VERIFYING TRANSCRIPTION FACTOR REGULATION OF ADV-1
PROTEIN RHYTHMS THROUGH COMPLEMENTATION

An Undergraduate Research Scholars Thesis

by

SARAH WESTLAKE

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Deborah Bell-Pedersen

May 2018

Major: Biology
# TABLE OF CONTENTS

## ABSTRACT

---

## ACKNOWLEDGMENTS

---

## NOMENCLATURE

---

## CHAPTER

### I. INTRODUCTION

---

#### Background

---

#### Preliminary Data

---

### II. METHODS

---

#### N. crassa Strains and Culture

---

#### Generating Constructs by PCR

---

#### Luciferase Assay

---

#### Data Analysis

---

### III. RESULTS

---

### IV. CONCLUSION

---

## REFERENCES

---

## APPENDIX

---
ABSTRACT

Verifying Transcription Factor Regulation of ADV-1 Protein Rhythms Through Complementation

Sarah Westlake
Department of Biology
Texas A&M University

Research Advisor: Dr. Deborah Bell-Pedersen
Department of Biology
Texas A&M University

Roughly half of the eukaryotic genome is under the control of a circadian clock at the level of transcript abundance, with rhythmic transcripts peaking at all possible phases of the day. However, the mechanism for phase regulation is not well understood. To determine how phase is regulated in *Neurospora crassa*, ChIP-seq was performed on the morning active transcription factor (TF) and core oscillator component, the White Collar Complex (WCC). The direct targets of WCC were found to be enriched for TFs. Further investigation of these first-tiered TFs revealed that they form a large and complex network made up of multiple feedback and feedforward loops. We hypothesized that this intricate TF network plays a role in phase regulation. To test this hypothesis, we focused our experiments on the network surrounding the first-tier TF, ADV-1. In preliminary data, we found that TF knockouts of CSP-1, CSP-2, SUB-1, and ZNF-21, altered the phase and/or period of ADV-1 protein rhythms. To confirm that these 4 TF knockouts, and not second site mutations, cause the changes in ADV-1 rhythms, genetic complementation will be performed. If the TF knockout causes the defect, then supplying the wild type gene back should rescue the ADV-1 rhythm back to wild type. This information will help us to validate which TFs in the network are necessary for generating proper phase control of ADV-1 protein rhythms.
ACKNOWLEDGEMENTS

I would like to thank my faculty advisor, Dr. Deborah Bell-Pedersen, and my graduate mentor, Jennifer Jung, for their guidance and support throughout the course of this research.

Thanks also go to my family, friends, colleagues, and the department faculty and staff for making my time at Texas A&M University a great experience.
CHAPTER I
INTRODUCTION

Background

Most eukaryotes have a molecular circadian clock that regulates daily rhythmic processes, including rhythmic gene expression, activity, and physiology (Pittendrigh, 1954, 1964). As a result of the Earth’s rotation on its axis, the environment has daily rhythms in light, temperature, and nutrient availability. As a result, most eukaryotes evolved a circadian clock that allowed anticipation of daily environmental cues (Baker, Loros, & Dunlap, 2012). The circadian clock free runs with an endogenous period of close to 24 h in constant environmental conditions and can be synchronized to precisely 24 h by environmental signals such as light (Baker et al., 2012; Pittendrigh, 1954). In humans, mutations in the clock or living against the clock, whereby the clock is out of phase with environmental cycles (e.g. shift-workers), increases the risk of developing diseases such as obesity, diabetes mellitus, cardiovascular disease, thrombosis, and inflammation (Maury, Ramsey, & Bass, 2010).

The circadian clock mechanism is comprised of 3 components: an oscillator that keeps time, input pathways that synchronize the oscillator, and output pathways from the oscillator that signal time of day information to control overt rhythms. The conserved eukaryotic oscillator mechanism consists of a negative feedback loop (Hardin, 2011). Because the clock mechanism is evolutionarily conserved, a simple eukaryotic model organism, like the filamentous fungus Neurospora crassa, can be a powerful tool for understanding the mechanisms of the clock.

N. crassa has historically been used to study complex genetic networks. It’s easy to manipulate in the lab due to its haploid genetic structure and rapid reproduction rate (Selker,
N. crassa grows by tube like structures called hyphae, that can that form interconnected networks called mycelia. Aerial hyphae made during asexual development, produce multinucleated asexual spores named conidia, which are used as an inoculum for many experiments. Additionally, N. crassa has been used as a genetic model for many decades {Beadle, 1941 #51; Davis, 2002 #52}, and thus, has many tools available for genetic modification, transformation, and mating.

In addition to its history as a genetic model, N. crassa was one of the first model systems used to study the circadian clock because it has a robust circadian rhythm in conidiation that can be easily monitored using the race tube assay (Kramer, 2007). Race tubes are long glass tubes that are bent upwards on both sides. The race tube is filled with minimal media and inoculated at one end with conidia. After synchronizing the clock by allowing the conidia to germinate for a day in natural light, the race tube is transferred to total darkness for about a week. Once every 22.5 hours conidia are formed from aerial hyphae, under control of the clock. The growth front is marked at the same time each day using a red safety light that does not affect the clock. Period is determined by the distance between the bands of conidia, and phase of the rhythm is measured using the light to dark transition as the phase reference point. However, measuring developmental rhythms is a proxy for assaying the core oscillator. To directly assay the core oscillatory or other specific genes, we can use a luciferase reporter {Gooch et al., 2014 {Morgan, 2003 #53}. This is done by fusing a codon-optimized luciferase sequence to the coding sequence of a gene of interest. In any case, the race tube assay provided an important tool for identifying mutations that altered the clock mechanism {Feldman, 1973 #54}, leading to the discovery of core components of the circadian oscillator.
The positive arm of the core oscillator in *N. crassa* is the morning-active White-Collar Complex (WCC), composed of WC-1 and WC-2. The negative arm contains FREQUENCY (FRQ) and FRQ-interacting RNA helicase (FRH) (Baker et al., 2012; Froehlich, Liu, Loros, & Dunlap, 2002; He et al., 2002). In the subjective morning, the blue-light photoreceptor WC-1 heterodimerizes with WC-2 to form the WCC (He & Liu, 2005b). The WCC activates the transcription of *frq* and other clock-controlled genes to influence daily rhythms in gene expression (Hurley et al., 2014). After FRQ protein is translated, it homodimerizes and forms a complex with FRQ-interacting helicase (FRH) (Cheng, He, He, Wang, & Liu, 2005). This FRQ-FRH complex (FFC) then translocates into the nucleus to recruit kinases to phosphorylate and inhibit the activity of the WCC (Larrondo, Olivares-Yanez, Baker, Loros, & Dunlap, 2015). Throughout the day, FRQ is phosphorylated and degraded, allowing the WCC to become active again the next morning (He & Liu, 2005a; Schafmeier et al., 2005; Schafmeier, Kaldi, Diernfellner, Mohr, & Brunner, 2006).

Importantly, almost half the eukaryotic genome is regulated at the transcriptional level by the clock with transcripts peaking at all possible times of the day (Hurley et al., 2014). However, how phase is differentially regulated is unknown. Understanding the mechanism controlling phase regulation of rhythmic gene expression has the potential to provide new therapies for treating circadian misalignment disorders including shift-work associated metabolic disease.

**Preliminary Data**

To study how the *N. crassa* oscillator regulates phase, genome-wide binding sites of WCC were identified using ChIP-seq (Smith et al., 2010). Among the approximately 200 target sites, an enrichment of targets encoding TFs was observed, including the TF ADV-1. Amidst the 24 TFs bound and potentially regulated by WCC, only the single deletion of ADV-1 abolished
overt rhythms in development and altered the expression of genes involved in metabolism (Dekhang et al., 2017). Therefore, we focused our attention on ADV-1 to determine how the clock controls the phase of ADV-1 rhythms and how rhythms in ADV-1 affect the phase of downstream clock outputs.

From ChIP-seq data of the TFs regulated by the WCC, we discovered that 13 TFs form a complicated network involving feedforward and feedback loops that have the potential to regulate ADV-1 expression (Figure 1). We hypothesized that this TF network plays a role in regulating the phase of transcriptional clock output. To test this idea, we initially measured the effect of single gene deletions in each of the 13 TFs in the network on ADV-1 protein rhythms. This was accomplished using strains that contained a luciferase reporter fused to the ADV-1 coding sequence, called a translational fusion (ADV-1::LUC). We found 4 TFs in the network, CSP-1, CSP-2, SUB-1, and ZNF-21, that may be required for proper rhythmic expression and phase of ADV-1 protein rhythms. In the deletions of CSP-1 and CSP-2, we observed a phase delay of about 4 and 2 hours, respectively, in ADV-1 protein rhythms. The deletion of SUB-1 shows a 1.4 hour decrease in the period of ADV-1::LUC protein rhythms while the absence of ZNF-21 shows an hour period increase (Figure 2). However, these conclusions assume that no secondary mutations that could affect ADV-1 protein rhythms occurred. To validate the knockout (KO) effects on ADV-1 protein rhythms, I will generate complementation strains that contain both the TF KO and a wild type copy of the TF gene and assay ADV-1::LUC rhythms. Once I verify that these TFs are necessary for ADV-1 phase control, I can further investigate if they alter other components in the network. The results of this project will be used to help generate a computational model that predicts the role of specific TFs, and the network as a whole, on phase.
Figure 1. TF network including ADV-1 protein. Each node represents a TF and each line represents ChIP-seq binding.
Figure 2. Representation of $\Delta csp-1$, $\Delta csp-2$, $\Delta znf-21$, and $\Delta sub-1$ on phase and period. The WT copy is represented with a black line while the KO is represented with the grey line. $\Delta csp-1$ shows about a 4 ± 0.3-hour phase delay (n=21). $\Delta csp-2$ shows about a 2 ± 0.5-hour phase delay (n=25). $\Delta znf-21$ shows about 1 ± 0.1-hour increase in period (n=7). $\Delta sub-1$ shows 1.4±0.2-hour decrease in period (n=9).
CHAPTER II
METHODS

N. crassa Strains and Culturing

*N. crassa* strains were purchased from The Fungal Genetics Stock Center (FGSC, Kansas State University, Manhattan, KS). ADV-1::LUC reporter strains were cultured on a minimal medium that contains 1x Vogels, 2% glucose, and 1.5% agar and grown for a week at 30° C and constant light (LL).

Generating Constructs by PCR

Oligonucleotide primers (Table 1) were designed to amplify the ORF and presumed complete control regions of each TF based on the known ChIP-seq binding sites within the network, as well as the flanking *csr-I* regions as diagrammed (Figure 3). CSP-1 and CSP-2 were amplified in three individual fragments, while ZNF-21 was amplified in four different fragments. First round PCR reactions generated the individual fragments 1-3 or 1-4, for each TF. PCR was carried out with Phusion Taq polymerase (ThermoFisher Scientific) and oligonucleotides generated from Integrated DNA Technologies (IDT) (Table 1). The thermal cycle was programmed for 2 minutes at 98° C to denature, 15 seconds at 98° C for additional denaturation, followed by 45 seconds at 61° C for annealing, 4 minutes at 72° C for elongation, followed by 29 cycles beginning at 15 seconds at 98° C, and then final extension at 72° C for 10 minutes. The PCR product for the individual fragments were then checked for size and amount on a 1% agarose gel in 1 X TAE buffer with a known concentration of a 1kb DNA ladder (ThermoFisher Scientific). The correct PCR product was gel purified using QIAquick (Quiagen, Valencia, CA). Fusion PCR (or assembly PCR) was then used to generate a fragment containing the 5’ flank of *csr-I*, the 5’ TF flank, the TF ORF, the 3’ TF flank, and the 3’ *csr-I* flank (Figure
3). This complete fragment targets recombination of the complementation construct to the *csr-1* genomic region during transformation, and in the process causes a deletion of the *csr-1* coding region, conferring cyclosporin A resistance to positive transformants (Bardiya & Shiu, 2007).

**Figure 3.** Example of a basic construct design that was used for CSP-1, CSP-2, ZNF-21, and SUB-1 with CSR-1 flanks on 5 and 3 prime ends. CSP-1 and CSP-2 had 3 fragments while ZNF-21 had 4 fragments.
<table>
<thead>
<tr>
<th>Fragment</th>
<th>Sequence</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR-1</td>
<td>Forward 1 (F1)</td>
<td>cgctggagaagctcattcca</td>
</tr>
<tr>
<td>CSP-1</td>
<td>Reverse 1 (R1)</td>
<td>atccggetctctgagttgacacctcaatgtccatctgatc</td>
</tr>
<tr>
<td></td>
<td>Forward 2 (F2)</td>
<td>gatcagatggacatggtgtaactcaagaggccgat</td>
</tr>
<tr>
<td></td>
<td>Reverse 2 (R2)</td>
<td>ccaatgacctgcacaactgtcagagactgacagtagagcagcca</td>
</tr>
<tr>
<td></td>
<td>Forward 3 (F3)</td>
<td>tggcgtttactgtcctgatcagttgacagttgcagctagt</td>
</tr>
<tr>
<td>CSP-2</td>
<td>R1</td>
<td>atataacatctctctgcagttgcagttgcagaagagcgagagatata</td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>gatcagatggacatggtggaagcagtgtagcagacgaagagttgtaata</td>
</tr>
<tr>
<td></td>
<td>R2</td>
<td>ccaatgacctgcacaactgtcagagactgacagtagagcagcca</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>cgagtaaaggtcaacacagaagactgacagtagagcagtcagtg</td>
</tr>
<tr>
<td>ZNF-21</td>
<td>R1</td>
<td>acatcgctacaagtccgtctcactcaatgtccatctgatc</td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>Gattcaagatggacatggtggaagcagtgtagcagcagagat</td>
</tr>
<tr>
<td></td>
<td>R2</td>
<td>Cacatgacctgcacaactgtcagagactgacagtagagcagtcagtg</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>agtaacagtggcacattttggacagttgcaggtcatgtaa</td>
</tr>
<tr>
<td>CSR-1</td>
<td>Reverse-final</td>
<td>caaccgagacacacaaactac</td>
</tr>
</tbody>
</table>

Table 1. Primer sequences that were used for each TF complementation construct.

Once the full constructs were amplified, the product was gel purified using QIAquick (Quiagen, Valencia, CA). The constructs were then cloned into pCR™ blunt vector and transformed into competent DH5a *E. coli* cells to obtain a large quantity of DNA. The constructs were then transformed into *N. crassa* via electroporation as described (Potter & Heller, 2010).

**Luciferase Assay**

A luciferase assay was used to determine the rhythmicity of ADV-1::LUC protein in strains containing the complementing CSP-1, CSP-2, SUB-1, or ZNF-21 genes. 96-well
Optiplates were filled with luciferase media containing 1x Vogel’s salts, 0.05mg/uL biotin, 0.03% glucose, 0.05% arginine, 1.9% Quinoic Acid, 1.5% BactoAgar, and 25µM of 10mM Luciferin.

To obtain cell samples to inoculate on the luciferase plates, dilutions of conidial suspensions were read for optical density, and normalized to a concentration of \(1 = 5\times10^6\) conidia/mL. Each well of the luciferase plate was then inoculated with 5ul of the suspension. A PerkinElmer Envision read the luciferase activity for 1 week before the data was analyzed.

**Data Analysis**

The PerkinElmer Envision outputs luciferase activity into a ‘.csv’ file, with luciferase counts for each timepoint, in each strain tested. The output counts file is reformatted into an excel workbook, as input to BioDare2 (https://biodare2.ed.ac.uk/welcome). BioDare2 implements a Fourier Transform Non-Linear Least Squares (FFT NLLS) algorithm, which works by fitting different harmonics of cosine waves to the luciferase traces. This outputs values and standard deviations in period, phase, and amplitude, which allows us to compare WT ADV-1::LUC traces to a KO.
CHAPTER III

RESULTS

I generated constructs using PCR that contained the WT copy of the desired gene, including all known binding sites of TFs within the network. I successfully obtained each individual fragment for all of the complementation constructs and purified them via gel purification (Figure 4). Using assembly PCR, the CSP-1 complementation construct was generated (Figure 5). Unfortunately, the other two full constructs only resulted in non-specific bands (not shown), when trying to fuse any combination of individual fragments. However, partial fusion of the fragments for CSP-2 and ZNF-21 could be made (Figure 6&7), with very low yield.

Once the full complementation construct is made for each TF, they will be transformed back into the deletion strain with an ADV-1 luciferase protein reporter. The KO of CSP-1 and CSP-2 has an observable defect in conidial separation (Brody & Martins, 1979). If the complementation is successful, then the Δcsp-1 or Δcsp-2 phenotype should be rescued. However, the function of ZNF-21 is not well characterized, and the KO has no obvious growth or conidial phenotype. In this case, to test success of ZNF-21 complementation, I will test the proper integration of the complementation construct by PCR. Data will be analyzed from the luciferase assay using BioDare2, as described above.
Figure 4. Gel photo showing individual fragments of CSP-2, CSP-1, and ZNF-21 obtained by PCR and after gel purification. Sizes of each fragment are represented in Table 1. The numbers above each lane represent the individual fragment for each TF. The box is highlighting the correct sized band for each fragment.

Figure 5. Gel photos showing CSP-1 partial fused fragments and the full construct by PCR before purification. (7.2 Kb for CSP-1 1+2, 7.4 Kb for CSP-1 2+3, and 8.4 Kb for CSP-1 full). Each box represents the correct sized band for each fragment.
Figure 6. Gel photo showing CSP-2 partial fused fragments by PCR before purification. (9.5 Kb for CSP-2 1+2 and 9.7 Kb for CSP-2 2+3). Each box represents the correct sized band for each partial fragment.

Figure 7. Gel photo of ZNF-21 partial fused fragments by PCR. (4.9 Kb for ZNF-21 1+2, 7.4 Kb for ZNF-21 2+3, and 4.7 Kb for ZNF-21 3+4). Each band highlighted represents the correct size.
I currently have the full construct made for CSP-1 and have made partial fragments for ZNF-21 and CSP-2. The CSP-1 full fragment is currently being transformed into the $\Delta csp-1$, ADV-1::LUC strain. Once the ADV-1::LUC complemented strains have been made and assayed for $\Delta csp-1$ and $\Delta csp-2$, I plan to proceed with ZNF-21 and SUB-1, using a similar method. The complementation constructs for CSP-2, ZNF-21, and SUB-1 have not been completed yet because the constructs were relatively large and technically difficult to assemble by PCR. To optimize PCR conditions for larger fragments, I designed new primers, added 3% dimethyl sulfoxide (DMSO) to aid in denaturing during PCR, and increased elongation times during PCR. Unfortunately, none of these methods produced positive results. However, I plan to take the partially fused fragments (fragments 1 and 2, fused and fragments 2+3, fused) and attempt to generate the full construct. Once these are generated I will begin to transform them into their appropriate KO; ADV-1::LUC strain. Luciferase assays will be performed and analyzed as explained above.

Finally, if the TF is the cause for altered phase of ADV-1 protein expression, then the complementation strain will produce ADV-1::LUC rhythms that match that of WT. However, if there is a secondary mutation that is causing the phase difference, then after supplying the complementation strains, the ADV-1::LUC rhythms will not match that of WT.

The goal for this project is to understand how the TF network is regulating phase, and to do so we need to understand the interactions between the TFs and ADV-1. To accomplish this, we are collaborating with computational and mathematical biologists to generate a model that
predicts the role of each TF, and the network as a whole, in phase regulation. Data from this thesis will verify the role of CSP-1, CSP-2, and ZNF-21 on ADV-1 rhythmicity and will be incorporated into the model.
REFERENCES


