we did not find enrichment at this promoter by ChIP-seq. The *al-3* PCR product band was much fainter than the *frq* and *vvd* bands (Appendix C Figure S1). These disparities may be explained by weak and/or transient binding in these regions.

Identification of novel WCC binding sites in promoters of transcription factors

We chose to focus on WCC binding near TF genes as the first step in unraveling transcriptional networks that respond to light and are controlled by the circadian clock. We initially predicted that light signaling by the WCC would involve the activation of a few key downstream TFs that would, in turn, control a network of target genes. Instead, we identified 28 known or putative TF genes with significant ($P<0.001$) WCC binding in

their promoters (Table 3-1; see Appendix C Figure S2). To discover if these TF genes were true targets for the WCC, we analyzed transcript levels by quantitative RT-PCR in response to a 15-min light exposure in *wc-2*$^{+}$ (*bd*) and *wc-2* deletion (*bd* Δ*wc-2)* strains because we expected that direct targets of WCC would be induced by light and that light induction requires WC-2, an obligatory subunit of the WCC (Cheng et al., 2002b; Froehlich et al., 2002; He and Liu, 2005b; Talora et al., 1999).

We found that of the 28 TF genes with significant WCC binding sites upstream of their transcriptional start sites, 21 were induced by light in a WC-2-dependent fashion (Table 3-1; Appendix C Figure S2A). Three genes (*nit-2*, NCU06534, and NCU08000) were not regulated by light under our conditions, but expression at wild-type levels was dependent on WC-2 (Appendix C Figure S2B). Four TF genes (NCU05994, NCU03273, NCU01871, and *cre-1*) were excluded as WCC targets, at least under our conditions (Appendix C Figure S2C). No target has been identified for the peaks in the promoters of NCU05994 and NCU03273, but for NCU01871 and *cre-1*, the neighboring genes (NCU01873 and *rhp55*, respectively) were light inducible in a WC-2-dependent manner, suggesting that they are the actual targets for the WCC.

Of the 21 TF genes we identified here as WCC targets, only 4 (*wc-1*, *sub-1*, *csp-1*, and *sah-1*) had been previously identified as light regulated (Chen et al., 2009), and *sub-1*, *csp-1*, and *sah-1* were not known to be involved in circadian output pathways. Conversely, two TF genes identified as light induced in the microarray study lacked

high-confidence WCC binding sites in their promoters in our study (NCU06407 *vad-3* and NCU03643, the gene encoding cutinase TF-1). However, both genes had a single perfect match to the WCC consensus site (GATCGA) and several sites with a single mismatch (GATCCA) within 1 kb of their predicted initiation codons. Thus, while the two studies were largely congruent, ChIP-seq identified more direct first-tier targets and was more suitable for detecting regulation of TF genes, which are often expressed at low levels, or whose expression levels are changed within a narrow range.

Examples of genes regulated by single or multiple WCC binding sites are shown in Figure 3-2, and data for the most significant peaks are summarized in Table 3-1. We found a single WCC binding site upstream of the gene for heat shock factor 2, *hsf-2*, and transcription of *hsf-2* was light inducible and dependent on WC-2 (Figure 3-2A). We found three WCC binding sites upstream of the *wc-1* gene (Figure 3-2B), which encodes the limiting subunit of the WCC. Transcription of *wc-1* was induced by light and dependent on WC-2, in agreement with previous reports (Ballario et al., 1996). Our data show that *wc-1* is directly regulated by the WCC in a positive-feedback loop, as had been suggested by previous work (Kaldi et al., 2006). Transcriptional start sites of the *wc-1* gene have been mapped at -924 and -1,222 bp (Kaldi et al., 2006). One WCC binding site is close to the transcriptional start site at 924. The other WCC binding sites at -3 and -4 kb are, however, located far upstream of the mapped transcriptional start sites. WCC binds to sites far upstream of predicted transcriptional start sites in additional TF genes, e.g., *sah-1*, *sub-1*, *nit-2*, *bek-1*, and NCU05964 (Appendix C Figure S2A). In

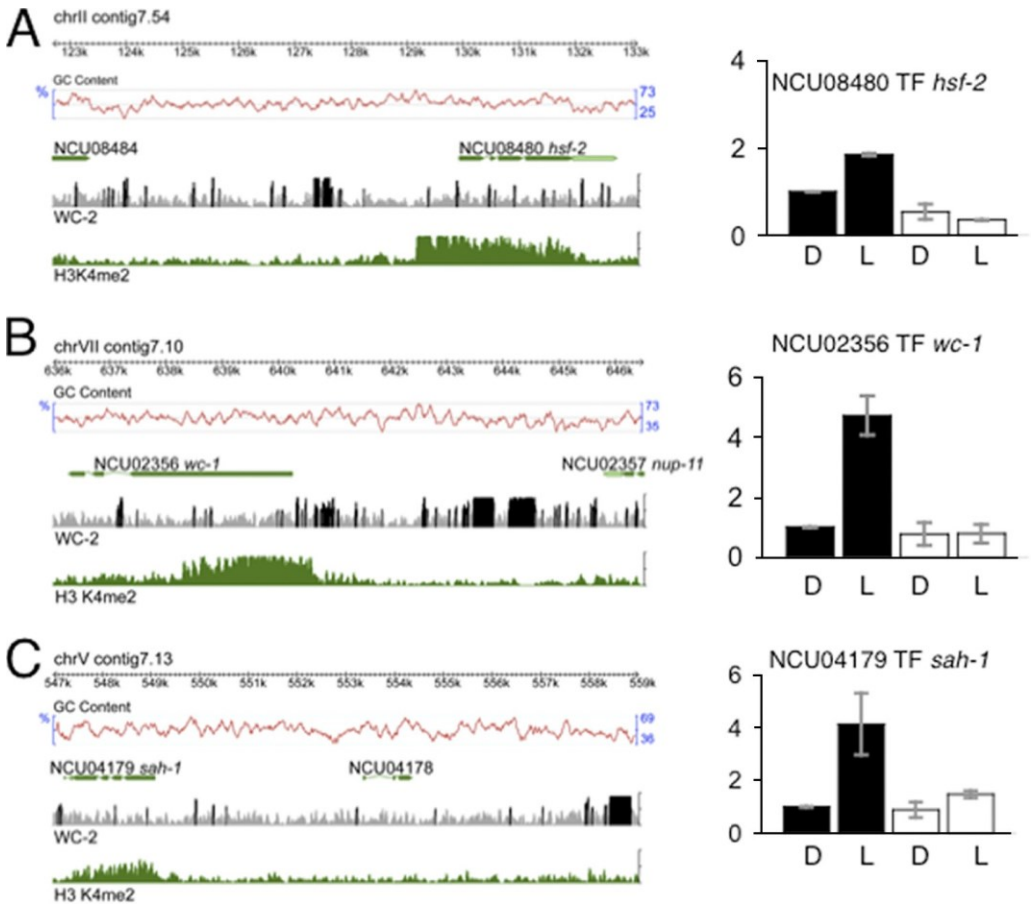


Figure 3-2. ChIP-seq reveals new targets of the WCC in the promoters of TF genes, and these genes are light inducible. (A) The *hsf-2* gene has a single WCC binding site and is light induced in a WC-2-dependent manner. (B) The *wc-1* gene, an early light-induced gene, has three upstream WCC binding sites, one close to and two 3 to 4 kb upstream of the transcriptional start site. (C) The *sah-1* gene has a single WCC binding site 9 kb upstream of its predicted transcriptional start site. Tracks, colors, and conditions for qRT-PCR are the same as for Figure 3-1 C and D. The error bars indicate standard deviations.

particular, a single strong WCC binding site is located about 10 kb upstream of the open reading frame (ORF) of *sah-1* (Figure 3-2C), which is light inducible in a WC-2-dependent manner. Thus, WCC binding appears to be capable of exerting long-range effects in the relatively compact *N. crassa* genome, where promoters are typically short and sometimes overlapping (Galagan et al., 2003). The precise mechanism of the enhancer-like effects of WCC merits further investigation.

Developmental stages that have been associated with light regulation, i.e., asexual sporulation (conidiation) and sexual development, are represented by nine TFs whose mutation results in developmental phenotypes (*hsf-2*, *adv-1*, *sub-1*, *sah-1*, *sah-2*, *bek-1*, *vad-2*, *csp-1*, and *ghh*) (Colot et al., 2006; Thompson et al., 2008). All nine genes are induced by light, and this response is abolished in a wc-2 deletion strain (Table 3-1; Appendix C Figure S2A).

WCC binding sites were associated with six TF genes involved in metabolism or stress response (*nit-2*, NCU05994, *sre* [regulator of iron uptake] (Zhou et al., 1998), *cre-1*, NCU08000, and NCU05964 [homolog of Aspergillus *vosA*] (Ni and Yu, 2007)). Three (*cre-1*, *vosA*, and *sre*) were light induced, and this response was abolished in the *wc-2* mutant, but the other three genes showed no light induction under our conditions (Appendix C Figure S2C). The expression levels of these genes in the *wc-2* strain were decreased. These results suggest that WCC directly impinges on metabolic pathways.

The downstream targets of 11 TFs identified here remain completely unknown. These putative TF genes (NCU00275, NCU01243, NCU03184, NCU03273, NCU04295, NCU06534, NCU07705, NCU07846, NCU08159, NCU09615, and NCU09829) have no previously described function, but they encode motifs that match well-studied DNA binding domains (Colot et al., 2006). NCU00275, NCU01243, NCU03184, NCU04295, NCU07705, NCU07846, NCU08159, NCU09615, and NCU09829 were light induced, a response abolished in the *wc-2* mutant (Appendix C Figure S2A).

The WCC activates an output pathway required for rhythmic spore development

To investigate the link between light and circadian clock pathways regulated by the WCC, I assayed available knockout mutants of the WCC target TFs for changes in circadian rhythms in development (data not shown). The most striking phenotype was observed in the *adv-1* mutant, so I investigated this mutant further for clock defects. Loss of *adv-1* had only minor effects on the development of spores, but development was no longer under the control of the clock (Figure 3-3A).

To determine if arrhythmicity of the *adv-1* mutant was due to an effect on the FRQ/WCC oscillator, I assayed rhythms in the accumulation of the FRQ protein in control and *adv-1* strains over the course of 2 days in cultures grown in the dark (Figure 3-3B and data not shown). No differences were observed in the FRQ rhythms in the mutant strains, suggesting that ADV-1 functions downstream of the oscillator within an

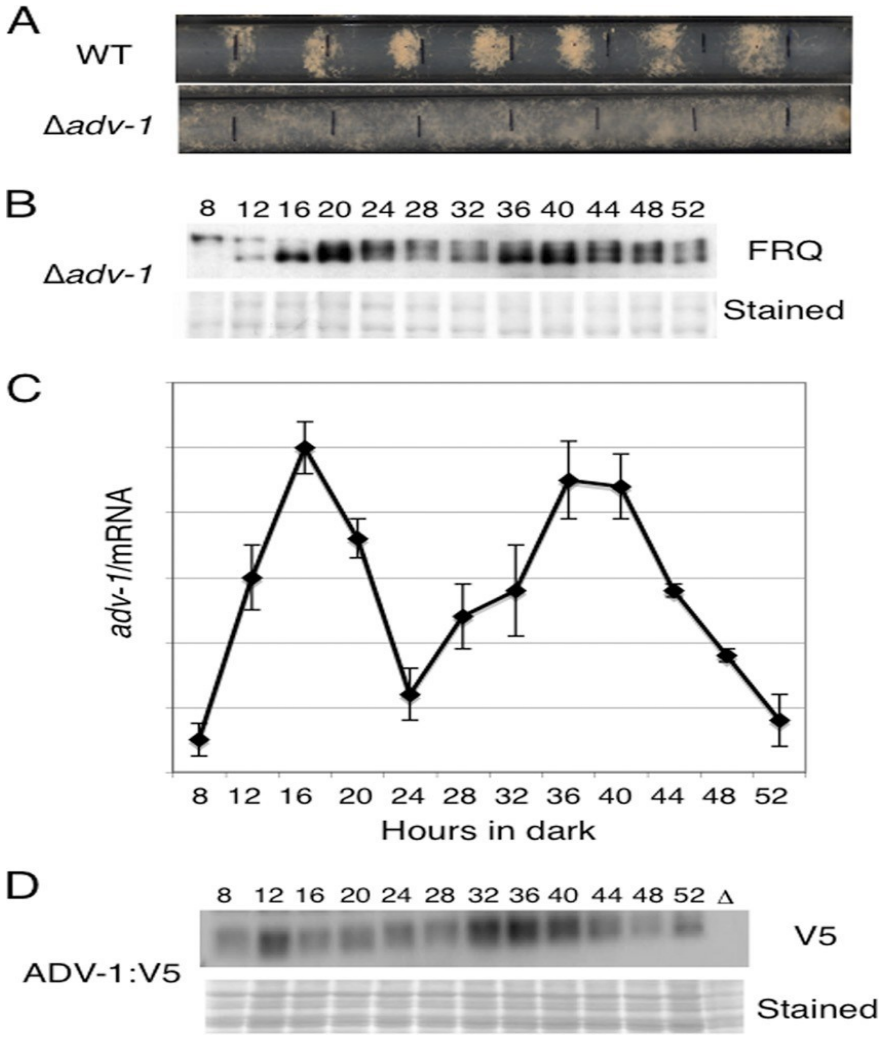


Figure 3-3. *adv-1* is a clock-controlled gene required for circadian rhythms in development. (A) Δ*adv-1* strains are arrhythmic. Shown is a race tube assay for rhythmic development in wild-type clock (WT) and Δ*adv-1* strains. Each strain was inoculated and grown in constant light for 24 h before transfer to constant darkness at 25°C, after which the growth front was marked every 24 h (black lines). (B) FRQ protein remains rhythmic in strains lacking ADV-1. Shown is a Western blot of FRQ protein in the indicated strain. The numbers of hours in constant darkness are shown above the blots. (C) *adv-1* mRNA accumulates with a circadian rhythm. Northern blots of *adv-1* mRNA levels were performed three times, and the average and SD for each time point were calculated and plotted. For each experiment, the *adv-1* mRNA from each time point was normalized to rRNA, and the lowest point was set to 1. The number of hours in constant darkness are shown on the x axis. (D) ADV-1 protein accumulates with a circadian rhythm. Shown is a Western blot of ADV-1–V5 probed with antibody to V5 from cultures harvested at the indicated times in the dark. Protein from the *adv-1* deletion strain (Δ) harvested after 16 h in the dark was used as a negative control to demonstrate the specificity of the antibody. Amido black-stained membranes are shown as loading controls in the Western blots in panels B and D (Stained). Each experiment was repeated at least two times with similar results.

output pathway from the clock. In support of this idea and consistent with previous microarray results (Dong et al., 2008), I observed that the accumulation of *adv-1* mRNA is rhythmic, peaking in the subjective morning (Figure 3-3C). This is similar to the time of peak WCC activity in the circadian clock (Schafmeier et al., 2005). Furthermore, I found that ADV-1 protein tagged with a V5 epitope accumulated with a circadian rhythm (Figure 3-3D). Future efforts will determine if the WCC consensus binding site present in the *adv-1* promoter (at -735 from the predicted start of transcription) is essential for circadian rhythmicity of *adv-1* mRNA accumulation and for the developmental rhythm.

DISCUSSION

In summary, we have identified a large number of direct targets of the WCC, the key regulator of the circadian clock. These targets include 24 TFs that have the potential to control downstream target genes on a second hierarchical level. Genome wide identification of target genes will allow us to build a detailed network of the early and late light responses, as well as circadian clock output pathways. Substantial effort over the past several years has been spent on identifying key components of circadian output pathways in *N. crassa* and other organisms, with only limited success (Cheng et al., 2002a; Taghert and Shafer, 2006; Vitalini et al., 2006). By applying WCC ChIP-seq to this effort, we have now uncovered excellent candidates for output pathway components, such as ADV-1, that lie directly downstream of the oscillator and that regulate distinct

overt rhythms. In addition to gene products that may mediate posttranslational clock and light effects, we also found many WCC binding sites in promoters of metabolic genes that act in various anabolic and catabolic pathways.

Most, but not all, of the genes that are direct targets of the WCC were found in our experiments to be light induced. This may not be surprising for genes that are subject to complex regulation. For example, a gene may be repressed under most growth conditions and be activated by light and the bound WCC only when these conditions are not met and repression is removed. Similarly, while we expect most, if not all, genes directly regulated by the WCC to be rhythmically transcribed, other regulatory elements may take precedence under certain growth conditions.

While many of the downstream genes regulated by the WCC are not well studied or are uncharacterized, most of them have homologs in plants and mammals. Thus, our work provides key information to refine the growing network of light- and clock-regulated genes in a genetically and biochemically tractable model organism.

MATERIALS AND METHODS

ChIP

N. crassa cultures (FGSC2489) were grown at 25°C in minimal medium (1X Vogel's salts, 2% glucose) in the light for 2 days, transferred to the dark for 12 h, light induced (160μE) for 8 min, and cross-linked in constant light with 1% paraformaldehyde for 15 min. We isolated nuclei (Luo et al., 1998) and performed ChIP as described previously (Johnson et al., 2002) on 6 mg of nuclear fraction with 2μg of WC-2 antibody (polyclonal antibody raised to a WC-2 protein fragment expressed in *Escherichia coli*) (Talora et al., 1999). Histone ChIPs were performed on germinated conidia with dimethylated H3 lysine 4 (H3K4me2) (Upstate; 07-030) antibodies as described previously (Tamaru et al., 2003). All ChIP experiments were validated by ChIP-PCR before (not shown) and after (Appendix C Figure S1) ChIP-seq library construction to verify previously described results.

ChIP-seq library construction and high-throughput sequencing

DNA was end repaired and ligated to adapters (Pomraning et al., 2009). The 200- to 500-bp fragments were gel purified and amplified with 20 cycles of PCR using Phusion polymerase (Finnzymes Oy). The PCR products were gel purified and sequenced on an

Illumina 1G sequencer in the Oregon State University Center for Genome Research and Biocomputing (OSU CGRB) core laboratories.

Region-specific ChIP PCR

Duplex PCR with [α^{32}P] dCTP was performed to determine enrichment in the ChIP samples relative to input DNA with region-specific oligonucleotide primers (Appendix D Table S1) (Pomraning et al., 2009; Tamaru et al., 2003). An *hH4-1* segment was used as a control. Phosphorimager screens were exposed to dried gels and analyzed with a GE Storm 820 imager.

Data analysis and visualization

Illumina reads were mapped to the *N. crassa* assembly (http://www.broadinstitute.org /annotation/genome/neurospora/MultiHome.html) genome with SOAP (Li et al., 2008b), with the following parameters: -s 10-c 41 -r 1 -g 1 -w 10,000 (seed size, 10; per-read incremental trimming of 3' basepairs to remove low-quality or nonaligning bases, choosing randomly for equal best locations for a read; only a single random location placement of reads; a maximum gap size of 1; and a maximum of 10,000 mapped locations).

A second analysis used CASHX (17) to map perfect 36-nucleotide (nt) matches to the genome, calculate reads per window across the whole genome, and assign significance values to areas defined as peaks. Only reads matching the genome 1 or 2 times were included in the analysis. We calculated the average read number per 500-bp sliding window, the standard deviation, and a *z* score for each window. ChIP-seq of the relatively compact *N. crassa* genome (~42 Mb) typically results in some background coverage (e.g., ChIP-seq data for H3K4me2 [Appendix C Figure S2]), which yields fewer false-negative calls than in larger, more complex genomes. Thus, complete absence of reads with respect to a given region of the reference genome strongly suggests that the region has been deleted from the resequenced strain (J. E. Stajich, K. M. Smith, and M. Freitag, unpublished data).

The updated *N. crassa* annotation tracks (Figure 3-1 and 3-2; Appendix C Figure S2) include extended transcript predictions of untranslated regions (UTRs) by using 279,323 *Neurospora tetrasperma* and 453,559 *Neurospora discreta* expressed sequence tags (ESTs) generated with 454 technology for the respective genome projects (Joint Genome Institute and J. E. Stajich, D. J. Jacobson, D. O. Natvig, N. L. Glass, and J. W. Taylor, unpublished data), as well as 84,309 *N. crassa* ESTs and full-length cDNAs available from GenBank and the NCBI Trace archive. This updated *N. crassa* annotation was produced with the PASA annotation system (Campbell et al., 2006) with assembly 7 and annotation from GenBank (AABX02000000; 10 September 2007). Visualization of

annotations and mapped read densities was done with Gbrowse (Stein et al., 2002) with wiggle track format.

Sequence analysis to find WCC consensus binding sites. Known WCC-bound promoters of light-induced genes (*frq*, *vvd*, and *al-3*), plus an additional 26 WCC binding sites found in this study at genes confirmed to be light induced, were extracted from assembly 7. One-kilobase regions centered on each peak were searched de novo for consensus cis elements with Weeder version 1.3.1 (parameters, NC large A M S T15) (Pavesi et al., 2004) and SCOPE (Carlson et al., 2007); http://genie.dartmouth.edu/scope/. The consensus motif logos in Figure 3-1B were generated using WebLogo version 2 (http://weblogo.berkeley.edu/logo.cgi).

Reverse transcription (RT)-PCR to assay light induction

We used standard laboratory *ras*^{bd} (Gorl et al., 2001) clock strains and Δ*wc2 ras*^{bd} (Collett et al., 2001) mutant strains for light induction experiments. The *ras*^{bd} mutation enhances the developmental rhythm in cultures (Sargent et al., 1966) but does not affect light responses or molecular rhythms in gene expression. The strains were grown for 24 h in 200 ml of Vogel's minimal medium (Davis, 2000) with 2% glucose and 0.5% arginine in constant dark and subjected to 15 min of full-spectrum light (160μE). We chose a 15-min time point to better compare our results with published results and to allow time for RNA accumulation. Total RNA was prepared using the peqGOLD

TriFAST kit (peqLab, Erlangen, Germany). We generated cDNA by reverse transcription with primers supplied with the QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Expression was analyzed by quantitative real-time PCR (Gorl et al., 2001) with the primers listed in Appendix D Table S2 in the supplemental material. We amplified cDNA for *act*, *frq*, and *frq* antisense RNA with specific primers (Appendix D Table S2).

Circadian rhythm assays

We crossed Δ*adv-1* to ras^{bd} to generate ras^{bd} Δ*adv-1* strains. The ras^{bd} Δ*adv-1* strain was assayed on standard race tube medium containing 1X Vogel's salts, 0.1% glucose, 0.5% arginine, and 1.5% agar. Race tube assays were performed in controlled environmental chambers in constant darkness at 25°C. For RNA and protein analyses, mycelial mats were grown in shaking liquid culture (100 rpm) in 25 ml of Vogel's minimal medium (1X Vogel's salts, 2% glucose, 0.5% arginine) and processed as described previously (Garceau et al., 1997). Transfer of cultures was performed so that the cultures were approximately the same age (within 8 h) at the time of harvest but the circadian times of the cultures varied (Correa et al., 2003). The ADV-1:V5-tagged strain was a generous gift from Jay Dunlap's laboratory and the Neurospora Program Project Grant and contained a C-terminal V5 tag. The tagged gene was recombined into the endogenous *adv-1* locus. Antibody for V5 was obtained from Invitrogen (Carlsbad, CA).

CHAPTER IV

THE TRANSCRIPTION FACTOR ADV-1 TRANSDUCES LIGHT SIGNALS, AND TEMPORAL INFORMATION TO CONTROL RHYTHMS IN CELL FUSION GENE EXPRESSION, IN *NEUROSPORA CRASSA*

INTRODUCTION

Circadian clocks, comprised of molecular oscillators, provide a mechanism for organisms to coordinate their physiology and behavior with daily environmental cycles, and to maintain internal temporal order (Bell-Pedersen et al., 2005). External cues, particularly light signals, entrain clocks to 24-h environmental cycles. As such, light and circadian clock signaling pathways are closely intertwined. Temporal information is relayed from circadian oscillators through output pathways to regulate daily rhythms in the expression of clock-controlled genes (ccgs) and overt rhythmicity. While substantial progress has been made in understanding the central clock machinery (Bell-Pedersen et al., 2005; Crane and Young, 2014), and cataloging rhythmic and light-responsive genes (Chen et al., 2009; Duffield, 2003; Hardin, 2000; Hurley et al., 2014; Pfeiffer et al., 2014; Vitalini et al., 2006; Wu et al., 2014), an important remaining challenge is to connect the circadian oscillator to the output pathways to determine how the clock regulates rhythms in biological processes.

N. crassa is one of the best-studied model systems for light responses and circadian rhythms (Vitalini et al., 2010). The core *N. crassa* FRQ/WCC circadian oscillator forms a characteristic negative feedback loop. In the FRQ/WCC oscillator, two PAS-domain containing GATA-type zinc finger transcription factors (TFs), White Collar-1 (WC-1) and White Collar-2 (WC-2) dimerize to form the White Collar Complex (WCC) (Ballario et al., 1996; Cheng et al., 2002b; Denault et al., 2001). WCC functions as a positive element in the oscillator, and activates transcription of the *frequency* (*frq*) gene (Froehlich et al., 2002; Lee et al., 2003; Vitalini et al., 2010). The negative component FRQ accumulates, enters the nucleus, interacts with FRH (FRQ interacting RNA helicase) (Cheng et al., 2005; Shi et al., 2010), and inhibits the activity of the WCC (He et al., 2005; Schafmeier et al., 2008; Schafmeier et al., 2005). Turnover of FRQ triggered by progressive phosphorylation and proteasome-dependent degradation, relieves WCC inhibition and reinitiates the cycle (He et al., 2003; Liu et al., 2000).

WC-1 is also a blue light photoreceptor (Cheng et al., 2002b; Cheng et al., 2003; He et al., 2002; Talora et al., 1999), and with its partner WC-2, functions to regulate light-responsive genes, as well as downstream *ccgs* (Chen et al., 2009; Smith et al., 2010). Upon light exposure, the WCC binds to light-responsive elements (LREs) in the promoters of target light-responsive genes (Froehlich et al., 2002; He and Liu, 2005b; Olmedo et al., 2010b; Smith et al., 2010). Recent RNA-seq studies revealed that at least 31% of expressed genes in *N. crassa* are regulated by light (Wu et al., 2014). In addition, ChIP-seq in cells given a short light pulse to activate the WCC revealed that WCC

binding increased at ~400 sites in the genome, including at the promoters of ~200 genes (Smith et al., 2010). Genes encoding TFs (24 TF genes) were enriched among these direct WCC targets. ChIP-seq data supported expression studies suggesting a hierarchical network in which light-activated WCC directly controls the expression of early light-induced genes, including some TF genes (Chen et al., 2009). These early light-induced TFs would in turn control the expression of late light-responsive genes. Consistent with this idea, the *sub-1* gene, encoding a GATA family TF, is an early light-induced gene whose promoter is bound by the WCC, and deletion of *sub-1* alters light responses for many late-light inducible genes (Chen et al., 2009; Smith et al., 2010). Furthermore, an additional connection between light-induced TFs controlled by the WCC and the clock is seen through the TF CSP-1. The levels of *csp-1* mRNA peak in the subjective morning (Sancar et al., 2011). CSP-1 functions primarily as a repressor, and as a result, some downstream targets of *csp-1* are ccgs that cycle antiphase to rhythmic CSP-1 levels, with peaks occurring in the subjective evening. However, how the other 22 TFs are linked to light signaling and/or circadian output pathways are not known. This information is necessary for generating models that will comprehensively describe the light and circadian TF network.

The *arrested development-1* (*adv-1*) gene promoter is bound by the WCC (Smith et al., 2010), and ADV-1 is a zinc finger TF necessary for vegetative and sexual development in N. crassa. Deleting *adv-1* impairs the formation of protoperithecia, the precursor to female reproductive structures (perithecia), rendering the strain female sterile (Colot et

al., 2006). Δ*adv-1* also lacks the ability to undergo somatic cell fusion during conidial germination, and during hyphal growth in the mature colony (Fu et al., 2011). Consistent with *adv-1* being a direct target of the WCC, the gene is both light- and clock-regulated, suggesting a role for ADV-1 in connecting downstream developmental processes to circadian clock and light signaling pathways (Smith et al., 2010). To elucidate the role of ADV-1 in controlling light and clock output pathways, we used ChIP-seq in WT cells, and RNA-seq in WT versus ΔADV-1 cells, to identify the ADV-1 direct and indirect target genes. Consistent with the Δ*adv-1* phenotype, ADV-1 direct targets were enriched for genes involved in sexual and asexual development, cell fusion, and metabolism.

Cell fusion between genetically identical cells is critical for normal development in organisms ranging from fungi to mammals (Aguilar et al., 2013; Fleissner et al., 2008). For example, in mammals, somatic cell fusion occurs between myoblasts during muscle development (Richardson et al., 2008), and between macrophages in osteoclasts during bone formation (Vignery, 2008). Filamentous fungi serve as simple models for understanding the mechanisms that control cell fusion (Lichius and Lord, 2014). In *N. crassa*, fusion occurs between conidial anastomosis tubes (CATs) during colony initiation, and between hyphae in the mature colony to form an interconnected hyphal network. This process allows the translocation of nuclei and other organelles, nutrients, and both macromolecules and small molecules throughout the colony, facilitating growth and reproduction (Simonin et al., 2012). Cell fusion occurs in three stages. These include (i) chemoattraction, involving mitogen-activated protein kinase (MAPK) pathways, (ii)

adhesion and cell wall remodeling, and (iii) membrane fusion. To date, more than 50 *N. crassa* proteins have been linked to intercellular communication and cell fusion; however, we still lack a detailed understanding of the molecular processes that promote cell fusion (Hickey et al., 2002). Here, we demonstrate that several critical cell fusion genes are clock-controlled. Furthermore, we show that rhythmic mRNA accumulation for some of the cell fusion genes, and phosphorylation and activity of the mitogen-activated protein kinase (MAPK) MAK-1 involved in controlling cell wall integrity and required for chemoattraction during cell fusion, is dependent on ADV-1. These data indicate an unforeseen role for the circadian clock in controlling the timing of cell fusion via rhythmic activation of ADV-1. Furthermore, while the light responses of the target genes directly under control by ADV- were altered as expected in ΔADV-1 cells, the temporal responses were not. These data indicate that, unlike the flat hierarchical TF network observed for controlling light signaling, the circadian TF network is more complex, with feedforward and feedback loops contributing to rhythmic transcription.

RESULTS

ADV-1 protein levels are clock-controlled

The promoter of *adv-1* is bound by the WCC, and *adv-1* mRNA levels are transiently induced by light in WT, but not in Δ*wc-1* cells (Smith et al., 2010). Consistent with WCC binding and activating the *adv-1* promoter, *adv-1* mRNA levels are low in Δ*wc-1*

cells as compared to WT cells, and *adv-1* mRNA accumulates rhythmically in WT, but not *Δwc-1* cells (Smith et al., 2010). To confirm that the rhythm in *adv-1* mRNA levels leads to rhythms in protein abundance, we created a translational fusion between ADV-1 and codon-modified luciferase (LUC) (Gooch et al., 2008). As expected, LUC activity was robustly rhythmic in constant darkness (DD) with a period of 20.3±0.3 h, and with peak levels during the late subjective day (DD20) (Figure 4-1), slightly lagging the DD16 peak previously observed in *adv-1* mRNA levels (Smith et al., 2010). ADV-1-LUC rhythms were abolished in Δ*frq* cells, demonstrating that the rhythm is dependent on a functional FRQ/WCC oscillator (Figure 4-1). These data prompted further investigation of ADV-1's role in light signal transduction pathways, and as a component of a circadian output pathway from the clock to regulate downstream *ccgs*.

ADV-1 direct targets, identified by ChIP-seq, are highly enriched for genes involved in metabolism, development, and cell fusion

To identify downstream genes controlled by ADV-1, we carried out ChIP-seq in a strain containing C-terminal V5 epitope-tagged ADV-1 (ADV-1::V5). WT and ADV-1::V5 cells, but not Δ*adv-1* cells, produced normal protoperithecia, demonstrating that the tagged protein was functional (Figure 4-2A). As expected, ADV-1::V5 levels were induced following light treatment (Figure 4-2B). Therefore, ADV-1 ChIP-seq was carried out on cells grown in the dark (LL0), and after light treatment (LL) for 15, 30 and 60 min to induce ADV-1 levels and genome-wide binding of ADV-1::V5. Sequence

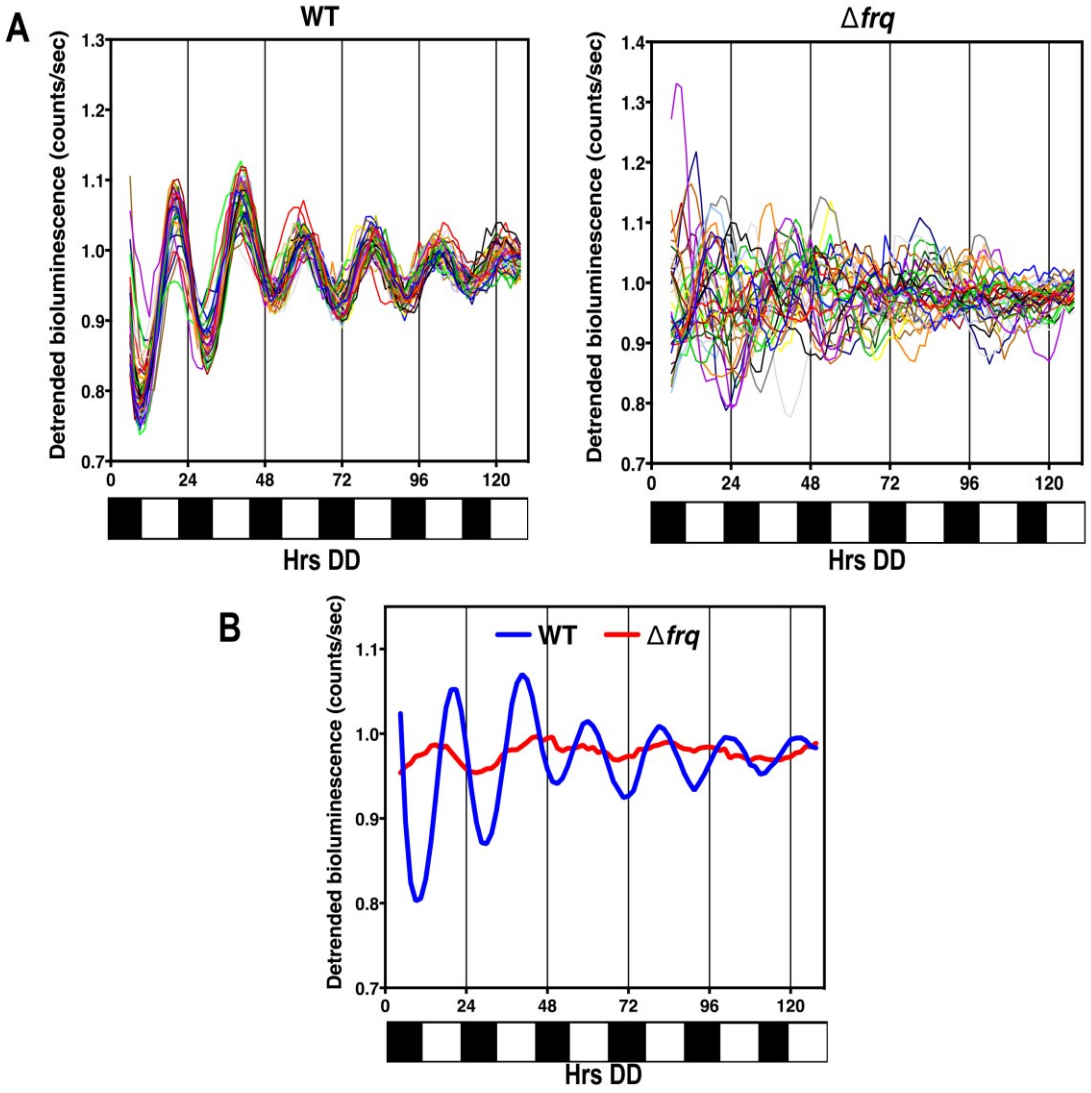


Figure 4-1. ADV-1 protein is regulated by the clock. (A) Individual baseline-detrended bioluminescence recordings of ADV-1-LUC in WT (n=42) (left), and ΔFRQ (n=34) (right) cells. (B) Average of all individual baseline detrended bioluminescence recordings of ADV-1-LUC in WT (blue) and ΔFRQ (red) cells. The time of recording in constant darkness (DD) is given on the x-axis (hrs). The bar graph below each plot represents subjective day (white) or subjective night (black) in DD.

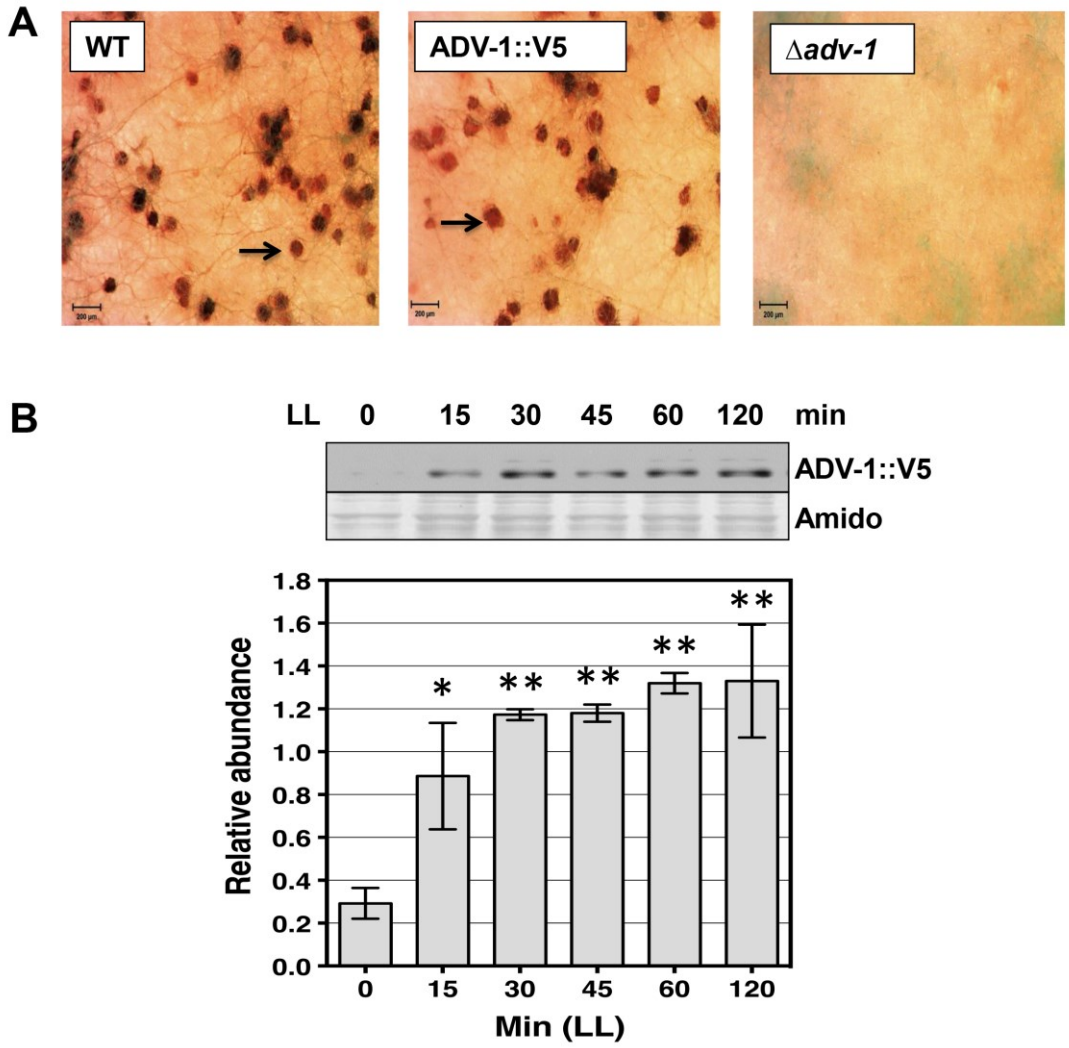


Figure 4-2. ADV-1::V5 is functional. (A) WT, Δ*adv-1* and ADV-1::V5 strains were grown on synthetic crossing medium (SCM) with 0.5% sucrose to initiate female sexual development, and photographed (25X; scale bar=200 µm). Protoperithecia (arrow) were observed in WT (left), and ADV-1::V5 (middle), but not in Δ*adv-1* (right) cells. (B) A representative western blot of ADV-1::V5 protein from cells that were exposed to light for the indicated times (min) and probed with V5 antibody (top panel). The total protein stained with amido black served as the loading control. The data are plotted below (± sd, n=4). The asterisks represent statistically significant differences in ADV-1::V5 protein levels in light treated cells, compared to dark samples (*P<0.05 and **P<0.005, Student's t test).

reads obtained for the ChIP samples (HTS 459-462), and the percent mapped to the *N. crassa* assembly are shown in Appendix E Table S1.

Using a stringent statistical cutoff value for peak calling (false discovery rate [FDR] <0.01), and stipulating that the ADV-1 binding site occurs within 10kb upstream of the transcription start site, 3 kb downstream of the transcription stop site, or within the coding region, 338, 673, 482, and 324 ADV-1-binding sites were identified in the genome from the 0, 15, 30 and 60 min light-treated ChIP samples, respectively. The number of genes assigned to ADV-1-binding sites was greater than the number of binding-site peaks due to the presence of binding sites between two divergently transcribed genes (Appendix H File S1). This ChIP-seq gene set (576 genes at LL0, 1064 genes at LL15, 776 genes at LL30, and 536 genes at LL60) was used for all subsequent analyses. Combining genes from all time points revealed that ADV-1 potentially controls the expression of 1187 unique gene targets (Appendix H File S1), representing ~12% of the predicted 9730 *N. crassa* genes. The distribution of binding peaks for each time point, based on their positions relative to the transcriptional start site from the annotated genes in Release 7 of the *N. crassa* genome (http://www.broad.mit.edu/annotation/genome/neurospora/), were examined (Figure 4-3). In all 4 dark or light conditions, ~60% of the ADV-1 binding sites were present upstream of the transcriptional start site, ~38% were downstream of the coding sequence, and the remaining binding sites (~2%) were located in gene coding regions.

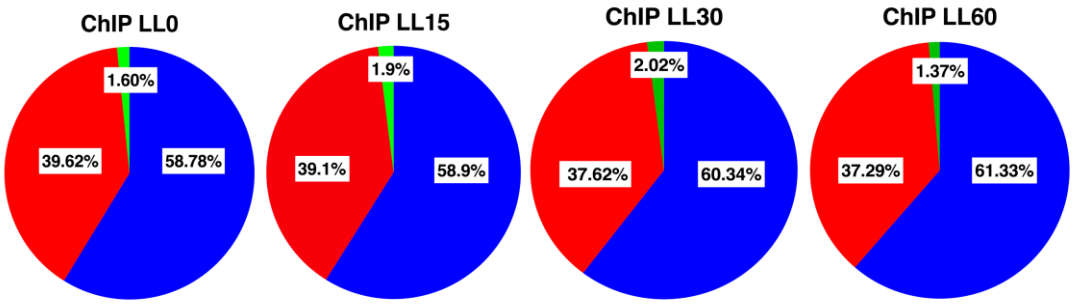


Figure 4-3. Distribution of ADV-1-bound peaks obtained from ChIP-seq. The pie chart shows the distribution of ADV-1 peaks from cells that were harvested for ChIP after growth in the dark (LL0), and after the indicated times (min) in the light (LL). The categories are based on the position of the peak from the nearest transcriptional start site (TSS) of an annotated gene: upstream of the TSS (blue), downstream of the TSS (red) and within the transcript coding sequence (green). The number of peaks in each category divided by the total number of peaks for each time point are given (%).

The presence of ADV-1-bound peaks in the dark suggested that light stimulation is not necessary for ADV-1 to be bound at some genomic regions. The number of ADV-1 binding sites increased ~2-fold after a 15 min light treatment (Appendix H File S1), consistent with the observed increase in ADV-1 protein levels compared to dark grown cultures (Figure 4-2B). Interestingly, a significant increase in binding was not observed in the 30- and 60-min light-treated samples relative to the 15-min sample, despite similar high levels of ADV-1 protein in these cells relative to cells exposed to light for 15-min (Figure 4-2B), suggesting that sufficient ADV-1 is present after 15 min of light to saturate these sites, or that longer light treatments reduce ADV-1 binding to its target sequences.

Gene Ontology (GO) analyses (Boyle et al., 2004) was used to identify the functional categories of genes bound by ADV-1 (Figure 4-4A and Appendix H File S2). An enrichment of GO categories was found for metabolic processes, hyphal growth, and asexual spore (conidia) development at every time point (Figure 4-4A). The light treated ChIP samples were also enriched for light-induced genes at 15 and 30 min, but not at 60 min. ADV-1 deletion strains are defective in the development of protoperithecia (Figure 4-2A; Colot et al., 2006). Consistent with this phenotype, there was enrichment for ADV-1 target genes involved in fruiting body development at all time points examined. Genes involved in cell fusion were also enriched at all time points, consistent with the absence of CATs in ΔADV-1 cells, since CAT formation requires cell fusion, and

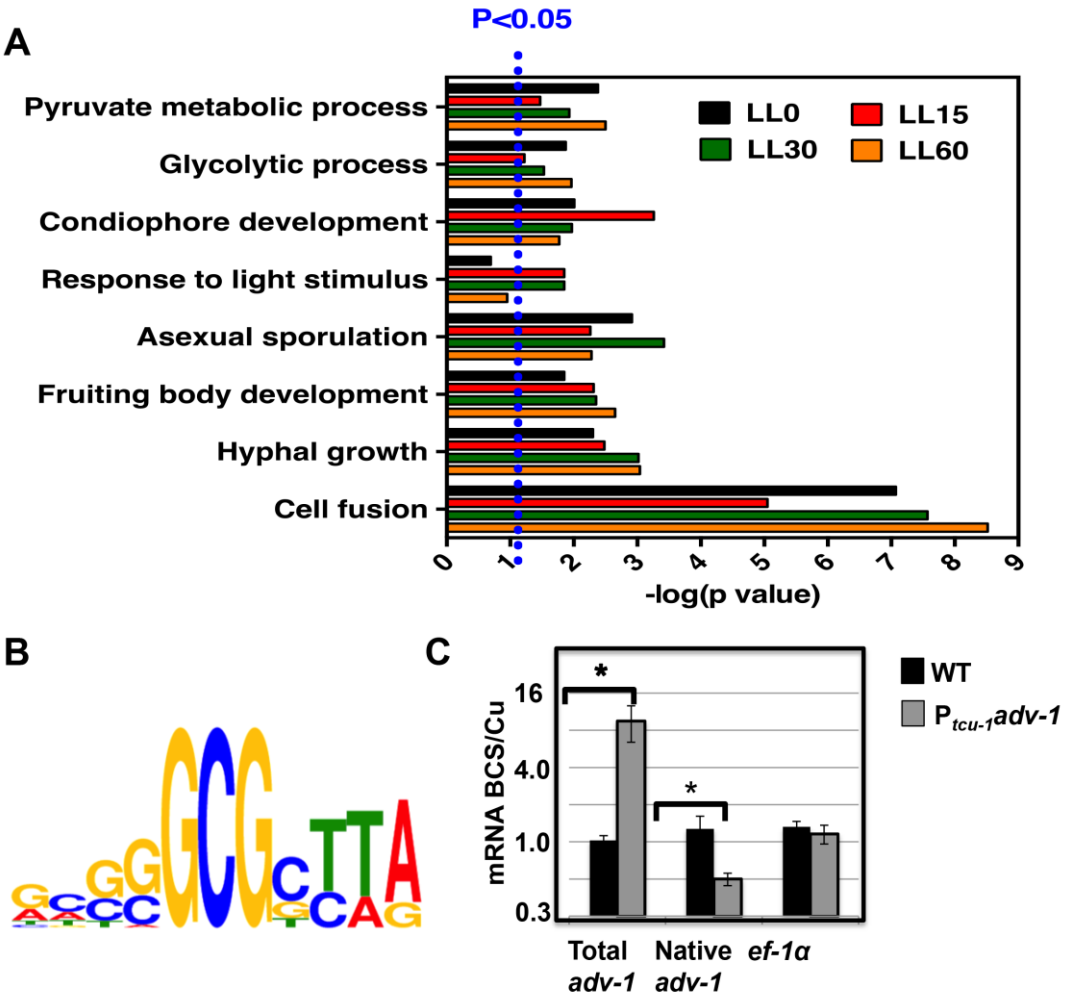


Figure 4-4. ADV-1 targets are enriched for genes involved in light responses, development, and metabolism, and ADV-1 negatively regulates its own expression. (A) Gene Ontology (GO) analysis of predicted ADV-1 direct target genes identified by ChIP-seq from cultures grown in the dark (LL0) and in LL for 15, 30 and 60 min. The GO terms (y axis) for significantly enriched categories are shown ($p \leq 0.05$; indicated by the dotted blue line) at one or more time points (x axis). For visualization, the p values are plotted as the -log. (B) Analyses of ADV-1 binding sites in the 30 most significant peaks identified by ChIP-seq revealed an ADV-1 consensus-binding site. The relative height of each nucleotide (shown in the 5' to 3' direction) reflects the degree of sequence conservation in the ADV-1 consensus-binding site. (C) Plot of *adv-1* and control *ef-1α* mRNA levels from WT cells (gray), and cells containing $P_{tcu\text{-}1}$*adv-1* at an ectopic locus, treated with BCS to induce $P_{tcu\text{-}1}$ promoter activity. The asterisks represent statistical significance by Student's t test ($P<0.002$ for total *adv-1* and native *adv-1*).

ADV-1 functions to regulate the expression of essential cell fusion genes (Fu et al., 2011; Leeder et al., 2013; Read et al., 2012).

A consensus ADV-1 binding motif was derived from the 30 most statistically significant ADV-1 ChIP peaks common to all growth conditions using the MEME motif discovery tool (Bailey et al., 2006) (Figure 4-4B). The core motif, consisting of GCGCTT, was enriched among all ADV-1 binding sites in the light, but only for the top 30 peaks in the dark, suggesting that for some binding sites, ADV-1 potentially recognizes different consensus sequences in the dark versus the light (Figure 4-5). A replicate ADV-1::V5 CHIP was used to validate the ChIP-seq results. ChIP-PCR was used to examine binding at 10 regions that were randomly selected among the predicted binding regions in the ChIP-seq studies. In all cases, enrichment shown by ChIP-PCR validated the ADV-1 ChIP-seq results (Figure 4-6). For NCU04847, two distinct ADV-1-bound peaks are located upstream of the gene; PCR validation confirmed ADV-1 binding at both sites. These experiments also validated binding of ADV-1 to sequences at the 3' end of the *adv-1* gene, suggesting feedback regulation. To determine if ADV-1 controls its own expression, we examined the effect of overexpression of *adv-1* mRNA at an ectopic site in the genome on the endogenous locus. $P_{tcu\text{-}1}$*adv-1* transformants were treated with BCS to induce transcription for the *tcu-1* promoter (Lamb et al., 2013). The total level of *adv-1* mRNA increased significantly following BCS treatment, while the levels of a control gene, *ef-1α*, not bound by ADV-1 were unchanged (Figure 4-4C).

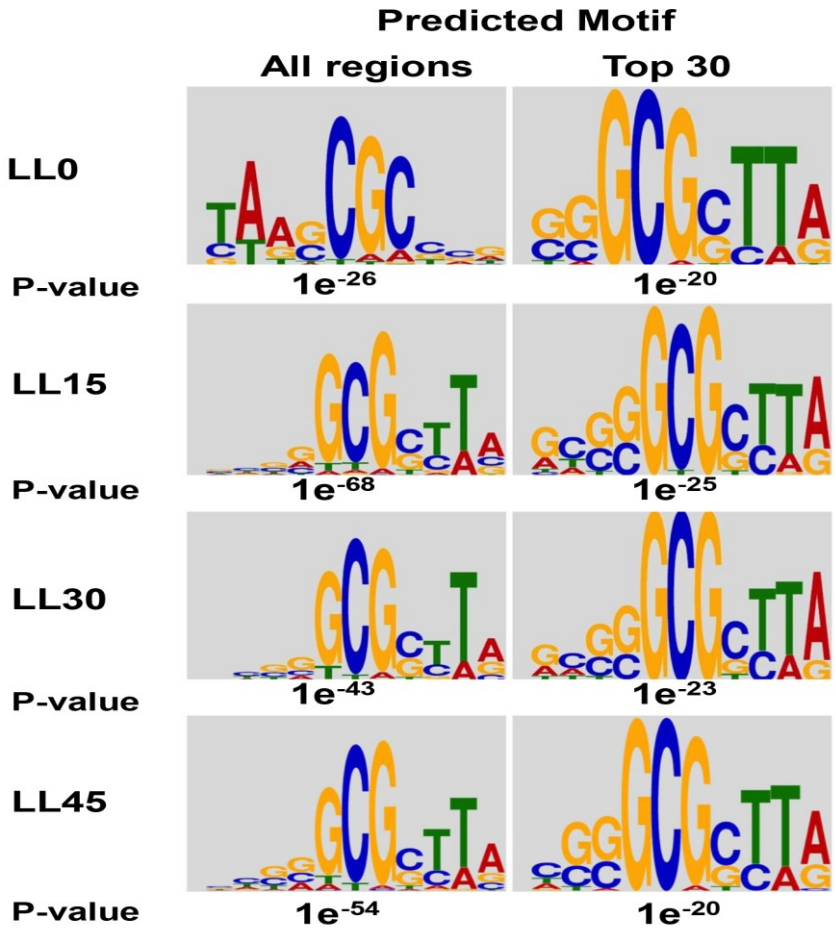


Figure 4-5. Predicted ADV-1 motifs. ADV-1 consensus binding motifs were identified from the entire set of ADV-1-bound peaks, and from the 30 most significant ADV-1 bound peaks under the indicated growth conditions for ChIP. The P-value for the motifs are indicated below.

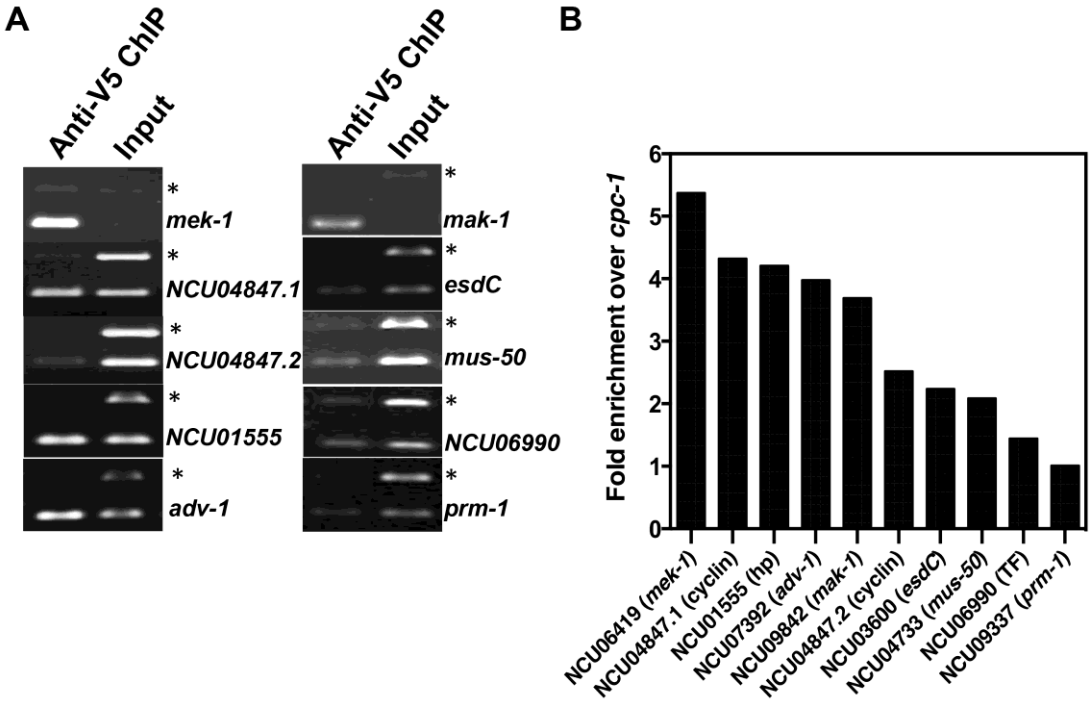


Figure 4-6. Validation of ADV-1 targets by independent ChIP-PCR. (A) Ethidium bromide stained gel of PCR products showing enrichment of ADV-1 bound signals at the promoters of the indicated genes compared to the negative control site *cpc-1* (Anti-V5 ChIP). The amplified *cpc-1* product is marked with an asterisk. Input DNA served as a control for the ChIP (Input). For NCU04847, two different primer sets were used (NCU04847.1 and NCU04847.2) to validate ADV-1 peaks located at two different sites upstream of the gene. (B) Plot of ADV-1 binding to gene promoters. The fold enrichment represents the amount of PCR product derived from amplification of ADV-1 target sites as compared to the negative control site *cpc-1*. TF= predicted TF; HP= hypothetical protein.

However, *adv-1* mRNA levels were significantly decreased at the native locus upon induction of ectopic *adv-1*, demonstrating that ADV-1 has the potential to negatively regulate its own expression.

Additional validation of the ChIP-seq data was accomplished by carrying out ADV-1 ChIP-seq on cells grown in Bird's medium (Appendix E Table S1, HTS 463-466) instead of Vogel's medium (which was used for the light-induction studies described above). A comparison between the two media were of interest because while Vogel's medium is routinely used for *N. crassa* growth in laboratories, Bird's medium was recommended for use in genomic studies due to enhanced pH stability as compared to Vogel's media (Metzenberg, 2004; Perkins, 2006). I found that the correlation between the ADV-1 ChIP binding peaks from cells grown in both Vogel's and Bird's media was high for each condition ($R^2 \geq 0.87$), further validating the ChIP-seq data and demonstrating that there are no major differences in ADV-1 function in the different media (Figure 4-7).

RNA seq reveals complex regulation of ADV-1 target genes

RNA isolated from WT and ΔADV-1 cells grown in the dark, and following light treatment for 15, 30, and 60 min was used for RNA-seq (Table S4). These experiments were carried out using 2 independent biological replicates, and statistical analyses of the

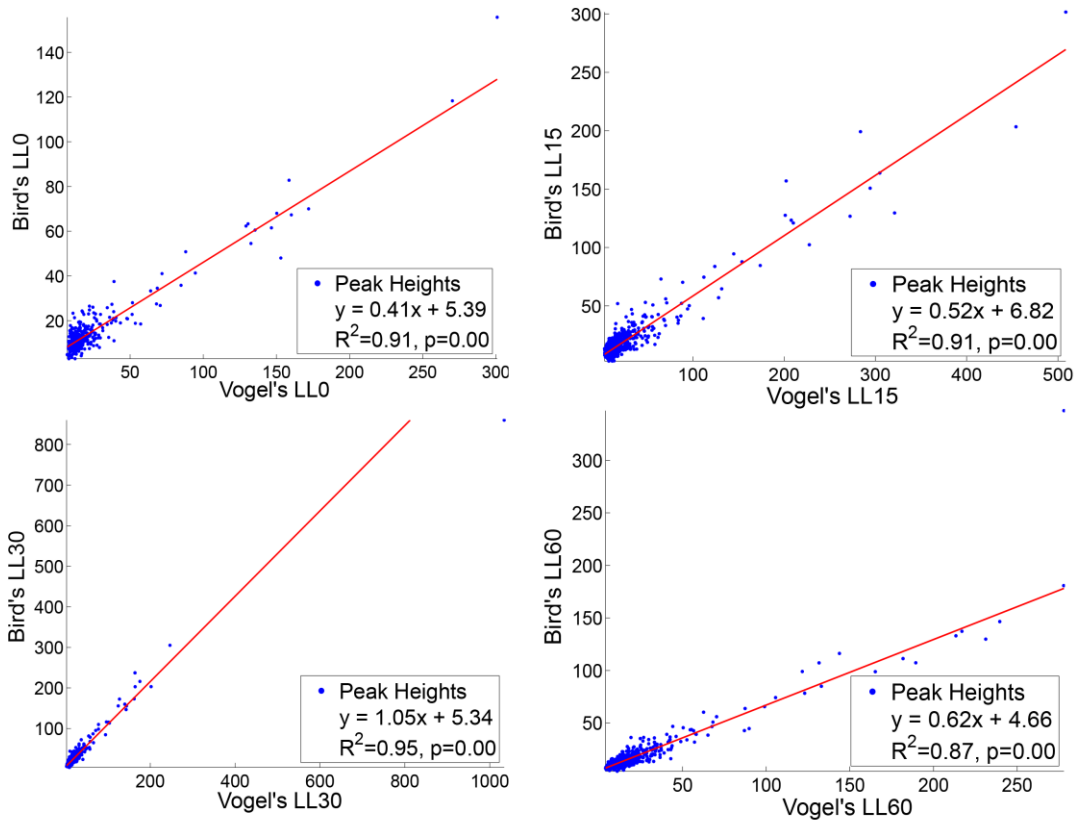


Figure 4-7. Correlation of ADV-1-bound peak heights obtained from ChIP-seq performed in Vogel's and Bird's media. The scatter plots show the coverage for peaks identified by ADV-1 ChIP-seq in Vogel's media (x-axis) against the coverage for peaks identified by ADV-1 ChIP-seq in Bird's media (y-axis) for each of the indicated growth conditions. The correlation (R^2) is given for each condition in the box.

differentially expressed genes was performed using the Cufflinks tool, Cuffdiff (Trapnell et al., 2013). Sequencing reads obtained for each replicate are listed in Appendix E Table S1. Comparison between WT and ΔADV-1 cells showed that, among 9681 predicted protein coding genes, there were differences of ≥2-fold in levels for 1702 mRNAs (17%) in the dark (LL0), 1711 mRNAs (17%) at LL15, and 2507 mRNAs (23%) at LL60. To increase the stringency, I applied a measure of the false discovery rate (q-value ≤ 0.2) to the data; 253 mRNAs (2%) at LL0, 216 mRNAs (2%) at LL15, and 187 mRNAs (2%) at LL60, met both criteria. While there were specific mRNAs that showed changes in levels that were dependent on the growth conditions of the cell, the overall number of differentially expressed genes was similar in the dark and light treated cells. Combining all time points, 358 unique genes had mRNA levels that changed ≥2-fold with a q-value of ≤0.2 in WT compared to ΔADV-1 cells in the dark and following light treatment (Appendix H File S3). This gene set represents ~4% of the 9681 expressed genes, and was used for all subsequent analyses.

Figure 4-8A shows hierarchical clustering of the 358 unique differentially expressed genes. Similar number of genes were activated or repressed by ADV-1, suggesting that ADV-1 has both positive and negative regulatory roles in gene expression. A comparison between the ADV-1 ChIP-seq data to the 358 ADV-1-regulated genes identified 65 genes that are directly controlled by ADV-1 (Appendix E Table S2). Of these 65 genes, 46 (70%) were activated by ADV-1, whereas 19 genes (30%) were repressed by ADV-1, indicating that ADV-1 primarily functions as a transcriptional

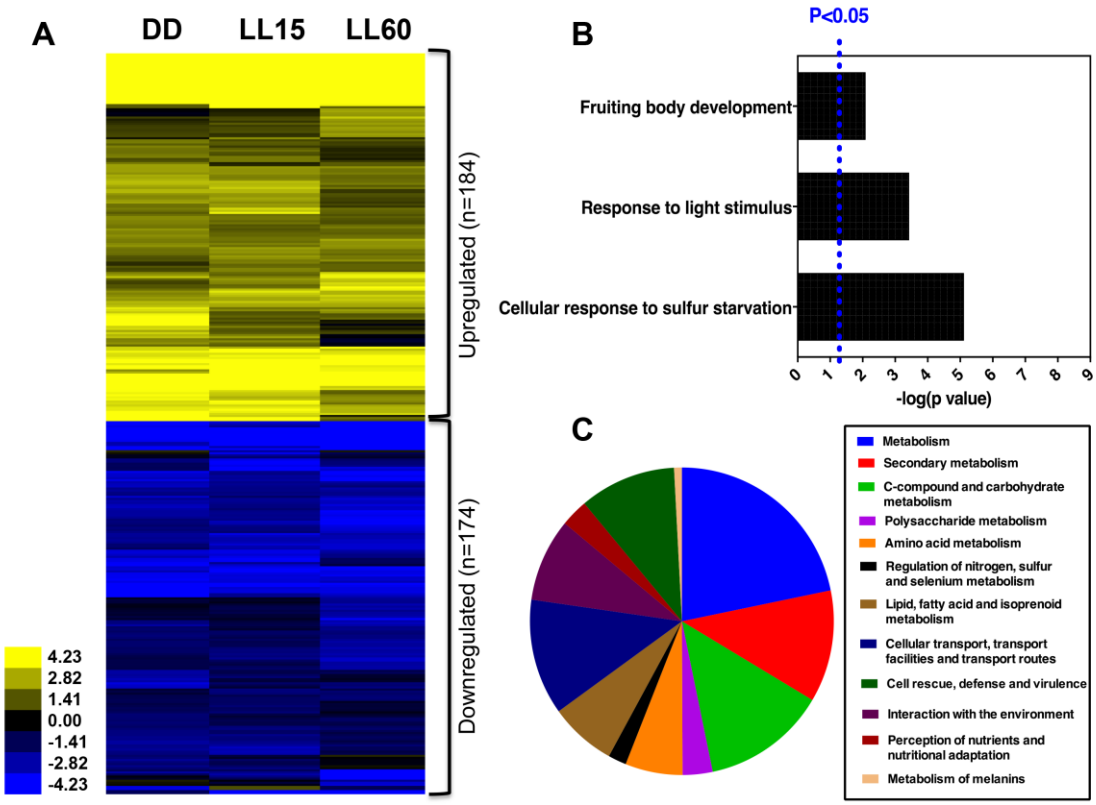


Figure 4-8. RNA-seq in WT versus Δ*adv-1* cells identified differentially expressed mRNAs. (A) Heat map of 358 genes that are differentially expressed between WT and Δ*adv-1* cells following 0, 15 and 60 min of light exposure. The genes were considered differentially expressed if they differed in at least one of the conditions with a fold-change ≥ 2 and q-value ≤ 0.2. The columns show the growth conditions, and the rows represent individual target genes. Blue represents reduced mRNA levels in Δ*adv-1* cells (174 genes), and yellow represents increased mRNA expression in Δ*adv-1* cells (184 genes), as compared to WT cells. (B) Gene Ontology (GO) analysis of the 358 genes regulated by ADV-1. The most highly significant enriched GO terms are shown ($p \leq 0.05$). The plot is labeled as described in Figure 2. (C) Functional Category (FunCat) enrichment analysis ($p \leq 0.05$) of the 358 genes regulated by ADV-1. The distribution of the top categories are shown on the pie chart.

activator of its direct targets. Furthermore, the poor overlap between the differentially expressed genes and the direct targets of ADV-1 revealed that binding does not necessarily equate with a change in mRNA levels. This may be due to binding and compensation by other TFs, or may reflect that regulatory mechanisms other than light are also important for controlling the expression of these ADV-1 targets.

GO analyses was used to determine functional categories of the 358 unique genes differentially controlled by ADV-1 (Figure 4-8B, Appendix E Table S3). As expected, an enrichment for light-responsive genes and genes involved in sexual development was observed. These data also revealed enrichment for genes involved in sulfur metabolism, based on differential expression of genes involved in sulfur assimilation, including NCU04433 (*cys-14*) encoding sulfate permease (Jarai and Marzluf, 1991), and NCU06041 (*ars-1*) encoding aryl sulfatase-1 (Marzluf, 1993) (Appendix E Table S3). For a more detailed view of the pathways potentially controlled by ADV-1 the same gene set was analyzed using FunCat (Ruepp et al., 2004), where enrichment for specific metabolic pathways influenced by ADV-1, including nitrogen, sulfur, amino acid and carbohydrate metabolism, was observed (Figure 4-8C and Appendix E Table S3). Furthermore, genes involved in transport and environmental signaling were enriched, as well as genes involved in melanin metabolism. Melanin helps protect fungal cells, particularly cells produced during sexual development in *N. crassa*, from UV light damage (Butler and Day, 1998), further supporting a link between ADV-1 and the sexual cycle.

ADV-1 regulates cell fusion gene expression

Comparisons among the 1187 unique direct gene targets of ADV-1 to the light-regulated genes identified by RNA-seq (Wu et al., 2014) revealed that at least 346 genes of the ADV-1 direct gene targets (29%) are potentially controlled by light. Of the light-regulated ADV-1 target genes, 178 (51%) are activated and 168 (49%) are repressed by light (Appendix H File S4). The same comparison was made to the genes that are differentially expressed in ADV-1 deletion cells (Appendix H File S4). Among the 358 unique genes that are differentially regulated by ADV-1, 171 genes (48%) are potentially controlled by light. Of the light-regulated genes differentially controlled by ADV-1, 88 (52%) are activated and 83 (48%) are repressed by light. Furthermore, of the 293 genes that are indirectly controlled by ADV-1, 141 (48%) are light-regulated, with 72 (51%) activated and 69 (49%) repressed by light. Of the 346 light-regulated genes that are bound by ADV-1, 271 (78%) had increased or decreased ADV-1 promoter binding in the light treated samples as compared to the dark samples (Figure 4-9A), and 151 of the 271 genes (55%) were differentially regulated by ADV-1. For genes that are direct targets of ADV-1 and are light-regulated, deletion of ADV-1 substantially altered the light response as predicted. For example, light repressed NCU00501 and NCU05194 mRNA levels in WT cells, but the mRNA levels were high in the light in ΔADV-1 cells (Figure 4-9B). Light activated expression of NCU05498 and NCU04343 in WT cells, but the mRNA levels were low in the light in ΔADV-1 cells. Together, these data

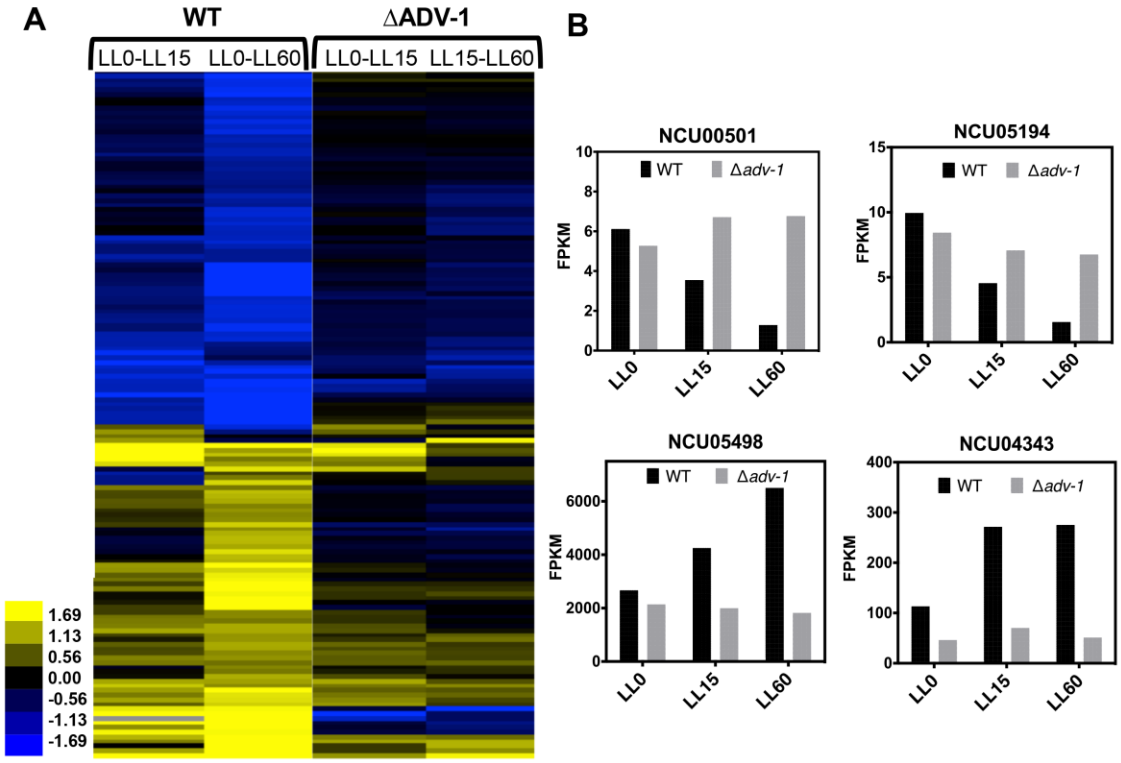


Figure 4-9. (A) Heat map of 151 ADV-1 direct target genes (rows) whose mRNA levels are light-regulated. The levels of mRNA from cells grown in the dark, or exposed to light for 15 and 60 min in WT and ΔADV-1 cells are shown where blue represents downregulation, and yellow represents upregulation in the light versus dark conditions. (B) The mRNA levels (FPKM values) of representative ADV-1 direct target genes from RNA-seq analyses in WT cells (black bars) and Δ*adv-1* cells (gray bar) grown in the dark (LL0) and harvested after light exposure for 15 (LL15) and 60 (LL60) min.

demonstrated that ADV-1 transduces light signals to some, but not all, of its direct target genes. The rhythmic accumulation of ADV-1 levels led us to predict that many of the ADV-1 targets would be under clock control. To identify rhythmic targets of ADV-1, a comparison between the ADV-1 direct targets at each time point to the 3,279 ccgs identified from rhythmic microarrays (Dong et al., 2008; Lewis et al., 2002; Nowrousian et al., 2003) and RNA-seq (Hurley et al., 2014) was made. Approximately one-third of all potential direct ADV-1 targets are predicted *ccgs* (391 *ccgs*) (Figure 4-10A and Appendix H File S4). Comparing the set of predicted direct ADV-1 targets to the sets of light-responsive genes and *ccgs* showed that 135 genes (12%) are both clock- and light-regulated and 587 genes (49%) are neither clock- nor light-controlled. In addition, 211 genes (18%) were light- but not clock-controlled, and 254 genes (21%) were clock- but not light-regulated (Appendix H File S4). Thus, many, but not all, of the direct ADV-1 target genes are *ccgs*. Furthermore, while a tight correlation was believed to exist between light-regulated genes and ccgs, our data suggest that not all light-regulated genes express rhythmic steady-state mRNA levels.

GO analysis was carried out to determine the functional categories enriched for *ccgs* directly controlled by ADV-1. The enriched categories included fruiting body development, circadian rhythms, light responses, metabolic processes and cell fusion (Figure 4-10B and Appendix E Table S4), similar to the enrichment observed for all ADV-1 direct targets (Figure 4-4A).

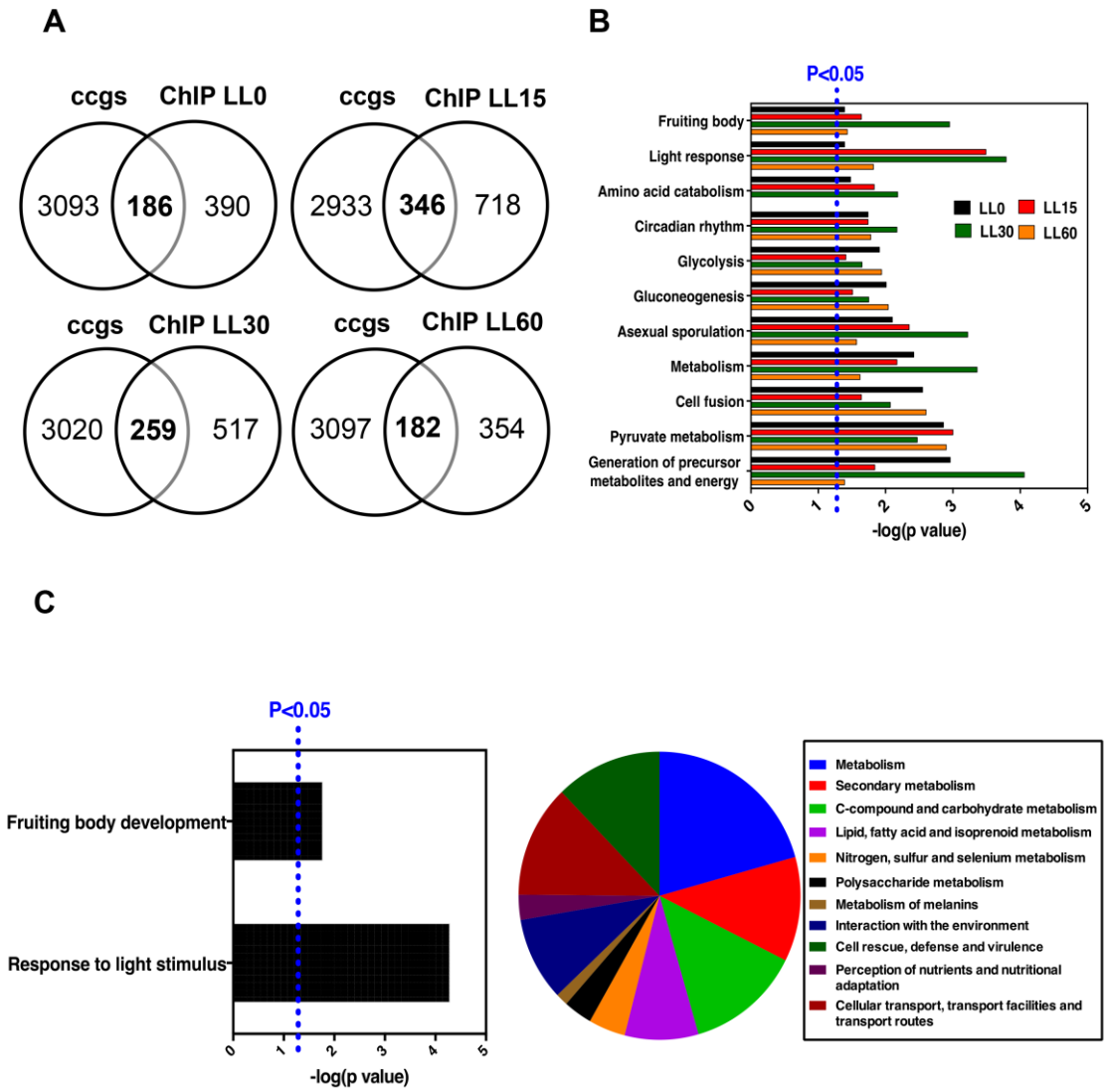


Figure 4-10. ADV-1 downstream targets identified by ChIP-seq and RNA-seq are enriched for ccgs. (A) Venn diagram comparing predicted ADV-1 direct target genes identified by ChIP-seq to predicted and known ccgs. The numbers in the overlapping regions denote genes that are both clock-controlled and bound by ADV-1. (B) Gene Ontology (GO) analysis of clock-controlled and ADV-1-bound genes. The most highly significant enriched GO terms categories are shown ($p \leq 0.05$). The plot is labeled as described in Figure 3-4. (C) Gene Ontology (GO) (left) and Functional Category (FunCat) (right) analyses of 101 genes that are both indirect targets of ADV-1 based on RNA-seq and are predicted or known *ccgs*. For both analyses, the most highly significant enriched categories are represented ($p \leq 0.05$), and the plots are as described in Figure 3-4.

To determine if temporal information is transduced downstream of the ADV-1 direct targets, I examined if any of the 293 genes that are indirect targets of ADV-1 are predicted *ccgs* (Appendix H File 4-S3). I identified 101 (34%) genes as predicted *ccgs*, about half of which are activated and half repressed by ADV-1 based on the RNA-seq data (Appendix E Table S5). Analyses of *ccgs* that are direct targets of ADV-1 revealed no correlation between the phase of *ccg* peak expression and activation versus repression by ADV-1 (Figure 4-11). Approximately half of the *ccgs* that are either activated or repressed by ADV-1 peak in the morning, and the remaining half are equally distributed between afternoon and nighttime peaks. These data suggested that the regulation of ccg phase is complex.

Pathway analyses of these ccgs (GO and FunCat) revealed enrichment for pathways (Figure 4-10C) that were similar to the full set of differentially expressed genes (Figure 4-8B&C). These data indicated that ADV-1 functions within a circadian output pathway, potentially controlling rhythmicity for both direct and indirect gene targets involved in development, environmental sensing, metabolism, and cell fusion. Clock control of development, environmental sensing, and metabolism are well established (Bass and Takahashi, 2010; Brown, 2014; Goldsmith and Bell-Pedersen, 2013; Summa and Turek, 2014). However, the data indicating that the clock regulates aspects of cell fusion was unexpected.

Deletion of cell fusion genes results in delays in or elimination of protoperithecia

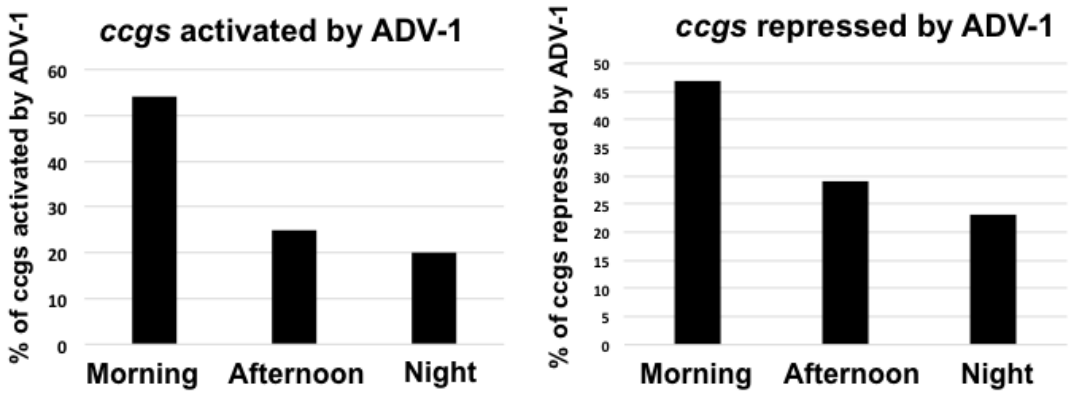


Figure 4-11. Phase of *ccgs* regulated by ADV-1. Phase distribution of *ccgs* activated (left) or repressed (right) by ADV-1. The x-axis denotes the peak phase of mRNA levels for the *ccgs* and the y-axis represents the percentage of *ccgs* activated or repressed by ADV-1 for each phase.

development, similar to the phenotype of ΔADV-1 cells (Fu et al., 2011; Lichius and Lord, 2014). Importantly, both ΔADV-1 cells, and many cell fusion gene deletion strains, lack CATs (Fleissner et al., 2009a; Fu et al., 2011; Leeder et al., 2013), which home toward each other to promote cell fusion. ADV-1 binding peaks were identified in the promoters of 15 cell fusion genes (Appendix E Table S6), and RNA-seq revealed that 9 out of the 15 genes showed at least a 2-fold change in mRNA levels in cells deleted for ADV-1 as compared to WT cells, and 4 of these met the more stringent criteria (q-value ≤0.2) (Figure 4-12). While changes in gene expression are apparent in the heat map for all of the cell fusion genes, some of the genes did not make the stringent RNA-seq cutoff. This may reflect low-level expression of these genes under the growth conditions used in RNA-seq. In any case, northern blot and qRT-PCR analyses confirmed ADV-1 regulation for 4 cell fusion genes that had a 2-fold change in mRNA levels (Figure 4-13). Interestingly, the expression profiles revealed both positive and negative regulation by ADV-1 for the cell fusion genes (Figures 4-12 & 4-13). For example low mRNA levels were observed for *ham-6*, *prm-1*, and *lfd-1*, and high mRNA levels were found for *ham-9*, at all time points in the absence of ADV-1. Some cell fusion gene mRNA levels were higher (e.g. *so* and *mek-1*) or lower (e.g. *pkr1* and *ham-5*) only in the light treated samples, although none of the cell fusion genes were previously identified as being light-responsive (Wu et al., 2014).

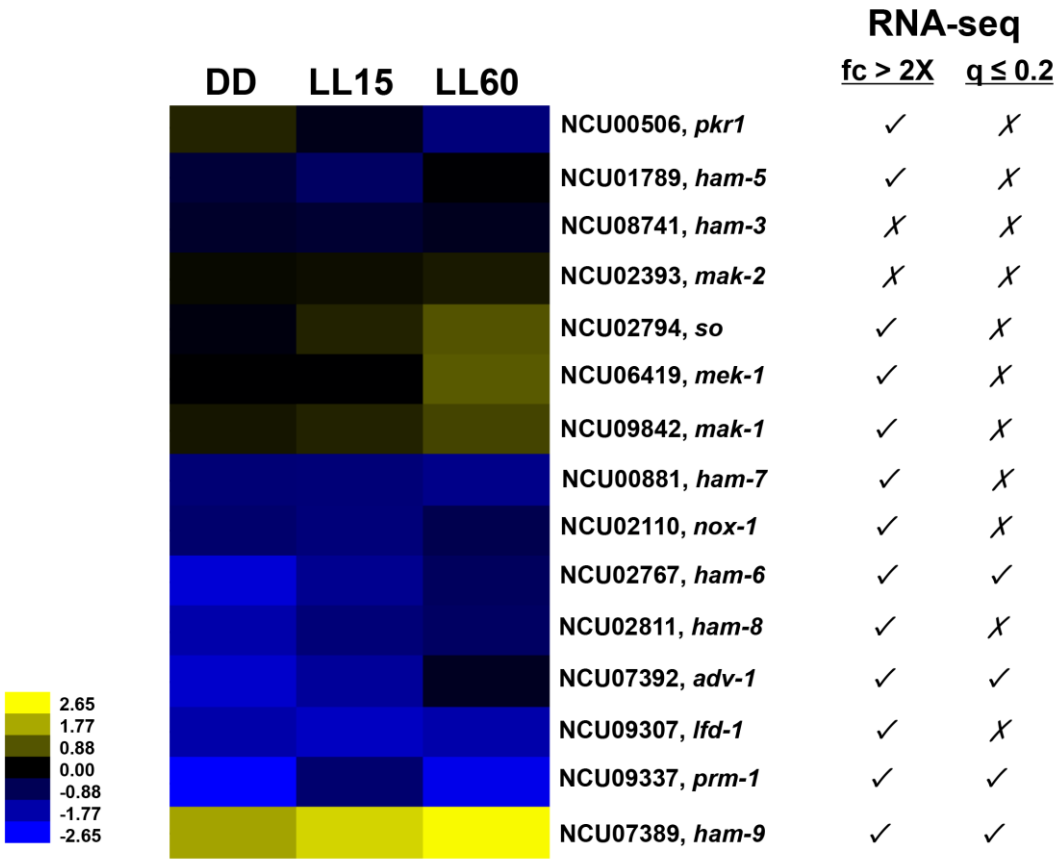


Figure 4-12. ADV-1 regulates genes involved in cell fusion. Heat map of 15 cell fusion genes identified as direct ADV-1 targets showing the fold change in mRNA levels in WT as compared to Δ*adv-1* cells following 0, 15 and 60 min of light exposure. Blue indicates reduced mRNA levels in Δ*adv-1* cells, and yellow represents increased mRNA expression in Δ*adv-1* cells, as compared to WT cells. The columns show the growth conditions, and the rows represent individual cell fusion genes in the heat map. Indicated on the right is whether or not the mRNAs had a fold change (fc) of 2 or more, and FDR value of $q \leq 0.2$ in the RNA-seq analyses.

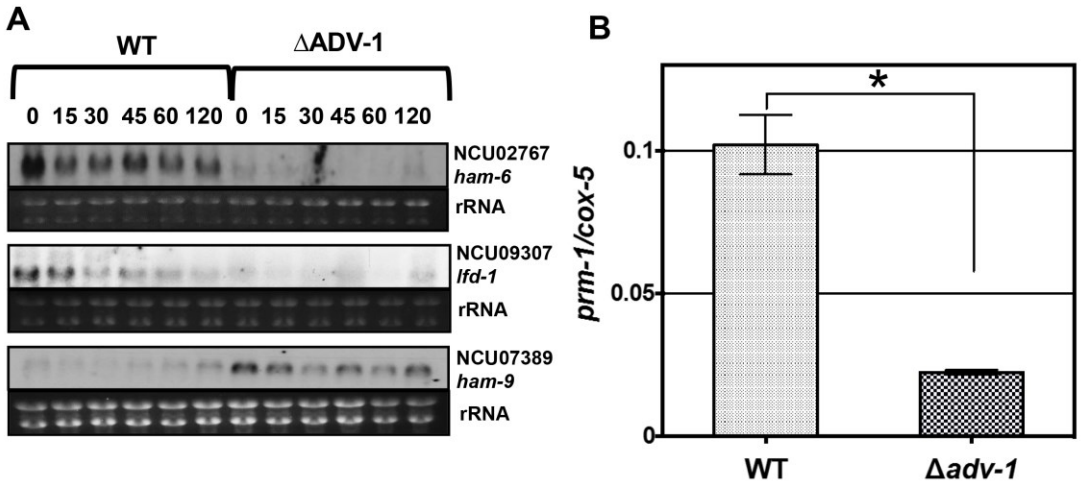


Figure 4-13. Validation of RNA-seq data. (A) RNA expression level analyses of the indicated cell fusion genes in WT and ΔADV-1 strains. Northern blots of RNA from tissues harvested in the dark (0) and after the indicated light treatment (LL) in min. rRNA is shown as a loading control. (B) Total RNA isolated from WT and Δ*adv-1* cells given a 60 min light treatment was used for qRT-PCR to measure *prm-1* mRNA levels. The mRNA levels were normalized to *cox-5*, a gene not under the control of ADV-1 (± SEM, n=2). The asterisk represents statistically significant differences in *prm-1* mRNA levels in Δ*adv-1* cells as compared to WT cells (*P<0.003, Student's t test).

ADV-1 transduces rhythmic signals from the clock to control rhythms in cell fusion gene mRNA accumulation

To determine if ADV-1 relays time of day information to its downstream targets, I asked if the cell fusion genes are *ccgs*, and if so, if their rhythms dependent upon ADV-1. First, promoter LUC fusion reporters were generated for NCU09337 (*prm-1*), NCU09307 (*lfd-1*), and NCU02767 (*ham-6*), three cell fusion genes that are activated by ADV-1, and for NCU07389 (*ham-9*), which is repressed by ADV-1. The expectation was that genes that are activated by ADV-1 would have reporter mRNA levels that peak during, or slightly after, the peak of ADV-1 levels in the late subjective day/early night, and genes that are repressed by ADV-1 would peak in antiphase, during the late subjective night/early morning. The promoters of all 4 genes were shown to drive robust rhythms in LUC expression, with the peak levels of *lfd-1*, *prm-1*, and *ham-9* occurring consistently at dusk (DD22), and peak levels of *ham-6* at in the early night (DD24) (Figure 4-14). While the activated genes peaked at the predicted time of day, *ham-9* did not. These data indicated that ADV-1 may not be sufficient to regulate the phase of rhythmic *ham-9* gene expression.

To determine if additional factors are required for rhythmicity of the cell fusion genes, I assayed LUC rhythms from the same constructs in ΔADV-1 cells. The rhythms in *prm-1*

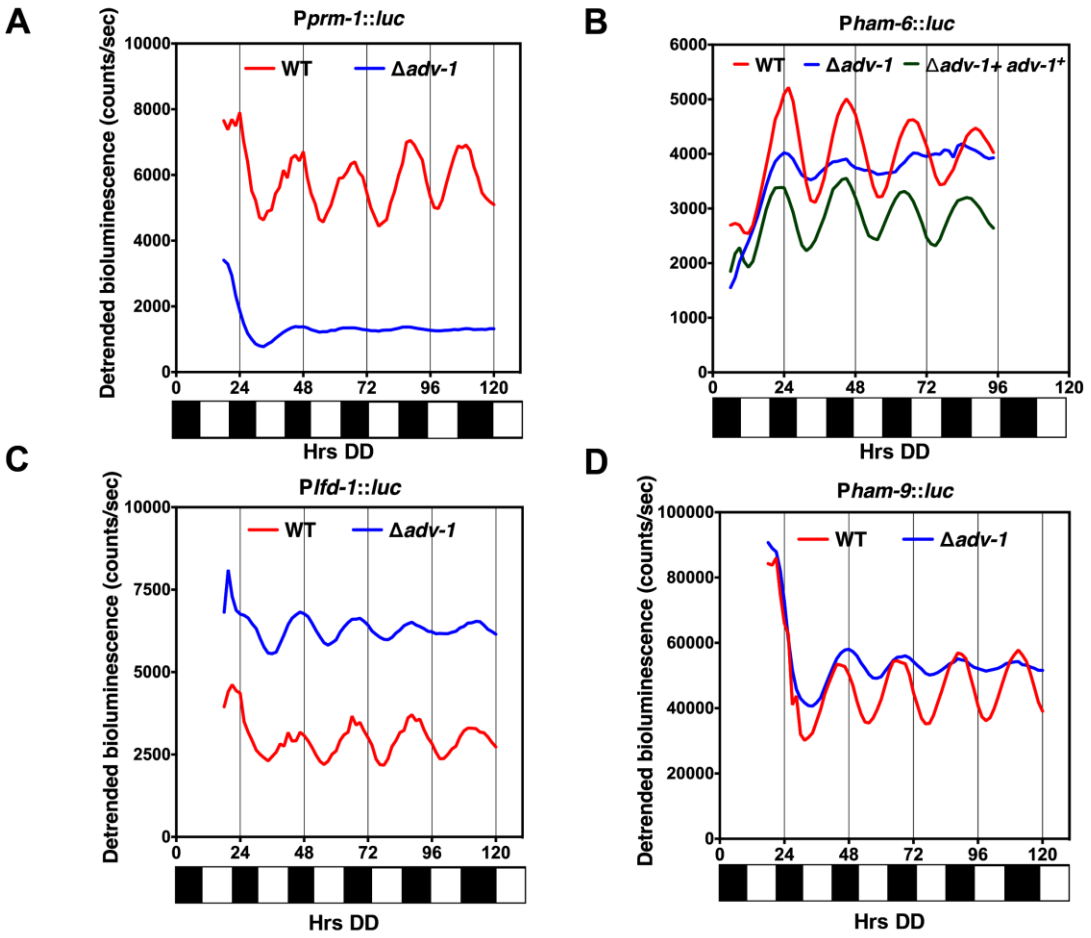


Figure 4-14. ADV-1 affects the levels and rhythmic expression of cell fusion genes. Representative bioluminescence records of WT (red), Δ*adv-1* (blue), and complemented Δ*adv-1*::*adv-1* (green) strains, containing the promoters of the indicated cell fusion genes driving luciferase expression. Detrended bioluminescence levels are indicated on the y-axis. The time of recording in constant darkness (DD) is shown on the x-axis (hrs). The bar graph below each plot represents subjective day (white) or subjective night (black) in DD.

and *ham-6* promoter activity were severely dampened in the mutant as compared to WT (Figure 4-14A and 4-14B), indicating that ADV-1 is necessary and sufficient to drive their mRNA rhythms. For *lfd-1* and *ham-9*, the overall levels of mRNA were increased in ΔADV-1 cells as compared to WT cells; however, mRNA rhythms with low amplitudes persisted in ΔADV-1 cells (Figure 4-14C and 4-14D). The higher mRNA levels observed for *lfd-1*::*luc* in ΔADV-1 cells contrasts with the RNA-seq and validation data (Figure 4-12 and 4-13A), and suggests that the *lfd-1*::*luc* construct lacks some regulatory elements. In any case, these data are consistent with ADV-1 being sufficient to transduce temporal information to some of its direct target genes, but that additional factors are needed to control rhythms in the expression of others (*ham-9* and possibly *lfd-1*). Finally, the rhythms in P*ham-6*::*luc* expression in ΔADV-1 cells were restored upon transformation and insertion of a wild type copy of the *adv-1* gene, confirming that the loss of rhythmicity was due to loss of ADV-1 (Figure 4-14B).

CAT chemoattraction requires the activity of MAPK pathway components (Fleissner et al., 2009b), demonstrating a link between cell fusion and MAPK pathway activation. In addition, similar to ADV-1, the circadian clock regulates daily MAPK activity in *N. crassa* (Bennett et al., 2013; de Paula et al., 2008; Lamb et al., 2011; Vitalini et al., 2007). Interestingly, components of the MAK-1 and MAK-2 MAPK pathways that are involved in cell wall integrity and pheromone responses, respectively, were found to be direct targets of ADV-1 (Appendix H File 4-S1). These included *mek-1*, encoding the MAPKK, *mak-1*, encoding the MAPK of the cell wall integrity pathway, and the *mak-2*

encoding the MAPK of the pheromone response pathway (Li et al., 2005; Maerz et al., 2008). The phenotypes of Δ*mek-1* and Δ*mak-1* deletion strains include defects in protoperithecia development, hyphal cell fusion, and asexual conidiation, similar to ΔADV-1 cells (Fu et al., 2011; Park et al., 2008; Park et al., 2011). Therefore, I investigated if the clock signals through ADV-1 to regulate the levels of MAPK pathway mRNAs, which in turn would affect the timing of MAK-1 activation (phosphorylation) for hyphal fusion. Circadian rhythms of phosphorylated MAK-1 (P-MAK-1) were examined in WT and ΔADV-1 cells. Consistent with previous data, the levels of P-MAK-1 were rhythmic ($p<0.0039$; period=22.6±0.7 h), peaking during the subjective day (Figure 4-15) (Bennett et al., 2013). As predicted, the rhythms in the levels of P-MAK-1 were abolished in ΔADV-1 cells, with generally high levels of P-MAK-1 over the cycle. I also examined the requirement for ADV-1 in P-MAK-2 rhythms. Robust rhythms in the levels of P-MAK-2 persisted in ΔADV-1 cells (data not shown), indicating that other mechanisms exist to control P-MAK-2 rhythmicity.

DISCUSSION

To determine how light and the circadian clock controls gene expression, I used ChIP-seq and RNA-seq to identify the direct and indirect gene targets of the *N. crassa* TF ADV-1. The ADV-1 target gene lists were compared to *N. crassa* genes identified in other genome-wide studies of light-regulated genes

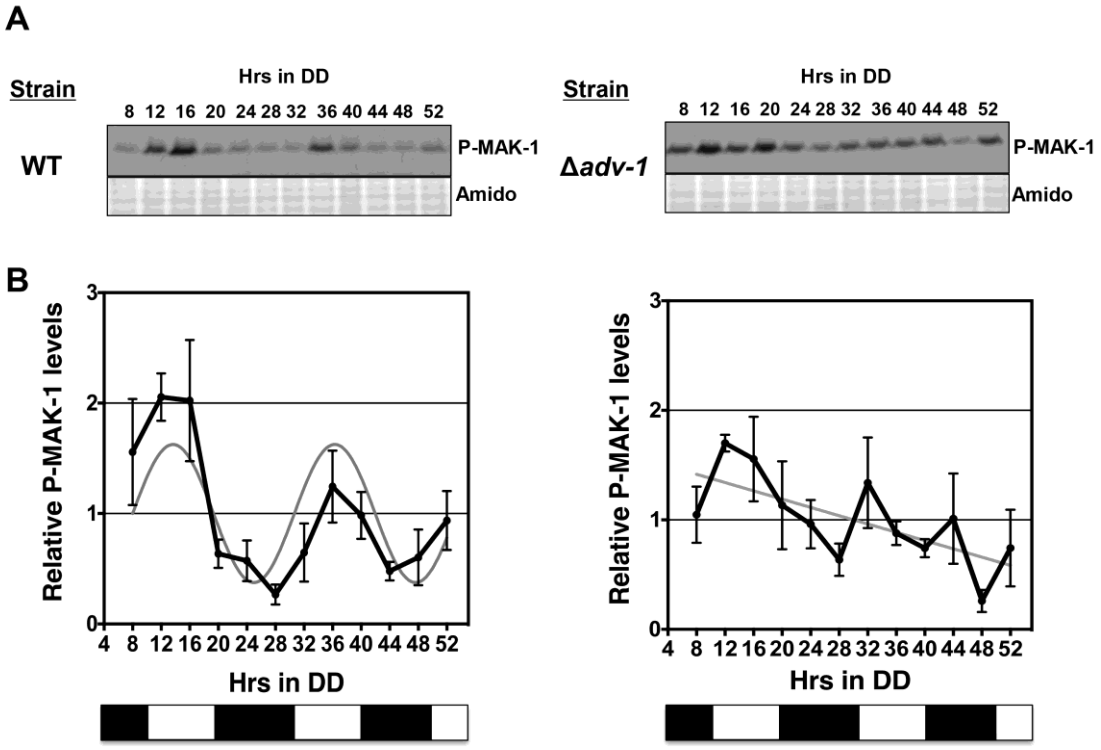


Figure 4-15. ADV-1 regulates rhythmic MAK-1 phosphorylation. (A) Representative western blot of phosphorylated-MAK-1 (P-MAK-1) from the indicated strains grown in constant dark (DD) and harvested every 4-h for 2 days. The blots were probed with an anti-phospho-p44/p42 antibody. The amido stained membranes were used as loading controls. (B) Plots showing normalized P-MAK-1 levels (± SEM, n=4). The time of harvest in constant darkness (DD) is shown on the x-axis (hrs). The bar graph below each plot represents subjective day (white) or subjective night (black) in DD. The levels of P-MAK-1 in WT cells (black) fit a sine wave (grey) (p=0.0039), whereas the levels of P-MAK-1 in Δ*adv-1* cells (black) fit better to a line (grey).

and *ccgs* (Dong et al., 2008; Hurley et al., 2014; Lewis et al., 2002; Wu et al., 2014). Together, these data sets provide a rich resource of information to unravel the role of ADV-1 in transducing light and temporal information to downstream targets, and to connect these gene targets to physiological processes. For example, this work led to the discovery of a role for ADV-1 in regulating temporal expression of cell fusion genes.

Consistent with a major function for ADV-1 in controlling gene expression, ~12% of all predicted *N. crassa* genes had ADV-1 binding sites within 10 kb upstream to 3 kb downstream, or within, the predicted open reading frame. Surprisingly, significantly fewer genes (4%) met out stringent criteria (≥2-fold change and q≤0.2) for being differentially regulated in ΔADV-1 cells based on RNA-seq; the overlap between ChIP-seq and RNA-seq data sets was 5%. Relaxing the RNA-seq stringency (≥2-fold change) significantly increased the overlap between the two data sets to 42%. In either case, not all genes showed differential expression demonstrating that ADV-1 binding to its target sequence is not always sufficient to activate or repress transcription. ChIP-seq analyses of several other TFs that are direct targets of the WCC showed binding of these TFs to many of the same ADV-1 target gene promoters (manuscript in preparation). Binding of these additional TFs to ADV-1 target sites may compensate for the loss of ADV-1. Furthermore, while ADV-1 primarily functions as an activator of transcription of its direct targets, it can also act as a repressor. No obvious differences in the binding sequences of repressed versus activated genes was identified. Therefore, it is likely that repression results from coordinate TF activity. Genome-wide studies in Drosophila

(Orian et al., 2005) and human cell lines (Cusanovich et al., 2014) found similar discrepancies in which the number of TF binding sites far exceeded the number of genes that are differentially regulated, supporting complex regulation of transcription.

ADV-1 binding to genomic loci increased 2-fold in cells given a 15 min light treatment compared to dark grown cells, consistent with rapid induction of ADV-1 levels in response to light stimuli (Figure 3-2B). The overall reduction in the number of ADV-1 binding sites with longer light treatment is not unexpected since its regulator, the WCC, undergoes photoadaptation (Chen et al., 2010a). However, because the levels of ADV-1 are still high with longer light exposure, I speculate that light-dependent posttranslational modifications reduce the binding of ADV-1 to certain gene promoters. For the majority of light-regulated genes bound by ADV-1 (78%), ADV-1 binding changes in the light-treated samples, and light regulation was dependent on ADV-1 in about half of these genes (Figure 4-9). These data confirmed that ADV-1 signals light information to some of its downstream targets, supporting a hierarchical network of TFs controlling light responses (Chen et al., 2009). Alternatively, at some loci, ADV-1 binding occurred both in the dark and in the light, and the dark consensus binding site differed slightly from the light bound ADV-1 consensus site (Figure 4-5), suggesting that ADV-1 functions in the dark to regulate distinct processes, including in relaying temporal information from the clock to downstream *ccgs*. Our data showing loss of rhythmicity for some ADV-1 direct targets confirms that ADV-1 is both necessary and sufficient to transduce temporal information to certain downstream *ccgs*. To fully

elucidate the impact of ADV-1 on the downstream circadian network, RNA-seq experiments are currently underway to examine changes in mRNA levels in ΔADV-1 cells as compared to WT cells harvested at different times of the day. Despite the major role of ADV-1 in light and circadian signaling, about half of the direct targets were predicted to be neither light- nor clock-controlled under our growth conditions. This may be due to mRNA stability precluding detection of changes in mRNA levels, or other modes of regulation, such as developmental control.

In *N. crassa*, cell-to-cell fusion plays an essential role in colony establishment, asexual development, and in the formation of protoperithecia (Fleissner et al., 2008). During colony establishment, CATs formed at the germ tube tips will fuse to establish an interconnected hyphal network. This branched hyphal network transports nutrients to aerial hyphae that, in turn, produce asexual conidiospores. Efficient transport of nutrients from the vegetative hyphae into developing protoperithecia is thought to be required for sexual development. Consistent with the role of cell-to-cell fusion in each of these processes, cell fusion mutants lack interconnected hyphal networks, produce short aerial hyphae and have reduced conidiospores, and they do not produce protoperithecia, rendering them female sterile (Lichius and Lord, 2014; Morgan et al., 2003). ΔADV-1 mutant cells do not form CATs (Fu et al., 2011), and therefore, not surprisingly, the mutant has developmental defects that mimic other cell fusion mutants (Lichius and Lord, 2014). Our genome-wide data are consistent with the phenotypes of the mutant, whereby the direct and/or indirect ADV-1 gene targets were found to function primarily

in asexual and sexual development, hyphal growth, and metabolism. Changes in metabolism are likely required to accommodate the increased energy demands for development (Adomas et al., 2010). In support of this idea, I found that ADV-1 controls the levels, and rhythmicity of the *fbp-1* gene encoding fructose-1.6-bisphsophatase (NCU04797), a key enzyme in gluconeogenesis (Figure 4-16). These data suggest coordinate regulation between genes involved in metabolism, growth, development, and cell fusion involving ADV-1. As ADV-1 is both light-induced and clock-controlled, and because light and the circadian clock influence the onset and timing of growth and development (Corrochano, 2007; Dunlap and Loros, 2006), our data suggested a role for the clock, through ADV-1, in regulating cell fusion gene expression. Indeed I found that ADV-1 signals temporal information from the clock to control rhythmic expression of several key cell fusion genes. Rhythms in *prm-1* and *ham-6* mRNA levels, and MAK-1 phosphorylation were abolished in ΔADV-1 cells, and the amplitude of the mRNA rhythms were reduced for *lfd-1* and *ham-9*, as compared to WT cells (Figure 4-14). It is now of interest to determine if the clock controls the timing of cell fusion, as would be predicted based on rhythmic expression of the cell fusion genes.

ADV-1 protein levels peak in the late day, several hours after the peak in WCC activity. I predicted that genes that are activated by ADV-1 would peak in mRNA levels close to the time of peak ADV-1 levels, and that genes that are repressed by ADV-1 would peak at the opposite phase, similar to what has been observed for genes controlled by the CSP-1 transcriptional repressor in *N. crassa* (Sancar et al., 2011). However, our results

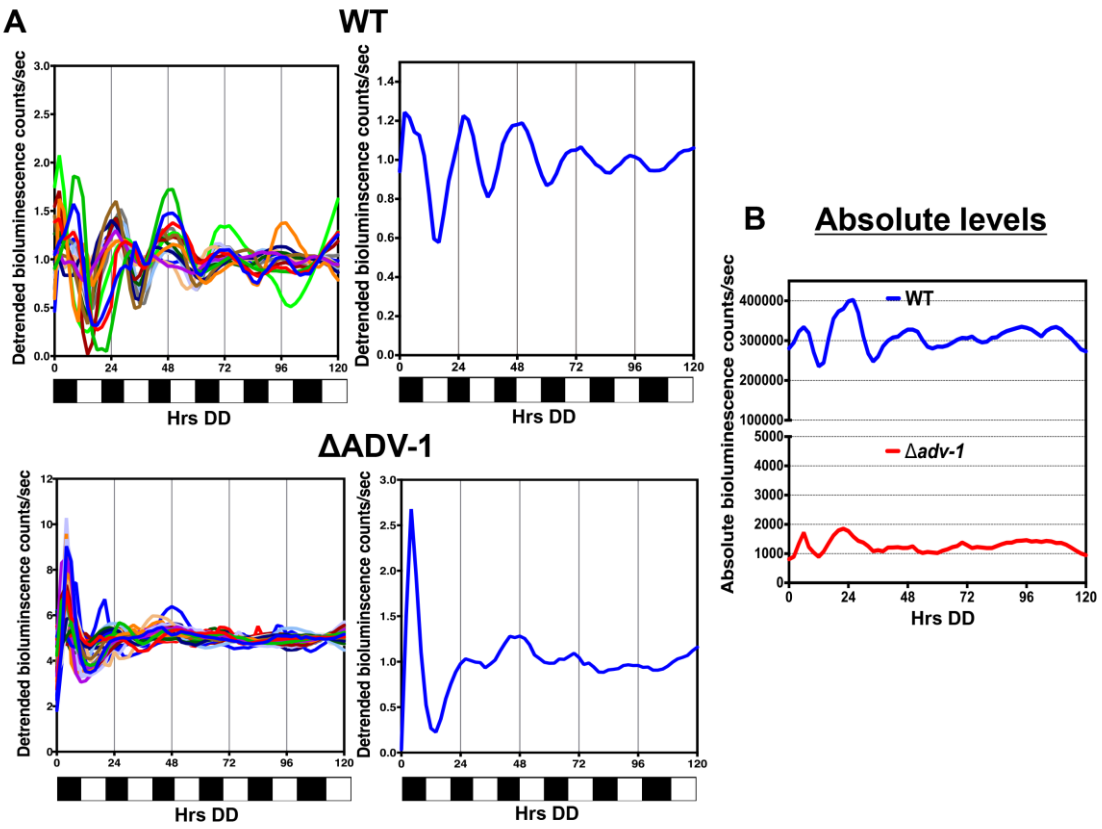


Figure 4-16. ADV-1 affects the levels and rhythmic expression *fbp-1* mRNA encoding fructose-1.6-bisphsophatase. (A) The average of all individual detrended bioluminescence recordings (left) and representative bioluminescence records (right) of WT (top) and Δ*adv-1* (bottom) cells, containing the *fbp-1* promoter driving luciferase. P*fbp-1*::*luc*. Detrended bioluminescence levels are indicated on the y-axis. The time of recording in constant darkness (DD) is shown on the x-axis (hrs). The bar graph below each plot represents subjective day (white) or subjective night (black) in DD. (B) Averaged raw bioluminescence recordings of *fbp-1* are shown together for WT (blue) and Δ*adv-1* (red).

are more complex. Specifically, *ham-9* is negatively controlled by ADV-1, and peaked in expression in the subjective early night (Figure 3-13D). This is similar to the time when the mRNA levels peaked for the positively controlled *prm-1*, *lfd-1*, and *ham-6* genes. Northern blot analyses to examine the mRNA levels of the cell fusion genes over the course of the day confirmed the LUC reporter assays (data not shown), ruling out discrepancies due to the use of the reporter. Furthermore, the phase inconsistency was not limited to a few *ccgs*, but was evident across *ccgs* regulated by ADV-1 (Figure 4-11). Part of the explanation for these results may stem from the regulatory network downstream of ADV-1, including ADV-1 autoregulation, and other TFs that are direct targets of the WCC and/or direct targets of ADV-1. ChIP-seq revealed that many of the ADV-1 direct targets, including *adv-1* itself, are also targets of the other WCC-controlled TFs (manuscript in preparation). These interactions form a nested set of feedforward and feedback loops that can have significant impacts on the dynamics of gene regulation (Mangan and Alon, 2003; Nakamichi et al., 2012; Ptacek et al., 2007; Ukai-Tadenuma et al., 2008). One function of the complex regulatory interactions that occur downstream of ADV-1 may be to generate distinct temporal dynamics of gene expression relative to the central clock. Taken together, our data supports a model whereby, unlike flat hierarchical network controlling light responses, the circadian regulatory network is much more complex.

In summary, using genome-wide approaches I demonstrate that ADV-1 functions to transduce light and temporal information, as well as other unknown signals, to target

genes involved in metabolism, cell growth, development, and cell fusion. These data provide an important resource for understanding the impact of light and the clock on regulating gene expression, and provides the platform for determining how interactions among TFs influence gene regulatory networks, including for example, how cross talk in the network controls the phase of rhythmic gene expression that is critical to temporal coordination.

MATERIAL AND METHODS

Strains and culture conditions

N. crassa strains, including WT 74OR23-1 (FGSC 2489) and ΔADV-1,A (FGSC 11042) were obtained from the Fungal Genetics Stock Center (FGSC, Kansas City, MO). The V5-tagged ADV-1 strain used for ChIP-seq was previously described (Smith et al., 2010). Protoperithecia assays were carried out on Westergaard and Mitchell's synthetic crossing medium containing 0.5% sucrose as a carbon source (Westergaard and Mitchell, 1947). Protoperithecia were scored from cultures grown 7-8 days in the dark at 20°C.

Primers Pvu I adv-1 F and Pvu I/Spe I adv-1 R were used to amplify full-length *adv-1* gene from WT *N. crassa* genomic DNA as the template. This DNA fragment was digested with Pvu I and then cloned into the corresponding sites of pCR blunt bar::P_{tcu-1}

(Lamb et al., 2013), to generate pDBP 506. Digestion of pDBP 506 with Spe I yielded a 4.9 kb *bar*-P_{tcu-1}-*adv-1* fragment that was ligated into Spe I cut pBM61 (*his-3* integrating vector) to generate pDBP 508. pDBP 508 was transformed into a *his-3*, *mat a* and *his-3*, mat *A* strain at the his locus to generate his-3^{+}::*bar*-P_{tcu-1}-*adv-1*, mat *a* and his-3^{+}::*bar*-P_{tcu-1}-*adv-1*, mat *A* strains respectively.

Promoter-driven luciferase reporter constructs were generated using recombinational cloning in yeast as previously described (Colot et al., 2006). Briefly, five PCR fragments (5' of *csr*, promoter region of gene, codon optimized *luc* sequence (Gooch et al., 2008), 3' UTR of gene, and 3' of *csr*) were co-transformed with gapped plasmid (pRS426) digested with XhoI and BamHI into yeast strain FY2 (MATα ura3-52). Following recombination, the full-length cassette was amplified and transformed into WT and ΔADV-1 cells.

Plasmid pCR blunt *bar adv-1* was used for complementation analyses of P*ham-6*::*luc*, Δ*adv-1* cells, was generated by PCR amplification of genomic DNA using primers to amplify ~3,500 bps upstream and 2000 bps downstream of the *adv-1* coding region. The primers contained Xho I restriction sites at their 5' and 3' ends. The Xho I digested PCR product was cloned into pCR blunt *bar* (Pall and Brunelli, 1993).

Luciferase assays

Luciferase levels were measured from conidia inoculated into 96-well microtitre plates containing 1X Vogel's salts, 0.01% glucose, 0.03% arginine, 0.1 M quinic acid, 1.5 % agar and 25μM firefly luciferin (LUCNA-300; Gold Biotechnology, St. Louis, MO), pH 6.0, To synchronize the cells to the same time of day, the plates were incubated in LL at 30°C for 24 h, and then transferred to DD 25°C. Bioluminescence was measured using a TopCount NXT™ Microplate Scintillation and Luminescence Counter (PerkinElmer Life Sciences, Boston, MA) with recordings taken every 90 mins for 4-5 days. Data were collected using the Import and Analysis (I&A) program (http://www.amillar.org/PEBrown/BRASS/BrassPage.htm). Raw luciferase data were detrended prior to analysis for period, phase, and amplitude determination using Biological Rhythms Analysis Software System (BRASS) running on BioDARE (Biological Data Repository) (http://www.biodare.ed.ac.uk) (Moore et al., 2014; Zielinski et al., 2014). The first day of recording was not included in the analyses to avoid potential artifacts caused by transfer of the plates to the TopCount.

RNA, protein, and ChIP analyses

Total RNA was extracted from cells ground in liquid nitrogen as described (Bell-Pedersen et al., 1996a). Transcript levels were detected with [α-32P]-UTP-labeled

riboprobes (Correa et al., 2003). Densitometry was performed using ImageJ software and normalized to ethidium bromide–stained ribosomal RNA. Quantitative RT-PCR was performed as described previously (Wu et al., 2014) and *cox-5* was used as an internal control for normalization. To detect ADV-1::V5 by western blot, ADV-1::V5 conidia (~1X10^7) were added to petri dishes containing 25 ml 1X Vogel's salts, 2% glucose and 0.5% arginine, pH 6.0, and incubated in LL 25°C for 24 h to prepare mycelial mats. Discs (5mm in diameter) were cut from the mats and a single disc was inoculated into each flask containing 65 ml of Vogel's media. The cultures were incubated in LL at 30°C for 24 h with shaking (200 rpm), and then transferred to DD 25°C for 24 h with shaking (200 rpm). The cultures were exposed to light (~21 μmole photons/m2/s or ~1500 lux) for 0, 15, 30, 45 and 60 min. Protein extraction and western blotting were carried out as previously described (Garceau et al., 1997) with a 1:5000 dilution of mouse monoclonal anti-V5 antibody (Invitrogen, Carlsbad, CA). ChIP was carried out as described (Tamaru et al., 2003), with some modifications. Briefly, conidia (~1X10^7) from ADV-1::V5 cells were inoculated into petri dishes containing either 25 ml of Vogel's or Bird's media (Metzenberg, 2004) containing 2% glucose, and the cultures were processed as described above for western blotting. Formaldehyde (1%) was added to cultures following light exposure for 0, 15, 30, 45 and 60 min, and the cultures were incubated for 30 minutes with shaking. Chromatin was sheared to an average size of 500 bp by sonication (Branson digital sonifier S-250D, microtip probe). The sheared chromatin was immunoprecipitated using 5 μg of mouse monoclonal anti-V5 antibody (Invitrogen,Carlsbad,CA). The immunoprecipitated DNA was extracted with phenol-

chloroform, and treated with RNase A for 2 hours at 50°C. ChIP DNA was quantified using Picogreen (Invitrogen,Carlsbad,CA) according to manufacturer's instructions. DNA libraries were prepared from immunoprecipitated chromatin (ChIP) and total chromatin (Input) for sequencing. An independent chromatin immunoprecipitation with anti-V5 antibody was carried out using ADV-1::V5 cells to validate ADV-1 binding sites on select genes. PCR primer pairs were designed to amplify specific ADV-1 binding regions using Primer3 (http://fokker.wi.mit.edu/primer3/input.htm). PCR primer pairs amplifying the coding region of *cpc-1* was used as a negative control.

ChIP-seq and data analyses

DNA was end repaired and ligated to adapters (Pomraning et al., 2009). Fragments (300 to 500 bp long) were gel-purified and amplified by 21-24 cycles of PCR with Phusion polymerase (Finnzymes Oy, NEB) and Illumina PCR primer (Pomraning et al., 2012). Libraries were sequenced on an Illumina GAII and processed with RTA1.8 and CASAVA1.7. Sequencing yields and percent mapped are listed in Appendix E Table S1.

ChIP-seq reads were sorted by adapter, and adapter sequences were removed. Quality scores were converted to Sanger format with the MAQ sol2sanger command (Li et al., 2008a). Fastq files were used as input for BWA (Li and Durbin, 2009) and aligned to a reformatted assembly 10 of the *N. crassa* genome (http://www.broadinstitute.org/annotation/genome/neurospora/GenomesIndex.html).

Sam-formatted alignment files from BWA were converted to bam format, sorted, and indexed with samtools (Li et al., 2009) for viewing in the gbrowse2 genome browser (Stein et al., 2002). Data are accessible at http://ascobase.cgrb.oregonstate.edu/cgi-bin/gb2/gbrowse/. HTS data from ChIP- and RNA-seq (below) were submitted to NCBI databases.

RNA seq and data analyses

RNA extraction, mRNA enrichment, cDNA preparation, and library preparation for sequencing was performed as previously described (Wu et al., 2014). Sequencing was carried out on a HiSeq2000 and analyzed with RTA 1.13.48 and CASAVA v1.8.2. Sequencing yields are listed in Appendix E Table S1. Following adapter sequence removal, reads were mapped to *N. crassa* Assembly 10 with Tophat using options -a 5 -m 1 -i 30 -I 2000 -p 5 --library-type fr-unstranded. Aligned reads can be viewed at http://ascobase.cgrb.oregonstate.edu/cgi-bin/gb2/gbrowse/ncrassa_public/. Tophat output was sorted and indexed with samtools and used as input for cuffdiff analysis to identify differentially expressed genes between WT and Δ*adv-1* strains. RNA-seq across a light induction time course in wildtype *N. crassa* has been previously described (Wu et al., 2014) and are available under GEO series accession GSE53534. Cuffdiff was run using two WT replicates and two Δadv-1 replicates (Appendix E Table S1) and the Broad version 6 reference transcriptome (http://www.broadinstitute.org/annotation/genome/neurospora/MultiDownloads.html).

Data analysis

The Gene identifiers assigned to different Gene Ontology (GO) terms were mapped using the GO Term Finder tool (http://go.princeton.edu/cgi-bin/GOTermFinder) with a *N. crassa* GO terms association file (go_for_nc12.tsv downloaded from http://www.broadinstitute.org/annotation/genome/neurospora/Downloads.html). Functional enrichment for direct and indirect ADV-1 targets was determined using the FunCat tool (FunCatDb). The statistical significance of enrichment of gene groups in functional categories relative to the whole genome is based on hypergeometric distribution (Ruepp et al., 2004). Heat maps were generated with log2 FPKM values using Cytoscape clusterMaker plugin (Morris et al., 2011).

Statistical analysis

Nonlinear regression to fit the rhythmic data to a sine wave (fitting period, phase, and amplitude) or a line (fitting slope and intercept) as well as Akaike's information criteria (AIC) to compare the fit of each data set to the 2 equations, were carried out using the Prism software package (GraphPad Software Inc, San Diego, CA). The p values reflect the probability that the sine wave fits the data better than a straight line. Statistical analyses were performed using the Student’s T-test. $P<0.05$ was considered to be statistically significant.

CHAPTER V

SUMMARY AND CONCLUSIONS

OVERVIEW

Circadian clocks allow organisms to anticipate daily environmental cycles, and regulate physiology and behavior to optimize the timing of resource allocation for improved fitness (Sharma, 2003). For example, our heart rate and blood pressure increase in anticipation of waking up each morning (Kyriacou and Hastings, 2010). In preparation for eating during the day, our liver and fat cells maximize the production of metabolizing and energy storage proteins (Froy, 2010; Green et al., 2008; Kovac et al., 2009). As a consequence of the widespread influence of the clock on physiology, disruption of the clock, either through mutation, or by living against the clock (e.g. shift work), significantly impacts our health (Karatsoreos, 2012), and through clock regulation of metabolism, drugs are more effective or more toxic depending on the time of administration (Griffett and Burris, 2013). Circadian clocks have a major influence on transcription. Between 10-40% of the eukaryotic genome is under control of the clock at the level of transcription initiation (Duffield, 2003; Koike et al., 2012; Menet et al., 2012; Nagel and Kay, 2012; Partch et al., 2014; Vitalini et al., 2006). These transcripts peak in abundance at all possible times of the day (phase) of the circadian cycle, implicating the action of a network of transcription factors controlling phase. However,

basic information on how the clock coordinates rhythmic gene expression to particular times of the day was lacking. Thus, the goal of my work was to begin to elucidate the complex transcriptional network that regulates circadian oscillations in gene expression, processes that are critical to understanding the molecular programs that drive circadian physiology and associated disease.

The purpose of this thesis was to determine the molecular connections that exist between the core clock components and the output pathways in the model organism *N. crassa*. We found that the core clock component WCC controls rhythms in the accumulation of 24 TFs, including both activators and repressors of transcription, whose levels peak in the morning. These data suggested a flat hierarchical network model whereby the morning active WCC regulates morning-expression and accumulation of a suite of TFs. Morning-specific activators would then control the expression of second tier genes that would peak a little later in the day, whereas morning-specific repressors of transcription would control the expression of the second tier genes that would peak in the opposite phase (evening). The set of second tier genes would likely encode additional TFs, which would positively and negatively control genes that peak slightly later in the day, and so on. In this way, ccgs that peak at all times of day would be represented in the flat hierarchical model. To test this model, I focused on identifying the direct and indirect ccg targets of the first tier TF ADV-1, which functions as both a positive and negative regulator of transcription, and determining what happens to rhythmicity of these targets when ADV-1 is deleted in cells. In contrast to the flat hierarchical model, I found that

the circadian output pathways are significantly more complex, with very little correlation between ADV-1 and phase regulation of target ccgs in the pathway. Thus, the original model was revised to include complex feedforward and feedback loops controlling ccg rhythmicity and phase. This network provides the platform for future studies combining both computational methods and experimental validation to unravel the mechanisms controlling the circadian output pathways.

LIGHT RESPONSIVE GENES ARE CONTROLLED BY A HIERARCHICAL NETWORK

Light and the circadian clock are intricately linked in *N. crassa*; therefore, in early work we used RNA-seq to discover what genes are controlled by light to better understand these connections. Prior to this study, light-responsive genes were identified using microarrays, which missed poorly expressed genes, and did not uncover any genes that were repressed by light. Two significant findings came from our light-regulated transcriptome study (Chapter II). First, we identified a novel group of genes that were repressed by light, and discovered that most of these genes encode proteins that are involved in ribosome biogenesis. These data suggested that ribosome production, and possibly increased translation, occurs at night. Consistent with this idea, our lab has shown that translation elongation is inhibited during the daytime as a result of rhythmic phosphorylation of eEF-2 (Caster and Bell-Pedersen, unpublished data). Furthermore, phosphorylation of translation initiation factor eIF2α, which inhibits translation

initiation, peaks during the day (Caster, Seifert, and Bell-Pedersen, unpublished data). Translation is one of the most energy demanding processes in the cell; therefore, based on these data, I predicted that the cell would have increased energy levels at night time to support a night time peak in translation, and that this would occur by the cell having increased glycolysis at night to produce ATP. In addition, I predicted that gluconeogenesis would peak at the opposite phase, during the day. In support of this idea, I found that the levels of *fbp-1* mRNA, encoding fructose 1-6 bisphosphatase, the key enzyme required for gluconeogenesis, cycles over the course of the day, peaking in the early morning. As a result of this work, it is now of interest to determine if the enzymatic activity of fructose 1-6 bisphosphatase, and phosphofructokinase, a key enzyme in glycolysis, cycles in antiphase.

A second major discovery from the light transcriptome studies was that the timing of the expression of light-responsive genes fits the flat hierarchical model for regulation of gene expression, with early light-responsive genes being direct targets of the WCC, and encoding TFs that are necessary for light induction of the late light-responsive genes. Most of the light regulated genes are also clock-controlled; therefore, these data led to the expectation that a similar hierarchical pathway would be responsible for regulating the circadian output gene network.

IDENTIFICATION OF CIRCADIAN OUTPUT PATHWAY COMPONENTS

To begin to determine if the circadian output network is hierarchical, I collaborated on a project to test the hypothesis that the WCC controls rhythmic expression of downstream target genes. This was an idea that had been proposed in the circadian field for many years, but had not yet been tested directly at a global level. It only became possible to map binding sites of transcription factors across the genomic landscape following the advent of high-throughput sequencing coupled to chromatin immunoprecipitation (ChIP-seq) (Johnson et al., 2007). We mapped the WC-2 binding sites across the *N. crassa* genome after a brief light pulse, which activates the WCC and stimulates DNA binding. Our expectation was that most, if not all, of the direct target genes would be regulated by both light and the clock. This was shown to be true for the majority of the WCC direct target genes, including the over represented class of proteins that function as transcription factors (TFs). A significant outcome of this work was the demonstration that the WCC directly controls the timing of expression of several first-tier TFs, which in turn, function to convey temporal information further downstream to amplify the number of ccgs. Following our initial study, similar experiments were performed in other organisms, including *Drosophila* (Abruzzi et al., 2011) and mice (Rey et al., 2011), identifying direct targets of the core clock TFs, CLK/CYC and CLOCK/BMAL1, respectively. In agreement with our findings, both studies found TFs to be significantly enriched as direct targets of the clock components. Altogether, data suggest TF-directed

hierarchical regulatory network to be a conserved feature of the circadian clock in eukaryotes.

CIRCADIAN CLOCK CONTROL OF PHASE

A major question that still needed to be addressed was how does the circadian clock determine the phase of target ccgs? Based on our discovery that rhythmic TFs are direct targets of the WCC, I originally proposed two models, which are not mutually exclusive, to explain how the circadian clock might influence ccg phase. The first model suggested that a clock-controlled transcriptional activator and a repressor would drive expression of ccgs in opposite phase. Supporting evidence in favor of the first model was provided by work on CSP-1 (Conidial Separation-1), one of the first-tier TFs in *N. crassa*. In this study, CSP-1 was shown to be a morning-specific repressor, and as predicted, a subset of direct CSP-1 targets were repressed in the morning and peaked in the evening (Sancar et al., 2011). However, not all evening-specific *ccgs* were found to be direct targets of CSP-1. One explanation for this finding is that other first-tier repressors exist besides CSP-1. Indeed, I discovered that ADV-1 acts both as an activator and a repressor of its direct targets. Therefore, the mechanism of transcriptional regulation for the remaining first-tier TFs warrants further investigation. This work is currently in progress using ChIP- and RNA-seq for all of the first tier TFs. The second model predicted that there are clock-controlled transcriptional activators at each tier in the hierarchical regulatory network that sequentially control the phase of ccgs. In this model, as we move down the

cascade, the peak phase of *ccg* expression will occur later than those at the top of the cascade. A prediction of this model is that an evening-specific *ccg* would be towards the end of the TF-directed hierarchy. Indirect evidence to support the second model is provided by SUB-1 (Submerged protopethecia-1), a first-tier WCC target TF shown to regulate late light-induced genes (Chen et al., 2009). Although this still needs to be tested, based on the regulation of late light-induced genes, SUB-1 may regulate ccgs that peak later in the day. At this point, I cannot rule out either model, and suspect based on our current data, that aspects of both models are involved in controlling the peak expression for some of the ccgs.

To further test the models for phase control of ccgs by the output network, I characterized the role of the first-tier TF ADV-1. Both *adv-1* mRNA and protein levels are rhythmic peaking during the day, and the rhythms are dependent on a functional clock. To understand how ADV-1 conveys temporal information to influence downstream rhythmic processes, I performed ChIP-seq to identify ADV-1 direct targets following light treatment of the cultures to induce ADV-1 levels, and RNA-seq to identify genes that are differentially regulated in WT compared to ΔADV-1 cells. Not surprisingly, many of the direct and indirect ADV-1 targets are predicted *ccgs* (Dong et al., 2008; Hurley et al., 2014; Lewis et al., 2002). Importantly, for the majority of light-regulated genes bound by ADV-1, ADV-1 binding increased in the light-treated samples as compared to the dark grown samples, and light regulation was abolished in ΔADV-1 cells as predicted. These data confirmed that ADV-1 is necessary for signaling light

information to some of its downstream targets, supporting the existence of a simple flat hierarchical network of TFs controlling light signaling pathways. On the contrary, ADV-1 control of ccg expression did not occur in predictable ways. For example, *ham-9*, encoding a protein containing a SAM (Sterile alpha motif) domain and two PH (Pleckstrin homology) domains, is repressed by ADV-1, whereas *prm-1*, encoding a transmembrane protein shown to be important for plasma membrane merging during fusion, is activated by ADV-1. I predicted that the *ccgs* that are activated by ADV-1 would have mRNA levels that peak during the subjective daytime, at or near the peak in ADV-1 levels. Alternatively, I predicted that genes that are repressed by ADV-1 would peak in the opposite phase, in the subjective night. Surprisingly, both *ham-9* and *prm-1* mRNA levels peaked at the same time, during the subjective night. Therefore, a flat hierarchical activation/repression model is not sufficient to explain the night-time expression of *ccgs* that are activated by ADV-1. One possible explanation for these results is that the promoter-driven luciferase reporter constructs used to monitor rhythmic mRNA levels lacked key *cis*-regulatory elements responsible for phase-specificity. While the luc reporter constructs included the ADV-1 binding site sequences identified by ChIP-seq, ADV-1 may function with other TFs to regulate downstream ccgs, and these regulatory elements might be lacking in the reporter constructs. If so, this could explain the discrepancy in phase. To test this possibility, I measured *ham-9* and *prm-1* mRNA levels in WT cells harvested over a circadian time course. Consistent with the luc reporter assays, both mRNAs peaked in the subjective evening, thus eliminating the lack of regulatory elements as the cause for our inability to predict phase. As such,

the major outcome of this work was the discovery that, unlike the flat hierarchical TF network controlling light responses, circadian output pathways are more complex, possibly involving several TFs functioning in a coordinated fashion to sculpt the rhythms in target genes. Recent reports of combinatorial TF activity at eukaryotic promoters are consistent with these results, and suggest that this phenomenon may be more common than previously appreciated (Fessele et al., 2002; Harbison et al., 2004; Martinez and Rao, 2012). As a result of this discovery, it became clear that a candidate gene approach using single deletion strains is not sufficient to gain a full understanding of the mechanistic details of the circadian gene regulatory network, and in particular how the phase of ccgs are determined.

SYSTEMS APPROACH TO MAP THE CIRCADIAN OUTPUT REGULATORY NETWORK

TFs can control gene expression in a combinatorial manner, rather than in isolation (Fessele et al., 2002; Harbison et al., 2004; Martinez and Rao, 2012). In support of this idea for circadian output networks, Drosophila CLK/CYK was recently shown to function synergistically with tissue-specific TFs bound to target gene promoters to control tissue-specific rhythmic gene expression (Meireles-Filho et al., 2014). Furthermore, in mammals, the timing of peak expression of a significant proportion of BMAL-1 target genes occurred much later than the daytime peak in BMAL-1 DNA

binding activity (Rey et al., 2011), supporting more complex regulatory circuits, similar to what I have observed for the downstream ADV-1 network in *N. crassa*.

Based on our TF ChIP-seq data, I found that ADV-1, along with a number of other first-tier TFs, often bind to the same regulatory elements on the genome, suggesting cooperative binding facilitated by protein-protein interaction between TFs. To test this hypothesis, we need to determine if these TFs physically interact with each other *in vivo* using co-immunoprecipitation (co-IP). To accomplish this, antibodies will need to be generated for each of the TFs, or strains containing multiply tagged TFs need to be generated. If we find that the TFs do not interact off of the DNA, we can use sequential chromatin immunoprecipitation (ChIP-reChIP) to test if they interact on the DNA. This will be accomplished using TFs tagged with different epitopes followed by PCR with primers for the regions containing binding sites for both TFs.

Deciphering phase determination mechanisms will likely require an integrative approach combining computational and experimental studies that allows for an iterative process of model-driven predictions, experimental validation, and refinement of the model. The benefits of such an approach applied to the circadian clock is best exemplified in Arabidopsis, where sophisticated models of the plant oscillator mechanism revealed the need for more components and regulatory feedback loops to explain the experimental data (Hubbard et al., 2009; Locke et al., 2006; Locke et al., 2005).

A major theme of the thesis was the emphasis on studying the dynamics of circadian output networks, instead of merely looking at individual components. A full description of the circadian output regulatory network would involve identifying all the components and mapping the interaction among these components and then analyzing their behavior in living cells. However, such an approach is likely not feasible. If, instead, we could generate computational models describing the interactions between a few of the components of the clock, and if these models are validated, the expectation is that they would provide a tool to describe the entire circadian network.

Although not presented in this dissertation, in an ongoing collaborative effort, I performed ChIP-seq on several of the first-tier WCC direct target TFs and identified candidate direct target genes. RNA-seq was also performed in WT versus TF-deletion cells to identify differentially expressed genes. These data enabled the group to construct a network of interactions among the first-tier TFs (Figure 5-1). The network is built from the first-tier TFs (nodes) that are connected to one another through "links", which signifies direct binding based on ChIP-seq data. The upstream network reveals that the promoter of *adv-1* is not only bound by the WCC, but also by several of the other first-tier TFs in feed forward loops. This regulation is thought to be important in controlling the phase of *adv-1* mRNA rhythms, and experiments are underway to test this prediction. In addition, ADV-1 binds to the promoters of several of the first-tier TFs forming feedback loops, and similarly is predicted to impart phase information on these TFs.

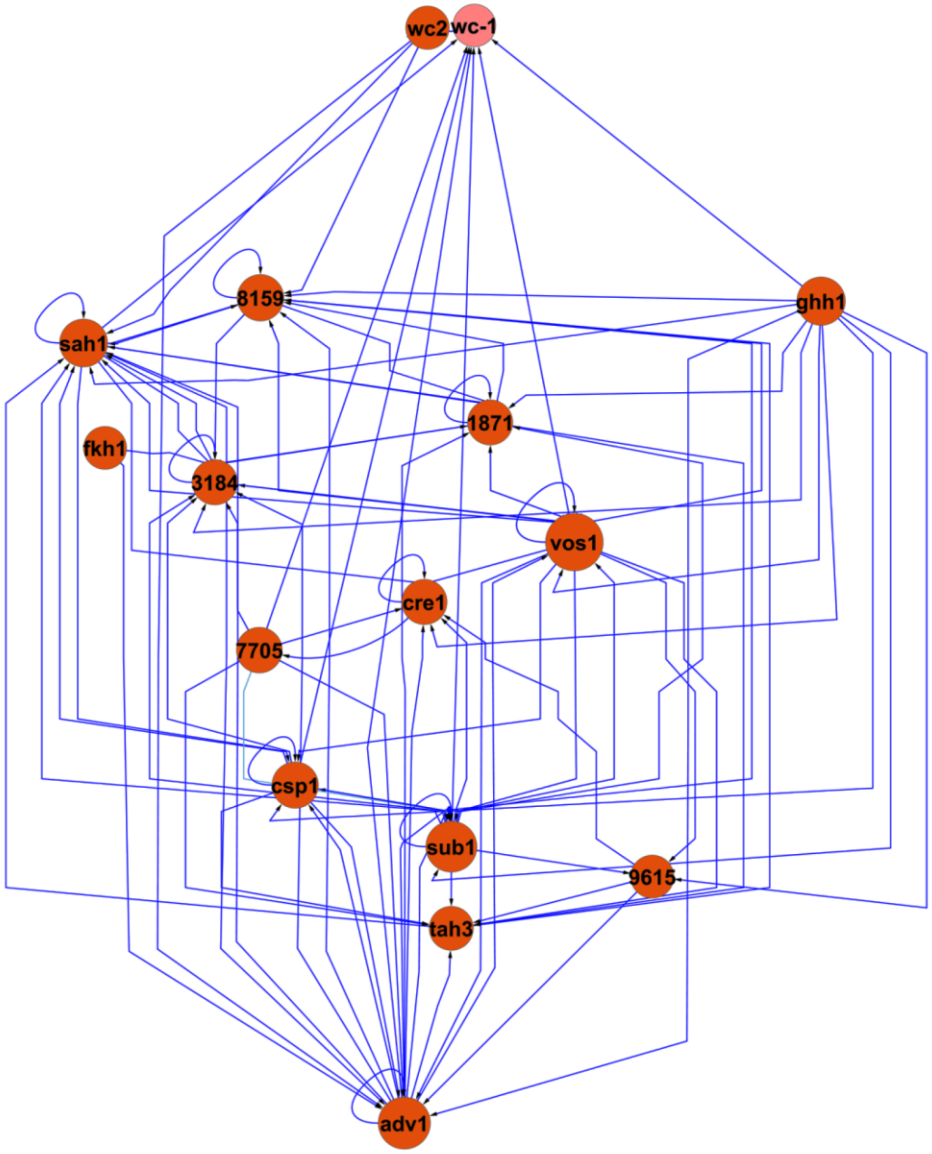


Figure 5-1. Schematic representation of the transcriptional network of the output pathway between WCC and ADV-1. 13 first-tier TFs that bind to the *adv-1* promoter based on ChIP-seq data are shown here. "Nodes" represent TFs and their regulators (TFs) joined together by "edges", which indicate ChIP-seq binding between a TF (regulator) and the promoter region of the target gene (TF). WCC is at the top and ADV-1 is at the bottom of the network.

The downstream ADV-1 network is similarly complex (Figure 5-2). Many of the first-tier TFs bind to the same gene targets as ADV-1, and this regulation may be critical to phase regulation of the ccgs. To test this idea, we have applied available rhythmic time course data from approximately 4000 *ccgs* identified in microarray studies (Dong et al., 2008). These data allowed us to model gene expression using frequency domain methods (Geva-Zatorsky et al., 2010). The frequency data set, along with TFs binding data set based on ChIP-seq were integrated to model the ADV-1 circadian output network (Figure 5-3). The model allows us to predict which perturbation in the regulatory network will influence the mRNA levels and rhythmicity of downstream ccgs, and to determine if there are any specific modules that control ccgs that peak at the same time of the day. Interestingly, I found that ADV-1 and VOS-1 share a large number of overlapping target genes that peak at the same time of the day (Figure 5-4). The model makes specific predictions on what would happen to the expression of these genes if one or both of the TFs are deleted in cells, and I am now testing these predictions. Finally, ADV-1 binds to the promoter of at least 8 additional TFs. Following ChIP-seq and RNA-seq over a circadian time course in WT and TF deletion cells for the ADV-1 targeted TFs, the data will be incorporated into the ADV-1 downstream transcription factor network model to predict which perturbations of the ADV-1 regulatory network impact the regulation of individual ADV-1 downstream targets. It is clear from my data that binding of a TF to DNA does not always result in a change in expression of the corresponding gene. Therefore, it will be important to also determine how deletion of a particular TF alters gene expression, and incorporate this data into the model. This work

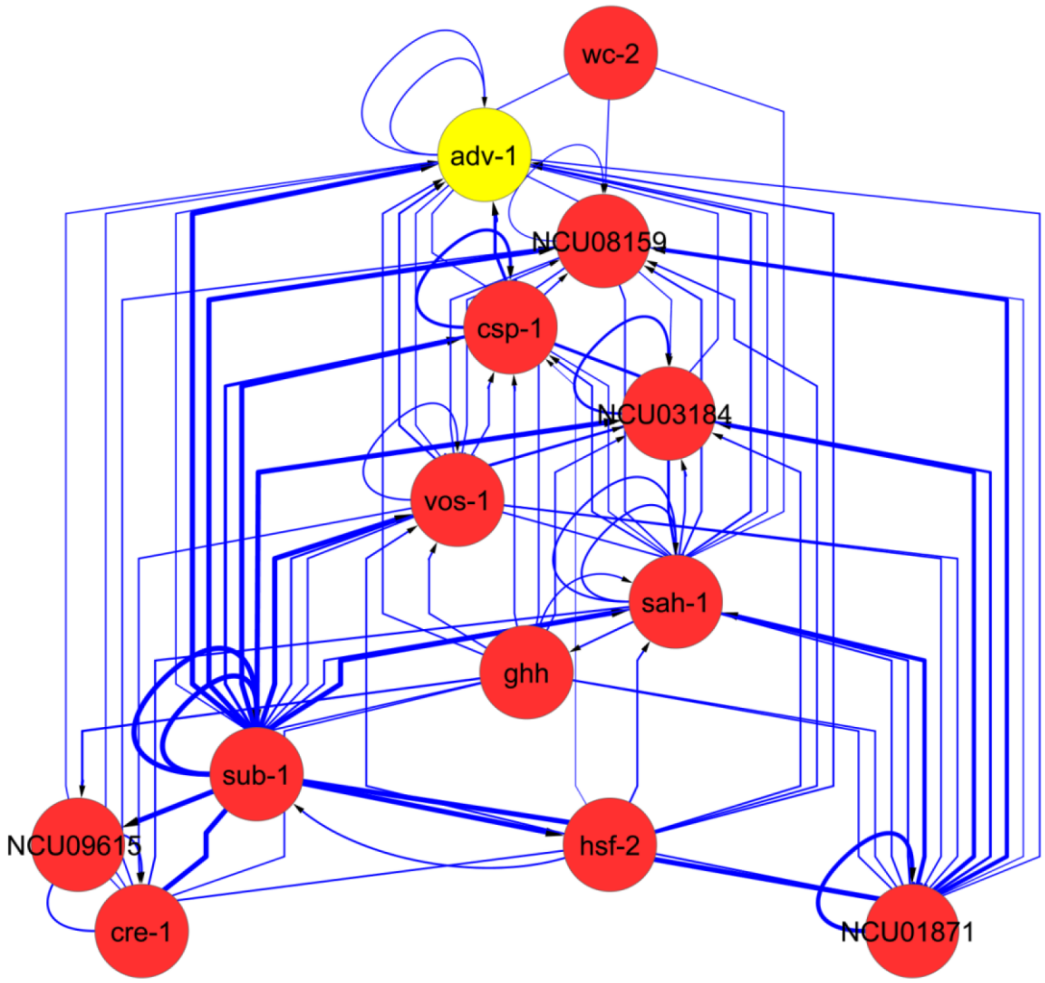


Figure 5-2. Schematic diagram of the circadian output network downstream of ADV-1. TFs that are bound by ADV-1 based on ChIP-seq data are shown. "Nodes" represent TFs and their regulators (TFs) joined together by "edges", which indicate binding between a TF (regulator) and the promoter region of the target gene (TF).

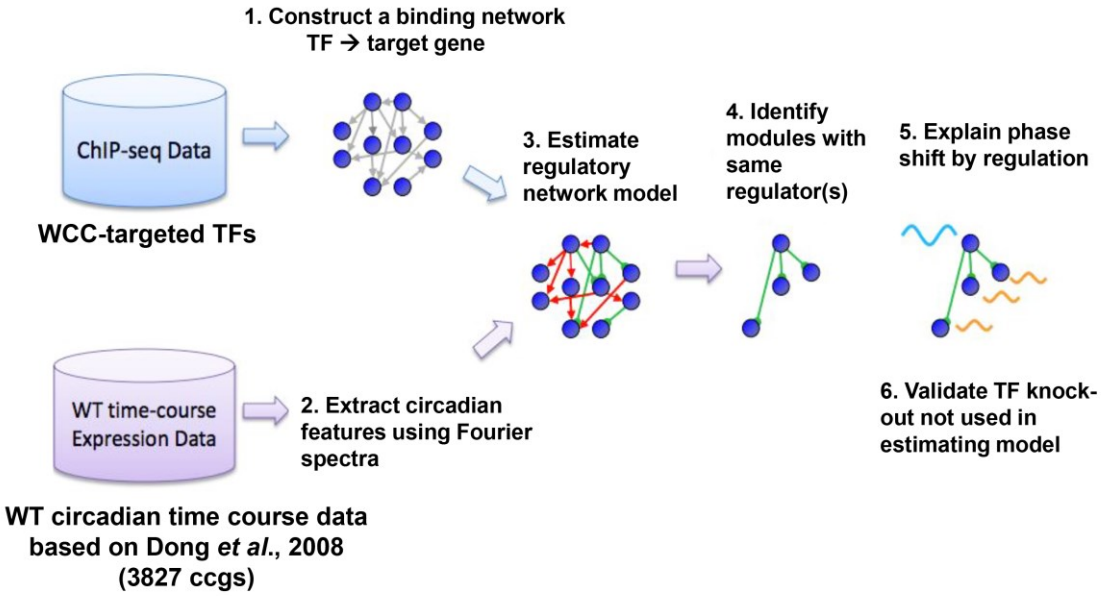


Figure 5-3. Scheme of mapping the circadian output regulatory network. The network is based on binding network (WCC-targeted TFs and their direct targets based on ChIP-seq) and clock-controlled genes (*ccgs*) rhythmic expression profiles (Dong et al., 2008).

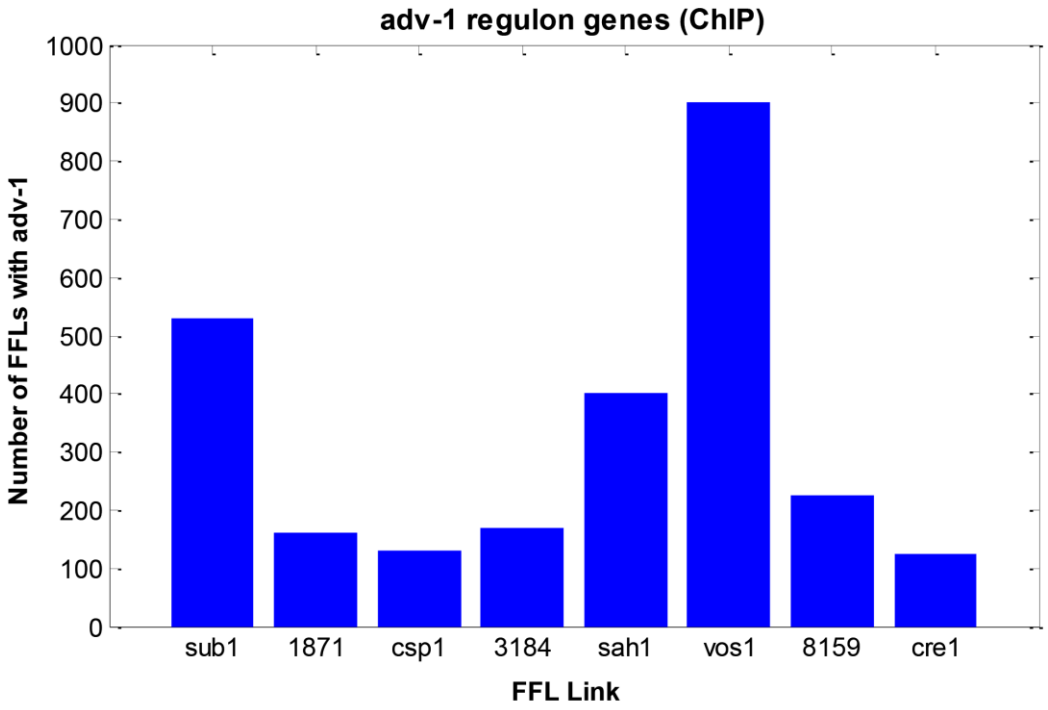


Figure 5-4. Feed forward loop involving ADV-1 and other first-tier TFs. A bar graph showing the total number of direct target genes on the y-axis that are co-regulated by ADV-1 and the indicated TFs on the x-axis.

determine if there are specific regulatory nodes used to control groups of genes that function in a particular biological process and that peak at the same time of day.

The ultimate goal is to be able to map the clock output regulatory network based on the design principles derived through the efforts of system identification, system analysis and system perturbation. Such studies will not only advance our understanding of circadian output pathways in *N. crassa*, but will help to uncover paradigms that can be applied to higher eukaryotes. The understanding of circadian output pathways is particularly of significance in humans because it provides the potential to develop interventions to diminish the deleterious effects of the clock disruption on human diseases, such as metabolic disorders that are prevalent in shift workers.

FUTURE DIRECTIONS

Clock control of cell fusion

Through the identification of ADV-1 direct target genes, I discovered an enrichment for genes involved in cell fusion. These data were consistent with the phenotype of ΔADV-1 cells, which have defects in asexual and sexual development, and are deficient in cell fusion (Colot et al., 2006; Fu et al., 2011). Because development requires cell fusion, it is likely that the developmental defects are the result of misregulation of the cell fusion genes in ΔADV-1 cells. Consistent with this idea, I found that the several key cell fusion

genes have constitutively low or high levels of mRNA in ΔADV-1 cells. I also predicted that the cell fusion genes would be rhythmically expressed based on rhythmic accumulation of ADV-1 protein. The cell fusion genes were indeed found to be rhythmic, however, it proved challenging to study the influence of clock on cell fusion because of the transient nature of the process, and inability to capture them at the moment when fusion occurs between hyphae. It would be interesting to develop more amenable experimental methods to monitor cell fusion, like tagging cell fusion proteins with fluorescent dye and tracking them as hyphae fuse with each other over a circadian time course.

Clock control of metabolism

The regulation of metabolic homeostasis is a central function of the clock (Bass and Takahashi, 2010). A large number of metabolic genes are regulated by the clock in mammals, including rate-limiting enzymes involved in glucose metabolism, and mice with defective clock genes are more prone to metabolic syndrome (Kovac et al., 2009). However, the mechanistic links between the clock and metabolism are not understood. ChIP-seq data suggested a potential role for ADV-1 in metabolism, with a significant number of genes bound by ADV-1 functioning in central carbon, amino acid, and nucleotide metabolism. These data provide the opportunity to unravel the mechanistic links between the clock and metabolism in a simple eukaryotic model system.

In order to understand the role of clock in metabolism, we are building on an existing genome-scale metabolic model in *N. crassa*. The model was able to accurately predict gene essentiality, nutrient requirements and synthetic lethals using a computational approach called flux-balance analysis (FBA) (Dreyfuss et al., 2013). FBA calculates the flow of metabolites through metabolic networks, thereby making it possible to make predictions (Orth et al., 2010). We are using an improved FBA method called E-flux, where the well-established and validated *N. crassa* metabolic model is integrated with mRNA expression data to increase the predictive power of the model (Brandes et al., 2012; Colijn et al., 2009). The gene expression data is used to constrain the availability of metabolic reactions. This approach was previously used to predict the impact of drugs on mycolic acid biosynthesis in *Mycobacterium tuberculosis* (Colijn et al., 2009). Conceptually, if the transcript level of a particular enzyme is low, then there will be more constraint on the maximum possible flux through the corresponding metabolic reaction and conversely, if the transcript level of a particular enzyme is high, then there will be less constraint on the maximum possible flux through the corresponding metabolic reaction (Colijn et al., 2009). However, this might not always be accurate as this approach does not take into account post-transcriptional and translational regulation that can influence enzymatic activity. In any case, E-flux can predict changes in metabolites following gene deletion, and in response to changes in mRNA expression. To understand temporal coordination of metabolism, we will use E-flux to study the impact of the clock, through ADV-1, on metabolism. First, integrating the metabolic model with high-resolution RNA-seq circadian data from WT cells will help predict the

changes in metabolite levels over the course of the day. Second, we will predict the influence of ADV-1 on metabolite levels during a circadian cycle based on gene expression data. The predictions will be validated by measuring the metabolites under genetic perturbations using commercially available colorimetric assays.

Circadian control of the cell cycle, DNA damage and DNA repair

Light and the circadian clock influence most major physiological processes in eukaryotes. Consistent with this idea, we identified several WC-2 binding sites in the promoters of known or predicted genes involved in cell cycle regulation, DNA replication and DNA repair. These data lend support to the known link between the clock and cell cycle in fungi and mammals (Hong et al., 2014; Masri et al., 2013; Pregueiro et al., 2006). As predicted for WCC targets, a majority of these proteins were found to be light-responsive, and some are predicted *ccgs* (Table 5-1). The link between light-response, circadian clock and DNA repair is not surprising in view of the mutagenic effects of light. In fact, one of the evolutionary advantages of the circadian clock is that it anticipates DNA damage and maintains genomic integrity through DNA repair in the face of damaging effects of the UV-light (Pittendrigh, 1993). Interestingly, some of the WC-2 direct targets were also identified as a direct target of ADV-1 (Chapter IV), reiterating the complexity of the transcriptional regulatory network controlling rhythmic outputs, in this case regulating DNA repair and cell cycle progression. Even though not included in the table, all these genes are bound by at least

Table 5-1. List of WCC direct target genes involved in cell cycle, DNA damage and DNA repair

NCU	Gene	Description	Light-regulated	CCG (RNA-seq)	ADV-1 direct target
NCU03072		DNA replication regulator	+	+	
NCU03968		DNA repair, SUMO ligase	+	+	+
NCU11300		anaphase promoting complex subunit 11	+		+
NCU05721		arrestin domain	+		
NCU01242		cyclinB3	+		
NCU00202		actin binding protein	+		
NCU00213	*bimA*	cdc27		+	
NCU07407	*msh-1*	DNA repair protein			+
NCU08290	*mus-51*	DNA repair protein	+	+	
NCU08806	*rhp-55*	DNA repair protein	+		+
NCU08606		cell cycle inhibitor	+		
NCU08626	*phr*	photolyase	+	+	
NCU08850	*mus-18*	UV damage induced	+	+	
NCU05781		spindle checkpoint protein			

two or more first-tier TFs along with WC-2. Once established, the systems approach to map the circadian output network, provides an exciting opportunity to map the regulation of these *ccgs* in *N*. *crassa*, and apply this information to higher eukaryotes.

Altogether, these proposed studies will provide important information concerning the design of circadian output networks and their influence on rhythmic processes like metabolism. The work presented in this thesis provides a foundation upon which these other studies can be built. The mechanistic understanding of circadian output pathways in eukaryotes and their relevance to clinical application will provide the opportunity to develop targeted interventions to mitigate the effects of clocks disruptions in humans such as metabolic disorders and cancer.

REFERENCES

Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S., and Rosbash, M. (2011). Drosophila CLOCK target gene characterization: implications for circadian tissue-specific gene expression. Genes & Development *25*, 2374-2386.

Adomas, A.B., Lopez-Giraldez, F., Clark, T.A., Wang, Z., and Townsend, J.P. (2010). Multi-targeted priming for genome-wide gene expression assays. BMC Genomics *11*, 477.

Aguilar, P.S., Baylies, M.K., Fleissner, A., Helming, L., Inoue, N., Podbilewicz, B., Wang, H., and Wong, M. (2013). Genetic basis of cell-cell fusion mechanisms. Trends in Genetics *29*, 427-437.

Akhtar, R.A., Reddy, A.B., Maywood, E.S., Clayton, J.D., King, V.M., Smith, A.G., Gant, T.W., Hastings, M.H., and Kyriacou, C.P. (2002). Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. Current Biology *12*, 540-550.

Aronson, B.D., Johnson, K.A., and Dunlap, J.C. (1994a). Circadian clock locus frequency: protein encoded by a single open reading frame defines period length and temperature compensation. Proceedings of the National Academy of Sciences of the United States of America *91*, 7683-7687.

Aronson, B.D., Johnson, K.A., Loros, J.J., and Dunlap, J.C. (1994b). Negative feedback defining a circadian clock: autoregulation of the clock gene frequency. Science *263*, 1578-1584.

Bailey, T.L., Williams, N., Misleh, C., and Li, W.W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Research *34*, W369-373.

Baker, C.L., Kettenbach, A.N., Loros, J.J., Gerber, S.A., and Dunlap, J.C. (2009). Quantitative proteomics reveals a dynamic interactome and phase-specific phosphorylation in the Neurospora circadian clock. Molecular Cell *34*, 354-363.

Ballario, P., Vittorioso, P., Magrelli, A., Talora, C., Cabibbo, A., and Macino, G. (1996). White collar-1, a central regulator of blue light responses in Neurospora, is a zinc finger protein. The EMBO Journal *15*, 1650-1657.

Bass, J., and Takahashi, J.S. (2010). Circadian integration of metabolism and energetics. Science *330*, 1349-1354.

Beadle, G.W., and Tatum, E.L. (1941). Genetic Control of Biochemical Reactions in Neurospora. Proceedings of the National Academy of Sciences of the United States of America *27*, 499-506.

Belden, W.J., Larrondo, L.F., Froehlich, A.C., Shi, M., Chen, C.H., Loros, J.J., and Dunlap, J.C. (2007). The band mutation in Neurospora crassa is a dominant allele of ras-1 implicating RAS signaling in circadian output. Genes & Development *21*, 1494-1505.

Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L., and Zoran, M.J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nature Reviews Genetics *6*, 544-556.

Bell-Pedersen, D., Dunlap, J.C., and Loros, J.J. (1996a). Distinct cis-acting elements mediate clock, light, and developmental regulation of the Neurospora crassa eas (ccg-2) gene. Molecular and Cellular Biology *16*, 513-521.

Bell-Pedersen, D., Garceau, N., and Loros, J. (1996b). Circadian rhythms in fungi. Journal of Genetics *75*, 387-401.

Bell-Pedersen, D., Lewis, Z.A., Loros, J.J., and Dunlap, J.C. (2001). The Neurospora circadian clock regulates a transcription factor that controls rhythmic expression of the output eas (ccg-2) gene. Molecular Microbiology *41*, 897-909.

Bell-Pedersen, D., Shinohara, M.L., Loros, J.J., and Dunlap, J.C. (1996c). Circadian clock-controlled genes isolated from Neurospora crassa are late night- to early morning-specific. Proceedings of the National Academy of Sciences of the United States of America *93*, 13096-13101.

Bennett, L.D., Beremand, P., Thomas, T.L., and Bell-Pedersen, D. (2013). Circadian activation of the mitogen-activated protein kinase MAK-1 facilitates rhythms in clock-controlled genes in Neurospora crassa. Eukaryotic Cell *12*, 59-69.

Berlin, V., and Yanofsky, C. (1985). Isolation and characterization of genes differentially expressed during conidiation of Neurospora crassa. Molecular and Cellular Biology *5*, 849-855.

Berman-Frank, I., Lundgren, P., and Falkowski, P. (2003). Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. Research in Microbiology *154*, 157-164.

Bieszke, J.A., Braun, E.L., Bean, L.E., Kang, S., Natvig, D.O., and Borkovich, K.A. (1999a). The nop-1 gene of Neurospora crassa encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. Proceedings of the National Academy of Sciences of the United States of America *96*, 8034-8039.

Bieszke, J.A., Li, L., and Borkovich, K.A. (2007). The fungal opsin gene nop-1 is negatively-regulated by a component of the blue light sensing pathway and influences conidiation-specific gene expression in Neurospora crassa. Current Genetics *52*, 149-157.

Bieszke, J.A., Spudich, E.N., Scott, K.L., Borkovich, K.A., and Spudich, J.L. (1999b). A eukaryotic protein, NOP-1, binds retinal to form an archaeal rhodopsin-like photochemically reactive pigment. Biochemistry *38*, 14138-14145.

Bobrowicz, P., Pawlak, R., Correa, A., Bell-Pedersen, D., and Ebbole, D.J. (2002). The Neurospora crassa pheromone precursor genes are regulated by the mating type locus and the circadian clock. Molecular Microbiology *45*, 795-804.

Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics *20*, 3710-3715.

Brandes, A., Lun, D.S., Ip, K., Zucker, J., Colijn, C., Weiner, B., and Galagan, J.E. (2012). Inferring carbon sources from gene expression profiles using metabolic flux models. PloS One *7*, e36947.

Brown, S.A. (2014). Circadian clock-mediated control of stem cell division and differentiation: beyond night and day. Development *141*, 3105-3111.

Butler, M., and Day, A. (1998). Fungal melanins: a review. Canadian Journal of Microbiology *44*, 1115-1136.

Campbell, M.A., Haas, B.J., Hamilton, J.P., Mount, S.M., and Buell, C.R. (2006). Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. BMC Genomics *7*, 327.

Carlson, J.M., Chakravarty, A., DeZiel, C.E., and Gross, R.H. (2007). SCOPE: a web server for practical de novo motif discovery. Nucleic Acids Research *35*, W259-264.

Ceriani, M.F., Hogenesch, J.B., Yanovsky, M., Panda, S., Straume, M., and Kay, S.A. (2002). Genome-wide expression analysis in Drosophila reveals genes controlling circadian behavior. The Journal of Neuroscience *22*, 9305-9319.

Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. Nature *447*, 497-500.

Chen, C.H., DeMay, B.S., Gladfelter, A.S., Dunlap, J.C., and Loros, J.J. (2010a). Physical interaction between VIVID and white collar complex regulates photoadaptation in Neurospora. Proceedings of the National Academy of Sciences of the United States of America *107*, 16715-16720.

Chen, C.H., Dunlap, J.C., and Loros, J.J. (2010b). Neurospora illuminates fungal photoreception. Fungal Genetics and Biology *47*, 922-929.

Chen, C.H., and Loros, J.J. (2009). Neurospora sees the light: light signaling components in a model system. Communicative & Integrative Biology *2*, 448-451.

Chen, C.H., Ringelberg, C.S., Gross, R.H., Dunlap, J.C., and Loros, J.J. (2009). Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora. The EMBO Journal *28*, 1029-1042.

Cheng, M.Y., Bullock, C.M., Li, C., Lee, A.G., Bermak, J.C., Belluzzi, J., Weaver, D.R., Leslie, F.M., and Zhou, Q.Y. (2002a). Prokineticin 2 transmits the behavioural circadian rhythm of the suprachiasmatic nucleus. Nature *417*, 405-410.

Cheng, P., He, Q., He, Q., Wang, L., and Liu, Y. (2005). Regulation of the Neurospora circadian clock by an RNA helicase. Genes & Development *19*, 234-241.

Cheng, P., Yang, Y., Gardner, K.H., and Liu, Y. (2002b). PAS domain-mediated WC-1/WC-2 interaction is essential for maintaining the steady-state level of WC-1 and the function of both proteins in circadian clock and light responses of Neurospora. Molecular and Cellular Biology *22*, 517-524.

Cheng, P., Yang, Y., Heintzen, C., and Liu, Y. (2001a). Coiled-coil domain-mediated FRQ-FRQ interaction is essential for its circadian clock function in Neurospora. The EMBO Journal *20*, 101-108.

Cheng, P., Yang, Y., and Liu, Y. (2001b). Interlocked feedback loops contribute to the robustness of the Neurospora circadian clock. Proceedings of the National Academy of Sciences of the United States of America *98*, 7408-7413.

Cheng, P., Yang, Y., Wang, L., He, Q., and Liu, Y. (2003). WHITE COLLAR-1, a multifunctional neurospora protein involved in the circadian feedback loops, light sensing, and transcription repression of wc-2. The Journal of Biological Chemistry *278*, 3801-3808.

Colijn, C., Brandes, A., Zucker, J., Lun, D.S., Weiner, B., Farhat, M.R., Cheng, T.Y., Moody, D.B., Murray, M., and Galagan, J.E. (2009). Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production. PLoS Computational Biology *5*, e1000489.

Collett, M.A., Dunlap, J.C., and Loros, J.J. (2001). Circadian clock-specific roles for the light response protein WHITE COLLAR-2. Molecular and Cellular Biology *21*, 2619-2628.

Colot, H.V., Park, G., Turner, G.E., Ringelberg, C., Crew, C.M., Litvinkova, L., Weiss, R.L., Borkovich, K.A., and Dunlap, J.C. (2006). A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proceedings of the National Academy of Sciences of the United States of America *103*, 10352-10357.

Correa, A., Lewis, Z.A., Greene, A.V., March, I.J., Gomer, R.H., and Bell-Pedersen, D. (2003). Multiple oscillators regulate circadian gene expression in Neurospora. Proceedings of the National Academy of Sciences of the United States of America *100*, 13597-13602.

Corrochano, L.M. (2007). Fungal photoreceptors: sensory molecules for fungal development and behaviour. Photochemical & Photobiological Sciences *6*, 725-736.

Corrochano, L.M. (2011). Fungal photobiology: a synopsis. IMA fungus *2*, 25-28.

Corrochano, L.M., Lauter, F.R., Ebbole, D.J., and Yanofsky, C. (1995). Light and developmental regulation of the gene con-10 of Neurospora crassa. Developmental Biology *167*, 190-200.

Crane, B.R., and Young, M.W. (2014). Interactive features of proteins composing eukaryotic circadian clocks. Annual Review of Biochemistry *83*, 191-219.

Crosthwaite, S.K., Dunlap, J.C., and Loros, J.J. (1997). Neurospora wc-1 and wc-2: transcription, photoresponses, and the origins of circadian rhythmicity. Science *276*, 763-769.

Crosthwaite, S.K., Loros, J.J., and Dunlap, J.C. (1995). Light-induced resetting of a circadian clock is mediated by a rapid increase in frequency transcript. Cell *81*, 1003-1012.

Cusanovich, D.A., Pavlovic, B., Pritchard, J.K., and Gilad, Y. (2014). The functional consequences of variation in transcription factor binding. PLoS Genetics *10*, e1004226.

Damiola, F., Le Minh, N., Preitner, N., Kornmann, B., Fleury-Olela, F., and Schibler, U. (2000). Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. Genes & Development *14*, 2950-2961.

Davis, R.H. (2000). Neurospora: contributions of a model organism, Vol 251 (Oxford University Press New York:).

Davis, R.H., and de Serres, F.J. (1970). [4] Genetic and microbiological research techniques for Neurospora crassa. Methods in Enzymology *17*, 79-143.

de Paula, R.M., Lamb, T.M., Bennett, L., and Bell-Pedersen, D. (2008). A connection between MAPK pathways and circadian clocks. Cell Cycle *7*, 2630-2634.

Denault, D.L., Loros, J.J., and Dunlap, J.C. (2001). WC-2 mediates WC-1-FRQ interaction within the PAS protein-linked circadian feedback loop of Neurospora. The EMBO Journal *20*, 109-117.

Dong, W., Tang, X., Yu, Y., Nilsen, R., Kim, R., Griffith, J., Arnold, J., and Schuttler, H.B. (2008). Systems biology of the clock in Neurospora crassa. PloS One *3*, e3105.

Dreyfuss, J.M., Zucker, J.D., Hood, H.M., Ocasio, L.R., Sachs, M.S., and Galagan, J.E. (2013). Reconstruction and validation of a genome-scale metabolic model for the filamentous fungus Neurospora crassa using FARM. PLoS Computational Biology *9*, e1003126.

Duffield, G.E. (2003). DNA microarray analyses of circadian timing: the genomic basis of biological time. Journal of Neuroendocrinology *15*, 991-1002.

Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271-290.

Dunlap, J.C., and Loros, J.J. (2004). The neurospora circadian system. Journal of Biological Rhythms *19*, 414-424.

Dunlap, J.C., and Loros, J.J. (2006). How fungi keep time: circadian system in Neurospora and other fungi. Current Opinion in Microbiology *9*, 579-587.

Dunlap, J.C., Loros, J.J., Colot, H.V., Mehra, A., Belden, W.J., Shi, M., Hong, C.I., Larrondo, L.F., Baker, C.L., Chen, C.H.*, et al.* (2007). A circadian clock in Neurospora: how genes and proteins cooperate to produce a sustained, entrainable, and compensated biological oscillator with a period of about a day. Cold Spring Harbor Symposia on Quantitative Biology *72*, 57-68.

Dunlap, J.C., Loros, J.J., and DeCoursey, P.J. (2004). Chronobiology: biological timekeeping (Sinauer Associates).

Fahlgren, N., Sullivan, C.M., Kasschau, K.D., Chapman, E.J., Cumbie, J.S., Montgomery, T.A., Gilbert, S.D., Dasenko, M., Backman, T.W., Givan, S.A.*, et al.* (2009). Computational and analytical framework for small RNA profiling by high-throughput sequencing. RNA *15*, 992-1002.

Faou, P., and Tropschug, M. (2003). A novel binding protein for a member of CyP40-type Cyclophilins: N.crassa CyPBP37, a growth and thiamine regulated protein homolog to yeast Thi4p. Journal of Molecular Biology *333*, 831-844.

Faou, P., and Tropschug, M. (2004). Neurospora crassa CyPBP37: a cytosolic stress protein that is able to replace yeast Thi4p function in the synthesis of vitamin B1. Journal of Molecular Biology *344*, 1147-1157.

Feldman, J.F., and Hoyle, M.N. (1973). Isolation of circadian clock mutants of Neurospora crassa. Genetics *75*, 605-613.

Fessele, S., Maier, H., Zischek, C., Nelson, P.J., and Werner, T. (2002). Regulatory context is a crucial part of gene function. Trends in Genetics *18*, 60-63.

Fleissner, A., Diamond, S., and Glass, N.L. (2009a). The Saccharomyces cerevisiae PRM1 homolog in Neurospora crassa is involved in vegetative and sexual cell fusion events but also has postfertilization functions. Genetics *181*, 497-510.

Fleissner, A., Leeder, A.C., Roca, M.G., Read, N.D., and Glass, N.L. (2009b). Oscillatory recruitment of signaling proteins to cell tips promotes coordinated behavior during cell fusion. Proceedings of the National Academy of Sciences of the United States of America *106*, 19387-19392.

Fleissner, A., Simonin, A.R., and Glass, N.L. (2008). Cell fusion in the filamentous fungus, Neurospora crassa. Methods in Molecular Biology *475*, 21-38.

Froehlich, A.C., Chen, C.H., Belden, W.J., Madeti, C., Roenneberg, T., Merrow, M., Loros, J.J., and Dunlap, J.C. (2010). Genetic and molecular characterization of a cryptochrome from the filamentous fungus Neurospora crassa. Eukaryotic Cell *9*, 738-750.

Froehlich, A.C., Liu, Y., Loros, J.J., and Dunlap, J.C. (2002). White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. Science *297*, 815-819.

Froehlich, A.C., Loros, J.J., and Dunlap, J.C. (2003). Rhythmic binding of a WHITE COLLAR-containing complex to the frequency promoter is inhibited by FREQUENCY. Proceedings of the National Academy of Sciences of the United States of America *100*, 5914-5919.

Froehlich, A.C., Noh, B., Vierstra, R.D., Loros, J., and Dunlap, J.C. (2005). Genetic and molecular analysis of phytochromes from the filamentous fungus Neurospora crassa. Eukaryotic Cell *4*, 2140-2152.

Froy, O. (2010). Metabolism and circadian rhythms--implications for obesity. Endocrine Reviews *31*, 1-24.

Fu, C., Iyer, P., Herkal, A., Abdullah, J., Stout, A., and Free, S.J. (2011). Identification and characterization of genes required for cell-to-cell fusion in Neurospora crassa. Eukaryotic Cell *10*, 1100-1109.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S.*, et al.* (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature *422*, 859-868.

Garceau, N.Y., Liu, Y., Loros, J.J., and Dunlap, J.C. (1997). Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell *89*, 469-476.

Geva-Zatorsky, N., Dekel, E., Batchelor, E., Lahav, G., and Alon, U. (2010). Fourier analysis and systems identification of the p53 feedback loop. Proceedings of the National Academy of Sciences of the United States of America *107*, 13550-13555.

Gin, E., Diernfellner, A.C., Brunner, M., and Hofer, T. (2013). The Neurospora photoreceptor VIVID exerts negative and positive control on light sensing to achieve adaptation. Molecular Systems Biology *9*, 667.

Goldsmith, C.S., and Bell-Pedersen, D. (2013). Diverse roles for MAPK signaling in circadian clocks. Advances in Genetics *84*, 1-39.

Gooch, V.D., Mehra, A., Larrondo, L.F., Fox, J., Touroutoutoudis, M., Loros, J.J., and Dunlap, J.C. (2008). Fully codon-optimized luciferase uncovers novel temperature characteristics of the Neurospora clock. Eukaryotic Cell *7*, 28-37.

Gorl, M., Merrow, M., Huttner, B., Johnson, J., Roenneberg, T., and Brunner, M. (2001). A PEST-like element in FREQUENCY determines the length of the circadian period in Neurospora crassa. The EMBO Journal *20*, 7074-7084.

Green, C.B., Takahashi, J.S., and Bass, J. (2008). The meter of metabolism. Cell *134*, 728-742.

Griffett, K., and Burris, T.P. (2013). The mammalian clock and chronopharmacology. Bioorganic & Medicinal Chemistry Letters *23*, 1929-1934.

Guo, J., Cheng, P., and Liu, Y. (2010). Functional significance of FRH in regulating the phosphorylation and stability of Neurospora circadian clock protein FRQ. The Journal of Biological Chemistry *285*, 11508-11515.

Hara, R., Wan, K., Wakamatsu, H., Aida, R., Moriya, T., Akiyama, M., and Shibata, S. (2001). Restricted feeding entrains liver clock without participation of the suprachiasmatic nucleus. Genes to Cells *6*, 269-278.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J.*, et al.* (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99-104.

Hardin, P.E. (2000). From biological clock to biological rhythms. Genome Biology *1*,1023-1.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science *290*, 2110-2113.

He, Q., Cha, J., He, Q., Lee, H.C., Yang, Y., and Liu, Y. (2006). CKI and CKII mediate the FREQUENCY-dependent phosphorylation of the WHITE COLLAR complex to close the Neurospora circadian negative feedback loop. Genes & Development *20*, 2552-2565.

He, Q., Cheng, P., Yang, Y., Wang, L., Gardner, K.H., and Liu, Y. (2002). White collar-1, a DNA binding transcription factor and a light sensor. Science *297*, 840-843.

He, Q., Cheng, P., Yang, Y.H., He, Q.Y., Yu, H.T., and Liu, Y. (2003). FWD1-mediated degradation of FREQUENCY in Neurospora establishes a conserved mechanism for circadian clock regulation. The EMBO Journal *22*, 4421-4430.

He, Q., and Liu, Y. (2005a). Degradation of the Neurospora circadian clock protein FREQUENCY through the ubiquitin-proteasome pathway. Biochemical Society Transactions *33*, 953-956.

He, Q., and Liu, Y. (2005b). Molecular mechanism of light responses in Neurospora: from light-induced transcription to photoadaptation. Genes & Development *19*, 2888-2899.

He, Q., Shu, H., Cheng, P., Chen, S., Wang, L., and Liu, Y. (2005). Light-independent phosphorylation of WHITE COLLAR-1 regulates its function in the Neurospora circadian negative feedback loop. The Journal of Biological Chemistry *280*, 17526-17532.

Heintzen, C., and Liu, Y. (2007). The Neurospora crassa circadian clock. Advances in Genetics *58*, 25-66.

Heintzen, C., Loros, J.J., and Dunlap, J.C. (2001). The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting. Cell *104*, 453-464.

Herrera-Estrella, A., and Horwitz, B.A. (2007). Looking through the eyes of fungi: molecular genetics of photoreception. Molecular Microbiology *64*, 5-15.

Hickey, P.C., Jacobson, D., Read, N.D., and Glass, N.L. (2002). Live-cell imaging of vegetative hyphal fusion in Neurospora crassa. Fungal Genetics and Biology *37*, 109-119.

Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. Nature Reviews Molecular Cell Biology *6*, 318-327.

Hong, C.I., Zamborszky, J., Baek, M., Labiscsak, L., Ju, K., Lee, H., Larrondo, L.F., Goity, A., Chong, H.S., Belden, W.J.*, et al.* (2014). Circadian rhythms synchronize mitosis in Neurospora crassa. Proceedings of the National Academy of Sciences of the United States of America *111*, 1397-1402.

Hsu, P.Y., Devisetty, U.K., and Harmer, S.L. (2013). Accurate timekeeping is controlled by a cycling activator in Arabidopsis. eLife *2*, e00473.

Hubbard, K.E., Robertson, F.C., Dalchau, N., and Webb, A.A. (2009). Systems analyses of circadian networks. Molecular Biosystems *5*, 1502-1511.

Hughes, M.E., Grant, G.R., Paquin, C., Qian, J., and Nitabach, M.N. (2012). Deep sequencing the circadian and diurnal transcriptome of Drosophila brain. Genome Research *22*, 1266-1281.

Hunt, S.M., Thompson, S., Elvin, M., and Heintzen, C. (2010). VIVID interacts with the WHITE COLLAR complex and FREQUENCY-interacting RNA helicase to alter light and clock responses in Neurospora. Proceedings of the National Academy of Sciences of the United States of America *107*, 16709-16714.

Hurley, J.M., Dasgupta, A., Emerson, J.M., Zhou, X., Ringelberg, C.S., Knabe, N., Lipzen, A.M., Lindquist, E.A., Daum, C.G., Barry, K.W.*, et al.* (2014). Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential. Proceedings of the National Academy of Sciences of the United States of America.

Ishii, C., Nakamura, K., and Inoue, H. (1991). A novel phenotype of an excision-repair mutant in Neurospora crassa: mutagen sensitivity of the mus-18 mutant is specific to UV. Molecular & General Genetics *228*, 33-39.

Jarai, G., and Marzluf, G.A. (1991). Sulfate transport in Neurospora crassa: regulation, turnover, and cellular localization of the CYS-14 protein. Biochemistry *30*, 4768-4773.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497-1502.

Johnson, L., Cao, X., and Jacobsen, S. (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Current Biology *12*, 1360-1367.

Jouffe, C., Cretenet, G., Symul, L., Martin, E., Atger, F., Naef, F., and Gachon, F. (2013). The circadian clock coordinates ribosome biogenesis. PLoS Biology *11*, e1001455.

Kaldi, K., Gonzalez, B.H., and Brunner, M. (2006). Transcriptional regulation of the Neurospora circadian clock gene wc-1 affects the phase of circadian output. EMBO Reports *7*, 199-204.

Karatsoreos, I.N. (2012). Effects of circadian disruption on mental and physical health. Current Neurology and Neuroscience Reports *12*, 218-225.

Kasuga, T., Townsend, J.P., Tian, C., Gilbert, L.B., Mannhaupt, G., Taylor, J.W., and Glass, N.L. (2005). Long-oligomer microarray profiling in Neurospora crassa reveals the transcriptional program underlying biochemical and physiological events of conidial germination. Nucleic Acids Research *33*, 6469-6485.

Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science *338*, 349-354.

Kojima, S., Sher-Chen, E.L., and Green, C.B. (2012). Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. Genes & Development *26*, 2724-2736.

Konopka, R.J., and Benzer, S. (1971). Clock mutants of Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America *68*, 2112-2116.

Kovac, J., Husse, J., and Oster, H. (2009). A time to fast, a time to feast: the crosstalk between metabolism and the circadian clock. Molecules and Cells *28*, 75-80.

Kyriacou, C.P., and Hastings, M.H. (2010). Circadian clocks: genes, sleep, and cognition. Trends in Cognitive Sciences *14*, 259-267.

Lamb, T.M., Finch, K.E., and Bell-Pedersen, D. (2012). The Neurospora crassa OS MAPK pathway-activated transcription factor ASL-1 contributes to circadian rhythms in pathway responsive clock-controlled genes. Fungal Genetics and Biology *49*, 180-188.

Lamb, T.M., Goldsmith, C.S., Bennett, L., Finch, K.E., and Bell-Pedersen, D. (2011). Direct transcriptional control of a p38 MAPK pathway by the circadian clock in Neurospora crassa. PloS One *6*, e27149.

Lamb, T.M., Vickery, J., and Bell-Pedersen, D. (2013). Regulation of gene expression in Neurospora crassa with a copper responsive promoter. G3: Genes| Genomes| Genetics *3*, 2273-2280.

Lauter, F.R., and Russo, V.E. (1991). Blue light induction of conidiation-specific genes in Neurospora crassa. Nucleic Acids Research *19*, 6883-6886.

Lauter, F.R., and Yanofsky, C. (1993). Day/night and circadian rhythm control of con gene expression in Neurospora. Proceedings of the National Academy of Sciences of the United States of America *90*, 8249-8253.

Lee, K., Dunlap, J.C., and Loros, J.J. (2003). Roles for WHITE COLLAR-1 in circadian and general photoperception in Neurospora crassa. Genetics *163*, 103-114.

Lee, K., Loros, J.J., and Dunlap, J.C. (2000). Interconnected feedback loops in the Neurospora circadian system. Science *289*, 107-110.

Leeder, A.C., Jonkers, W., Li, J., and Glass, N.L. (2013). Early colony establishment in Neurospora crassa requires a MAP kinase regulatory network. Genetics *195*, 883-898.

Levine, J.D., Funes, P., Dowse, H.B., and Hall, J.C. (2002). Resetting the circadian clock by social experience in Drosophila melanogaster. Science *298*, 2010-2012.

Lewis, Z.A., Correa, A., Schwerdtfeger, C., Link, K.L., Xie, X., Gomer, R.H., Thomas, T., Ebbole, D.J., and Bell-Pedersen, D. (2002). Overexpression of White Collar-1 (WC-1) activates circadian clock-associated genes, but is not sufficient to induce most light-regulated gene expression in Neurospora crassa. Molecular Microbiology *45*, 917-931.

Li, D., Bobrowicz, P., Wilkinson, H.H., and Ebbole, D.J. (2005). A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in Neurospora crassa. Genetics *170*, 1091-1104.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, H., Ruan, J., and Durbin, R. (2008a). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research *18*, 1851-1858.

Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008b). SOAP: short oligonucleotide alignment program. Bioinformatics *24*, 713-714.

Lichius, A., and Lord, K.M. (2014). Chemoattractive Mechanisms in Filamentous Fungi. Open Mycology Journal *8*, 28-57.

Linden, H., Ballario, P., and Macino, G. (1997a). Blue light regulation in Neurospora crassa. Fungal Genetics and Biology *22*, 141-150.

Linden, H., and Macino, G. (1997). White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in Neurospora crassa. The EMBO Journal *16*, 98-109.

Linden, H., Rodriguez-Franco, M., and Macino, G. (1997b). Mutants of Neurospora crassa defective in regulation of blue light perception. Molecular & General Genetics *254*, 111-118.

Lindquist, S. (1981). Regulation of protein synthesis during heat shock. Nature *293*, 311-314.

Liu, Y. (2005). Analysis of posttranslational regulations in the Neurospora circadian clock. Methods in Enzymology *393*, 379-393.

Liu, Y., and Bell-Pedersen, D. (2006). Circadian rhythms in Neurospora crassa and other filamentous fungi. Eukaryotic Cell *5*, 1184-1193.

Liu, Y., Loros, J., and Dunlap, J.C. (2000). Phosphorylation of the Neurospora clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. Proceedings of the National Academy of Sciences of the United States of America *97*, 234-239.

Liu, Y., Merrow, M., Loros, J.J., and Dunlap, J.C. (1998). How temperature changes reset a circadian oscillator. Science *281*, 825-829.

Liu, Y., Tsinoremas, N.F., Johnson, C.H., Lebedeva, N.V., Golden, S.S., Ishiura, M., and Kondo, T. (1995). Circadian orchestration of gene expression in cyanobacteria. Genes & Development *9*, 1469-1478.

Locke, J.C., Kozma-Bognar, L., Gould, P.D., Feher, B., Kevei, E., Nagy, F., Turner, M.S., Hall, A., and Millar, A.J. (2006). Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. Molecular Systems Biology *2*, 59.

Locke, J.C., Millar, A.J., and Turner, M.S. (2005). Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana. Journal of Theoretical Biology *234*, 383-393.

Loros, J.J., Denome, S.A., and Dunlap, J.C. (1989). Molecular cloning of genes under control of the circadian clock in Neurospora. Science *243*, 385-388.

Loros, J.J., and Dunlap, J.C. (1991). Neurospora crassa clock-controlled genes are regulated at the level of transcription. Molecular and Cellular Biology *11*, 558-563.

Loros, J.J., and Feldman, J.F. (1986). Loss of temperature compensation of circadian period length in the frq-9 mutant of Neurospora crassa. Journal of Biological Rhythms *1*, 187-198.

Luo, C., Loros, J.J., and Dunlap, J.C. (1998). Nuclear localization is required for function of the essential clock protein FRQ. The EMBO Journal *17*, 1228-1235.

Maerz, S., Ziv, C., Vogt, N., Helmstaedt, K., Cohen, N., Gorovits, R., Yarden, O., and Seiler, S. (2008). 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 *179*, 1313-1325.

Malzahn, E., Ciprianidis, S., Kaldi, K., Schafmeier, T., and Brunner, M. (2010). Photoadaptation in Neurospora by competitive interaction of activating and inhibitory LOV domains. Cell *142*, 762-772.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences of the United States of America *100*, 11980-11985.

Martinez, G.J., and Rao, A. (2012). Immunology. Cooperative transcription factor complexes in control. Science *338*, 891-892.

Marzluf, G.A. (1993). Regulation of sulfur and nitrogen metabolism in filamentous fungi. Annual Review of Microbiology *47*, 31-55.

Masri, S., Cervantes, M., and Sassone-Corsi, P. (2013). The circadian clock and cell cycle: interconnected biological circuits. Current Opinion in Cell Biology *25*, 730-734.

McCarthy, J.J., Andrews, J.L., McDearmon, E.L., Campbell, K.S., Barber, B.K., Miller, B.H., Walker, J.R., Hogenesch, J.B., Takahashi, J.S., and Esser, K.A. (2007). Identification of the circadian transcriptome in adult mouse skeletal muscle. Physiological Genomics *31*, 86-95.

McClung, C.R. (2006). Plant circadian rhythms. The Plant Cell Online *18*, 792-803.

McClung, C.R., Fox, B.A., and Dunlap, J.C. (1989). The Neurospora clock gene frequency shares a sequence element with the Drosophila clock gene period. Nature *339*, 558-562.

McColl, D., Valencia, C.A., and Vierula, P.J. (2003). Characterization and expression of the Neurospora crassa nmt-1 gene. Current Genetics *44*, 216-223.

McNally, M.T., and Free, S.J. (1988). Isolation and characterization of a Neurospora glucose-repressible gene. Current Genetics *14*, 545-551.

Meireles-Filho, A.C., Bardet, A.F., Yanez-Cuna, J.O., Stampfel, G., and Stark, A. (2014). cis-regulatory requirements for tissue-specific programs of the circadian clock. Current Biology *24*, 1-10.

Menet, J.S., Rodriguez, J., Abruzzi, K.C., and Rosbash, M. (2012). Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. Elife *1*, e00011.

Metzenberg, R.L. (2004). Bird medium: an alternative to Vogel medium. Fungal Genetics Newsletter *51*, 19-20.

Michael, T.P., and McClung, C.R. (2003). Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis. Plant Physiology *132*, 629-639.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P*., et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Moore, A., Zielinski, T., and Millar, A.J. (2014). Online period estimation and determination of rhythmicity in circadian data, using the BioDare data infrastructure. Methods in Molecular Biology *1158*, 13-44.

Morgan, L.W., Greene, A.V., and Bell-Pedersen, D. (2003). Circadian and light-induced expression of luciferase in Neurospora crassa. Fungal Genetics and Biology *38*, 327-332.

Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., and Ferrin, T.E. (2011). clusterMaker: a multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics *12*, 436.

Nagel, D.H., and Kay, S.A. (2012). Complexity in the wiring and regulation of plant circadian networks. Current Biology *22*, R648-657.

Nakamichi, N., Kiba, T., Kamioka, M., Suzuki, T., Yamashino, T., Higashiyama, T., Sakakibara, H., and Mizuno, T. (2012). Transcriptional repressor PRR5 directly regulates clock-output pathways. Proceedings of the National Academy of Sciences of the United States of America *109*, 17123-17128.

Ni, M., and Yu, J.H. (2007). A novel regulator couples sporogenesis and trehalose biogenesis in Aspergillus nidulans. PloS One *2*, e970.

Nowrousian, M., Duffield, G.E., Loros, J.J., and Dunlap, J.C. (2003). The frequency gene is required for temperature-dependent regulation of many clock-controlled genes in Neurospora crassa. Genetics *164*, 923-933.

Olmedo, M., Ruger-Herreros, C., and Corrochano, L.M. (2010a). Regulation by blue light of the fluffy gene encoding a major regulator of conidiation in Neurospora crassa. Genetics *184*, 651-658.

Olmedo, M., Ruger-Herreros, C., Luque, E.M., and Corrochano, L.M. (2010b). A complex photoreceptor system mediates the regulation by light of the conidiation genes con-10 and con-6 in Neurospora crassa. Fungal Genetics and Biology *47*, 352-363.

Orian, A., Grewal, S.S., Knoepfler, P.S., Edgar, B.A., Parkhurst, S.M., and Eisenman, R.N. (2005). Genomic binding and transcriptional regulation by the Drosophila Myc and Mnt transcription factors. Cold Spring Harbor Symposia on Quantitative Biology *70*, 299-307.

Orth, J.D., Thiele, I., and Palsson, B.O. (2010). What is flux balance analysis? Nature Biotechnology *28*, 245-248.

Ouyang, Y., Andersson, C.R., Kondo, T., Golden, S.S., and Johnson, C.H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. Proceedings of the National Academy of Sciences of the United States of America *95*, 8660-8664.

Pall, M.L., and Brunelli, J. (1993). A series of six compact fungal transformation vectors containing polylinkers with multiple unique restriction sites. Fungal Genetics Newsletter *40*, 59-62.

Panda, S., Antoch, M.P., Miller, B.H., Su, A.I., Schook, A.B., Straume, M., Schultz, P.G., Kay, S.A., Takahashi, J.S., and Hogenesch, J.B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. Cell *109*, 307-320.

Park, G., Pan, S., and Borkovich, K.A. (2008). Mitogen-activated protein kinase cascade required for regulation of development and secondary metabolism in Neurospora crassa. Eukaryotic Cell *7*, 2113-2122.

Park, G., Servin, J.A., Turner, G.E., Altamirano, L., Colot, H.V., Collopy, P., Litvinkova, L., Li, L., Jones, C.A., Diala, F.G.*, et al.* (2011). Global analysis of serine-threonine protein kinase genes in Neurospora crassa. Eukaryotic Cell *10*, 1553-1564.

Partch, C.L., Green, C.B., and Takahashi, J.S. (2014). Molecular architecture of the mammalian circadian clock. Trends in Cell Biology *24*, 90-99.

Pavesi, G., Mereghetti, P., Mauri, G., and Pesole, G. (2004). Weeder Web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes. Nucleic Acids Research *32*, W199-203.

Perkins, D.D. (2006). How to choose and prepare media. 1-15.

Perlman, J., Nakashima, H., and Feldman, J.F. (1981). Assay and Characteristics of Circadian Rhythmicity in Liquid Cultures of Neurospora crassa. Plant Physiology *67*, 404-407.

Pfeiffer, A., Shi, H., Tepperman, J.M., Zhang, Y., and Quail, P.H. (2014). Combinatorial complexity in a transcriptionally centered signaling hub in Arabidopsis. Molecular Plant *7*, 1598-1618.

Pittendrigh, C.S. (1960). Circadian rhythms and the circadian organization of living systems. Cold Spring Harbor Symposia on Quantitative Biology *25*, 159-184.

Pittendrigh, C.S. (1993). Temporal organization: reflections of a Darwinian clock-watcher. Annual Review of Physiology *55*, 16-54.

Pomraning, K.R., Smith, K.M., Bredeweg, E.L., Connolly, L.R., Phatale, P.A., and Freitag, M. (2012). Library preparation and data analysis packages for rapid genome sequencing. Methods in Molecular Biology *944*, 1-22.

Pomraning, K.R., Smith, K.M., and Freitag, M. (2009). Genome-wide high throughput analysis of DNA methylation in eukaryotes. Methods *47*, 142-150.

Pregueiro, A.M., Liu, Q., Baker, C.L., Dunlap, J.C., and Loros, J.J. (2006). The Neurospora checkpoint kinase 2: a regulatory link between the circadian and cell cycles. Science *313*, 644-649.

Ptacek, L.J., Jones, C.R., and Fu, Y.H. (2007). Novel insights from genetic and molecular characterization of the human clock. Cold Spring Harbor Symposia on Quantitative Biology *72*, 273-277.

Purschwitz, J., Muller, S., Kastner, C., and Fischer, R. (2006). Seeing the rainbow: light sensing in fungi. Current Opinion in Microbiology *9*, 566-571.

Read, N.D., Goryachev, A.B., and Lichius, A. (2012). The mechanistic basis of self-fusion between conidial anastomosis tubes during fungal colony initiation. Fungal Biology Reviews *26*, 1-11.

Rerngsamran, P., Murphy, M.B., Doyle, S.A., and Ebbole, D.J. (2005). Fluffy, the major regulator of conidiation in Neurospora crassa, directly activates a developmentally regulated hydrophobin gene. Molecular Microbiology *56*, 282-297.

Rey, G., Cesbron, F., Rougemont, J., Reinke, H., Brunner, M., and Naef, F. (2011). Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. PLoS Biology *9*, e1000595.

Richardson, B.E., Nowak, S.J., and Baylies, M.K. (2008). Myoblast fusion in fly and vertebrates: new genes, new processes and new perspectives. Traffic *9*, 1050-1059.

Rodriguez, J., Tang, C.H., Khodor, Y.L., Vodala, S., Menet, J.S., and Rosbash, M. (2013). Nascent-Seq analysis of Drosophila cycling gene expression. Proceedings of the National Academy of Sciences of the United States of America *110*, E275-284.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M.*, et al.* (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Research *32*, 5539-5545.

Sachs, M.S., and Yanofsky, C. (1991). Developmental expression of genes involved in conidiation and amino acid biosynthesis in Neurospora crassa. Developmental Biology *148*, 117-128.

Sancar, G., Sancar, C., Brugger, B., Ha, N., Sachsenheimer, T., Gin, E., Wdowik, S., Lohmann, I., Wieland, F., Hofer, T.*, et al.* (2011). A global circadian repressor controls antiphasic expression of metabolic genes in Neurospora. Molecular Cell *44*, 687-697.

Sargent, M.L., and Briggs, W.R. (1967). The effects of light on a circadian rhythm of conidiation in neurospora. Plant Physiology *42*, 1504-1510.

Sargent, M.L., Briggs, W.R., and Woodward, D.O. (1966). Circadian nature of a rhythm expressed by an invertaseless strain of Neurospora crassa. Plant Physiology *41*, 1343-1349.

Schafmeier, T., Diernfellner, A., Schafer, A., Dintsis, O., Neiss, A., and Brunner, M. (2008). Circadian activity and abundance rhythms of the Neurospora clock transcription factor WCC associated with rapid nucleo-cytoplasmic shuttling. Genes & Development *22*, 3397-3402.

Schafmeier, T., Haase, A., Kaldi, K., Scholz, J., Fuchs, M., and Brunner, M. (2005). Transcriptional feedback of Neurospora circadian clock gene by phosphorylation-dependent inactivation of its transcription factor. Cell *122*, 235-246.

Schwerdtfeger, C., and Linden, H. (2000). Localization and light-dependent phosphorylation of white collar 1 and 2, the two central components of blue light signaling in Neurospora crassa. European Journal of Biochemistry *267*, 414-422.

Schwerdtfeger, C., and Linden, H. (2001). Blue light adaptation and desensitization of light signal transduction in Neurospora crassa. Molecular Microbiology *39*, 1080-1087.

Schwerdtfeger, C., and Linden, H. (2003). VIVID is a flavoprotein and serves as a fungal blue light photoreceptor for photoadaptation. The EMBO Journal *22*, 4846-4855.

Shalgi, R., Hurt, J.A., Krykbaeva, I., Taipale, M., Lindquist, S., and Burge, C.B. (2013). Widespread regulation of translation by elongation pausing in heat shock. Molecular Cell *49*, 439-452.

Sharma, V.K. (2003). Adaptive significance of circadian clocks. Chronobiology International *20*, 901-919.

Shi, M., Collett, M., Loros, J.J., and Dunlap, J.C. (2010). FRQ-interacting RNA helicase mediates negative and positive feedback in the Neurospora circadian clock. Genetics *184*, 351-361.

Shimura, M., Ito, Y., Ishii, C., Yajima, H., Linden, H., Harashima, T., Yasui, A., and Inoue, H. (1999). Characterization of a Neurospora crassa photolyase-deficient mutant generated by repeat induced point mutation of the phr gene. Fungal Genetics and Biology *28*, 12-20.

Shrode, L.B., Lewis, Z.A., White, L.D., Bell-Pedersen, D., and Ebbole, D.J. (2001). vvd is required for light adaptation of conidiation-specific genes of Neurospora crassa, but not circadian conidiation. Fungal Genetics and Biology *32*, 169-181.

Simonin, A., Palma-Guerrero, J., Fricker, M., and Glass, N.L. (2012). Physiological significance of network organization in fungi. Eukaryotic Cell *11*, 1345-1352.

Smith, K.M., Sancar, G., Dekhang, R., Sullivan, C.M., Li, S., Tag, A.G., Sancar, C., Bredeweg, E.L., Priest, H.D., McCormick, R.F.*, et al.* (2010). Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. Eukaryotic Cell *9*, 1549-1556.

Spriggs, K.A., Bushell, M., and Willis, A.E. (2010). Translational regulation of gene expression during conditions of cell stress. Molecular Cell *40*, 228-237.

Stein, L.D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., Nickerson, E., Stajich, J.E., Harris, T.W., Arva, A.*, et al.* (2002). The generic genome browser: a building block for a model organism system database. Genome Research *12*, 1599-1610.

Stokkan, K.A., Yamazaki, S., Tei, H., Sakaki, Y., and Menaker, M. (2001). Entrainment of the circadian clock in the liver by feeding. Science *291*, 490-493.

Storch, K.F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F.C., Wong, W.H., and Weitz, C.J. (2002). Extensive and divergent circadian gene expression in liver and heart. Nature *417*, 78-83.

Summa, K.C., and Turek, F.W. (2014). Chronobiology and obesity: Interactions between circadian rhythms and energy regulation. Advances in Nutrition *5*, 312S-319S.

Suzuki, S., Sarikaya Bayram, O., Bayram, O., and Braus, G.H. (2013). conF and conJ contribute to conidia germination and stress response in the filamentous fungus Aspergillus nidulans. Fungal Genetics and Biology *56*, 42-53.

Taghert, P.H., and Shafer, O.T. (2006). Mechanisms of clock output in the Drosophila circadian pacemaker system. Journal of Biological Rhythms *21*, 445-457.

Talora, C., Franchi, L., Linden, H., Ballario, P., and Macino, G. (1999). Role of a white collar-1-white collar-2 complex in blue-light signal transduction. The EMBO Journal *18*, 4961-4968.

Tamaru, H., Zhang, X., McMillen, D., Singh, P.B., Nakayama, J., Grewal, S.I., Allis, C.D., Cheng, X., and Selker, E.U. (2003). Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in Neurospora crassa. Nature Genetics *34*, 75-79.

Tang, C.T., Li, S., Long, C., Cha, J., Huang, G., Li, L., Chen, S., and Liu, Y. (2009). Setting the pace of the Neurospora circadian clock by multiple independent FRQ phosphorylation events. Proceedings of the National Academy of Sciences of the United States of America *106*, 10722-10727.

Tauber, E., Last, K.S., Olive, P.J., and Kyriacou, C.P. (2004). Clock gene evolution and functional divergence. Journal of Biological Rhythms *19*, 445-458.

Thompson, S., Croft, N.J., Sotiriou, A., Piggins, H.D., and Crosthwaite, S.K. (2008). Neurospora crassa heat shock factor 1 Is an essential gene; a second heat shock factor-like gene, hsf2, is required for asexual spore formation. Eukaryotic Cell *7*, 1573-1581.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature Biotechnology *31*, 46-53.

Ueda, H.R., Chen, W., Adachi, A., Wakamatsu, H., Hayashi, S., Takasugi, T., Nagano, M., Nakahama, K., Suzuki, Y., Sugano, S.*, et al.* (2002). A transcription factor response element for gene expression during circadian night. Nature *418*, 534-539.

Ukai-Tadenuma, M., Kasukawa, T., and Ueda, H.R. (2008). Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. Nature Cell Biology *10*, 1154-1163.

Verma, S., and Idnurm, A. (2013). The Uve1 endonuclease is regulated by the white collar complex to protect cryptococcus neoformans from UV damage. PLoS Genetics *9*, e1003769.

Vignery, A. (2008). Macrophage fusion: molecular mechanisms. Methods in Molecular Biology *475*, 149-161.

Vitalini, M.W., de Paula, R.M., Goldsmith, C.S., Jones, C.A., Borkovich, K.A., and Bell-Pedersen, D. (2007). Circadian rhythmicity mediated by temporal regulation of the activity of p38 MAPK. Proceedings of the National Academy of Sciences of the United States of America *104*, 18223-18228.

Vitalini, M.W., de Paula, R.M., Park, W.D., and Bell-Pedersen, D. (2006). The rhythms of life: circadian output pathways in Neurospora. Journal of Biological Rhythms *21*, 432-444.

Vitalini, M.W., Dunlap, J., Heintzen, C., Liu, Y., Loros, J., and Bell-Pedersen, D. (2010). Circadian rhythms. In Cellular and Molecular Biology of Filamentous Fungi, K.A. Borkovich, and D.J. Ebbole, eds. (Washington, DC: ASM Press), pp. 442-466.

Vitalini, M.W., Morgan, L.W., March, I.J., and Bell-Pedersen, D. (2004). A genetic selection for circadian output pathway mutations in Neurospora crassa. Genetics *167*, 119-129.

Wang, N., Yoshida, Y., and Hasunuma, K. (2007a). Catalase-1 (CAT-1) and nucleoside diphosphate kinase-1 (NDK-1) play an important role in protecting conidial viability under light stress in Neurospora crassa. Molecular Genetics and Genomics *278*, 235-242.

Wang, N., Yoshida, Y., and Hasunuma, K. (2007b). Loss of Catalase-1 (Cat-1) results in decreased conidial viability enhanced by exposure to light in Neurospora crassa. Molecular Genetics and Genomics *277*, 13-22.

Westergaard, M., and Mitchell, H.K. (1947). Neurospora V. A synthetic medium favoring sexual reproduction. American Journal of Botany, 573-577.

Woelfle, M.A., and Johnson, C.H. (2006). No promoter left behind: global circadian gene expression in cyanobacteria. Journal of Biological Rhythms *21*, 419-431.

Wu, C., Yang, F., Smith, K.M., Peterson, M., Dekhang, R., Zhang, Y., Zucker, J., Bredeweg, E.L., Mallappa, C., Zhou, X.*, et al.* (2014). Genome-Wide Characterization of Light-Regulated Genes in Neurospora crassa. G3: Genes| Genomes| Genetics *4*, 1731-1745.

Yamashita, K., Shiozawa, A., Watanabe, S., Fukumori, F., Kimura, M., and Fujimura, M. (2008). ATF-1 transcription factor regulates the expression of ccg-1 and cat-1 genes in response to fludioxonil under OS-2 MAP kinase in Neurospora crassa. Fungal Genetics and Biology *45*, 1562-1569.

Yang, Y., He, Q., Cheng, P., Wrage, P., Yarden, O., and Liu, Y. (2004). Distinct roles for PP1 and PP2A in the Neurospora circadian clock. Genes & Development *18*, 255-260.

Yoshida, Y., Iigusa, H., Wang, N., and Hasunuma, K. (2011). Cross-talk between the cellular redox state and the circadian system in Neurospora. PloS One *6*, e28227.

Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. Nature Reviews Genetics *2*, 702-715.

Zhou, L.W., Haas, H., and Marzluf, G.A. (1998). Isolation and characterization of a new gene, sre, which encodes a GATA-type regulatory protein that controls iron transport in Neurospora crassa. Molecular & General Genetics *259*, 532-540.

Zhu, H., Nowrousian, M., Kupfer, D., Colot, H.V., Berrocal-Tito, G., Lai, H., Bell-Pedersen, D., Roe, B.A., Loros, J.J., and Dunlap, J.C. (2001). Analysis of expressed sequence tags from two starvation, time-of-day-specific libraries of Neurospora crassa reveals novel clock-controlled genes. Genetics *157*, 1057-1065.

Zielinski, T., Moore, A.M., Troup, E., Halliday, K.J., and Millar, A.J. (2014). Strengths and limitations of period estimation methods for circadian data. PloS One *9*, e96462.

Zvonic, S., Ptitsyn, A.A., Conrad, S.A., Scott, L.K., Floyd, Z.E., Kilroy, G., Wu, X., Goh, B.C., Mynatt, R.L., and Gimble, J.M. (2006). Characterization of peripheral circadian clocks in adipose tissues. Diabetes *55*, 962-970.

APPENDIX A

SUPPORTING FIGURES FOR CHAPTER II

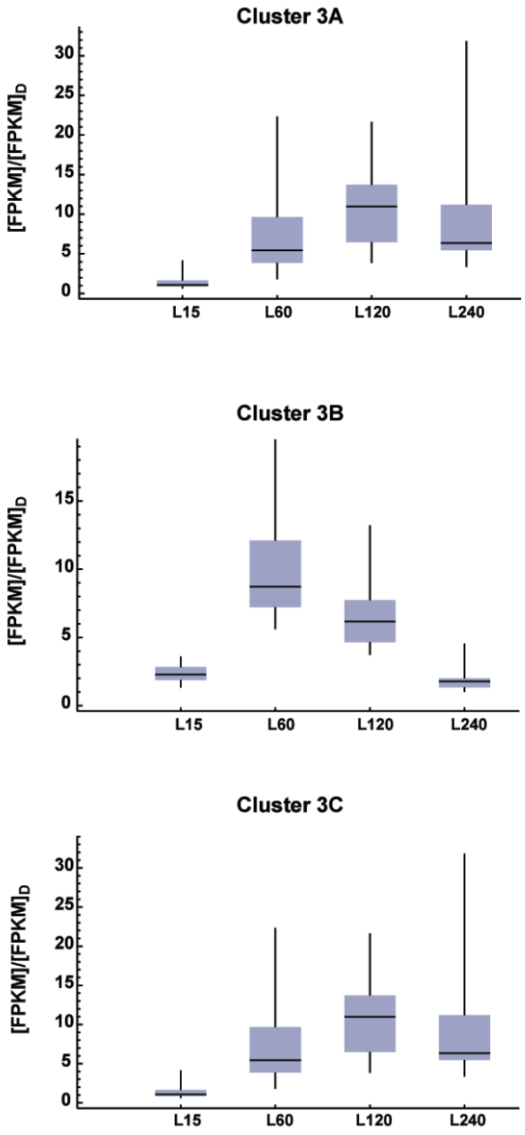


Figure S1. Expression-changes for transcripts in each of Cluster 3 subclusters A, B, and C demarcated in (Figure 2-3). Values for each time-point (L15, L60, L120 and L240) are normalized to expression in the dark. The horizontal black bar is the median, the box top and bottom are the 75% and 25% quantiles, and the whiskers extend to the maximum and minimum values.

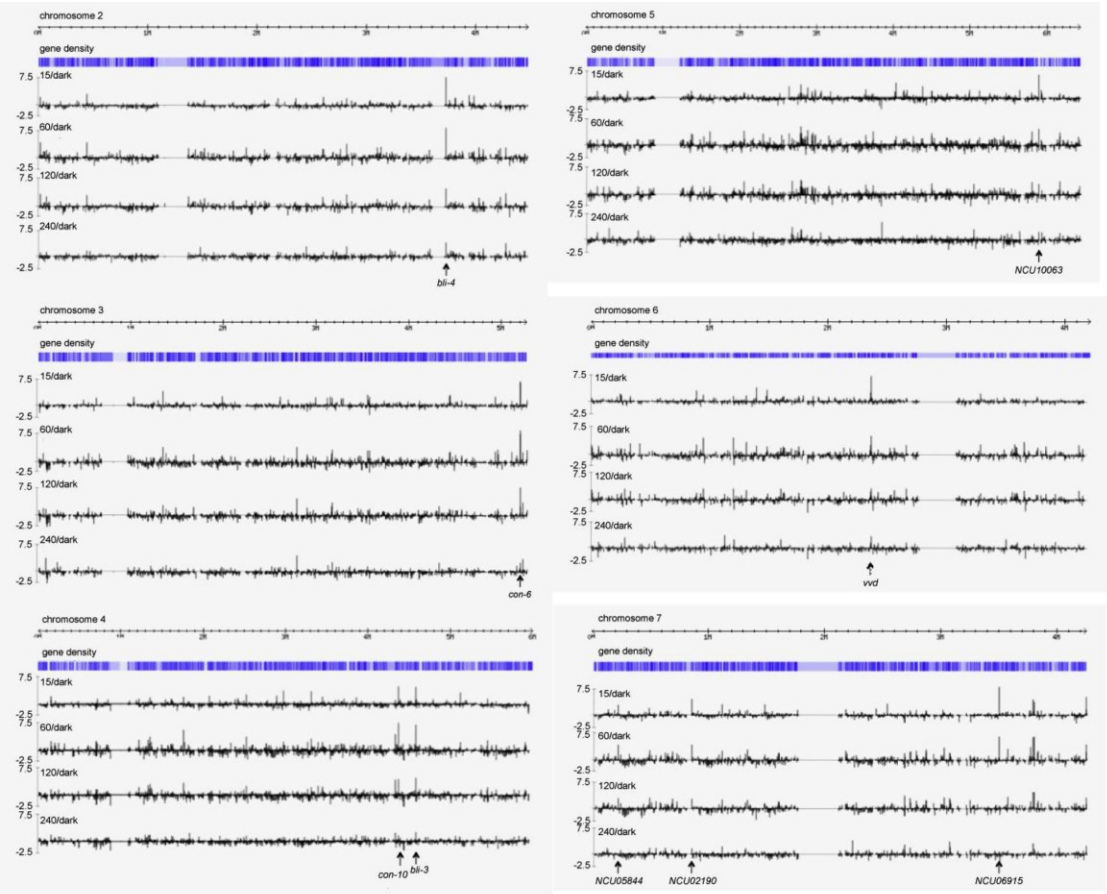


Figure S2. Pattern of light-regulation of genes on Linkage Groups II-VII (Chromosomes 2-7). The log-2 change in expression in the light versus the dark is given on the Y-axis for each time-point in the light (15, 60, 120 and 240 min).

APPENDIX B

SUPPORTING TABLES FOR CHAPTER II

Table S1. Summary of read-depth from Illumina RNA-seq

HTS number	Replicate	Light	Reads
HTS760	1	0	36787743
HTS761	1	15	34876232
HTS762	1	60	40079277
HTS763	1	120	37470890
HTS764	1	240	38931156
HTS920	2	0	11803108
HTS921	2	15	10583565
HTS922	2	60	11132975
HTS923	2	120	10635334
HTS924	2	240	6019056

APPENDIX C

SUPPORTING FIGURES FOR CHAPTER III

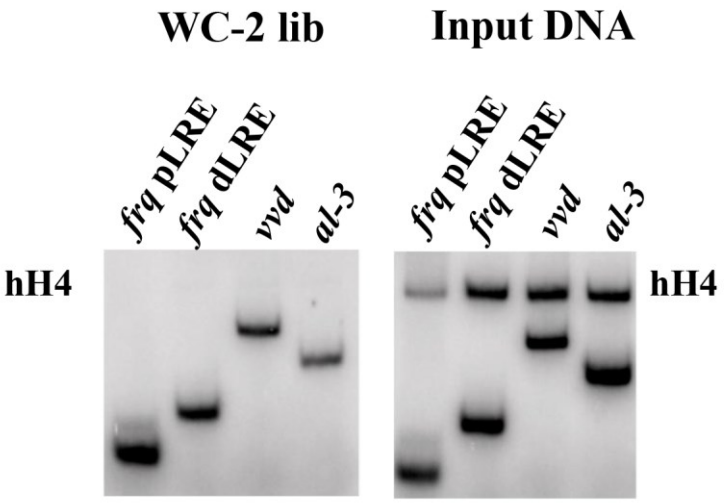


Figure S1. Validation of ChIP-seq libraries. After ChIP-seq library construction, PCR shows enrichment of four known WCC binding sites (*frq* pLRE and dLRE, *vvd* and *al-3*) in the WC-2 library.

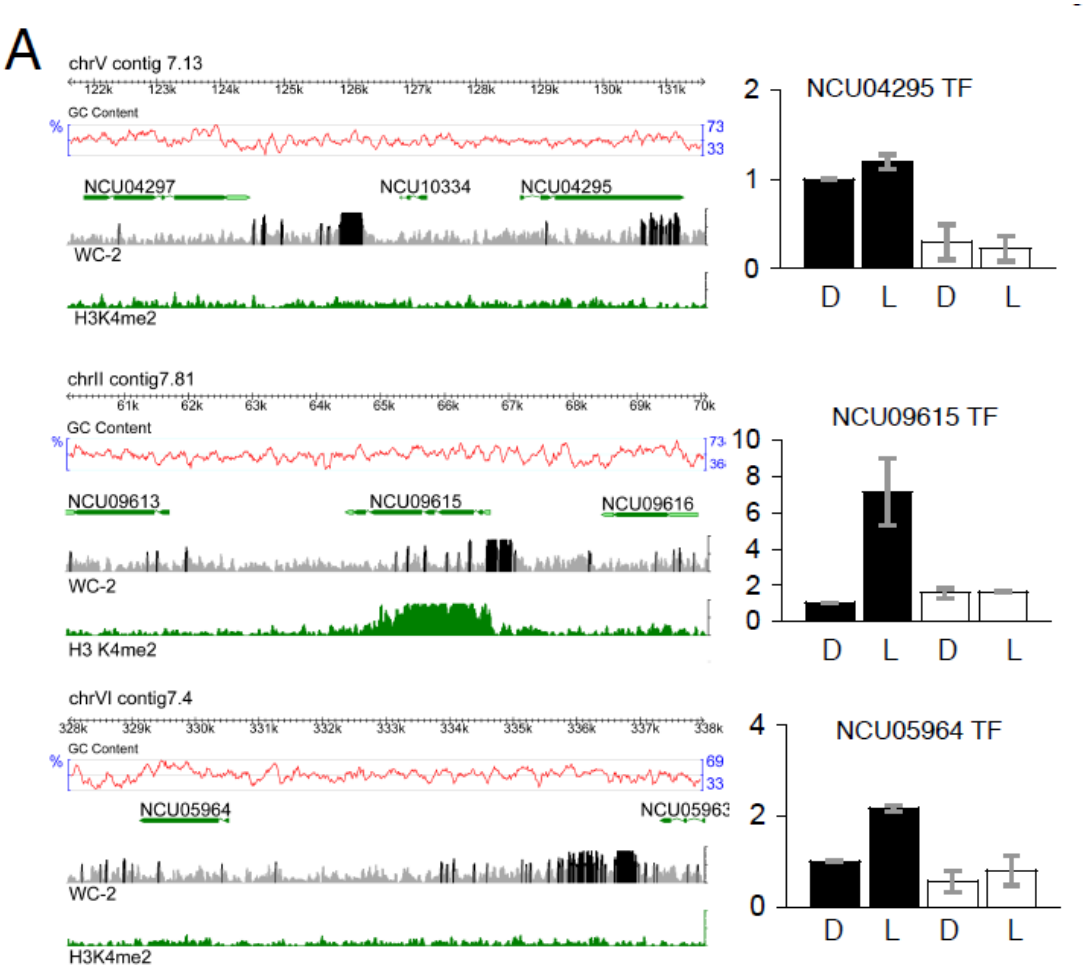


Figure S2. ChIP-seq reveals WCC regulation of TF genes. TFs with WCC peaks in their promoters and whose expression is: A. light induced and dependent on WC-2, B. not light induced but dependent on WC-2, and C. not light induced and independent of WC-2 (summarized in Table 2-1B; TF genes shown in Figure 2-2 are omitted). SOAP mapped sequence reads from the WC-2 ChIP library are plotted as a histogram of sequence coverage per base along fragments containing loci of interest. The y-axis has a maximum value of 20 reads per base and regions with >10 reads are shown in black, others in grey. Gene annotations are based on known or predicted cDNAs (UTRs are shown in light green, coding regions in dark green and introns are shown as thin lines). Quantitative RT-PCR was performed on RNA isolated from control (black bars) and Δ*wc-2* (white bars) strains grown in the dark (D) or following a 15-min light induction (L). Values shown are the average of two experiments (y-axis).

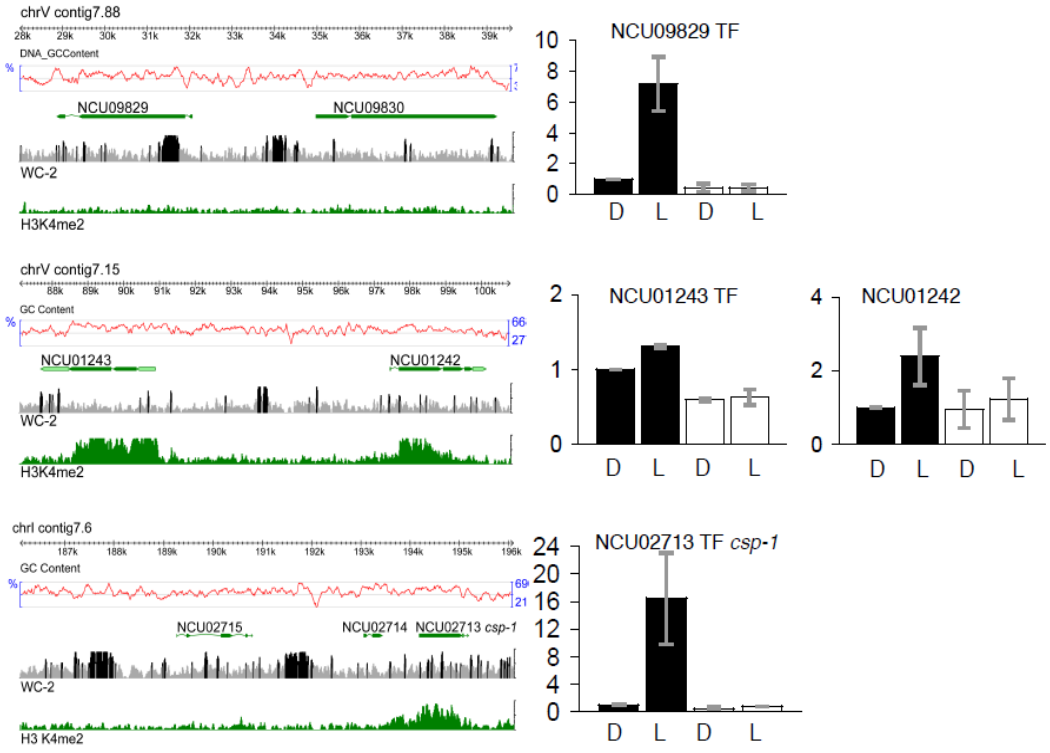


Figure S2. Continued.

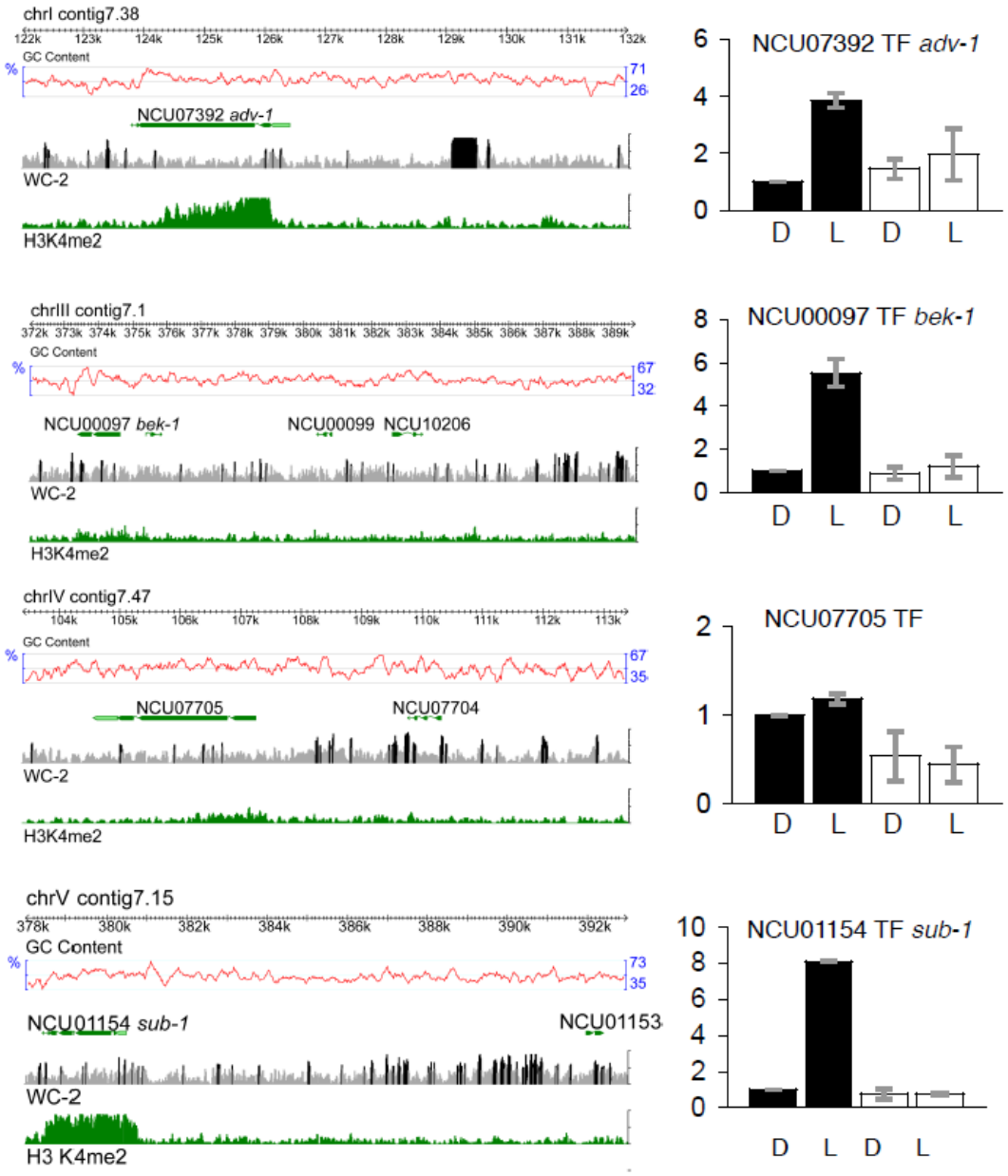


Figure S2. Continued.

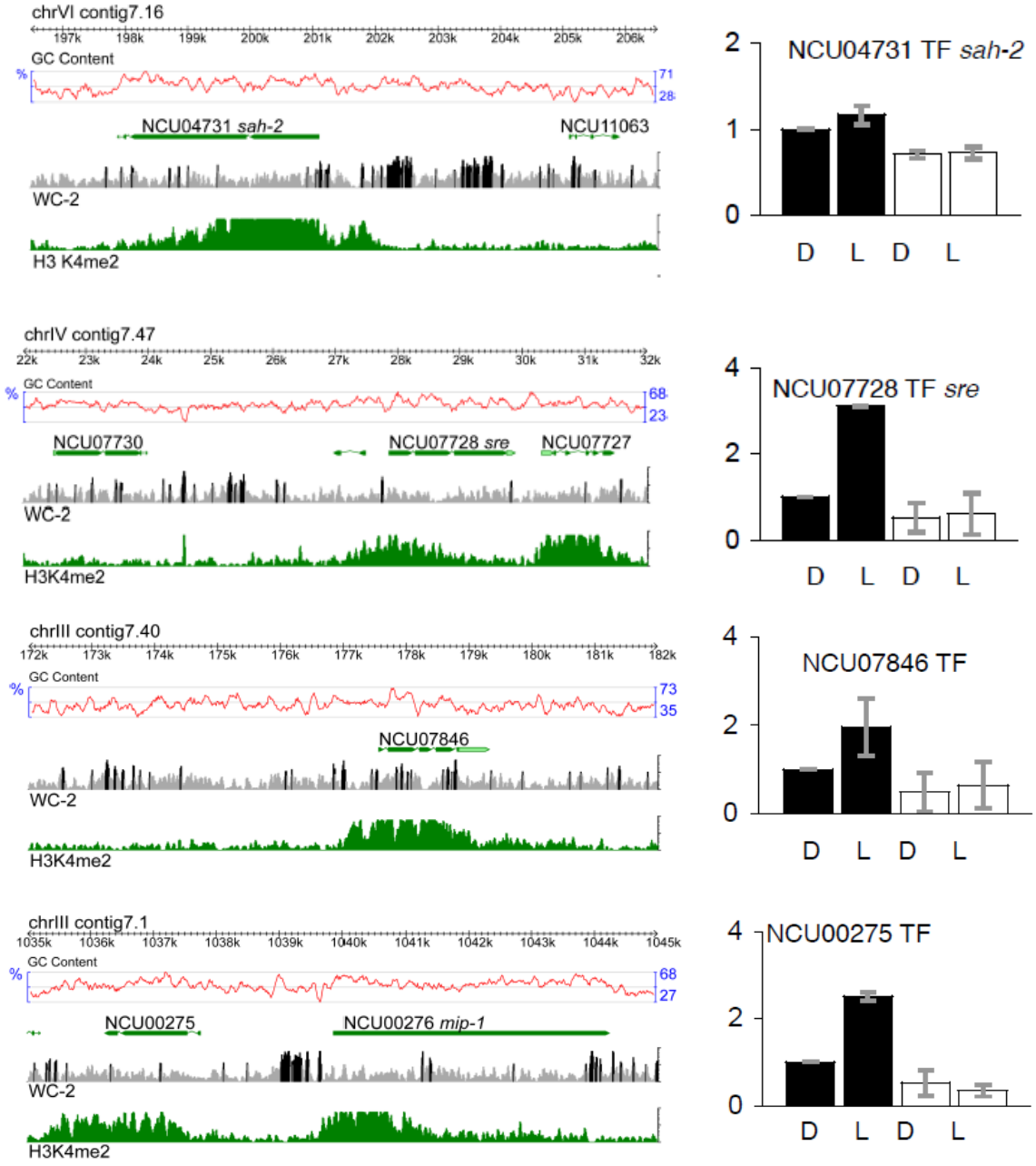


Figure S2. Continued.

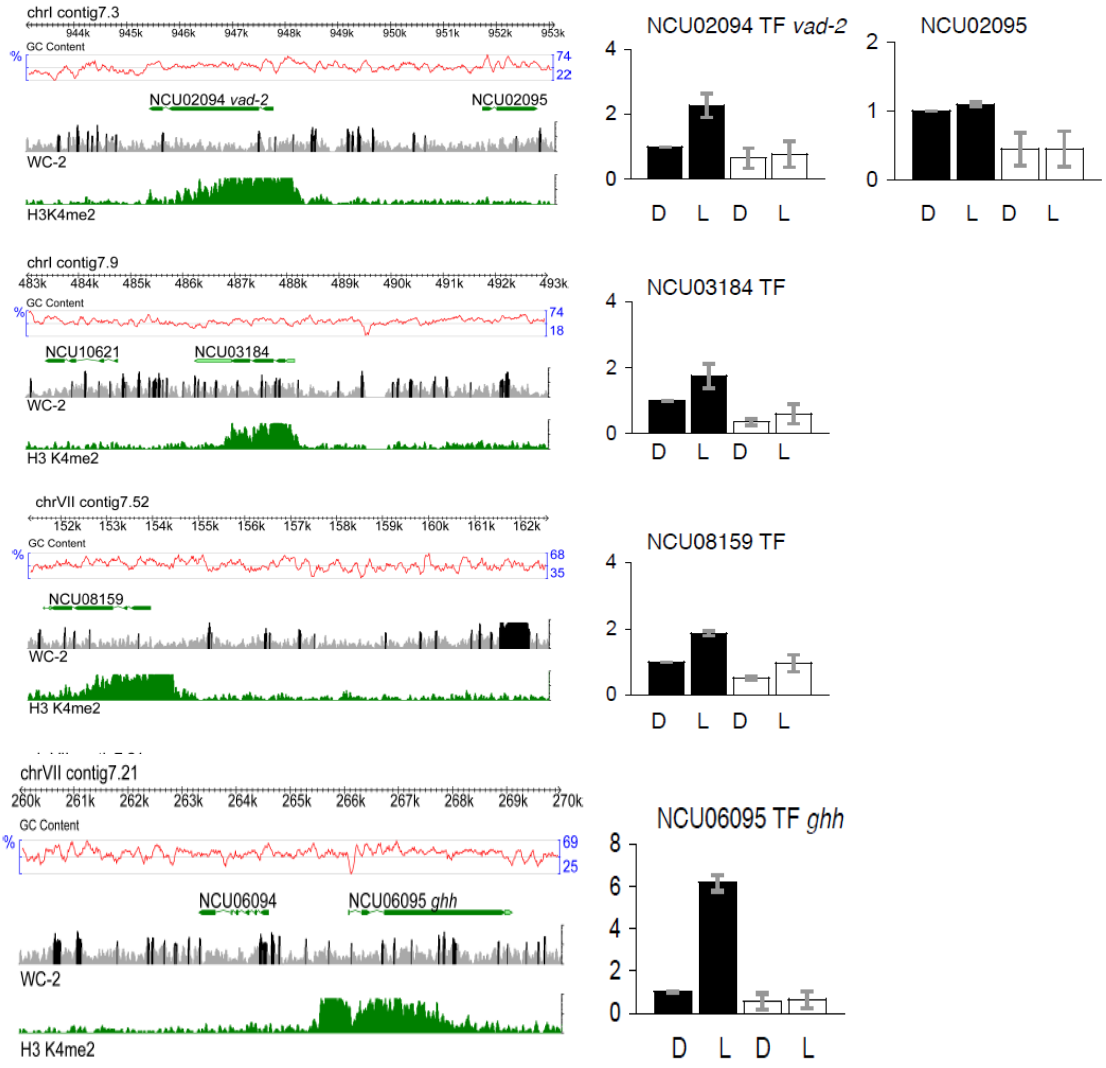


Figure S2. Continued.

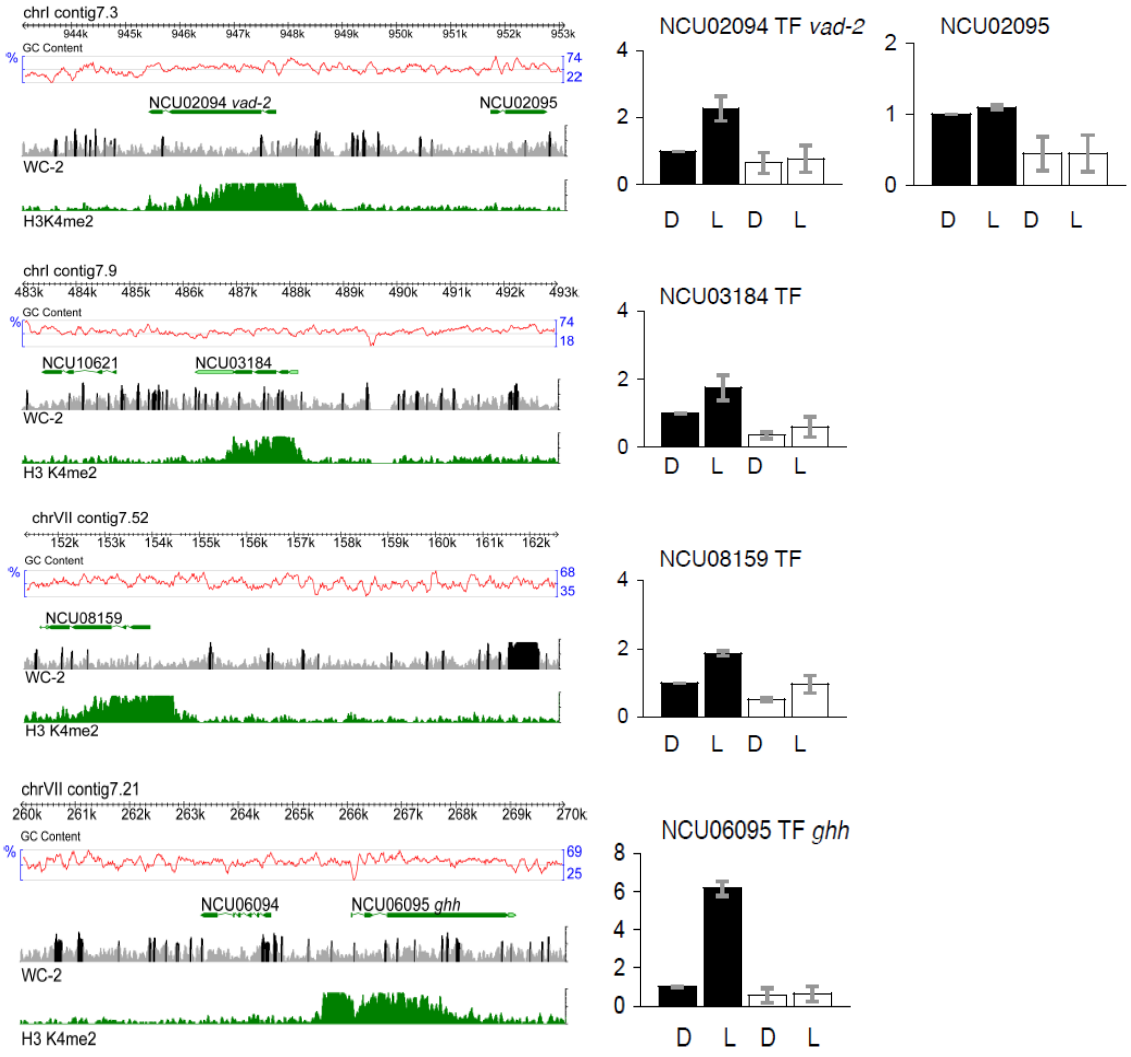


Figure S2. Continued.

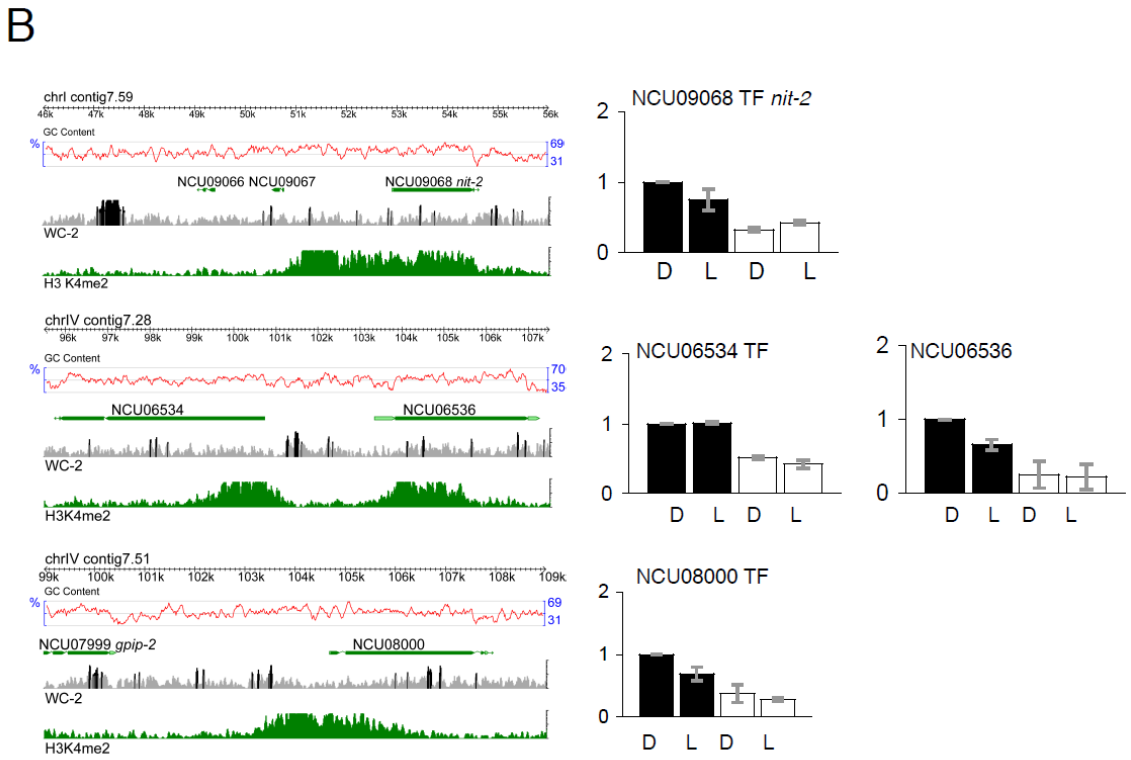


Figure S2. Continued.

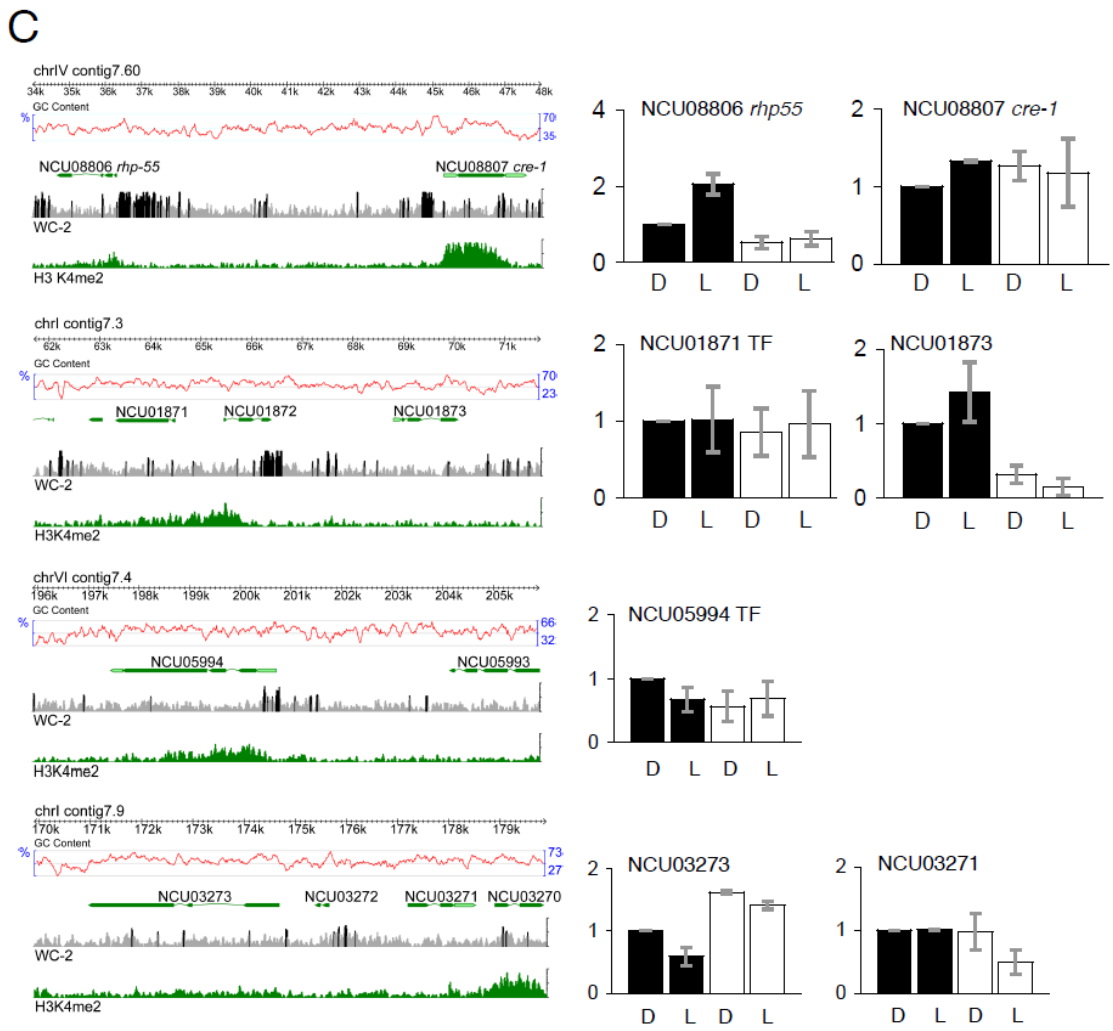
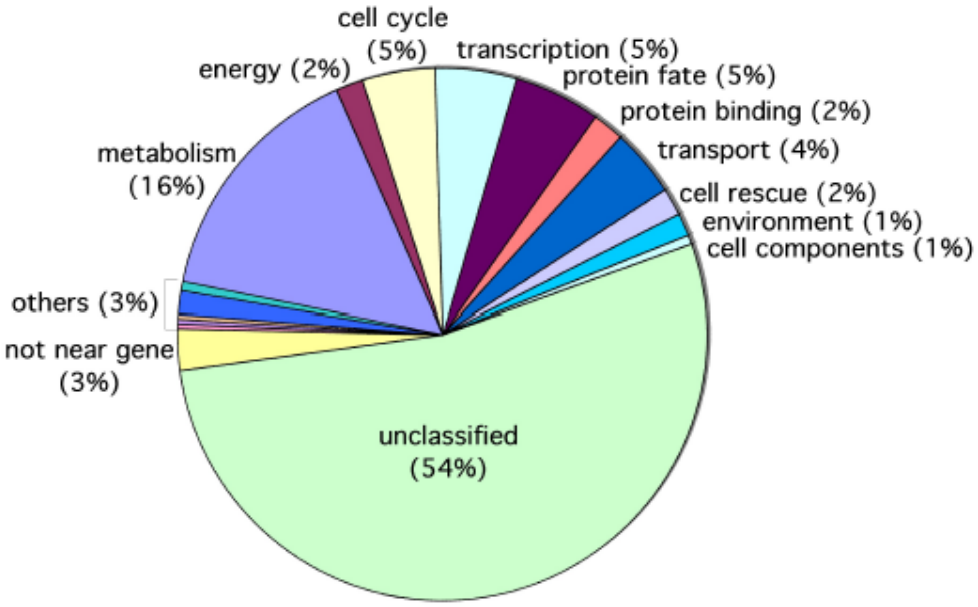


Figure S2. Continued.



All significant peaks (see above)

Functional category	#	%
Metabolism	91	16%
Energy	9	2%
Cell Cycle	27	5%
Transcription	30	5%
Protein Fate	32	5%
Protein Binding	11	2%
Transport	24	4%
Cell Rescue	11	2%
Environment	8	1%
Cellular Components	4	1%
Unclassified	316	54%
Not Near Gene	15	3%
Cell Fate	1	0%
Signal Transduction	1	0%
Differentiation	1	0%
Subcellular localization	3	1%
Protein Synthesis	7	1%
Regulation of Metabolism	4	1%
Total	**584**	**100%**

Peaks with z score >5 (Fig. 1A)

Functional category	#	%
Metabolism	18	17%
Energy	1	1%
Cell Cycle	11	10%
Transcription	9	8%
Protein fate	7	6%
Protein Binding	7	6%
Transport	2	2%
Cell Rescue	1	1%
Environment	3	3%
Cellular Components	3	3%
Unclassified	46	42%
Not Near Gene	1	1%
Total	**109**	**100%**

Figure S3. Genes with WCC binding sites fall into diverse functional categories. The pie chart shows functional categories for all 584 regions. The tables show a comparison of the most significant 109 WC-2 binding sites (z-score >5, $p<2.9e-7$) with all 584 regions (z-score 3.09, $p<0.001$). This comparison reveals enrichment for factors involved in cell cycle control, transcription, protein binding and interactions with the environment.

APPENDIX D

SUPPORTING TABLES FOR CHAPTER III

Table S1. Oligonucleotide pairs

A. Primers used for ChIP-PCR.

Name	Sequence (5' to 3')	expected size (bp)
frq_pLREF	GCAGAGGACCCTGAACTTTTC	
frq_pLRER	TCTCTTGCTCACTTTCCCACAG	100
frq_dLREF	GGTCGGGTTGCCGTGACT	
frq_dLRER	GTCATCGGTCTTGCAGCGG	120
vvdF	CGCACCCAAGATTACTGTTTCCACC	
vvdR	TAGAGTCGAGGATGAGATGGAG	280
al-3F	CTCAGCAGGCAAGCAGGGCAAATC	
al-3R	CTAACCCCGTGCGTCGTCTGATCC	225
hH4-1F	AACCACCGAAACCGTAGAGGGTAC	
hH4-1R	ATCGCCGACACCGTGTGTTGTAAC	400

B. Primers used for isolation of cDNA.

Name	Sequence (5' to 3')
act	CTTGATGTCACGAACGATTTCG
frq	AAGACGAAATCCTCGTCATTC
frq antisense	GTATCTCAATCTGCTTTGTAACCTGGC

C. Primers used for Taqman probe realtime PCR.

Name	Sequence (5' to 3')
NCU01242f	TCCGCACCGGTCAAGAAA
NCU01242r	GGAAGTGGGCGCATGAGAT
NCU01242probe	CGACCAACGTATACCGCGACAACAAAGTC
NCU04731f	TCCTGCAGCGAACATCCACTA
NCU04731r	GTCATCACAGTCCTGCTCGAGAA
NCU04731probe	ATGTCCTCCCGGAGCATCTTGCCA
NCU01871f	CCACGACAACACAAGACCTTAACTT
NCU01871r	CATCCATGACTTGAGACTGGAACA
NCU01871probe	TCGAAGGCACAAATACCAACTCGCCA
NCU01243f	CAATGCCGATGATGAGGAGAA
NCU01243r	GGGAGAGGGACTGCGGAAT
NCU01243probe	TCCAGACCCCCTCTGCGAACATCA
NCU04295f	TCTTGTTGAGGCTCCCGAGTT
NCU04295r	CGAGCCGCTCTGGAAGAAG
NCU04295probe	CTCAGTCCAGCAGCGTTACCCGTTGA
NCU03968f	CATCTTTCCAAATCCGCTACCTACT

Table S1. Continued.

NCU03968r	CCCGTCAGCTGTGATTTACGA
NCU03968probe	TGGCGCCATCAACGATACCCTGTTT
NU00582f	CAGAGACCGTCAGAATGGAACTCT
NCU00582r	GCGGAAAAGCTTGTCACCATAC
NCU00582probe	TGGCGCGATTACATGAGGTTGTGTCA
NCU03071f	TGCAGCATGAATGGATCATGA
NCU03071r	CGCCTTGACTCGTCGTTGA
NCU03071probe	TCGAGC CTGCCACCCCTAGTAGCG
NCU03072f	CGATGGTGGGTGACTCACAAG
NCU03072r	AGCCGAGATGTGGTGTTTGG
NCU03072probe	CAAGAGGATGAGAAAATGGTCGCGGC
NCU06017f	AGCTATGGATCGAGGAGGGT
NCU06017r	GTCCTCATTCATCTCGGGAA
NCU06017probe	CGTCCAGACGTTTCCGGATTCG
NCU07389f	ACGTTCGCCAATTTCTTGAC
NCU07389r	GAGCACGTAATCTCGCATGA
NCU07389probe	CCAACACCGCCGTTCGAGCT
NCU02095f	TTGATGCGCGTAAAGTCTTG
NCU02095r	AGCTCTTGTCGTTTCTCCCA
NCU02095probe	TCCATCTCCTCTTCGTATCTGCGCTC
NCU03271f	CATCGATTTGCTTCCGTACA
NCU03271r	AAACCCGTGAATGTGAATGG
NCU03271probe	CGCTGCGCCATTCTTCACGA
NCU02800f	CGATGCCTACAAGGAAGAGC
NCU02800r	TTTCTGGGACATAGCGTTCC
NCU02800probe	TCAGGCCAAGATGGCACAAACG
NCU001873f	GAGTCTGTGGCAGTTCACGA
NCU01873r	GATTGCGATTCGGAAGGTTA
NCU01873probe	CGCCGAAACTGGCCTCACCT
NCU06536f	ACGAGACAGAATTGGGCAAG
NCU06536r	CTTGTACTTGCCATCGGGTT
NCU06536probe	TATGGCGCAGTCCGCAGCC
NCU09068f	GCGGAATGGACGTTGATAGT
NCU09068r	GTTGAAGGTGGGCATTGTTT
NCU09068probe	CCCGCCAACTCGACAGGATCT
NCU05964f	ATGCGCCATCAAATACACAA
NCU05964r	CATCGTCGTGCTCGGTACTA
NCU05964probe	CGTTTGCAAGCTTGTCGCTCCA
NCU09829f	AAACAGACGTTGGATTTGGC
NCU09829r	ACCGACAGGAAAGCAAACAC
NCU09829probe	AGATCACCGCCAAGGCGTGC
NCU00097f	CCAACACCAGCCCATATACC
NCU00097r	GCCCATGACGGTGTTCTAGT
NCU00097probe	CCCTCCTCACGGGTATCCTCAAGA
NCU07705f	TGGACCTCTCCAACACTTCC

Table S1. Continued.

NCU07705r	CGATAGTGGCTCCTCCATGT
NCU07705probe	CAGGCGTACCAAACGTTGACACGA
NCU06534f	GCTGATCCAGTCCTTGCTTC
NCU06534r	GTAGCCATGCACGCTCACTA
NCU06534probe	CGCATCCAAGGCCCTCACTGA
NCU07728f	AACCGTTGTTATTGCTTGCC
NCU07728r	CGCATTGCAAATGGTATGAC
NCU07728probe	TTCGTCCCTTCGCCAGAGCG
NCU00275f	CGACAATTTCATGGATGTGG
NCU00275r	CTAGGCTTCGCTCGCATACT
NCU00275probe	CGGGCACCTCTTCCGTATCCG
NCU02094f	TATACCGAACCTTGGCCTTG
NCU02094r	CCACCCATGGTAGTTGATCC
NCU02094probe	CCGCCTCGAACGGATCCATC
NCU05994f	AGGCACGCAACTCGTACTCT
NCU05994r	GCGCTTTGAAAGGTTGAGTC
NCU05994probe	CCGAGTCAGGCAACCCGCTG
NCU03273f	AACGTCAACTGGTCCCAGAC
NCU03273r	GGCTAGACTGTGCAGGAAGG
NCU03273probe	CGCCCGACGAGAAGCGTCTC
NCU03184f	TAGTCCTCCCATGTCTTCGG
NCU03184r	AGTAGTTGGAGCTGCCCTGA
NCU03184probe	TCGAGCAAGCCAGCAAGCCA
NCU08000f	GTGGAAGCGTATCTGGTGGT
NCU08000r	ATATGACAAGGTCGCCCAAG
NCU08000probe	ACCGATCGACAGCTGTCGCG
NCU08159f	TTTCCAATGGAGGAGACGAC
NCU08159r	CAAGACCCGCTCTCTGTTTC
NCU08159probe	CCCTCTGCCTTTCCCTCCACG
NCU06095f	AGCACACCAGAGAGGAGGAA
NCU06095r	GAGAAACCGTCAAACGAAGC
NCU06095probe	CGATGCGCGTCCTCTTGTCG

APPENDIX E

SUPPORTING TABLES FOR CHAPTER IV

Table S1. The total number of sequenced reads obtained for each ChIP-seq and RNA-seq time points and the percent mapped to the Neurospora assembly.

HTS#	**Name**	**Media**	**Flowcell**	**Lane**	**Adapter sequence**	**Total reads**	**Mapped reads**	**Pct aligned**	
ChIP-seq:									
HTS459	ADV-1::V5 LL0	Vogel's	248	6	AGTCT	8074857	7311499	90.55	
HTS460	ADV-1::V5 LL15	Vogel's	248	6	CTAGT	8215955	7044127	85.74	
HTS461	ADV-1::V5 LL30	Vogel's	248	6	GACTT	7465356	6888393	92.27	
HTS462	ADV-1::V5 LL60	Vogel's	248	6	TCGAT	8691971	7673593	88.28	
HTS463	ADV-1::V5 LL0	Bird's	248	7	ACGTT	7054866	6270359	88.88	
HTS464	ADV-1::V5 LL15	Bird's	248	7	CGTAT	8108212	7490777	92.39	
HTS465	ADV-1::V5 LL30	Bird's	248	7	GTACT	7955625	7199917	90.5	
HTS466	ADV-1::V5 LL60	Bird's	248	7	TACGT	8929262	8164865	91.44	
HTS#	**Name**	**Media**	**Flowcell**	**Lane**	**Index**	**Adapter sequence**	**Total Reads**	**Mapped reads**	**Pct aligned**
RNA-seq:									
HTS640	ΔADV-1 LL0 OdT	Bird's	FC1060	1	AD001	ATCACG	22152176	22152176	100
HTS641	ΔADV-1 15	Bird's	FC1060	1	AD003	TTAGGC	21195283	21195283	100

Table S1. Continued.

HTS#	Name	Media	Flowcell	Lane	Index	Adapter sequence	Total Reads	Mapped reads	Pct aligned
HTS642	ΔADV-1 LL60 OdT	Bird's	FC1060	1	AD008	ACTTGA	20840266	20840266	100
HTS655	ΔADV-1 LL0 N6 Rep 1	Bird's	FC1060	3	AD009	GATCAG	20100831	20100831	100
HTS656	ΔADV-1 LL15 N6 Rep 1	Bird's	FC1060	3	AD010	TAGCTT	26728753	26728753	100
HTS657	ΔADV-1 LL60 N6 Rep 1	Bird's	FC1060	3	AD011	GGCTAC	27382616	27382616	100
HTS849	ΔADV-1 LL0 N6 Rep 2	Bird's	FC1090	6	AR011	GGCTAC	30785833	30785833	100
HTS850	ΔADV-1 LL15 N6 Rep 2	Bird's	FC1090	6	AR020	GTGGCC	31744647	31744647	100
HTS851	ΔADV-1 LL60 N6 Rep 2	Bird's	FC1090	6	AR021	GTTTCG	31946411	31946411	100

Table S2. List of 65 unique genes (listed by their unique NCU# identifier) that were identified as ADV-1 direct targets by ChIP-seq, and are differentially expressed with q≤0.2, fold change (fc) ≥2X in Δadv-1 cells compared to WT cells by RNA-seq at one or more time points (LL0, LL15 and LL60).

Gene	Broad annotation	DDlog2 (fold change)	DD q value	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU00985	hydrolase	-1.79769	1	-1.79769	1	-6.70057	0.00118125
NCU00994	endothiapepsin	-3.52382	0.146661	-3.97729	0.0591233	-4.33467	0.0544314
NCU00995	hypothetical protein	-2.40832	0.147075	-2.38672	0.117461	-3.74542	0.0275177
NCU02297	hypothetical protein	4.06445	0.0266811	1.73517	0.954377	1.69311	0.933245
NCU02988	hypothetical protein	-4.02474	0.00023138	-3.8922	0.00028704	-3.89994	0.00652876
NCU03641	beta-glucosidase 2	-2.37899	0.410675	-3.23556	0.0375295	-4.20976	0.0250452
NCU04249	hypothetical protein	2.98698	0.0103513	2.37502	0.115352	1.97346	0.610487
NCU04342	hypothetical protein	-0.632059	0.999991	-0.664756	0.999999	-2.82847	0.117891
NCU04525	hypothetical protein	1.03417	0.999991	1.82653	0.557968	2.87398	0.168155
NCU04645	DUF124 domain-containing protein	-1.79961	0.339772	-1.88211	0.296714	-1.24961	0.999917
NCU04732	hypothetical protein	-3.65947	0.00054548	-4.28446	2.0163E-05	-4.13186	0.00261585
NCU04816	hypothetical protein	-1.57694	0.537721	-1.94528	0.255247	-3.49548	0.01832
NCU05137	non-anchored cell wall protein 1	-4.35874	0.00013811	-4.48786	2.8677E-05	-5.03591	0.00355955
NCU05712	hypothetical protein	-5.13642	0.00024775	-4.2232	0.0313691	-4.94758	0.0339594
NCU05835	hypothetical protein	-3.11266	0.0266526	-3.97575	0.00323912	-2.50338	0.438059
NCU05841	UMTA	-2.1777	0.190337	-3.1401	0.00734315	-3.62423	0.0191696
NCU05915	hypothetical protein	-4.27746	0.00268629	-4.47113	0.00387067	-4.24835	0.0219139
NCU06005	glycerol kinase	2.1262	0.137914	3.30183	0.00148812	4.00645	0.00355955
NCU06261	uracil phosphoribosyltransferase	2.18838	0.116383	2.83091	0.015052	1.47433	0.879924

Table S2. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD q value	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU06467	hypothetical protein	-2.91261	0.279374	-2.51923	0.147342	-1.44294	0.999917
NCU06945	hypothetical protein	6.56684	5.1095E-10	5.84699	3.9842E-08	5.99286	7.1002E-05
NCU07222	hypothetical protein	-1.60578	0.490762	-1.64145	0.506776	-2.18349	0.430865
NCU07392	transcriptional regulatory protein Pro-1	-2.12613	0.147075	-1.60409	0.510116	-0.352964	0.999917
NCU07787	CCG-14	-3.5641	0.0118025	-2.48984	0.20678	-4.82797	0.039155
NCU08447	hypothetical protein	1.84966	0.374106	1.75724	0.426876	1.37111	0.933245
NCU00606	short-chain alcohol dehydrogenase	3.92137	0.0116388	2.19456	0.573366	2.21875	0.691641
NCU00876	ferric reductase transmembrane component	-1.47458	0.874544	-2.34919	0.139039	-2.17218	0.511181
NCU01897	hypothetical protein	-1.87739	0.285675	-1.90138	0.272381	-2.71969	0.150603
NCU01898	hypothetical protein, variant	-2.66299	0.0933937	-2.70518	0.0840978	-3.6686	0.135005
NCU02184	hypothetical protein	3.0734	0.016256	2.44736	0.136477	2.0148	0.801291
NCU02767	PRO41 protein	-2.23238	0.184386	-1.50097	0.74288	-0.976435	0.999917
NCU03035	hypothetical protein	-2.83601	0.0207621	-1.92878	0.243397	-1.08015	0.999917
NCU03494	heterokaryon incompatibility protein	-3.86189	0.112189	-4.60087	0.0777472	-1.79769e+308	0.600129
NCU03643	cutinase transcription factor 1 beta	-2.99341	0.0264963	-1.53328	0.761703	-1.81811	0.729779
NCU03696	nitroreductase	3.3367	0.00369791	3.35626	0.00487498	2.11433	0.539372
NCU04557	hypothetical protein	-3.06439	0.158712	-2.71622	0.255042	-3.40719	0.107516
NCU0513	endo alpha	-2.498	0.0547	-2.39742	0.0748166	-2.62862	0.196201

Table S2. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD q value	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU05788	hypothetical protein	-3.52876	0.0163081	-1.96573	0.704519	-3.16162	0.206009
NCU06912	hypothetical protein	-2.83494	0.0143444	-3.10217	0.00444062	-3.48831	0.0219139
NCU07192	hypothetical protein	-2.83009	0.0182129	-3.86832	0.00012477	-2.98948	0.0882796
NCU07337	hypothetical	0.647506	0.999991	-0.154849	0.999999	-2.626	0.191413
NCU07395	DNA-binding protein SMUBP-2	-1.01696	0.999991	-1.17566	0.971703	-2.97197	0.168629
NCU07452	NADPH dehydrogenase	2.44282	0.274326	3.51945	0.00487498	3.15511	0.237425
NCU07696	hypothetical protein	-2.43474	0.0690471	-2.03458	0.210528	-1.42375	0.933245
NCU07736	PEP5	1.24128	0.97283	1.53014	0.58596	2.81516	0.117664
NCU07748	hypothetical protein	0.405709	0.999991	2.6176	0.117461	0.9432	0.999917
NCU07804	hypothetical protein	-2.23141	0.163217	-2.03398	0.2803	-2.23654	0.452641
NCU07817	non-anchored cell wall protein 3	-2.26679	0.498965	-2.49986	0.334024	-4.38966	0.272804
NCU08056	ABC drug exporter AtrF	2.22194	0.20287	1.83207	0.516467	2.36339	0.457187
NCU08223	hypothetical protein	-4.18802	0.110568	-2.68006	0.889572	-1.79769e+308	0.720884
NCU08655	hypothetical protein	-1.52511	0.732246	-1.8548	0.408625	-3.00402	0.0964484
NCU08720	hypothetical protein	-4.82713	9.3617E-07	-4.6154	1.5771E-06	-5.01069	0.00010995
NCU08739	endothiapepsin	-2.15949	0.189262	-1.06424	0.999999	-2.63547	0.305746
NCU09024	hypothetical protein	3.25523	0.00706245	3.48669	0.00258557	3.28829	0.0524218
NCU09337	plasma membrane fusion protein prm-1	-3.40675	0.0846685	-1.15678	0.999999	-2.4437	0.612438
NCU09506	hypothetical protein	-2.81035	0.0176713	-3.10618	0.0064441	-1.81205	0.753775
NCU09702	endo-beta-1,6-galactanase	-1.16068	0.999991	-1.60222	0.881398	-2.8998	0.17639
NCU09764	hypothetical protein	-2.27132	0.158712	-0.915678	0.999999	-0.0669638	0.999917

Table S2. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD q value	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU09775	alpha-N-arabinofuranosidase	-0.953502	0.999991	-0.751792	0.999999	-4.20358	0.00192203
NCU09791	beta-1,3-exoglucanase	2.56324	0.0765864	2.24649	0.190542	1.71415	0.796442
NCU09848	hypothetical protein	-2.91137	0.0578152	-2.67153	0.217298	-1.76826	0.865002
NCU09906	hypothetical protein	-2.24469	0.154931	-2.4605	0.0807213	-2.11739	0.518494
NCU10865	hypothetical protein	-6.88294	4.664E-05	-5.18018	0.00148812	-7.45767	0.00010995
NCU11763	hypothetical protein	-2.14702	0.178306	-1.57462	0.561984	-1.48339	0.886589

Table S3. Gene Ontology (GO) and Functional Category (FunCat) analyses of 358 ADV-1 regulated genes.

GO ID	Term	P value	List # gene annotations	List size	Total # gene annotation	Population size
GO:0010438	Cellular response to sulfur starvation	7.85E-06	4	358	5	10000
GO:0009416	Response to light stimulus	0.000384383	26	358	353	10000
GO:0030582	Fruiting body development	0.008336034	9	358	98	10000

FUNCTIONAL CATEGORY	abs SET	rel SET	abs GENOME	rel GENOME	rel SET/rel GENOME	P-VALUE
01.20 secondary metabolism	85	27.5	1225	12.1	2.27272727	8.04E-14
01.05 C-compound and carbohydrate metabolism	95	30.8	1641	16.3	1.88957055	8.91E-11
20.09.18.07 non-vesicular cellular import	22	7.14	136	1.35	5.28888889	1.25E-10
01.02 nitrogen, sulfur and selenium metabolism	40	12.9	426	4.23	3.04964539	1.80E-10
01 METABOLISM	157	50.9	3495	34.7	1.46685879	2.01E-09
20.03 transport facilities	48	15.5	754	7.48	2.07219251	7.94E-07
32.05.05 virulence, disease factors	24	7.79	279	2.77	2.81227437	4.61E-06
34.11.12 perception of nutrients and nutritional adaptation	21	6.81	225	2.23	3.05381166	5.23E-06
34.01.01 homeostasis of cations	32	10.3	445	4.42	2.33031674	5.69E-06
32.05 disease, virulence and defense	43	13.9	706	7.01	1.9828816	9.90E-06
34.01.01.01 homeostasis of metal ions (Na, K, Ca etc.)	25	8.11	322	3.19	2.54231975	1.75E-05
20.01.03 C-compound and carbohydrate transport	25	8.11	329	3.26	2.48773006	2.52E-05
34.01 homeostasis	34	11	525	5.21	2.11132438	2.67E-05
01.05.03 polysaccharide metabolism	23	7.46	299	2.97	2.51178451	4.42E-05
01.20.05 metabolism of acetic acid derivatives	13	4.22	121	1.2	3.51666667	8.06E-05
20.01.15 electron transport	30	9.74	467	4.63	2.10367171	9.43E-05
20.09.18 cellular import	34	11	565	5.61	1.96078431	0.0001156

Table S3. Continued.

FUNCTIONAL CATEGORY	abs SET	rel SET	abs GENOME	rel GENOME	rel SET/rel GENOME	P-VALUE
20.01 transported compounds (substrates)	77	25	1696	16.8	1.48809524	0.00014276
01.01 amino acid metabolism	44	14.2	824	8.18	1.73594132	0.00018071
01.05.02 sugar, glucoside, polyol and carboxylate metabolism	27	8.76	418	4.15	2.11084337	0.00019492
20.01.23 allantoin and allantoate transport	6	1.94	29	0.28	6.92857143	0.00020466
01.02.07 regulation of nitrogen, sulfur and selenium metabolism	14	4.54	153	1.51	3.00662252	0.00024415
01.06 lipid, fatty acid and isoprenoid metabolism	51	16.5	1022	10.1	1.63366337	0.00028277
20.01.01 ion transport	29	9.41	474	4.7	2.00212766	0.00028453
20.01.01.01 cation transport (H+, Na+, K+, Ca2+ , NH4+, etc.)	27	8.76	429	4.26	2.05633803	0.00029611
20.01.07 amino acid/amino acid derivatives transport	12	3.89	121	1.2	3.24166667	0.0003219
01.20.05.11 metabolism of polyketides	8	2.59	58	0.57	4.54385965	0.00035464
32.07 detoxification	28	9.09	459	4.55	1.9978022	0.00038036
34 INTERACTION WITH THE ENVIRONMENT	63	20.4	1358	13.4	1.52238806	0.00038088
20.09.16.01 Type I protein secretion system (ABC-type transport systems)	10	3.24	91	0.9	3.6	0.00044494
20.01.25 vitamine/cofactor transport	10	3.24	92	0.91	3.56043956	0.00048595
01.07 metabolism of vitamins, cofactors, and prosthetic groups	31	10	538	5.34	1.87265918	0.00049646
20.01.01.07 anion transport	12	3.89	127	1.26	3.08730159	0.00050238
34.11 cellular sensing and response to external stimulus	42	13.6	814	8.08	1.68316832	0.00052726
32.10 degradation / modification of foreign (exogenous) compounds	11	3.57	113	1.12	3.1875	0.00066522
01.01.09 metabolism of the cysteine - aromatic group	21	6.81	316	3.13	2.17571885	0.00069667

Table S3. Continued.

FUNCTIONAL CATEGORY	abs SET	rel SET	abs GENOME	rel GENOME	rel SET/rel GENOME	P-VALUE
16.21.07 NAD/NADP binding	20	6.49	295	2.93	2.21501706	0.00072574
20.01.27 drug/toxin transport	13	4.22	152	1.5	2.81333333	0.00077168
01.05.02.07 sugar, glucoside, polyol and carboxylate catabolism	21	6.81	319	3.16	2.15506329	0.00078698
01.06.10 regulation of lipid, fatty acid and isoprenoid metabolism	10	3.24	99	0.98	3.30612245	0.00086994
01.20.35 metabolism of secondary products derived from L-phenylalanine and L-tyrosine	13	4.22	155	1.53	2.75816993	0.0009265
20.01.01.01.01 heavy metal ion transport (Cu+, Fe3+, etc.)	13	4.22	157	1.55	2.72258065	0.00104356
32.07.05 detoxification by export	10	3.24	106	1.05	3.08571429	0.00147426
20.01.11 amine / polyamine transport	7	2.27	56	0.55	4.12727273	0.00150162
20 CELLULAR TRANSPORT, TRANSPORT FACILITIES AND TRANSPORT ROUTES	88	28.5	2165	21.5	1.3255814	0.00181341
01.01.09.05.02 degradation of tyrosine	4	1.29	19	0.18	7.16666667	0.00231463
01.20.35.01 metabolism of phenylpropanoids	12	3.89	152	1.5	2.59333333	0.00241252
32.05.05.01 toxins	9	2.92	96	0.95	3.07368421	0.00262688
01.20.37 metabolism of peptide derived compounds	7	2.27	62	0.61	3.72131148	0.00272469
16.21.05 FAD/FMN binding	12	3.89	158	1.56	2.49358974	0.00332154
16.21 complex cofactor/cosubstrate/vitamin e binding	29	9.41	558	5.54	1.69855596	0.00357331
01.20.36 non-ribosomal peptide synthesis	5	1.62	35	0.34	4.76470588	0.00396404
20.01.13 lipid/fatty acid transport	14	4.54	206	2.04	2.2254902	0.00434091
20.03.22 transport ATPases	11	3.57	144	1.43	2.4965035	0.00464217
01.20.33 metabolism of	4	1.29	23	0.22	5.8636363	0.00480279

Table S3. Continued.

FUNCTIONAL CATEGORY	abs SET	rel SET	abs GENOME	rel GENOME	rel SET/rel GENOME	P-VALUE
34.01.03 homeostasis of anions	5	1.62	37	0.36	4.5	0.00506456
32 CELL RESCUE, DEFENSE AND VIRULENCE	74	24	1830	18.1	1.32596685	0.00536604
32.07.03 detoxification by modification	10	3.24	128	1.27	2.5511811	0.00583363
01.02.02.09 catabolism of nitrogenous compounds	5	1.62	39	0.38	4.26315789	0.00636566
01.20.35.01.01 metabolism of melanins	6	1.94	55	0.54	3.59259259	0.00642296
20.03.02.03 antiporter	10	3.24	130	1.29	2.51162791	0.00649568
01.25.01 extracellular polysaccharide degradation	7	1.94	56	0.55	3.52727273	0.00701352
34.11.03 chemoperception and response	20	6.49	360	3.57	1.81792717	0.00730646
02.07 pentose-phosphate pathway	8	2.59	94	0.93	2.78494624	0.00799293
32.10.07 degradation / modification of foreign (exogenous) polysaccharides	5	1.62	42	0.41	3.95121951	0.00873175
34.01.03.01 homeostasis of sulfate	2	0.64	5	0.04	16	0.0087765
20.03.25 ABC transporters	10	3.24	136	1.35	2.4	0.00883641
16.17.05 sodium binding	5	1.62	43	0.42	3.85714286	0.00963939
01.02.03 sulfur metabolism	4	1.29	29	0.28	4.60714286	0.01115941
01.05.08.07 C-4 compound catabolism	4	1.29	29	0.28	4.60714286	0.01115941
16.11 amino acid/amino acid derivatives binding	3	0.97	16	0.15	6.46666667	0.01181419
34.01.07 cholesterol homeostasis	2	0.64	6	0.05	12.8	0.01289995
01.01.09.01 metabolism of glycine	5	1.62	47	0.46	3.52173913	0.01391741
01.01.09.01.01 biosynthesis of glycine	4	1.29	32	0.31	4.16129032	0.01573128
01.06.02.02 glycolipid metabolism	7	2.27	87	0.86	2.63953488	0.01699621
01.05.08 C-4 compound metabolism	4	1.29	33	0.32	4.03125	0.01748097
01.25 extracellular metabolism	8	2.59	109	1.08	2.39814815	0.01846681
01.20.15 chorismate met.	7	2.27	89	0.88	2.57954545	0.01904158

Table S3. Continued.

FUNCTIONAL CATEGORY	abs SET	rel SET	abs GENOME	rel GENOME	rel SET/rel GENOME	P-VALUE
01.01.05 metabolism of urea cycle, creatine and polyamines	6	1.94	70	0.69	2.8115942	0.0198912
34.01.01.03 homeostasis of protons	8	2.59	111	1.1	2.35454545	0.02037832
01.01.09.04.02 degradation of phenylalanine	5	1.62	52	0.51	3.17647059	0.02085412
01.05.09 aminosaccharide metabolism	6	1.94	71	0.7	2.77142857	0.02118852
20.03.02.03.01 proton driven antiporter	6	1.94	71	0.7	2.77142857	0.02118852
02.16 fermentation	12	3.89	202	2	1.945	0.02135364
32.05.03 defense related proteins	9	2.92	135	1.34	2.17910448	0.02266939
16.16 antigen binding	2	0.64	8	0.07	9.14285714	0.02312458
32.05.01.03.01 organic chemical agent resistance	4	1.29	38	0.37	3.48648649	0.02801459
16.05 polysaccharide binding	5	1.62	57	0.56	2.89285714	0.02972949
20.01.09 peptide transport	4	1.29	39	0.38	3.39473684	0.03048875
01.01.09.03.01 biosynthesis of cysteine	3	0.97	23	0.22	4.40909091	0.03193565
01.20.37.03 metabolism of peptide antibiotics	3	0.97	24	0.23	4.2173913	0.03569407
32.01.04 pH stress response	3	0.97	24	0.23	4.2173913	0.03569407
01.06.02.03 sphingolipid metabolism	2	0.64	10	0.09	7.11111111	0.03569765
01.02.07.01 regulation of nitrogen metabolism	4	1.29	41	0.4	3.225	0.03581248
01.01.09.05 metabolism of tyrosine	5	1.62	60	0.59	2.74576271	0.03604399
32.05.01 resistance proteins	16	5.19	320	3.17	1.63722397	0.03683595
32.05.01.03 chemical agent resistance	10	3.24	173	1.71	1.89473684	0.03975466
01.05.07 C-3 compound metabolism	2	0.64	11	0.1	6.4	0.04276419
02.45 energy conversion and regeneration	10	3.24	177	1.75	1.85142857	0.045271
01.06.05 fatty acid metabolism	13	4.22	253	2.51	1.6812749	0.04714236
01.08.04 regulation of the metabolism of intercellular mediators	4	1.29	45	0.44	2.93181818	0.04797192

Table S4. Gene Ontology (GO) analyses of known *ccgs* (p-value≤0.05) with significant ADV-1-binding sites identified by ADV-1 ChIP-seq. GO analyses were carried out for each of the time points (LL0, LL15, LL30 and LL60).

Dark ChIP/ccgs	**GO term**	**P-value**	**List #gene annotations**	**List size**	**Total #gene annotations**	**Population size**
GO:0006091	generation of precursor metabolites and energy	0.00108378	5	186	41	10000
GO:0006090	pyruvate metabolic process	0.00136940	3	186	12	10000
GO:0006949	syncytium formation	0.00281424	4	186	31	10000
GO:0007154	cell communication	0.00364106	6	186	77	10000
GO:0008152	metabolic process	0.00373465 7	25	186	733	10000
GO:0030436	asexual sporulation	0.00778705	6	186	90	10000
GO:0006094	gluconeogenesis	0.00961147 2	2	186	8	10000
GO:0006096	glycolytic process	0.01220168 5	2	186	9	10000
GO:0042752	regulation of circadian rhythm	0.01788708 3	3	186	29	10000
GO:0009063	cellular amino acid catabolic process	0.03299166 9	2	186	15	10000
GO:0009416	response to light stimulus	0.04074739 8	12	186	353	10000
GO:0030582	fruiting body development	0.04091464 2	5	186	98	10000
GO:0070787	conidiophore development	0.04807324 2	4	186	71	10000
15LL ChIP/ccgs						
GO:0009416	response to light stimulus	0.00032314	26	346	353	10000
GO:0070787	conidiophore development	0.00346527	8	346	71	10000
GO:0030436	asexual sporulation	0.00445861	9	346	90	10000
GO:0008152	metabolic process	0.00664676	39	346	733	10000
GO:0030448	hyphal growth	0.01319377	11	346	144	10000
GO:0006091	generation of precursor metabolites and energy	0.01422359	5	346	41	10000
GO:0009063	cellular amino acid catabolic process	0.01457987	3	346	15	10000

Table S4. Continued.

15LL ChIP/ccgs GO term		P-value	List #gene annotations	List size	Total #gene annotations	Population size
GO:0042752	regulation of circadian rhythm	0.01821917	4	346	29	10000
GO:0030582	fruiting body development	0.02262322	8	346	98	10000
GO:0006949	syncytium formation	0.02284407	4	346	31	10000
GO:0006094	gluconeogenesis	0.03037744	2	346	8	10000
GO:0009649	entrainment of circadian clock	0.03037744	2	346	8	10000
GO:0009651	response to salt stress	0.03406777	4	346	35	10000
GO:0006096	glycolytic process	0.03815632	2	346	9	10000
GO:0009648	photoperiodism	0.04659924	2	346	10	10000
GO:0006090	pyruvate metabolic process	0.0076262	3	346	12	1000
30LL ChIP/ccgs						
GO:0006091	generation of precursor metabolites and energy	8.60E-05	7	259	41	10000
GO:0009416	response to light stimulus	0.000161884	22	259	353	10000
GO:0008152	metabolic process	0.000433053	35	259	733	10000
GO:0030436	asexual sporulation	0.0005 93395	9	259	90	10000
GO:0030582	fruiting body development	0.001101217	9	259	98	10000
GO:0006090	pyruvate metabolic process	0.0033 53632	3	259	12	10000
GO:0009063	cellular amino acid catabolic process	0.006541431	3	259	15	10000
GO:0042752	regulation of circadian rhythm	0.006705832	4	259	29	10000
GO:0006949	syncytium formation	0.0085 25843	4	259	31	10000
GO:0030433	ER-associated ubiquitin-dependent protein catabolic process	0.009709059	2	259	6	10000
GO:0070787	conidiophore development	0.010929988	6	259	71	10000
GO:0005977	glycogen metabolic process	0.013357269	2	259	7	10000

Table S4. Continued.

30LL ChIP/ccgs	**GO term**	**P-value**	**List #gene annotations**	**List size**	**Total #gene annotations**	**Population size**
GO:0009649	entrainment of circadian clock	0.017501923	2	259	8	10000
GO:0006096	glycolytic process	0.022114467	2	259	9	10000
GO:0009648	photoperiodism	0.027167502	2	259	10	10000
GO:0044042	glucan metabolic process	0.044711983	2	259	13	10000
GO:0006364	rRNA processing	0.047482742	3	259	31	10000
GO:0006090	pyruvate metabolic process	0.00125008	3	182	12	10000
GO:0006949	syncytium formation	0.00251005	4	182	31	10000
GO:0006094	gluconeogenesis	0.00904343	2	182	8	10000
GO:0006096	glycolytic process	0.01148513	2	182	9	10000
60LL ChIP/ccgs						
GO:0009416	response to light stimulus	0.01516932	13	182	353	10000
GO:0042752	regulation of circadian rhythm	0.01645149	3	182	29	10000
GO:0030436	asexual sporulation	0.02656311	5	182	90	10000
GO:0006006	glucose metabolic process	0.03514406	2	182	16	10000
GO:0042440	pigment metabolic process	0.03514406	2	182	16	10000
GO:0030582	fruiting body development	0.03648945	5	182	98	10000
GO:0006091	generation of precursor metabolites and energy	0.0408158	3	182	41	10000
GO:0070787	conidiophore development	0.14723632	3	182	71	10000
GO:0044237	cellular metabolic process	0.02361676	21	182	708	10000

Table S5. List of 101 indirect targets of ADV-1 that are predicted *ccgs*. The list includes (log2) fold change for each gene, where fold change is the difference in the normalized reads (FPKM) of a gene between WT and ΔADV-1 cells, and the corresponding FDR-adjusted p-value (q-value).

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU00473	hypothetical protein	-0.998814	0.999991	-1.42661	0.715112	-2.61382	0.191413
NCU00701	N,O-diacetylmuramidase	-0.571529	0.999991	-1.16153	0.999999	-3.18479	0.0694578
NCU00729	hypothetical protein	2.8638	0.0541072	2.67789	0.0888378	2.2125	0.62545
NCU00830	ctr copper transporter	-1.71007	0.547007	-1.96223	0.365556	-0.813464	0.999917
NCU00955	monooxygenase	-3.65164	0.00044922	-3.94238	7.5791E-05	-3.0516	0.07382
NCU01055	chromate ion transporter	6.16759	0.00027835	5.7408	0.100025	6.28775	0.999917
NCU01193	hypothetical protein	-3.7402	0.00026699	-3.91717	7.1714E-05	-2.81722	0.119947
NCU01595	SOF1	-0.564293	0.999991	0.85724	0.999999	2.79601	0.191413
NCU01861	short chain dehydrogenase/reductase family	3.19496	0.00432575	2.43608	0.0569322	-0.332617	0.999917
NCU01904	2-deoxy-D-gluconate 3-dehydrogenase	-2.90831	0.135221	-2.67973	0.939937	-4.12149	0.510339
NCU01957	AR2	0.93816	0.999991	-2.77431	0.0185261	0.568314	0.999917
NCU02027	acyl-protein thioesterase 1	2.14127	0.13938	2.11437	0.150473	0.742557	0.999917
NCU02344	fungal cellulose binding domain-containing	-4.05396	0.00023771	-3.80255	0.00144418	-3.46555	0.0683108
NCU02369	polygalacturonase	-5.00959	0.188769	-4.74965	0.562225	-4.71355	0.999917
NCU02480	short-chain dehydrogenase/reductase	2.15315	0.519134	3.36027	0.0358559	1.9417	0.887722
NCU02636	ubiquitin conjugating enzyme	-1.16188	0.999991	-2.34497	0.0762154	-1.56432	0.815827
NCU02651	hypothetical protein	-1.9989	0.268925	-2.3187	0.0769418	-1.92697	0.600129

Table S5. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU02773	hypothetical protein	2.71978	0.0204629	2.76431	0.0191396	1.6702	0.783074
NCU02915	hypothetical protein	1.46311	0.874544	2.57332	0.200772	3.31254	0.0848417
NCU02946	mitochondrial carrier domain-containing protein	2.33349	0.12829	1.83386	0.450642	1.27464	0.999917
NCU03026	chitinase	-1.80163	0.999991	-4.34841	0.0357374	-2.30475	0.761465
NCU03208	hypothetical protein	-1.97433	0.205304	-2.03126	0.20678	-2.34649	0.33854
NCU03358	ketoreductase	2.49611	0.0999137	2.74405	0.0591233	1.59629	0.887758
NCU03383	hypothetical protein	-2.36639	0.588517	-4.41127	0.030101	-2.2701	0.747849
NCU03813	formate dehydrogenase	3.63736	0.00048676	5.2128	5.5094E-07	3.29688	0.052342
NCU03949	FMN-dependent 2-nitropropane dioxygenase	-1.3172	0.999991	-0.895693	0.999999	-2.68819	0.178082
NCU04371	peptidyl-prolyl cis-trans isomerase fkr-3	2.13747	0.147075	1.66862	0.45547	0.955476	0.999917
NCU04384	hypothetical protein	3.28095	0.0207621	3.20313	0.0419524	2.34931	0.476332
NCU04451	hypothetical protein	-2.34451	0.204885	-1.82973	0.455466	-2.47586	0.324055
NCU04455	hypothetical protein	0.0586375	0.999991	0.546974	0.999999	2.865	0.142851
NCU04488	hypothetical protein	-0.483536	0.999991	-1.52473	0.758608	-2.83219	0.164041
NCU04635	hypothetical protein	3.39214	0.00219393	2.9408	0.0185261	-0.0699191	0.999917
NCU04695	methyltransferase-UbiE family protein	-3.82368	0.00263071	-2.06086	0.445412	-2.26938	0.511181
NCU04765	hypothetical protein	-0.153115	0.999991	0.633218	0.999999	2.98649	0.112342
NCU04865	polyketide synthase 3	2.45281	0.150407	1.97984	0.562225	1.60232	0.933245
NCU04931	hypothetical protein	-2.09035	0.150407	-2.55153	0.0365493	-3.90633	0.00479916
NCU04950	hypothetical protein	-2.05345	0.208662	-1.32675	0.999999	-2.53547	0.29393

Table S5. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU05154	calcium P-type ATPase	-2.04547	0.326886	-2.36827	0.0994572	-2.87655	0.114768
NCU05013	hypothetical protein	1.31749	0.977615	1.4648	0.757604	3.49754	0.02948
NCU05354	hypothetical protein	-3.29703	0.184386	-2.76036	0.534897	-4.23282	0.168629
NCU05567	hypothetical protein	-1.79769e+308	0.554005	-3.00198	0.46396	-5.59598	0.0208055
NCU05598	rhamnogalacturonase B	-1.19132	0.955574	-0.583394	0.999999	-2.65562	0.167418
NCU05627	high affinity glucose transporter ght1	-2.34115	0.0978475	-1.6523	0.533684	-0.272256	0.999917
NCU05707	hypothetical protein	1.79769e+308	0.378351	6.21502	0.0203791	4.11299	0.160734
NCU05768	mating response protein POI2	-2.28974	0.562572	-3.11474	0.136156	-4.4338	0.00282564
NCU05822	hypothetical protein	1.72655	0.763598	1.98887	0.345962	3.62969	0.048795
NCU05828	hypothetical protein	2.2651	0.17934	2.88959	0.0125358	2.62953	0.191413
NCU05885	flavin-binding monooxygenase	9.00073	0	8.76579	0	7.1412	1.4828E-09
NCU05886	DUF895 domain membrane protein	13.0343	1.211E-10	10.3869	0	9.08644	7.5287E-12
NCU05887	oxidoreductase	8.56831	0	9.10384	0	8.65621	1.2906E-12
NCU06040	hypothetical protein	9.58646	0	9.01553	0	8.23704	7.0868E-11
NCU06235	hypothetical protein	1.72319	0.384466	2.03481	0.194757	1.94538	0.600129
NCU06326	pectate lyase 1	-1.14247	0.999991	-2.22501	0.197805	-2.06037	0.619113
NCU06603	ThiJ/PfpI family protein	2.14474	0.179561	2.78132	0.0233906	1.15967	0.999917
NCU06607	hypothetical protein	-3.40713	0.00398285	-2.58256	0.0800486	-2.25685	0.487428
NCU06824	hypothetical protein	-2.57246	0.230827	-2.90436	0.171267	-2.19423	0.703198

.

Table S5. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU07067	mannosyl-oligosaccharide alpha-1,2-mannosidase	-1.54375	0.915764	-1.65397	0.747266	-2.89434	0.130466
NCU07068	K(+)/H(+) antiporter 1	4.03283	0.00014091	3.12704	0.00749275	2.59913	0.235739
NCU07150	hypothetical protein	1.79769e+308	0.322443	6.85507	0.00533762	1.79769e+308	0.511181
NCU07325	conidiation-specific protein con-10	4.74233	0.00040971	3.55474	0.00040424	0.665492	0.999917
NCU07345	hypothetical protein	2.69344	0.0266526	2.0829	0.162651	-0.419626	0.999917
NCU07375	MFS phosphate transporter	-2.55407	0.0369477	-2.06176	0.191027	0.383517	0.999917
NCU07441	hypothetical protein	-1.40242	0.750296	-1.14652	0.983789	-3.18462	0.0575499
NCU07474	ergot alkaloid biosynthetic protein A	4.03773	0.00023602	3.48268	0.00171795	1.84882	0.660407
NCU07820	pantothenate transporter	13.9191	0	13.7962	1.2281E-13	13.3969	1.0516E-11
NCU07976	hypothetical protein	-4.06487	0.135221	-3.70536	0.150989	-4.09646	0.0431494
NCU08017	hypothetical protein	3.70164	0.0248465	1.74155	0.87016	0.3389	0.999917
NCU08146	hypothetical protein	3.9137	0.00938541	3.9597	0.00533762	3.22488	0.167418
NCU08226	hypothetical protein	2.30841	0.110096	0.707018	0.999999	1.83131	0.747181
NCU08364	choline sulfatase	5.06184	3.2082E-07	4.64558	1.3264E-06	5.61514	6.3614E-06
NCU08439	leptomycin B resistance protein pmd1	-2.45296	0.0959041	-3.24517	0.00444062	-1.38775	0.97497
NCU08603	ankyrin repeat protein	1.42038	0.778228	2.45529	0.0992472	1.95878	0.697007

Table S5. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU08880	neutral amino acid permease	2.55029	0.0364848	2.43292	0.0534018	2.66752	0.168629
NCU08907	BYS1 domain-containing protein	-3.94857	0.00207217	-2.76192	0.0923853	-5.62435	0.0098342 6
NCU09040	oxidoreductase	4.38656	1.0193E-05	4.02777	4.8683E-05	3.16701	0.101552
NCU09094	pre-rRNA-processing protein ipi-1	-2.19031	0.699673	-2.46138	0.347789	0.121546	0.999917
NCU09147	hypothetical protein	5.031	0.0002686 4	5.88215	5.3225E-05	4.61374	0.0035595 5
NCU09181	hypothetical protein	-5.30201	0.0541072	-3.60029	0.190542	-3.86433	0.235739
NCU09230	cystathionine gamma-lyase	3.91109	0.0002477 5	4.91689	5.5094E-07	4.46156	0.0010456 5
NCU09263	hypothetical protein	-4.72898	0.0001232 6	-3.78397	0.0019264 2	-3.86202	0.0578836
NCU09277	hypothetical protein	-2.40563	0.252676	-2.52802	0.142625	-2.65089	0.247938
NCU09352	far upstream element-binding protein 2	0.873922	0.999991	1.14635	0.976624	2.66084	0.167418
NCU09403	NmrA family protein	2.00551	0.190337	1.79327	0.375478	-0.395063	0.999917
NCU09514	hypothetical protein	2.43962	0.0551923	1.99603	0.223244	1.86669	0.631791
NCU09674	O-methyltransferase family 3	-1.57587	0.534893	-1.43899	0.704964	-2.9114	0.160734
NCU09693	hypothetical protein, variant	-1.6237	0.670668	-1.96234	0.32384	-2.29989	0.457187
NCU09698	hypothetical protein	3.7592	0.0446913	4.1044	0.076039	2.14206	0.722206
NCU09734	hypothetical protein	3.34049	0.0144133	3.0104	0.0595998	2.99774	0.168155
NCU09738	alpha-ketoglutarate dependent xanthine	1.51828	0.588517	2.45258	0.0586016	1.8313	0.661311
NCU09767	membrane transporter	2.82446	0.113681	2.87432	0.396256	1.38386	0.999917

Table S5. Continued.

Gene	Broad annotation	DDlog2 (fold change)	DD qvalue	LL15log2 (fold change)	LL15 qvalue	LL60log2 (fold change)	LL60 qvalue
NCU09830	ABC multidrug transporter	0.929437	0.999991	1.97769	0.278449	1.95598	0.595839
NCU09959	hypothetical protein	4.27307	0.00096752	4.86803	0.00017101	4.04114	0.0137586
NCU09968	hypothetical protein	-1.79769e+308	0.30156	-12.2558	0.334024	-9.26201	7.5375E-06
NCU10014	hypothetical protein	-1.73744	0.370839	-1.99527	0.209853	-2.12636	0.437406
NCU10055	opsin-1	2.40101	0.143896	2.31813	0.13549	0.894462	0.999917
NCU09773	oligopeptide transporter	-3.88388	0.00018021	-2.95824	0.0126642	-4.28014	0.00125957

Table S6. List of cell fusion genes identified as direct ADV-1 targets by ChIP-seq. The list also includes if ADV-1 binding sites were present in the dark (LL0) and light-treated samples (LL15, LL30 and LL60). 15 of the cell fusion genes were present in both Vogel's and Bird's media. 2 additional cell fusion genes were identified only in Bird's media.

S. No.	Cell fusion	Gene symbol	Broad annotation	LL 0	LL15	LL30	LL60	Bird's medium	References
1	NCU06419	*mek-1*	MAP kinase kinase	✓	✓	✓	✓	present	Maerz *et. al*.,2008
2	NCU08741	*ham-3*	Striatin Pro11	✓	✓	✓	✓	present	Simonin *et.al.*,2010
3	NCU02110	*nox-1*	NADPH oxidase-1	✓	✓	✓	✓	present	Read *et.al.*,2012
4	NCU02794	*ham-1/so*	Fso1	✓	✓	✓	✓	present	Fleißner *et.al*.,2005
5	NCU01789	*ham-5*	IDC1 protein	✓	✓	✓	✓	present	Aldabbous *et.al*.,2010
6	NCU00881	*ham-7*	Hypothetical protein	✓	✓	✓	✓	present	Fu *et.al.*,2011, Maddi *et.al*.,2012
7	NCU07389	*ham-9*	SAM and PH domain-containing protein	✓	✓	✓	✓	present	Fu *et.al.*,2011, Fu *et. al.*,2014
8	NCU00506	*pkr-1*	ER membrane protein	✗	✗	✓	✗	present	Fu *et.al*.,2011
9	NCU07392	*adv-1*	Transcriptional regulatory protein Pro-1	✓	✓	✓	✓	present	Fu *et.al.*,2011
10	NCU09842	*mak-1*	Mitogen-activated protein kinase MKC1	✓	✓	✓	✓	present	Maerz *et. al*.,2008
11	NCU09337	*prm-1*	Pheromone-regulated membrane protein 1-like	✓	✓	✓	✓	present	Fleißner *et.al*.,2009
12	NCU02393	*mak-2*	Mitogen-activated protein kinase-2	✓	✓	✓	✓	present	Pandey *et.al*.,2004
13	NCU02811	*ham-8*	Hypothetical protein	✓	✓	✓	✓	present	Fu *et.al*.,2011
14	NCU02767	*ham-6*	Pro41	✗	✓	✓	✓	present	Fu *et.al.*,2011, Leeder *et.al.*,2013
15	NCU09307	*lfd-1*	Late fusion defect-1	✓	✓	✓	✓	present	Palma-Guerrero *et. al.*,2014
16	NCU01545	*atg-8*	Autophagy protein-8	✗	✗	✗	✗	present	Fu *et.al.*,2011

Table S6. Continued.

S. No.	Cell fusion	Gene symbol	Broad annotation	LL 0	LL1 5	LL3 0	LL6 0	Bird's medium	References
17	NCU02896	*ada-3*	All development altered-3, TF	✗	✗	✗	✗	✗	Fu *et.al.*,2011
18	NCU09263	*ham-11,acw-4*	Anchored cell wall protein-4	✗	✗	✗	✗	present	Leeder *et.al.*,2013
19	NCU06842	*rcm-1*	Regulator of conidiation and morphology-1	✗	✗	✗	✗	✗	Aldabbous *et.al.*,2010

APPENDIX F

ONLINE SUPPORTING FILES FOR CHAPTER II

A total of thirteen online supporting files are available for Chapter II along with electronic copy of this dissertation.

File S1. CuffDiff analyses of FPKM (fragments per kilobase per million reads) for the combined RNA-seq datasets from two biological replicates. The ratio values given are the log2 ratio values.

File S2. CuffDiff analyses of FPKM for biological replicate 1. The ratio values given are the log2 ratio values.

File S3. CuffDiff analyses of FPKM for biological replicate 2. The ratio values given are the log2 ratio values.

File S4. Predicted genes that are not expressed in the combined biological replicates (FPKM<1) that are not designated as hypothetical proteins. The ratio values given are the log2 ratio values.

File S5. FunCat analyses for predicted genes that are highly expressed in the dark in the combined biological replicates (FPKM>400 and FPKM>1000).

File S6. CuffDiff analyses for the subset of transcripts determined to be regulated 2-fold in response to light. The ratio values given are the log2 ratio values.

File S7. FunCat analysis of the 999 mRNAs most up-regulated in response to light.

File S8. FunCat analysis of the 999 mRNAs most down-regulated in response to light.

File S9. CuffDiff analyses for the subsets of transcripts determined to respond to light with q≤0.2, 0.1 or 0.05. The ratio values given are the log2 ratio values.

File S10. Gene Ontology (GO) analyses of genes that were 2-fold regulated by light and genes whose light-regulation met the q£0.2 stringency requirement.

File S11. Funcat and GO enrichment analyses of the 5 major clusters delineated in Figure 2-3A.

File S12. RNA-seq data (extracted from File S1) for the 27 TFs analyzed as WCC targets in (Smith et al. 2010). The ratio values given are the log2 ratio values.

File S13. FunCat analysis of mRNAs whose levels change more than 16-fold in response to light (Table 2-1).

APPENDIX G

ONLINE SUPPORTING FILES FOR CHAPTER III

One online supporting file is available for Chapter III along with electronic copy of this dissertation.

File S1. All regions of WCC enrichment at $p<0.001$ and z score >3.09

APPENDIX H

ONLINE SUPPORTING FILES FOR CHAPTER IV

A total of four online supporting files are available for Chapter IV along with electronic copy of this dissertation.

File S1. ADV-1 ChIP-Seq data. Regions of the genome significantly enriched ($p< 0.01$) for ADV-1-binding are reported. The file contains worksheets for peaks identified by ADV-1 ChIP-seq from cells grown in the dark (LL0) and after a light pulse (LL15, LL30 and LL60). The gene designations (NCU#), gene symbol, annotation, and contig for called genes are given in columns B, C, D, and E. The location of the binding site is given in columns F, G, and H. The peak coordinates (reg_summit) (column H) and ranks based on coverage over the mean (cvg_over_mean) (column I), and the position and distance of the peaks relative to the gene are shown in columns J and K. The unique genes worksheet gives the 1187 unique predicted ADV-1 target genes that had statistically significant binding sites at one or more time points. Note that in the first column "S. No.", some of the numbers are repeated. This indicates that more than one ADV-1-bound peak was assigned to that gene.

File S2. Gene Ontology (GO) analyses of direct targets identified by ADV-1 ChIP-seq. LL0 contains enriched GO terms with $p\leq0.05$ for ADV-1 direct targets identified by

ChIP-seq at time point 0. Similarly, LL15, LL30 and LL60 contain enriched GO terms with p≤0.05 at time points 15, 30 and 60 min after light treatment, respectively.

File S3. Transcripts detected by RNA-Seq of total RNA from WT and ΔADV-1 strains. The file contains worksheets for differential expression results comparing ΔADV-1 cells to WT cells. The all genes worksheet contains the cuffdiff outputs for all loci. The unique 2X & q≤0.2 folder contains all genes called differentially expressed with ≥2-fold changes, and q-values of ≤0.2 in ΔADV-1 cells compared to WT cells at one or more time points (LL0 ,LL15 and LL60). The differentially expressed genes from each condition that had a ≥2-fold change and q-values of ≤0, or that had a ≥2-fold change are in separate worksheets. The (log2) fold changes for each gene are presented, where fold change is the difference in the normalized reads (FPKM) of a gene between WT and ΔADV-1 cells. A positive (log2) fold change values represents increased mRNA levels in Δ*adv-1* compared to WT cells (negative regulation). A negative (log2) fold change value represents decreased mRNA levels in Δadv-1 compared to WT cells (positive regulation). The gene designation (NCU#), annotation, gene locus, FPKM values, and statistical data, are given in the all genes worksheet for reference.

File S4. Comparison of ADV-1 downstream targets identified by ChIP- and RNA-seq to all known light- and clock- regulated genes.

APPENDIX I

Genome wide mapping of TFs that are direct WCC targets

A suite of 24 TFs were identified as direct WCC targets based on WC-2 ChIP-seq after a brief light pulse. Most of these TF were light- and clock-regulated (Chapter III). To understand the clock regulatory network, it was important to identify the downstream targets of these TFs. The data generated were used to build the network map of the first-tier TFs and their targets genes. The TFs were epitope-tagged with V5, and their protein levels were checked by Western blot in tissues harvested in dark and after light treatment (Figure A-1). After confirming that the protein levels could be detected, ChIP was performed using ChIP-grade anti-V5 antibody, followed by high-throughput sequencing as described in Chapter III and Chapter IV.

Circadian clock and sexual development in fungi

There have been reports of rhythms in sexual spore production and release in fungi (Bell-Pedersen et al., 1996b). Furthermore, it was shown that the circadian clock regulates mating behavior by controlling the rhythmic expression of pheromone genes in *N. crassa* (Bobrowicz et al., 2002). ADV-1 knockout cells do not form protoperithecia that are required to form the reproductive structures, perithecia, suggesting a role for a

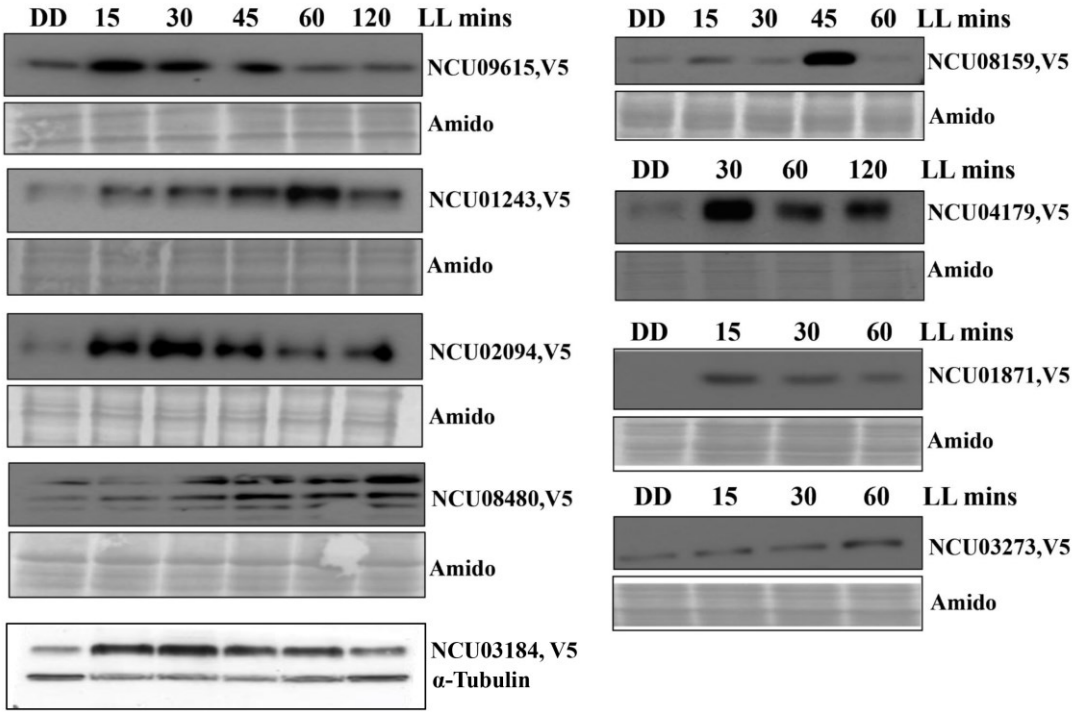


Figure A-1. Validation of the V5-tagged TFs used for ChIP-seq. Western blots of V5-tagged TFs, which are direct WCC targets, probed with anti-V5 antibody in the dark (DD), and after exposure to light (LL) for the indicated times in min. The total protein stained by amido is shown as loading controls, or probed with antibody to tubulin.

clock-controlled TF in sexual development (Colot et al., 2006). To further investigate the link between the clock and sexual development, I carried out assays to determine if perithecia are produced in a circadian manner. Conidia from one mating type (either *mat a* or mat A) were inoculated on a nitrogen-limiting medium, which favors sexual development, synchronized in light for 24 h and then grown in constant dark at 25°C for 5-7 days to initiate protoperithecia formation. Using a red safety light, conidial suspension from the opposite mating type (either *mat A* or *mat a*) were inoculated, and incubated for another 5-7 days allowing the protoperithecia to mature into perithecia (Figure A-2). The results from these experiments were not conclusive as to whether perithecia formation is regulated by the circadian clock. There were dark patches (perithecia), alternating with clear zones (undifferentiated hyphae) resembling circadian pattern, however these results were not reproducible.

Phosphorylation of ADV-1

Upon light treatment, ADV-1 might get phosphorylated, which can then activate ADV-1 and regulate expression of light-responsive target genes. To test this possibility, I extracted proteins from tissues harvested in dark and after light treatment for 30 and 120 mins, treated with phosphatase and then performed western blot to detect ADV-1,V5 protein. The untreated and mock treatment samples showed a slower mobility band, which shifted downward upon phosphatase treatment, suggesting the possibility of ADV-1 undergoing phosphorylation (Figure A-3). Furthermore, the slower mobility

Figure A-2. Assays for circadian rhythms in perithecia formation. Wild type *N. crassa* of opposite mating types grown under limiting nitrogen conditions in constant darkness (DD), to initiate female sexual development. The black patches are fertilized perithecia, which alternate with vegetative hyphae.

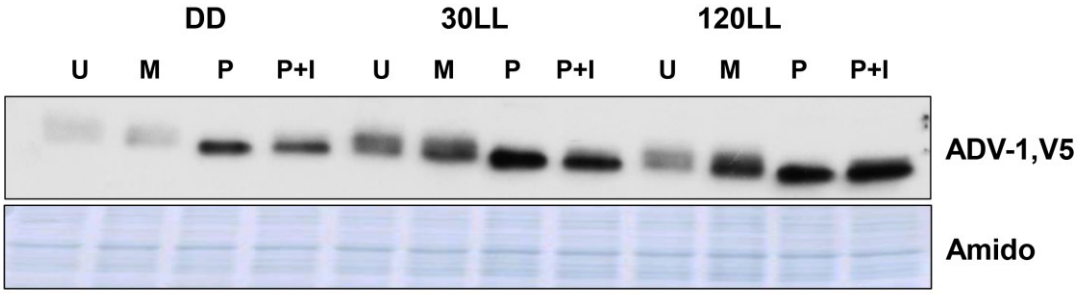


Figure A-3. Phosphorylation of ADV-1. Protein was extracted from ADV-1::V5 cells harvested in the dark or after light treatment for 30 and 120 mins. Western blot of protein from ADV-1::V5 cells that were untreated (U), mock treated (M), treated with λ-phosphatase (P), or treated with λ-phosphatase plus phosphatase inhibitors (P+I); and probed with anti-V5 antibody. The total protein stained by amido is shown as loading controls.

band was observed in both dark and light treated samples, and only differences in overall levels are observed between dark and light treated samples. As a further confirmation of phosphorylation events, protein samples were incubated with phosphatase inhibitors prior to phosphatase treatment, which did not fully restore the upper band. This can be due to inefficiency of the phosphatase inhibitors. In any case, the experiment needs to be repeated to confirm ADV-1 phosphorylation.

Circadian clock and cell cycle

A predicted cyclin (NCU04847) was identified as a direct target of ADV-1 by ChIP-seq (Figure A-4). ADV-1-binding to the promoter of NCU04847 after a light pulse was validated by an independent ChIP-PCR (Figure 4-6). Disruption of the clocks has been linked to numerous pathological disorders including cancer, suggesting the link between circadian clock and cell cycle progression. Since ADV-1 is a clock-controlled first-tier TF, I further characterized the role of ADV-1 in the regulation of the predicted cyclin. Preliminary studies indicated that NCU04847 was light-induced and that the light induction was dependent on ADV-1. NCU04847 mRNA levels were also shown to be rhythmic, however, the rhythms persisted in ΔADV-1 cells. It would be interesting to determine if NCU04847 is targeted by other first-tier TFs, and if there are other TFs that bind to the promoters of NCU04847, that might explain the results.

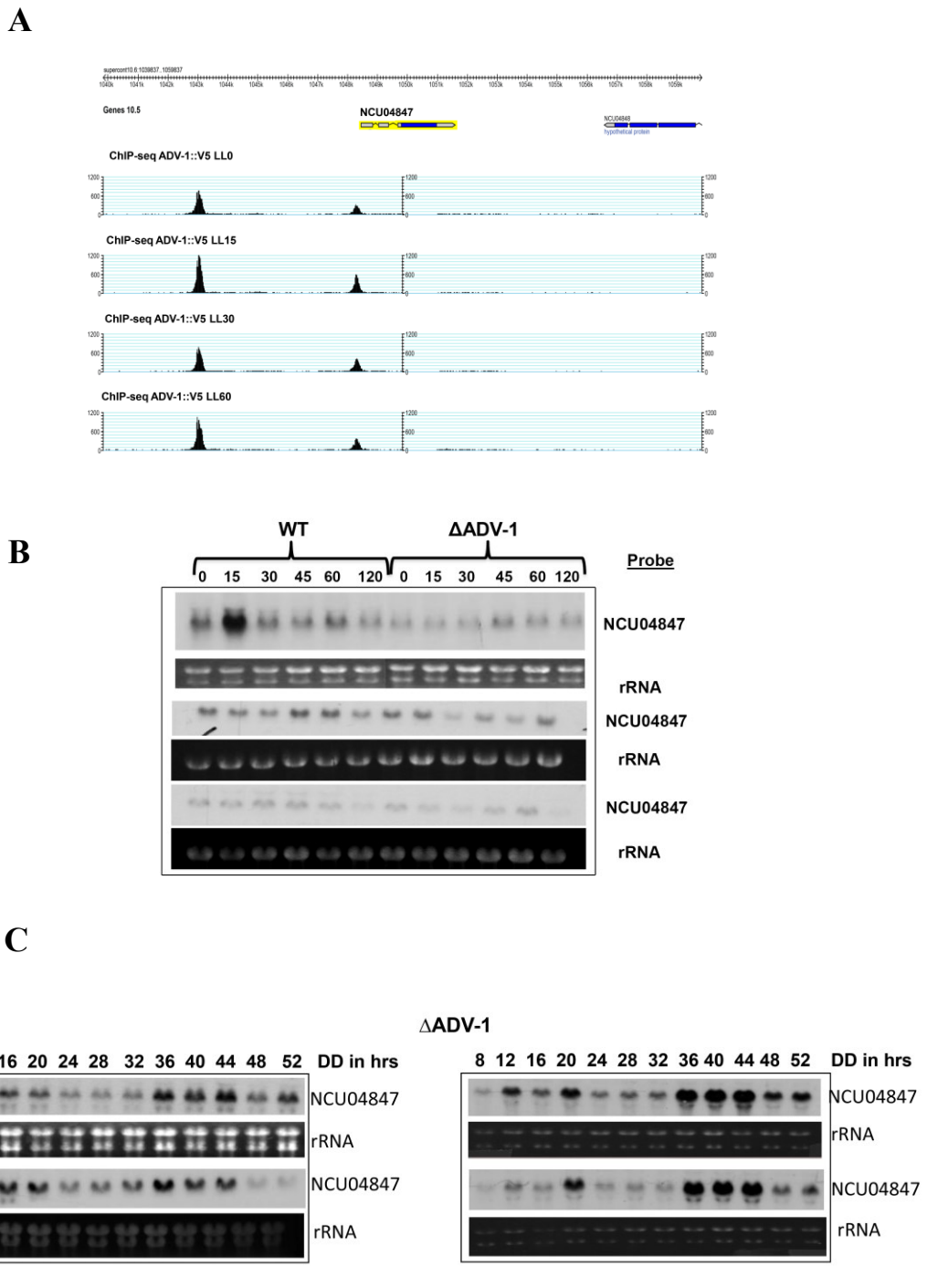


Figure A-4. (A) ChIP-seq reveals NCU04847 (predicted cyclin) as an ADV-1 target. Two ADV-1-binding sites are located upstream of the NCU04847 transcriptional start site, and both peaks are enriched after light treatment. (B) Northern blots of NCU04847 mRNA levels from WT and ΔADV-1 cells in the dark (0) and after light treatment for 15, 30, 45, 60 and 120 mins. (C) Northern blots of NCU04847 mRNA levels isolated from WT and ΔADV-1 cells harvested after the indicated times in the dark (DD). rRNA is shown as a loading control in B and C.

Circadian clock and ROS homeostasis

N. crassa displays rhythmic ROS (Reactive oxygen species) production in constant darkness (Yoshida et al., 2011). Enzymes with opposite functions, CAT-1 (Catalase-1) and NOX-1 (NADPH oxidase), involved in scavenging and generating ROS respectively are likely to be responsible for oscillations observed in cellular ROS levels. Indeed, CAT-1 activity was found to display a low-amplitude rhythm. It has been shown that WCC regulates the expression of *cat-1* via rhythmic activation of the OS-2 MAPK pathway (Lamb et al., 2012). NOX-1 is a major source of ROS and NOX-1 activity is required for rhythmic ROS accumulation (Yoshida et al., 2011). However, the link between the circadian clock and NOX-1 is not known. ADV-1 ChIP-seq revealed that ADV-1 binds to the promoters of NOX-1 and NOR-1 (NADP oxidase regulator-1) (Figure A-5). To understand the role of ADV-1 in regulating ROS oscillation, I worked to establish a method to detect ROS levels in *N. crassa* cells. One method to detect ROS uses the NBT assay, where nitro blue tetrazolium (NBT) is reduced by superoxides into the reaction product, formazan. Formazan is water insoluble, and therefore precipitates in solution. I examined ROS levels using the NBT assay in WT, ΔADV-1, ΔNOX-1 and ΔNOR-1 (heterokaryon) strains under standard laboratory conditions (constant light at 30°C) (Figure A-5). As expected, ROS levels indicated by the black precipitate were higher in WT compared to ΔNOX-1 and ΔNOR-1 (het), whereas no change in ROS levels was observed in WT versus ΔADV-1 cells. NBT assays are semi-quantitative,

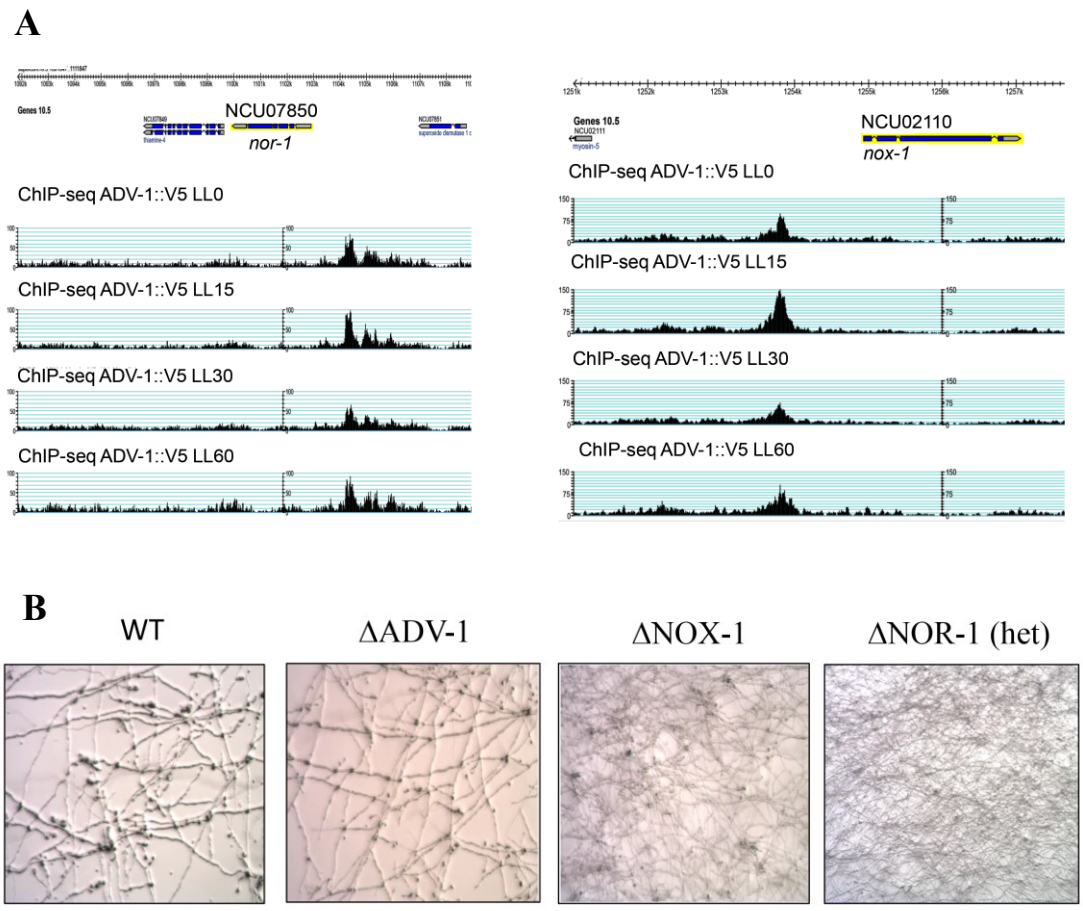


Figure A-5. (A) ChIP-seq reveals *nor-1* and *nox-1* as an ADV-1 target. ADV-1-binding sites are located upstream of the *nor-1* and *nox-1* transcriptional start site in the dark (LL0) and after light treatment as indicated for 15, 30 and 60 mins. (B) Nitro blue tetrazolium (NBT) to formazan (black precipitate) by superoxide anion (O_2^-) was used to measure ROS levels in WT, ΔADV-1, ΔNOX-1 and ΔNOR-1 (het) strains.

therefore, in the future, it would be best to use a more sensitive and quantitative method to measure ROS levels.

Do first-tier TF feedback to regulate wc-1 mRNA levels?

To determine if any of the first-tier TFs feedback to regulate *wc-1* expression, we compared the light induction of *wc-1* mRNA between WT and the TF knockouts. No significant differences in light induction of *wc-1* mRNA were observed in ΔADV-1, ΔVOS-1, ΔSAH-1 and ΔNCU00275 cells (Figure A-6). In ΔCSP-1 cells, *wc-1* mRNA levels were light-induced, but consistent with its role as a repressor, *wc-1* mRNA levels were comparably higher in ΔCSP-1 cells than in WT cells. In addition, in ΔCSP-1 cells, *wc-1* mRNA levels increased with long exposures to light. These data suggested that CSP-1 alters photoadaptation (Chen et al., 2010a).

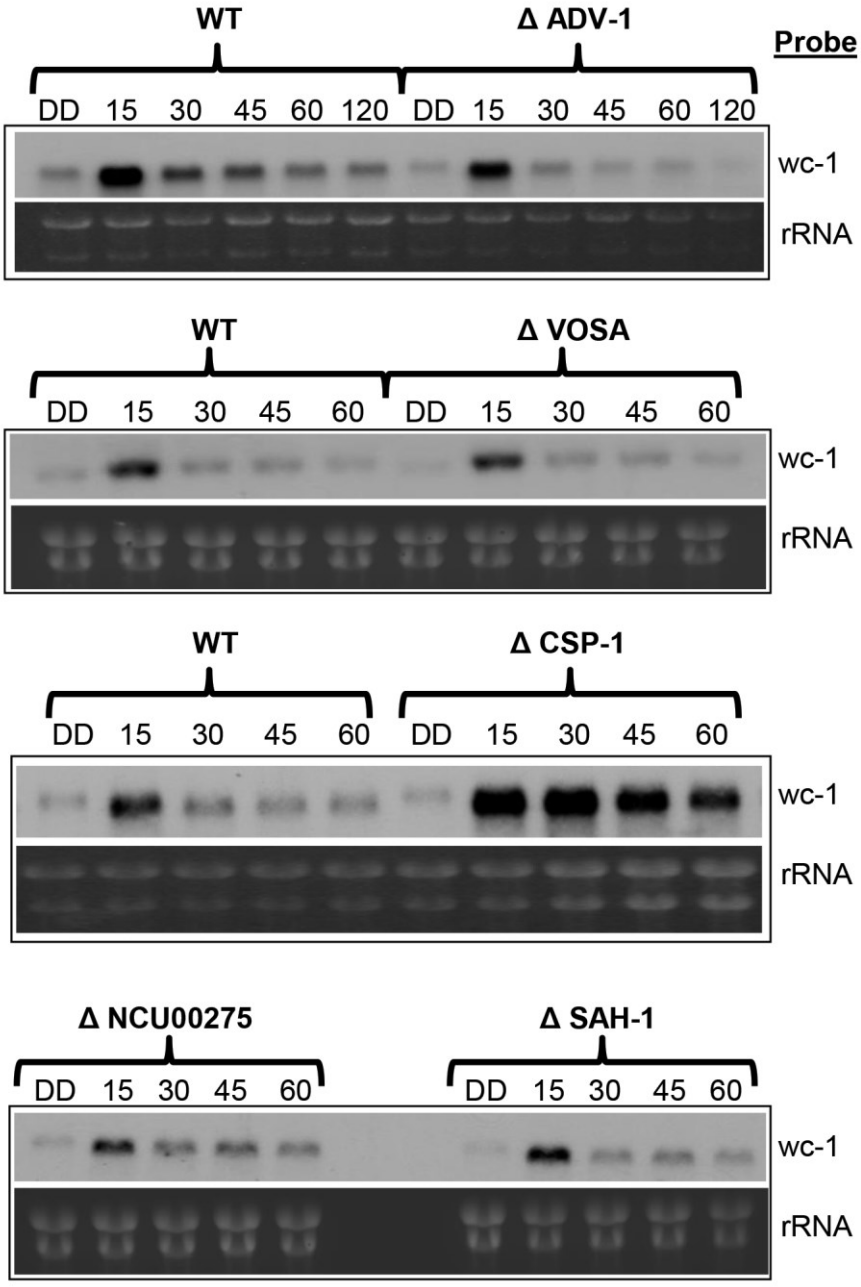


Figure A-6. *wc-1* mRNA levels in WT and TF knockout cells in the dark and after light treatment. Northern blots of *wc-1* mRNA levels isolated from WT and the indicated TF KOs in the dark (DD) and after the indicated light treatments (min). rRNA is shown as loading controls.