UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE ROLES AND MECHANISTIC FUNCTIONS OF NUCLEAR FACTOR Y TRANSCRIPTION FACTORS IN PLANTS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ZACHARY A. MYERS
Norman, Oklahoma
2019

THE ROLES AND MECHANISTIC FUNCTIONS OF NUCLEAR FACTOR Y TRANSCRIPTION FACTORS IN PLANTS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF MICROBIOLOGY AND PLANT BIOLOGY

BY

Dr. Anne Dunn, Chair

Dr. Ben F. Holt III, co-Chair

Dr. John P. Masly

Dr. Mariëlle Hoefnagels

Dr. Laura Bartley

Dedication

This dissertation is dedicated to my wife, Rachel, and my son, Charles.
Thank you for making the journey as exciting as the destination.

Acknowledgements

I have enjoyed constant support and encouragement from so many people during my time at OU. I would like to thank my friend and advisor, Dr. Ben Holt, for the opportunity to work with in his lab. I am so grateful for his consistent mentoring and advice, and for always leading by example. Members of my committee, Dr. Anne Dunn, Dr. Mariëlle Hoefnagels, Dr. Laura Bartley, and Dr. JP Masly, have been completely supportive of me and my development, and I am very thankful for their encouragement. Past mentors and committee members, Dr. Sharon Kessler, Dr. Ben Smith, Dr. David Durica, and Dr. Marc Libault, were equally supportive and important for my success as a graduate student. In particular, I owe a great deal to Dr. Sharon Kessler for her mentoring and support while Ben was out of state during my early years in the program.

My tenure in the Holt lab has been shared with a number of amazing people, each of whom has contributed to my growth. Dr. Rod Kumimoto, Dr. Chamindika Siriwardana, and Dr. Swadhin Swain, all postdocs during my time in the lab, taught me many of the fundamentals I have built upon. Fellow graduate students, including Daniel Jones, Krystal Gayler, Emily Carpenter, and Yanan Bai, have given valuable assistance in how I think about and frame my projects. Daniel – thank you for your friendship, you helped me keep everything in perspective when I truly needed it, and it was a great comfort to know that I always had an afternoon-beer buddy. Finally,

many undergraduates in the lab have helped breathe life our day-to-day activities. I had the pleasure of mentoring many of these students in my time in the lab, and would like to thank Calvin Coffin, Younji Sohn, Andrew Willoughby, Shannan Bialek, Sam Keltner, and Reid Selby. Mentoring is one of the main reasons that I love what I do, and I certainly learned as much from them as they did from me. Andrew, thanks for always being there to bounce wacky ideas around and for always wanting to lend a hand.

Finally, I would like to acknowledge the enormous love and support I have received from my immediate and extended family – even when it meant I had to miss a few family dinners to get work done. My wife, Rachel, and my son, Charles, are an inspiration for me. My father, Andrew, my grandmother, Patricia, and my brother, Alexander, have been supportive of every endeavor I've ever undertaken, and I couldn't ask for a better family. My mother- and father-in-law, Nancy and John Breon, as well as my grandmother-in-law, Rogene Miller, have encouraged and cheered me along every step of the way. Last, but certainly not least, my faithful canine companion, Kaylee 'Bear' Myers, has been with me since my first semester of my graduate studies, and is always good for a walk when science hits the fan.

TABLE OF CONTENTS

List of Tables

List of Figures

Statement on Personal Contributions

Generally speaking, I was responsible for all first drafts that were refined into the chapters contained in this dissertation. I received editing advice and assistance with proofreading for each chapter, primarily from my advisor (Dr. Ben Holt) and from other co-authors and lab members. My contributions to each specific chapter were as follows:

Chapter 1

I drafted and wrote the entirety of this chapter.

Chapter 2

I collected or analyzed the majority of the data presented, including: hypocotyl elongation phenotyping of the *nf-yc triple hy5* mutant in all light and dark conditions, development of the novel FRET-based protein-protein interaction assay, functional analysis of the RNA sequencing data, hypocotyl cell count and elongation data, and the GUS expression analysis. Further, I assisted with the NF-YA2 active degradation analysis and with the generation of various complementation lines.

Chapter 3

I was responsible for all of the data presented **with the following exceptions**: Dr. Swadhin Swain produced all of the publication-quality western blots, and undergraduate assistants Shannan Bialek and Sam

Keltner assisted with cloning on select projects. Under my guidance, Shannan was responsible for cloning the point mutations to break the T2A sites in the effector module, while Sam assisted with cloning of binding site permutations used in the reconstruction of NF-Y DNA binding specificity.

Chapter 4

I was responsible for data collection of red and far red grown seedlings, and conducted the auxin inhibition treatments. Undergraduate assistant Reid Selby helped with plating and data collection of white and blue light root elongation experiments.

Chapter 5

I drafted and wrote the entirety of this chapter.

Abstract

Transcriptional regulation is a central component of biology whose importance cannot be overstated. From basal transcription of genes necessary for the physical make-up of a cell, to differential gene expression in response to internal or external signaling, life itself is built upon transcription. Transcription factors (TFs) modulate where and to what extent transcription occurs within the genome, and because of this important role in facilitating life, TFs have been intensely studied since the advent of molecular biology. This dissertation explores the developmental roles and molecular functions of the *Arabidopsis thaliana* NUCLEAR FACTOR-Y (NF-Y) family of transcription factors, a heterotrimeric TF complex composed of three evolutionarily-diverged subunits, NF-YA, NF-YB, and NF-YC. Interactions between NF-Y subunits have been explored for a number of years, and the interaction of the NF-Y complex with its cognate DNA binding sequence, *CCAAT*, has been further described through crystallization of a DNA-bound NF-Y complex. Because of this deep understanding, the NF-Y are an excellent candidate for modeling complex TF interactions and their impact on DNA binding.

In Chapter 1, I introduce and describe research into the NF-Y in two parts. In Part 1, I introduce the importance of transcriptional regulation and provide a broad history and description of the NF-Y complex in plants. In Part 2, I provide an exhaustive description of recent plant NF-Y research and future

perspectives. Taken together, this chapter aims to establish the importance of studying transcription factors and provide the reader with an up-to-date understanding of what is known regarding the developmental roles and mechanistic functions of the NF-Y.

One of the developmental roles of the NF-Y is in mediating seedling development in response to light, a process known as photomorphogenesis. In Chapter 2, I provide the first description of photomorphogenic roles for *NF-YC3*, *NF-YC4*, and *NF-YC9*. This work describes broad regulation of blue, red, and far red light signaling networks for these NF-YC subunits, where they coordinate light-dependent inhibition of hypocotyl elongation. Further, I identify that these NF-YC subunits function partially independent of the well-described ELONGATED HYPOCOTYL5 (HY5), despite confirming a direct physical interaction between NF-YC9 and HY5 through a novel microscopy-based approach combining Fluorescence Resonance Energy Transfer, Acceptor Photobleaching, and Fluorescence Lifetime Imaging. I further explore the interactions between NF-YC and HY5 in Chapter 4, where I identify a suite of mutually-exclusive phenotypic relationships in seedling primary root elongation in different light conditions.

In Chapter 3 I describe the development and implementation of a yeast synthetic reporter system, the **D**ynamic **I**nterdependent TF binding

Molecular **R**eporter (DIMR) system. The DIMR system probes oligomeric TF DNA binding *in vivo*, and has been able to completely recapitulate all tested aspects of NF-Y DNA binding. While my proof-of-concept is limited to the NF-Y, the DIMR system is designed as modularly as possible, and should be readily adaptable to study any TF DNA interaction with minor modifications.

Chapter 1: Introduction

The second section of this introductory chapter was previously published at Current Opinions in Plant Biology.

Myers, Zachary A. and Holt, Ben F. (2018) NUCLEAR FACTOR-Y: still complex after all these years?. Current Opinions in Plant Biology.

Chapter 1, Section I – historical perspective on the study and composition of the NUCLEAR FACTOR-Y transcription factors

The existence and persistence of life is borne on the back of the so-called Central Dogma of molecular biology, a concept encompassing the transfer of information between DNA, RNA, and protein molecules. Transcription factors, proteins that influence the transfer from DNA to RNA through transcription, are one of several fundamental components of the Central Dogma, and significant amounts of research have been invested into understanding their many roles and activities. Reflecting their importance, transcription factors are estimated to account for up to 10% of total encoded genes in some sequenced genomes, though many model organisms encode fewer – e.g., *Drosophila melanogaster* and *Caenorhabditis elegans* have an estimated 4.7% and 3.6%, respectively (Mitsuda and Ohme-Takagi 2009). The complexity of how transcription factors integrate signals and coordinate into functional complexes has been a consistent and fruitful direction for researchers for many decades.

Transcription is the process of creating an RNA molecule from a DNA blueprint, and along with translation, is a key connection between the genome and protein machinery. In eukaryotes, transcription is accomplished through the function of an RNA polymerase complex that simultaneously reads and copies a DNA template strand into a

complementary RNA molecule. Of the five RNA polymerase complexes found across eukaryotes (two of which are plant-specific), RNA polymerase II is responsible for the synthesis of pre-messenger RNA that underlies gene expression (Dignam et al. 1983). This synthesis is tightly and precisely controlled by transcription factors, which ultimately function to either directly recruit or to alter the activity of RNA polymerase and influence the frequency of transcription of a gene (Mitchell and Tjian 1989). This regulation by transcription factors is critically important, and is the means through which external and endogenous signals can be integrated at the cellular level.

General transcription factors facilitate assembly of the preinitiation complex that stabilizes and directs RNA polymerase, and are broadly required for all RNA polymerase II-mediated transcription (Roeder 1991). This includes the well-known TATA BINDING PROTEIN (TBP), as well as a suite of TATA ASSOCIATED FACTORS (TAFs) and RNA polymerase II-specific transcription factors (TFIIs). With the notable exception of TBP, most general transcription factors do not directly bind DNA, and instead function within the larger RNA polymerase complex; however, not all gene promoters contain the TATA box recognized by TBP. Instead, some eukaryotic promoters rely on other TFs to appropriately position the preinitiation complex through physical interaction with TBP/TFIID or TFIIB.

Most transcription factors function more indirectly by acting on other transcription factors or impacting the accessibility of chromatin along a target gene promoter (Levine and Manley 1989; Serfling et al. 1985). These transcription factors are broadly described as enhancers or repressors, depending on their mode of action, and function to fine-tune transcriptional activity in a context-dependent manner. Repressors and enhancers can assemble at a variety of locations, though the distinctions of 'proximal' and 'distal' promoters are used to describe the region of preinitiation complex assembly and the regions of enhancer/repressor binding, respectively (Harbison et al. 2004). Within the distal promoter, transcription factors can bind DNA and regulate transcription up to thousands of bases upstream or downstream of the transcriptional start site (Palstra 2009); however, the majority of enhancer and repressor binding occurs within the first few hundred base pairs upstream of the transcription start site (Tirosh et al. 2009). The physical interaction between transcription factors and DNA is of particular interest to the scientific community, as this is the mechanism through which many transcription factors integrate signals into a cohesive transcriptional output. Many forces influence the interactions between transcription factors and DNA, with epigenetic modification playing a significant part in directing transcriptional activity (see reviewed here, (Pikaard and Scheid 2014)). Beyond the epigenetic landscape, sequence-specific DNA-binding by transcription factors is a key means to modulate active recruitment of RNA polymerase to a specific genomic locus, and as

such, a better understanding of the basic biology of transcription factor DNA binding could facilitate future practical applications in agriculture or health.

One of the earliest-identified interactions between transcription factors and DNA was that of the NUCLEAR FACTOR-Y (NF-Y) complex and its association with the *CCAAT* box in the late 1980s in humans and yeast (Dorn et al. 1987; Jones et al. 1985). The thirty years following the initial identification and description of the NF-Y has yielded a detailed understanding of the mechanistic function of the NF-Y. In this introduction, I will specifically focus on the NF-Y genes of the model plant *Arabidopsis thaliana* to better understand complex transcription factor-DNA interactions.

The NF-Y consist of three distinct families of genes (NF-YA, NF-YB, and NF-YC) that function as heterotrimeric transcription factor complexes (Mantovani 1999). The functional NF-Y complex directly binds DNA at conserved *CCAAT* pentamers found in many eukaryotic gene promoters (Figure 1.1, (Maity and de Crombrugghe 1998)). While the NF-Y can be found in nearly all eukaryotes, NF-Y lineages in higher plants have evolved to maintain multiple copies of each subunit (with 10 of each subunit encoded in Arabidopsis, (Petroni et al. 2012; Siefers et al. 2009)). This duplication and retention appears to have led to significant functional redundancy, and as such, plant NF-Y research has progressed at a slower pace compared to its mammalian and fungal counterparts.

NF-Y complex formation is a tightly-regulated process that has been studied extensively in animal systems, and while a few key differences have been reported in plant systems, the general scheme appears conserved. NF-YB and NF-YC dimerize in the cytoplasm prior to nuclear localization, and this NF-YB/NF-YC dimer then recruits NF-YA to form the mature NF-Y complex (Figure 1.1, (Sinha et al. 1996; Sinha et al. 1995)). This complex can then bind DNA at the *CCAAT* box and influence transcription either positively or negatively. In both plant and animal systems, the NF-Y have been reported to recruit other transcription factors to mediate transcriptional activity (Liu and Howell 2010; Siriwardana et al. 2014; Huang et al. 2015; Myers et al. 2016; Tang et al. 2017; Siriwardana et al. 2016), though this is not thought to be a necessity for basal NF-Y mediated gene regulation (i.e., the mature NF-Y complex alone is thought to possess the necessary properties to regulate transcription).

NF-YA features and family members

Though the mature NF-Y complex is composed of three different components, most research has focused on understanding the roles of individual families of NF-Y proteins. The NF-YA family members share highly conserved domains necessary for both DNA binding and interaction with the NF-YB/NF-YC dimer, though their NH_2- and COOH- terminal regions are considerably more divergent (Mantovani et al. 1994; Petroni et

al. 2012). Phylogenetic analysis of the genomic coding sequence of each NF-YA gene of *Arabidopsis thaliana* suggests that each is more closely related to a single other gene than the entire family (Figure 1.2). This observation is supported by the identification of several pairs of NF-YA genes with redundant functions in embryogenesis, root developmental architecture, and phytohormone responses (Mu et al. 2012; Siriwardana et al. 2014; Sorin et al. 2014; Siriwardana et al. 2016). Interestingly, the DNA binding domain of NF-YA is strikingly similar to the conserved domain of the plant-specific CCT (CONSTANS, CONSTANS-LIKE, and TOC1) protein family, an observation that has been a driving force behind much recent NF-Y research ((Wenkel et al. 2006), discussed in Part II below).

Sequence alignments of NF-YA conserved domains of *Arabidopsis thaliana* show a ~60 amino acid stretch of conserved residues arranged into a pair of alpha helices with a joining linker region. Many of the residues in the second alpha helix have been shown to make direct, sequence-specific physical content within the minor groove of human NF-Y bound DNA (Figure 1.3, (Nardini et al. 2013)). Interestingly, AtNF-YA2, which is one of the most heavily-studied subunits of the family, is rather atypical in its conserved domain, encoding a particularly long linker between its NF-YB/NF-YC interaction domain and its DNA binding domain. While the functional significance of this extended linker has not been thoroughly investigated, it is worth considering the fact that a majority of our understanding of the

mechanism of plant NF-Y function is through the lens of this somewhat-atypical NF-YA.

NF-YB features and family members

NF-YB and NF-YC family members each encode conserved histone fold domains, with similarity to H2B and H2A type histones, respectively (Frontini et al. 2002; Ceribelli et al. 2008; Romier et al. 2003). Phylogenetic analysis of the genomic coding sequence of each NF-YB gene of Arabidopsis identified a number of small 2-3 member clades, several of which have been functionally linked (Figure 1.4). In particular, NF-YB2 and NF-YB3 have been extensively studied for their roles in photoperiod-dependent floral induction (Kumimoto et al. 2008; Gnesutta et al. 2017a), while NF-YB6 and NF-YB9, also known as LEAFY COTYLEDON1 (LEC1) and LEC1-LIKE (L1L), are well-described regulators of embryogenesis (Yamamoto et al. 2009; Vicient et al. 2000; Baumlein and Junker 2012; Hilioti et al. 2014).

The conserved domain of the NF-YB family is nearly 100 amino acids in length, but shows remarkably few divergences between family members (Figure 1.5). Early analyses on the NF-YB subunit of rat identified multiple classes of mutations that impair various aspects of NF-Y complex function, including (1) dimerization of NF-YB/NF-YC, (2) trimerization of the NF-YB/NF-YC dimer with NF-YA, or (3) the ability of the full NF-Y complex to

bind DNA (Sinha et al. 1996). The existence of these various classes of mutations demonstrate that DNA binding of the NF-Y complex is an intricate process where transcription factor-DNA interactions, including sequence specificity, stabilization to DNA, and transcriptional regulation, cannot be completely attributed to a single subunit of the complex.

Another important aspect of NF-YB structure and function is the mono-ubiquitination of a lysine reside within the conserved domain of human NF-YB (Nardini et al. 2013). This ubiquitination was found to be required for stable DNA binding by the human NF-Y complex. Interestingly, only 3 of the 10 Arabidopsis *NF-YB* genes have retained this lysine residue (*NF-YB3*, *NF-YB4*, and *NF-YB7*), while the remaining 7 *NF-YB* genes have an arginine in the same relative position. While both residues are chemically similar (positively-charged side chains), only lysine can be covalently modified with a ubiquitin moiety, suggesting that the 7 arginine-encoding NF-YBs have evolved a slightly different mechanism of DNA binding and stabilization. Whether there is any biological significance to this observation remains to be tested.

NF-YC features and family members

As mentioned above, all NF-YC proteins encode a conserved H2A-like histone fold domain; however, the sequences flanking this domain tend to be fairly divergent. Functionally-verified clades can be identified through

phylogenetic analysis of genomic coding sequences (Figure 1.6, (Kumimoto et al. 2010; Liu et al. 2016; Myers et al. 2016)). Less work has been aimed at understanding the structure and function of NF-YC subunits than those of NF-YB; however, a similar trend has been observed, where several conserved residues have been identified to be necessary for aspects of NF-Y complex formation and DNA binding (Figure 1.7, (Sinha et al. 1995; Kim et al. 1996)). While mutation of these conserved residues has primarily been examined in animal NF-YCs, several mutations have been verified in Arabidopsis NF-YC subunits.

One aspect of NF-YC research that has surpassed that of other NF-Y subunits is the identification and description of interaction with other proteins. NF-YC proteins from both human and Arabidopsis have been identified to physically interact with a suite of other transcription factors, including members of the basic Leucine Zipper (bZIP) and basic Helix-Loop-Helix (bHLH) families (Figure 1.8, (Yoshida et al. 2001; Huang et al. 2015; Myers et al. 2016; Tang et al. 2017)). The regulatory consequences of these interactions still remains relatively unexplored, though a significant co-occurrence of CCAAT boxes and G boxes (which are bound by both bZIP and bHLH transcription factors) has been described (Zambelli and Pavesi 2017). These relationships are further explored in Section II below.

Chapter 1, Section II – recent advances in understanding the molecular and developmental roles of the NUCLEAR FACTOR-Y

Due in large part to their sessile nature, plants have developed sophisticated signaling mechanisms to perceive and respond to the environment, allowing them to fine-tune their growth, development, and stress responses. Alterations in gene regulation underlie this plasticity, and, for a number of reasons described below, NUCLEAR FACTOR-Y (NF-Y) transcription factors are particularly interesting research candidates to better understand how plants integrate endogenous and environmental signals so effectively.

NF-Y transcription factors consist of protein subunits from three distinct families (NF-YA, NF-YB, and NF-YC) that are traditionally understood to function as a heterotrimeric transcription factor complex. The functional NF-Y complex recognizes and binds the consensus sequence *CCAAT*; however, not all *CCAAT* pentamers are bound by NF-Y complexes, suggesting that additional influences of flanking nucleotides, local chromatin structure, and interactions with other trans-acting factors also contribute to NF-Y binding. NF-YB and NF-YC subunits contain histone-fold domains that facilitate their dimerization and subsequent translocation from the cytosol to the nucleus, where NF-YA is then recruited to form a complete NF-Y complex that can bind DNA and affect transcription (Figure 1.1).

Protein crystallography studies have shown that NF-YA subunits make all of the direct physical contacts with *CCAAT*, while NF-YB and NF-YC form a binding platform for NF-YA and make stabilizing contacts with the DNA backbone (Nardini et al. 2013). Individual NF-Y subunits are not generally found to independently bind DNA with high affinity or regulate transcriptional activity, though recent discoveries have suggested that functional, non-canonical trimeric complexes are also a natural fate for histone-fold containing NF-YB/NF-YC subunits (Gnesutta et al. 2017a).

While NF-Y are found in all eukaryotes, the individual NF-Y families have significantly expanded across the plant lineage – for example, the *Arabidopsis thaliana* (Arabidopsis) genome has ~10 genes encoding each subunit type (Petroni et al. 2012). The implications of this expansion are likely quite significant when considering that the NF-Y function as a heterotrimeric complex, with as many as 1000 theoretical NF-Y complexes being able to form in Arabidopsis (Siefers et al. 2009). Unsurprisingly, the NF-Y have been implicated in many developmental and physiological processes and have recently become a target of translational approaches to create more resilient and nutritious crops (Li et al. 2015; Qu et al. 2015; Yadav et al. 2015). This review focuses on the most impactful developmental, physiological, and mechanistic NF-Y research of the last few years, and we encourage interested readers to look through any of a number of high-quality NF-Y reviews to obtain a more complete picture

(Brambilla and Fornara 2016; Swain et al. 2017; Zanetti et al. 2017; Gnesutta et al. 2018).

Roles of the NF-Y in development and physiology

Arguably the best understood role of the NF-Y in plants is their regulation of the principle florigen in Arabidopsis, *FLOWERING LOCUS T (FT)*, during photoperiodic flowering. While this line of research can be traced back over a decade (Wenkel et al. 2006; Kumimoto et al. 2008; Kumimoto et al. 2010; Ben-Naim et al. 2006), a more precise understanding of how the NF-Y regulate *FT* has only recently emerged. Prior to 2016, a series of observations indicated that NF-Y regulation of *FT* was mediated through CONSTANS (CO), a transcription factor previously suspected to directly regulate *FT (Putterill et al. 1995; Samach et al. 2000)*. Two schools of thought developed to describe the relationship between NF-Y and CO — the 'recruitment model,' positing that a complete NF-Y trimeric complex recruited CO to the *FT* promoter, and the 'replacement model,' suggesting that CO replaced NF-YA to form a CO/NF-YB/NF-YC trimer (NF-CO) that could bind to and regulate the *FT* promoter. This debate has largely been settled, and the standing model supports aspects of both ideas.

The prevailing hypothesis is that NF-Y regulates *FT* through the action of a pair of protein complexes (NF-Y and NF-CO) that are brought into close proximity of each other while also binding the *FT* promoter at two distinct

binding sites separated by over 5kb (Cao et al. 2014; Siriwardana et al. 2016; Gnesutta et al. 2017a; Adrian et al. 2010; Turck et al. 2008; Liu et al. 2014; Tiwari et al. 2010). The interaction of these two complexes is facilitated by the formation of chromatin loops, with a canonical NF-Y complex binding a distal *CCAAT* box ~5.3kb upstream of the *FT* start and a non-canonical NF-CO complex binding at a proximal CORE (CO Responsive Element) site immediately upstream of the *FT* start (Cao et al. 2014; Gnesutta et al. 2017a). The key insight, which addressed the original confusion, is that both NF-YA and CO appear to be separately utilizing NF-YB/NF-YC dimers to facilitate DNA binding. In addition to advancing our specific, mechanistic understanding of NF-Y/CO regulation of *FT* during flowering, these observations are likely to be broadly paradigmatic for describing the actions of the larger CCT (CO, CO-LIKE, and TOC1) family of proteins, of which CO is just one member (further discussed below).

Most research into NF-Y regulation of photoperiodic flowering has focused on the NF-Y/CO relationship, though a few reports have emerged suggesting that this is not the only relevant NF-Y role. One alternative role is as a regulator of histone H3K27me3 de-methylation in the *SUPPRESSION OF OVEREXPRESSION OF CO 1 (SOC1)* promoter through interaction with RELATIVE OF EARLY FLOWERING 6 (REF6, (Hou et al. 2014)). This regulation appears to also integrate gibberellic acid (GA) signaling, as binding of NF-Y to the *SOC1* promoter is enhanced in

response to application of exogenous GA. As only NF-YA specifically bound the SOC1 promoter fragment and the putative binding site did not contain a canonical *CCAAT* box, the exact mechanism of this regulation requires additional examination. In rice, NF-Y actually inhibit flowering in response to long days (Kim et al. 2016), though several rice NF-YB and NF-YC orthologs are able to complement Arabidopsis late-flowering mutants (Hwang et al. 2016).

A recent report identified and characterized NF-Y regulation of miR156, a central component of the floral induction aging pathway in chrysanthemum (Wei et al. 2017). This is the first published instance of an NF-Y modulating miRNA biogenesis, though previous research demonstrated that several Arabidopsis NF-YA proteins were themselves targeted by miR169, including a recently identified relationship between miR169 and several NF-YA subunits in regulating lateral root development (Sorin et al. 2014; Luan et al. 2015). The relationship between NF-YA and miR169 also extends to nodule formation during legume-rhizobia symbiotic interactions, where NF-YA is an important factor regulating the function and persistence of nodule meristems (Combier et al. 2006). More recent work has shown that NF-YA proteins are also necessary for the earliest stages of nodule formation, and that a conserved suite of NF-YB and NF-YC proteins are also important in these developmental processes (Laloum et al. 2014; Baudin et al. 2015). A more precise understanding of this process is beginning to emerge, where

it appears that NF-Y mediated regulation begins following the first cortical cell divisions preceding nodule formation (Hossain et al. 2016).

While some aspects of NF-Y research have matured to the point of understanding mechanistic function, others remain less concrete. Of particular note, several groups have recently reported roles for the NF-Y during light- and dark-mediated seedling development. In particular, NF-YB9, also known as LEAFY COTYLEDON 1 (LEC1), has been implicated as a positive regulator of dark-mediated development (Huang et al. 2015), while NF-YC1, 3, 4, and 9 appear to be positive regulators of light-mediated development (Myers et al. 2016; Tang et al. 2017). In both cases some uncertainty remains regarding the mechanistic function of the relevant NF-Y subunit(s). NF-YB9 regulation of dark-mediated development appears to be mediated through interaction with PHYTOCHROME INTERACTING FACTOR 4 (PIF4, (Huang et al. 2015)); however, because a) hypocotyl elongation was used to measure dark-mediated development, b) the hypocotyl is of embryonic origin, and c) the hypocotyl was shorter prior to germination in the observed *nf-yb9* mutants, decoupling these pre- and post-germination phenotypes remains challenging (Pelletier et al. 2017; Baumlein and Junker 2012).

NF-Y regulation of light-mediated development is somewhat less complex to interpret as the relevant NF-YC subunits do not appear to be involved in

the same processes as NF-YB9 during embryogenesis and seed maturation. Higher-order *nf-yc* loss of function mutants are unable to completely inhibit hypocotyl elongation in most light conditions, with more drastic phenotypes associated with lower light intensities (Myers et al. 2016). This regulation is independent of the action of ELONGATED HYPOCOTYL 5 (HY5), a heavily studied regulator of light-mediated development that is nevertheless thought to physically interact with several NF-YCs. Additionally, overexpression of some NF-YA subunits produces a constitutively photomorphogenic response, and NF-YA2 appears to be degraded in a dark- and proteasome-dependent manner (Myers et al. 2016). This type of degradation is a hallmark of the regulation of photomorphogenic transcription factors, suggesting that the NF-Y complex could function analogously to other known light signaling components while also regulating a unique set of downstream targets. This is consistent with earlier research showing NF-Y binding to cellular elongation and auxin biosynthesis genes (Junker et al. 2012), and reductions in histone acetylation on hypocotyl elongation related genes in *nf-y* mutants (Tang et al. 2017). Further study of the genetic relationship between the NF-Ys and other light signaling components, as well as mechanistic dissection of the degradation of NF-YA proteins, will be necessary to fully understand the action of the NF-Y in photomorphogenesis.

Finally, a significant number of reports continue to identify the NF-Y as crucial regulators of plant stress and hormone responses. The initial characterizations of the roles of the NF-Y in these responses fall outside the purview of this work (reviewed in (Swain et al. 2017)), but recent work has continued to identify important roles of the plant NF-Y in regulating ABA responses (Xuanyuan et al. 2017; Bi et al. 2017; Liu et al. 2016; Siriwardana et al. 2014), salt stress (Manimaran et al. 2017; Zhang et al. 2015; Ma et al. 2015), drought stress (Quach et al. 2015; Lee et al. 2015; Alam et al. 2015), freezing stress (Shi et al. 2014), and various combinations of each (Li et al. 2016; Alam et al. 2015; Xu et al. 2014b; Yang et al. 2017). Several groups have found that orthologs of the Arabidopsis NF-Ys significantly impact stress responses when ectopically expressed in Arabidopsis (Lee et al. 2015; Ma et al. 2015; Zhang et al. 2015; Li et al. 2016), and while the biological significance of these types of approaches must be carefully determined in each case, the combined observations suggest that the roles of the NF-Y are fairly conserved across plants. In most of these cases, relatively few specific details are known; however, a few key hubs have been proposed. In particular, many of the stress responses observed in NF-Y mutants are linked to ABA signaling, where several recent publications have proposed that the NF-Y modulate ABA signaling either through epigenetic regulation of *ABI5* or through formation of an NF-Y/ABI5 complex to regulate ABA-responsive genes (Liu et al. 2016; Bi et al. 2017). Further, a recent large-scale analysis of over 20 ABA-responsive transcription

factors, including NF-YB2 and NF-YC2, found that multiple components assemble onto ABA-responsive elements, and that many of these large assemblages contain NF-Y bound *CCAAT* boxes (Song et al. 2016). Interestingly, when looking at the regulatory scheme of the ABA response through hierarchical clustering analyses, NF-YB2 and NF-YC2 functioned upstream of most responsive elements, suggesting that they are of fundamental importance in mediating ABA responses.

Mechanisms of NF-Y mediated transcriptional regulation

Significant progress has been made in the past two years toward understanding the mechanism of NF-Y mediated transcriptional regulation in plants (Figure 1.8). Historically, the NF-Y have been described as a heterotrimeric complex that can recognize and bind to the *CCAAT* box (Figure 1.8A). This mechanism was elucidated from a combination of mutational analyses and protein-protein interaction assays over the last 30 years, culminating with the crystal structure of the mammalian NF-Y trimer bound to DNA (Nardini et al. 2013). This crystal structure showed that NF-YA is responsible for all physical contacts to the *CCAAT* box by inserting an alpha helix into the DNA minor groove, while the NF-YB/NF-YC subunits bind to the sugar-phosphate backbone flanking the *CCAAT* box and facilitate NF-YA insertion. Plant NF-Ys share the same essential DNA-binding amino acids, are affected by the same point mutations, and appear to bind *CCAAT* in the same basic manner (Siriwardana et al. 2016; Gnesutta et al. 2017a).

More recently, a novel mechanism describing NF-Y mediated transcriptional regulation has emerged in the form of the above-mentioned NF-CO complex, where CO can take the place of NF-YA in interaction with the NF-YB/NF-YC dimer (Figure 1.8B, (Gnesutta et al. 2017a)). The existence of this variant of the NF-Y complex was predicted over a decade ago, following the observation that the CCT domain of CO shares significant sequence similarity to the conserved DNA-binding domain of NF-YA (Wenkel et al. 2006; Ben-Naim et al. 2006). Through mutational studies, electrophoretic mobility shift assays, and chromatin immunoprecipitation assays, it appears that complex formation with either NF-YA or CO is a natural fate for the NF-YB/NF-YC dimer (Gnesutta et al. 2017a; Siriwardana et al. 2016). Interestingly, while all NF-YA variants share a set of conserved amino acids that are essential for physical contact with the *CCAAT* box, several of these residues are not shared in the DNA binding domain of either CO or the CCT family (Petroni et al. 2012; Wenkel et al. 2006). Further, the NF-CO complex cannot bind the canonical *CCAAT* box that the NF-Y complex recognizes, and instead binds with high affinity to the related pentanucleotide *CCACA* (Gnesutta et al. 2017a). Whether this relationship is exclusive to CO or extends to other CCT proteins remains unexplored; however, a few additional CCT domain containing proteins have been identified to interact with NF-Y subunits (Cao et al. 2011a; Li et al. 2011; Ben-Naim et al. 2006), suggesting that this relationship and mechanism

could be conserved. With more than 40 CCT domain-containing proteins encoded in Arabidopsis, this mode of NF-CCT (CCT/NF-YB/NF-YC) mediated transcriptional regulation is poised to become an important new research frontier.

While a generic NF-Y complex appears to already possess the necessary components to independently regulate transcription, there have been many reports of the NF-Y recruiting additional transcription factors (Figure 1.8C, (Liu and Howell 2010; Huang et al. 2015)). In most instances, interactions between the NF-Y complex and other proteins appears to be mediated through NF-YC. Additionally, the nature of the recruited protein appears to affect binding specificity, with several proposed NF-Y bound *CCAAT* boxes harboring adjacent G-boxes that are bound by known NF-Y-interacting partners (primarily bZIP and bHLH transcription factors across both plant and animal lineages, (Huang et al. 2015)). In fact, the exact spacing between *CCAAT* and G-boxes, at least in mammals, appears to be significantly enriched from what would be expected in random spacing (Song et al. 2016; Zambelli and Pavesi 2017), suggesting some significance underlies their co-occurrence. While this type of co-association is readily testable, other types, such as that seen in the multi-component assemblage of NF-Y and NF-CO complexes at the *FT* promoter (discussed above, and Figure 8D), could prove more difficult to identify *de novo*.

In addition to interacting with other transcription factors, the NF-Y also appear to be able to recruit and direct chromatin remodeling enzymes to affect transcription. The NF-Y have long been described as pioneer transcription factors that bind otherwise inaccessible DNA and initiate a transition towards a more permissive chromatin landscape (Oldfield et al. 2014), a property that is thought to be a consequence of the histone fold-like properties of NF-YB and NF-YC. While these properties have been explored in more depth in animal systems, several recent reports have observed similar activity in Arabidopsis in regulation of flowering time, photomorphogenesis, ABA-mediated seed germination, and vernalization (Tang et al. 2017; Liu et al. 2016; Hou et al. 2014; Tao et al. 2017). The most complete picture of this aspect of NF-Y function was just recently described, where LEC1 was shown to be essential for the resetting of floral-inducing epigenetic modifications following reproduction, allowing the progeny to delay flowering until appropriate environmental conditions are present (Tao et al. 2017). Further research in plants will likely continue to support what has been learned in animal systems in terms of directing epigenetic modifications; however, whether the plant NF-Ys themselves are post-translationally modified similar to their animal orthologs remains unexplored.

The question of how the NF-Y discriminate between given *CCAAT* boxes is important and remains unaddressed. The crystal structure of the DNA-

bound mammalian complex showed that only NF-YA makes sequence-specific contacts, and all of the residues that physically interact with the *CCAAT* box are completely conserved between animals and plants (Nardini et al. 2013). Because of these observations, it seems likely that the ability to discriminate between *CCAAT* boxes does not reside within NF-YA itself. One possible explanation is that chromatin accessibility could influence the ability to bind a given *CCAAT* box; however, the pioneering nature of the NF-Y complex could mitigate the effects of inaccessible chromatin. Another likely possibility is that the residues flanking the *CCAAT* box are important for NF-Y binding. In this scenario, specific NF-YB/NF-YC combinations might provide specificity for different flanking sequences around the *CCAAT* box – an effect already observed when comparing the atypical LEC1/NF-YC dimer crystal structure to that of mammalian NF-Y (Gnesutta et al. 2017b). While this remains a difficult prospect to test, a thorough and systematic analysis of the effect different NF-YB/NF-YC dimers have on DNA binding specificities would significantly deepen our understanding of NF-Y regulatory capacity.

Conclusions

While the NF-Y have been actively studied for over three decades, the last few years have seen an enormous leap in our understanding of their function. Researchers have steadily identified an increasing number of developmental and physiological roles of the NF-Y, and we are beginning

to understand the mechanistic function of the NF-Y complex. Propelled by a high-resolution crystal structure of the NF-Y complex, we have developed a firm grasp of the biochemistry of its DNA binding capacity. As more is learned, more complex and nuanced questions become increasingly tenable, and the field is now poised to address whether non-canonical NF-Y complexes, such as NF-CO or NF-CCT, are the exception or the rule.

.

Figures and Tables

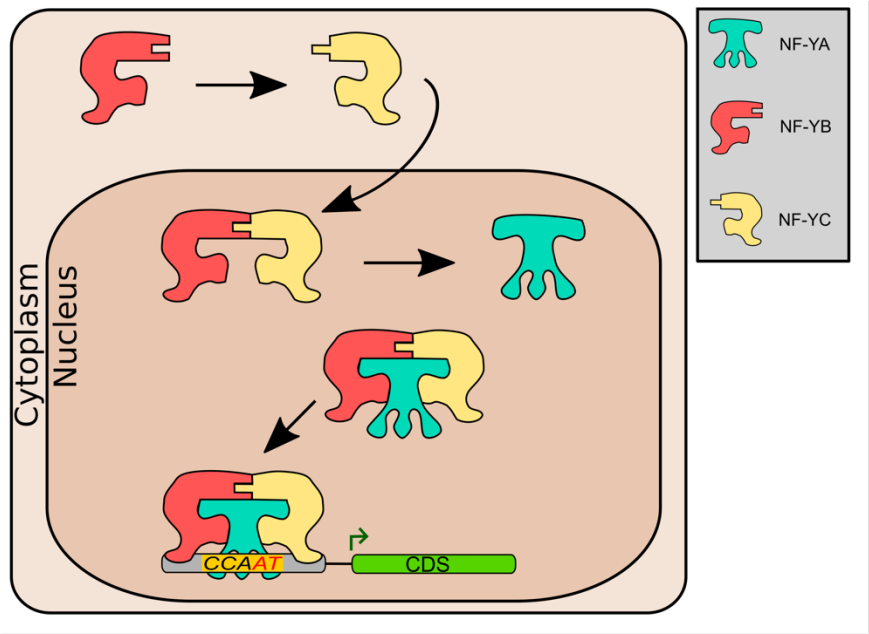

Figure 1.1. Stepwise formation of the NF-Y complex. Research in animal systems have shown that assembly of the complete NF-Y complex occurs in a stepwise manner, where NF-YB and NF-YC dimerize in the cytoplasm, move into the nucleus, recruit an NF-YA component, and subsequently bind DNA and affect transcription.

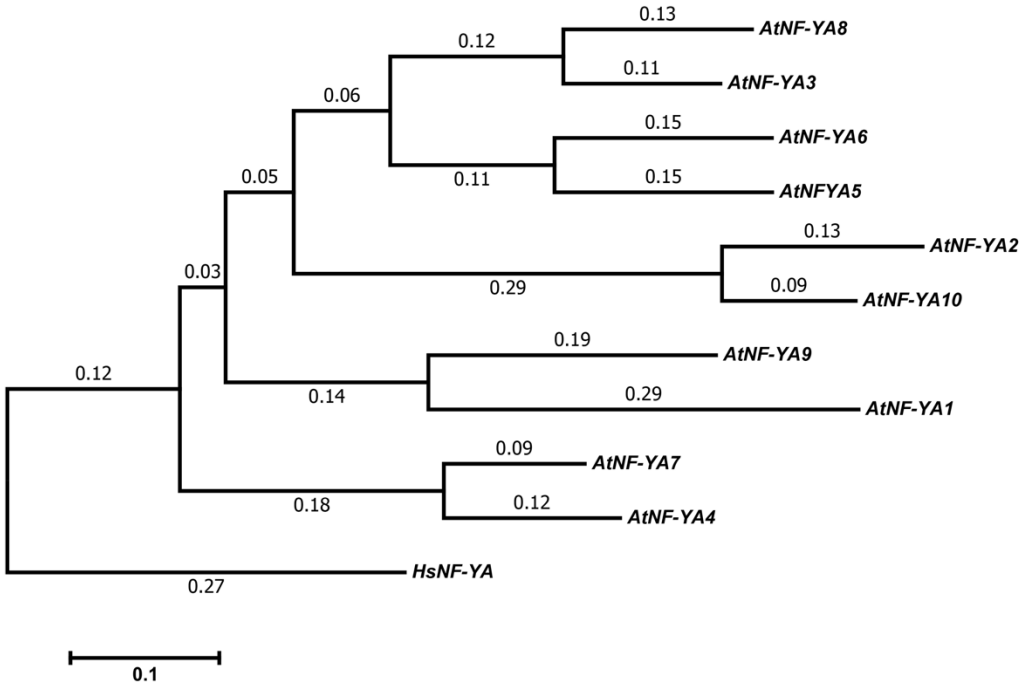

Figure 1.2. Arabidopsis NF-YA family phylogeny. Full-length genomic coding sequences were used for phylogenetic analysis. Multiple sequence alignments were generated through MUSCLE (Edgar 2004), while the evolutionary relationship was inferred through the Neighbor Joining method in MEGA7 (Kumar et al. 2016). Evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004) and are in the units of the number of base substitutions per site. The tree is rooted on *Homo sapiens* NF-YA genomic coding sequence.

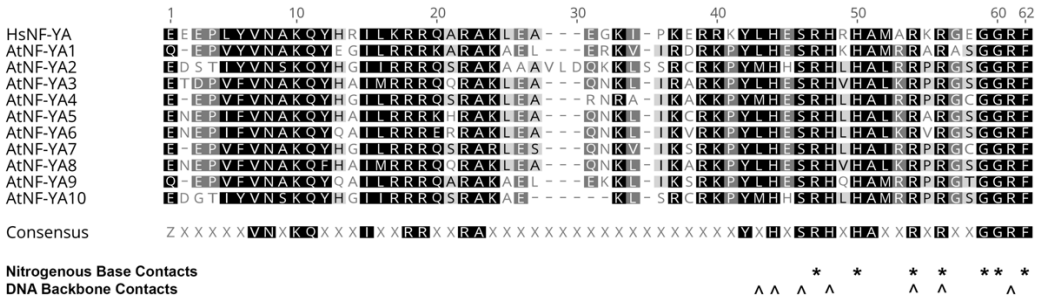

Figure 1.3. Multiple sequence alignment of NF-YA conserved domain. This alignment was generated using MUSCLE in Geneious R9 (https://www.geneious.com). The Consensus line shows completely-conserved resides, asterisks denote residues that make physical contact with DNA nitrogenous bases (i.e., the CCAAT box), and carats denote residues that make physical contact with the DNA backbone.

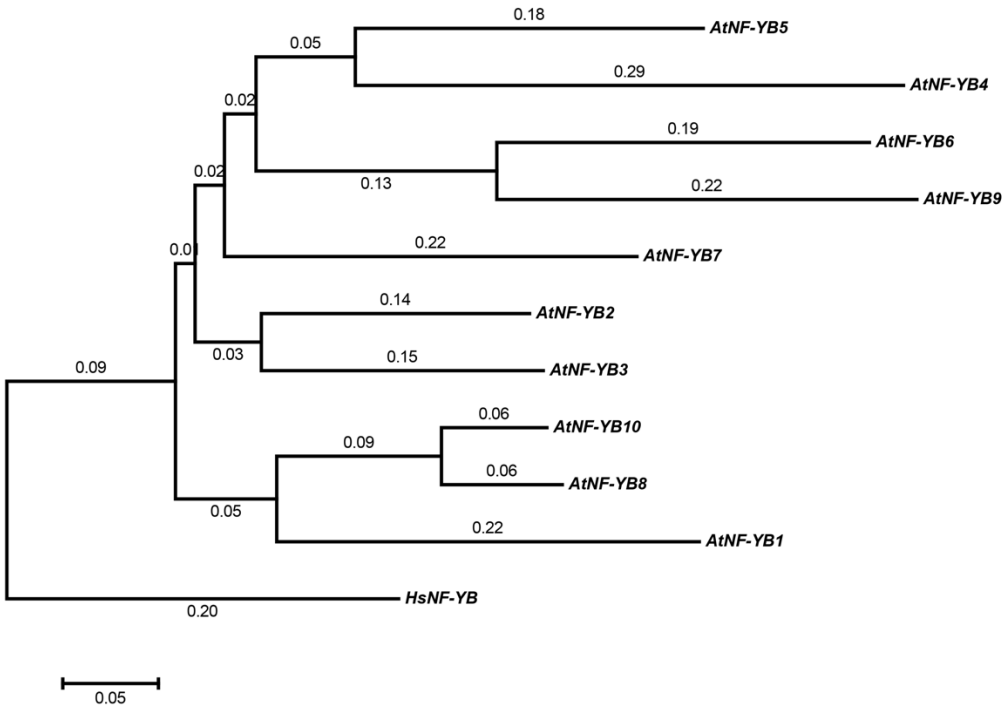

Figure 1.4. Arabidopsis NF-YB family phylogeny. Full-length genomic coding sequences were used for phylogenetic analysis. Multiple sequence alignments were generated through MUSCLE (Edgar 2004), while the evolutionary relationship was inferred through the Neighbor Joining method in MEGA7 (Kumar et al. 2016). Evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004) and are in the units of the number of base substitutions per site. The tree is rooted on *Homo sapiens* NF-YA genomic coding sequence.

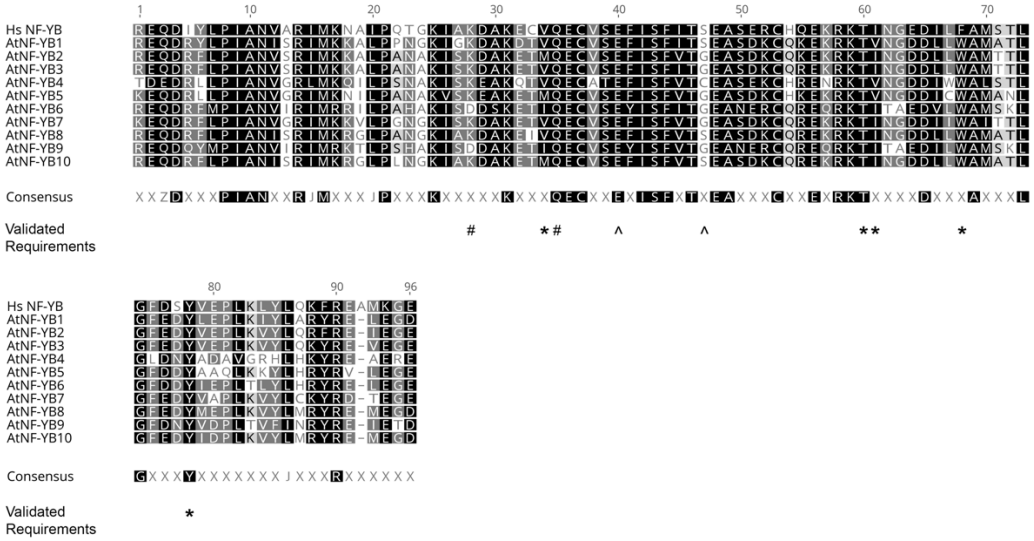

Figure 1.5. Multiple sequence alignment of NF-YB conserved domain. This alignment was generated using MUSCLE in Geneious R9 (https://www.geneious.com). The Consensus line shows completely-conserved resides, asterisks denote residues that are required for NF-YB/NF-YC dimerization, carats denote residues necessary for interaction of the NF-YB/NF-YC dimer with NF-YA, and pound signs denote residues necessary for the complete NF-Y trimer to bind DNA.

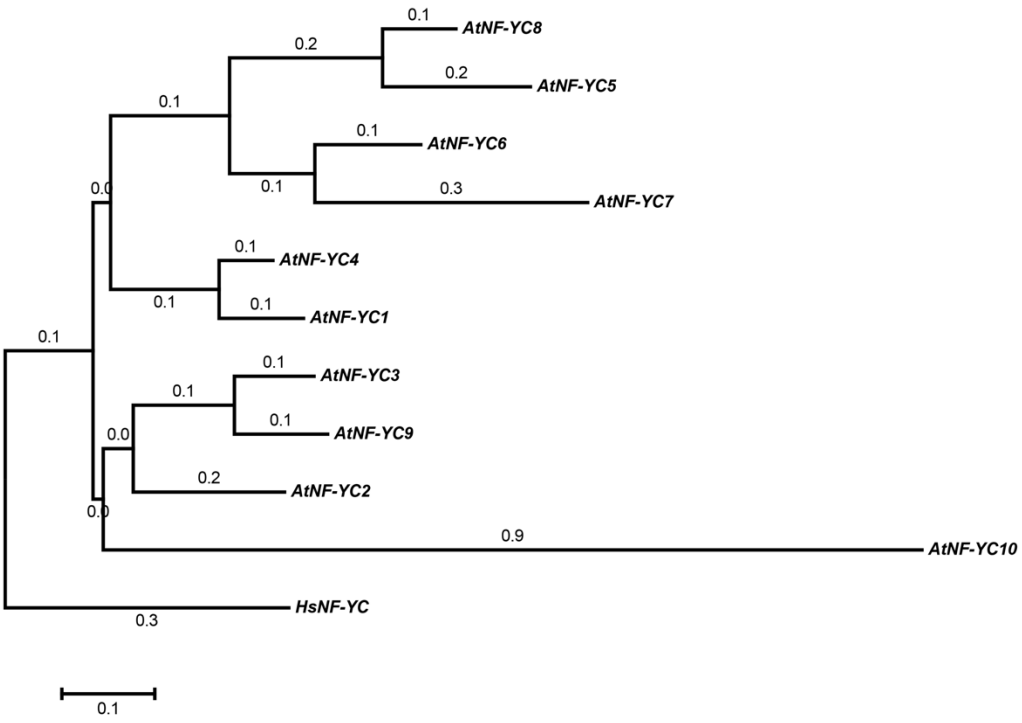

Figure 1.6. Arabidopsis NF-YC family phylogeny. Full-length genomic coding sequences were used for phylogenetic analysis. Multiple sequence alignments were generated through MUSCLE (Edgar 2004), while the evolutionary relationship was inferred through the Neighbor Joining method in MEGA7 (Kumar et al. 2016). Evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004) and are in the units of the number of base substitutions per site. The tree is rooted on *Homo sapiens* NF-YA genomic coding sequence.

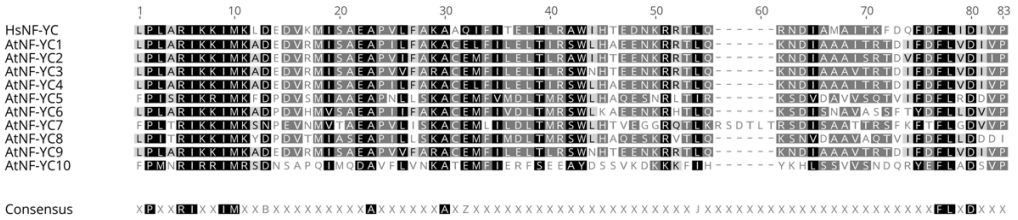

Figure 1.7. Multiple sequence alignment of NF-YC conserved domain. This alignment was generated using MUSCLE in Geneious R9 (https://www.geneious.com). The Consensus line shows completely-conserved resides.

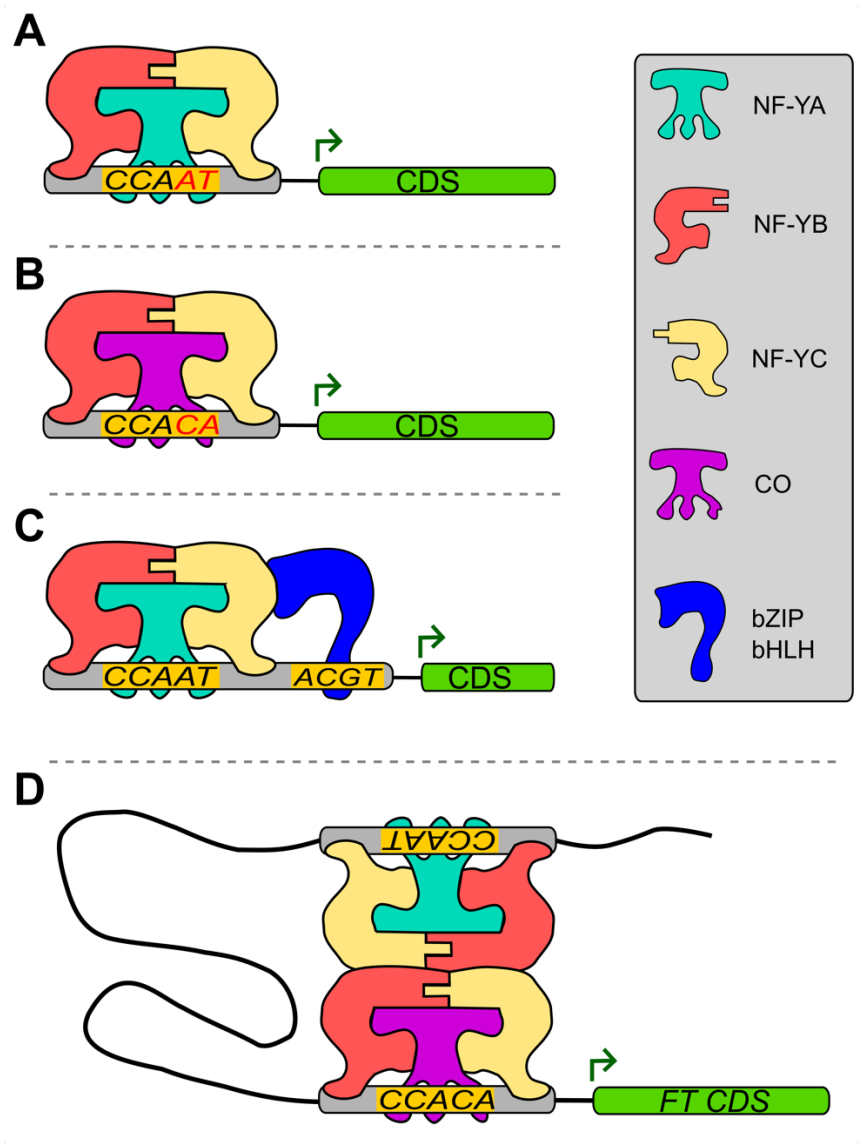

Figure 1.8. Mechanisms of NF-Y mediated transcriptional regulation. (**A**) The canonical NF-Y complex consists of single NF-YA, NF-YB, and NF-YC subunits that bind with high affinity to the *CCAAT* box, (**B**) while the recently-described NF-CO complex substitutes the NF-YA component for CO and binds the sequence *CCACA*. (**C**) The canonical NF-Y complex has been described to recruit additional sequence-specific transcription factors, such as bZIPs and bHLHs, potentially modulating the ability of the complex to bind specific CCAAT boxes. (**D**) The NF-Y complex is also capable of mediating long-distance interactions such as that observed in the regulation of flowering time at the *FT* promoter, where an NF-Y and an NF-CO complex bind at consensus sequences approximately 5kb apart.

Chapter 2: NUCLEAR FACTOR Y, subunit C (NF-YC) transcription factors are positive regulators of photomorphogenesis in *Arabidopsis thaliana*

Zachary A. Myers[1¶], Roderick W. Kumimoto[1,#a¶], Chamindika L. Siriwardana[1], Krystal K. Gayler[1], Jan R. Risinger[2,#b], Daniela Pezzetta[1,#c], Ben F. Holt III[#d]

This chapter was published in *PLoS Genetics*.

Myers, Zachary A, Kumimoto, RW., Siriwardana, CL., Gayler, KK., Risinger, JR., Pezzetta, Daniela, and Holt, BF III. (2016). NUCLEAR FACTOR Y, Subunit C (NF-YC) Transcription Factors Are Positive Regulators of Photomorphogenesis in Arabidopsis thaliana. *PLoS Genetics*

[1]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, United States of America
[2]Bioo Scientific, Austin, Texas, United States of America
[#a]Current Address: Department of Plant Biology, University of California Davis, Davis, California, United States of America
[#b]Current Address: Invitae Corporation, San Francisco, California, United States of America
[#c]Current Address: Department of Applied Genetics, Freie Universität, Berlin, Germany
[#d]Current Address: AgBiome Scientific, Research Triangle Park, North Carolina, United States of America
[¶]These authors contributed equally to this work.

Abstract

Recent reports suggest that NF-Y transcription factors are positive regulators of skotomorphogenesis in *Arabidopsis thaliana*. Three *NF-YC* genes (*NF-YC3*, *NF-YC4*, and *NF-YC9*) are known to have overlapping functions in photoperiod dependent flowering and previous studies demonstrated that they interact with basic leucine zipper (bZIP) transcription factors. This included ELONGATED HYPOCOTYL 5 (HY5), which has well-demonstrated roles in photomorphogenesis. Similar to *hy5* mutants, we report that *nf-yc3 nf-yc4 nf-yc9* triple mutants failed to inhibit hypocotyl elongation in all tested light wavelengths. Surprisingly, *nf-yc3 nf-yc4 nf-yc9 hy5* mutants had synergistic defects in light perception, suggesting that NF-Ys represent a parallel light signaling pathway. As with other photomorphogenic transcription factors, *nf-yc3 nf-yc4 nf-yc9* triple mutants also partially suppress the short hypocotyl and dwarf rosette phenotypes of CONSTITUTIVE PHOTOMORPHOGENIC 1 (*cop1*) mutants. Thus, our data strongly suggest that NF-Y transcription factors have important roles as positive regulators of photomorphogenesis, and in conjunction with other recent reports implies that the NF-Y are multifaceted regulators of early seedling development.

Introduction

Plants utilize multiple properties of light, such as intensity, quality, and direction, to guide growth and development (Neff et al. 2000). The effects of light on plant development are exemplified by the transition of seedlings from dark growth (where they exhibit skotomorphogenesis) to light growth (photomorphogenesis). This transition is crucial for plant viability and is characterized by the inhibition of hypocotyl elongation, the expansion of cotyledons, and the accumulation of photosynthetic pigments. In *Arabidopsis thaliana*, several different classes of photoreceptors mediate light perception, including the phytochromes which perceive red and far red light, cryptochromes, phototropins, and LOV (Light, Oxygen, or Voltage) domain proteins, which are blue light receptors, and UV RESISTANCE LOCUS 8, the most-studied morphogenic photoreceptor for UV-B light (Chory 2010; Kliebenstein et al. 2002; Rizzini et al. 2011; Sharrock and Quail 1989; Zoltowski and Imaizumi 2014). Of these receptors, the photomorphogenic transition is primarily controlled through the actions of the phytochromes and the cryptochromes (Fankhauser and Chory 1997). Through their combined actions, signaling cascades are initiated that significantly modify the expression of at least two thousand genes in *Arabidopsis* (Ma et al. 2001).

While sustained photomorphogenic growth requires the actions of multiple phytochromes and cryptochromes, the initial signaling cascade is

established primarily through phyA, which accumulates to high levels in darkness (Somers and Quail 1995). Upon activation by far red light, phyA is imported into the nucleus through interactions with FAR RED ELONGATED HYPOCOTYL 1 (FHY1) and FHY1-LIKE (FHL) (Hiltbrunner et al. 2006). The physical interaction of phyA with FHY1/FHL is also necessary for phyA to bind further downstream transcription factors that regulate light signaling (Li et al. 2010; Yang et al. 2009). The function of phyA, as well as the photomorphogenic downstream transcription factors, is modulated at multiple levels, including through phyA-mediated protein phosphorylation and targeted, proteasome-mediated degradation. The ubiquitination and targeting of many photomorphogenic proteins for proteasome degradation is regulated through the actions of CONSTITUTIVE PHOTOMORPHOGENESIS 1 (COP1). In the light and in response to the initiation of phytochrome-mediated signal transduction, COP1 protein is excluded from the nucleus, allowing the accumulation of photomorphogenesis-promoting transcription factors (Holm et al. 2002; Stacey et al. 1999; von Arnim and Deng 1994).

One of the most significant targets of COP1 is HY5, a relatively small bZIP transcription factor that regulates photomorphogenesis by activating a large number of further downstream transcription factors (Chattopadhyay et al. 1998; Lee et al. 2007; Oyama et al. 1997). HY5 has also been identified as an integrator of pathways not directly related to light signaling, including

hormone signaling (abscisic acid (ABA) and brassinosteroids), apoptosis, and temperature acclimation (Chen and Xiong 2008; Li and He 2015; Catala et al. 2011; Toledo-Ortiz et al. 2014). Unlike many other genes involved in the light signaling cascade, the phenotypes of *hy5* mutants are not wavelength-specific (Osterlund et al. 2000; Oyama et al. 1997). Other COP1 targets, including the bHLH transcription factor LONG HYPOCOTYL IN FAR-RED 1 (HFR1) and the MYB transcription factor LONG AFTER FAR-RED LIGHT 1 (LAF1), function in a wavelength-dependent manner; while both *hfr1* and *laf1* mutants have reduced responses to far red light, only *hfr1* has a visible phenotype in blue light, and neither mutant exhibits phenotypes in red or white light (Ballesteros et al. 2001; Fairchild et al. 2000; Fankhauser and Chory 2000; Soh et al. 2000; Jang et al. 2005). Further elucidation of the light-signaling cascade has revealed a handful of other transcription factors whose function is necessary for normal photomorphogenic growth and are also regulated by COP1, including the B-box (BBX) containing proteins SALT TOLERANCE HOMOLOG 2 (STH2/BBX21) and LIGHT REGULATED ZINC FINGER1/STH3/BBX22, the bHLH proteins PHYTOCHROME RAPIDLY REGULATED 1 (PAR1) and PAR2, and the *Mutator* transposase-like FAR-RED ELONGATED HYPOCOTYL 3 (FHY3) and FAR-RED IMPAIRED RESPONSE 1 (FAR1) (Datta et al. 2007; Datta et al. 2008; Lin et al. 2007; Zhou et al. 2014). Thus, light perception, and the associated photomorphogenic signaling cascades, converges at a small suite of transcription factors just downstream of a

COP1-mediated hub. From these terminal transcription factors it appears that the signal cascade immediately fans out to thousands of light-regulated genes (Zhang et al. 2011; Shi et al. 2013).

While significant progress has been made identifying and characterizing transcription factors functioning at this COP1-mediated hub, there are likely to be undiscovered pieces in this puzzle. No combination of downstream transcription factor mutants that can phenocopy the *phyA* mutant has been identified – e.g., when grown in far red light, the hypocotyls of *hy5 hfr1 laf1* triple mutants are ~60% as long as the *phyA* mutant (Jang et al. 2013). While this triple mutant has significantly longer hypocotyls than any of the single mutants or double mutant combinations, residual far red light perception is clearly still present. This is in contrast to *fhy1 fhl* double mutants which appear phenotypically identical to both *phyA* mutants and dark grown plants for hypocotyl elongation (Genoud et al. 2008; Hiltbrunner et al. 2006; Zhou et al. 2005). One explanation is that the downstream transcription factor components are already known, but the right combination of mutations has yet to be assembled in a single genotype. For example, as with *hy5 hfr1 laf1* plants, *hy5 sth2 sth3* mutants are also additively defective in light perception (Datta et al. 2008), but no *hy5 hfr1 laf1 sth2 sth3* higher order mutant has been reported. Alternatively, additional transcription factor components may remain unknown. Following two recent publications ((Junker et al. 2012; Huang et al. 2015), we report

here additional strong evidence for the involvement of NUCLEAR FACTOR Y (NF-Y) transcription factors in light perception.

NF-Y transcription factors consist of three unique proteins, called NF-YA, NF-YB, and NF-YC, and each is encoded by a small family of ~10 genes in Arabidopsis (this expansion is mirrored in other sequenced plant species, including monocots and dicots; (Laloum et al. 2013; Petroni et al. 2012)). None of the NF-Y subunits is thought to regulate transcription independently; instead, the mature NF-Y transcription factor is composed of one of each subunit type and all three subunits contribute to DNA binding. NF-YB and NF-YC initially form a dimer in the cytoplasm that translocates to the nucleus where a trimer is formed with NF-YA (Sinha et al. 1996; Sinha et al. 1995; Frontini et al. 2004; Kahle et al. 2005; Liu and Howell 2010). Thus, regulation of any one NF-Y subunit can alter the function of the entire complex. Following nuclear assembly of the mature complex, NF-Ys bind DNA at *CCAAT*-containing *cis* regulatory elements and are typically positive regulators of gene expression (Ceribelli et al. 2008). Although the generalized characterization of NF-Ys (largely from animal and yeast systems) describes them as binding DNA in the proximal regions of promoters, recent data suggests that they also bind more distal regions of promoters to regulate gene expression (Fleming et al. 2013; Cao et al. 2014). In the animal lineage, each NF-Y is usually encoded by only one or two genes and the functional consequences of expanded *NF-Y* gene

families in the plant lineage remains only modestly explored. Nevertheless, much progress has been made in recent years describing the roles of individual NF-Y subunits in the control of specific processes, especially the control of photoperiod-dependent flowering through interactions with CONSTANS (CO) (Kumimoto et al. 2008; Kumimoto et al. 2010; Wenkel et al. 2006), various functions in the development of nitrogen-fixing root nodules in legume species (Baudin et al. 2015; Combier et al. 2006; Zanetti et al. 2011), and abscisic acid signaling during germination and early seedling establishment, often mediated by interactions with bZIP transcription factors (Suzuki et al. 2007; Kumimoto et al. 2013; Siriwardana et al. 2014; Huang et al. 2015; Junker et al. 2012).

Relevant to NF-Y roles in photomorphogenesis and light perception, little is currently known. However, NF-YA5 and NF-YB9 were previously implicated in regulating blue light-dependent transcript accumulation for LIGHT-HARVESTING CHLOROPHYL A/B BINDING PROTEIN (Warpeha et al. 2007) and NF-Y complexes were also shown to bind and regulate the expression of the spinach photosynthetic gene AtpC (Kusnetsov et al. 1999). Further, the promoters of a number of light signaling components were bound by LEC1/NF-YB9 (LEAFY COTYLEDON 1 (Meinke et al. 1994; West et al. 1994)) in chromatin immunoprecipitation experiments, including light harvesting and chlorophyll binding proteins (e.g., LHCA1 and LHCB5) and transcriptional regulators of light perception (e.g., HY5, HY5

HOMOLOG (HYH), and HFR1) (Junker et al. 2012). Finally, alterations in hypocotyl elongation resulting from both *NF-YB* loss of function and inducible overexpression have been observed (Junker et al. 2012; Meinke 1992), including the recent report that LEC1/NF-YB9 regulates skotomorphogenesis through physical interaction with PHYTOCHROME-INTERACTING FACTOR 4 (Huang et al. 2015).

Previous work in our lab identified physical interactions between NF-YC and HY5, as well as other bZIP proteins. Here we extend these initial observations to show that these same NF-YC proteins (NF-YC3, 4, and 9) are broad spectrum regulators of light perception. Interestingly, in the same way that HY5, HFR1, and LAF1 can physically interact, but still appear to signal through independent pathways, *hy5 nf-yc* mutants also show additive - even synergistic - light perception defects. This manuscript characterizes several photomorphogenesis-related phenotypes of NF-YC mutants and proposes that the NF-Y complexes constitute a novel component of the light signaling cascade, functioning at least partially independent of HY5, HFR1, and LAF1. We further demonstrate that *nf-yc* mutants can partially suppress several *cop1* mutant phenotypes and that proteasome regulation of NF-Y complexes during light perception is mitigated through NF-YA subunits. Similar to the multiple regulatory roles of HY5 in light perception and abscisic acid (ABA) signaling, our cumulative research on these three NF-

Y proteins demonstrates that they have essential roles in photoperiod-dependent flowering, ABA perception, and light perception.

Results

Inhibition of hypocotyl elongation in short day, white light, and several individual light wavelengths requires NF-YC

We initially observed slightly elongated hypocotyls in plate grown *nf-yc3-1 nf-yc4-1 nf-yc9-1* triple mutants (hereafter *nf-yc triple*, (Kumimoto et al. 2010)). These visual differences primarily appeared in plants grown for shorter day lengths. To quantify these observed differences, we compared *nf-yc triple* mutants to their parental Columbia (Col-0) ecotype under continuous white light (cWL), short day (SD, 8hrs light/16hrs dark), and continuous dark (cD) conditions (Figure 2.1A-D). While continuous light and continuous dark grown *nf-yc triple* mutants were not significantly different from Col-0, short day grown seedlings had moderately elongated hypocotyls (~50-60% longer, Figure 2.1B-C). To de-convolute the contributions of individual *NF-YC* genes, we additionally examined hypocotyl elongation for the six possible single and double mutants from the three mutant alleles. Modest differences were observed for only the *nf-yc3 nf-yc9* double mutant (Figure 2.1B), although we note that this mutant phenotype was inconsistent in additional experiments. Overall, the data suggested that *NF-YC3, NF-YC4, and NF-YC9* were collectively necessary for the suppression of hypocotyl elongation. Supporting the genetic data showing overlapping functions, all three genes were strongly expressed in the hypocotyl with peak expression in the vascular column (Figure S2.1).

To determine whether the hypocotyl elongation defects were wavelength specific, we additionally examined the same suite of mutants grown in continuous blue (cB), far red (cFR), and red (cR) light conditions (Fig 2.2A-C). In cFR conditions, no significant differences were observed. However, in cR and cB light the *nf-yc triple* mutants were ~50% longer than Col-0. Additionally, significant hypocotyl elongation defects were observed in some single and double mutants (ranging from ~18-29% longer that Col-0). As with the short day white light measurements, differences in the single and double mutants in cB and cR light were less robust between repeated experiments than for the *nf-yc triple* mutants. Interestingly, longer hypocotyls were always associated with the presence of the *nf-yc9* mutant allele – somewhat unexpected as the *nf-yc3-1* and *nf-yc4-1* alleles are strong knockdowns while the *nf-yc9-1* allele retains ~20-25% normal expression levels (Figure S2.2 and (Kumimoto et al. 2010)). Collectively, these data demonstrate that NF-YCs are broad spectrum regulators of light perception.

FLIM-FRET analysis confirms the HY5 by NF-YC9 physical interaction

NF-Y complexes are known to associate with bZIP transcription factors in both plants and animals (Yamamoto et al. 2009; Liu and Howell 2010; Singh et al. 2011; Yoshida et al. 2001). Relevant to light perception, we previously

reported a modest yeast two-hybrid (Y2H) interaction between NF-YC4 and NF-YC9 with HY5 and a non-interaction with NF-YC3, although we assumed the inability to detect an NF-YC3 interaction was likely due to its autoactivation problems in the Y2H system. To further confirm this Y2H data, we performed transient interaction assays in *Nicotiana benthamiana* (Fig 2.3A-C). We utilized fluorescence lifetime imaging (FLIM) and fluorescence recovery after photobleaching (FRAP) to detect fluorescence resonance energy transfer (FRET) between epitope-tagged NF-YC and HY5 proteins. In these experiments, HY5 or NF-YB2 (positive control for interaction with NF-YC) were translationally fused to enhanced yellow fluorescent protein (YFP, (Nagai et al. 2002)) and assayed for FRET against NF-YC3, 4, and 9 fused to modified cerulean 3 (mCer3, (Markwardt et al. 2011)). By comparing fluorescence lifetimes of the donor (mCer3), pre- and post-photobleaching of the acceptor (YFP), we could infer whether or not chosen protein pairs were closely physically associated. Direct physical interaction between proteins was indicated by a significant increase in the lifetime of mCer3 upon YFP photobleaching (Wouters and Bastiaens 1999). Fluorescence recovery of both mCer3 and YFP was monitored during acceptor photobleaching, and was used as an internal control for balancing the destruction of YFP signal and the preservation of mCer3 signal during experimentation (Figure 2.3B). After identifying a proper photobleaching regimen, in pairs of known interacting proteins we observed that mCer3 signal would increase over the course of the initial photobleaching event,

but not over subsequent treatments (Figure 2.3B). This was consistent with what is expected when observing FRET, as a significant majority of the acceptor (YFP) is destroyed in the initial photobleaching event, and further photobleaching events have a reduced effect on the already diminished pool of YFP.

As a positive interaction control, we initially tested NF-YB2:YFP by NF-YC3, 4, and 9:mCer3 and were able to consistently detect significantly increased mCer3 fluorescence lifetimes after YFP photobleaching (Figure 2.3C). This is consistent with previous publications showing strong Y2H and *in vivo* NF-YB by NF-YC interactions (Wenkel et al. 2006; Hackenberg et al. 2012; Kumimoto et al. 2010; Cao et al. 2011a; Cao et al. 2011b). As a negative control for each interaction test, we demonstrated that when NF-YB2 lacked the YFP fusion, mCer3 lifetimes were not altered after a photobleaching treatment (Figure 2.3C). Substituting HY5:YFP for NF-YB2:YFP demonstrated that NF-YC9 could consistently physically interact with HY5; however, no FLIM-FRET interaction was detected between NF-YC3 or NF-YC4 and HY5. Thus, it remains possible that HY5 only interacts with a subset of the light perception-regulating NF-YC proteins described here (see Discussion).

NF-YC and HY5 genetically interact to suppress hypocotyl elongation in white light

With the knowledge that at least some NF-YCs can physically interact with HY5, we generated *nf-yc triple hy5* mutants and examined them for hypocotyl elongation phenotypes in both SD and cWL (Figure 2.4A-C). Surprisingly, in SD conditions the *nf-yc triple hy5* mutants were considerably longer than either parental mutant line, suggesting that the previously observed NF-YC roles in hypocotyl elongation were at least partially independent of HY5. Even more striking was the strongly synergistic increase in hypocotyl elongation in cWL in the *nf-yc triple hy5* mutants over both mutant parents (Figure 2.4C). Compared to parental Col-0, dark grown plants showed no differences in hypocotyl elongation for any of the mutant genotypes (Figure 2.4D). Rescue assays confirmed that each gene (*NF-YC3*, *4*, *9* and *HY5*) was capable of significantly suppressing the *nf-yc triple hy5* elongated hypocotyl phenotype (Figure S2.3). Collectively, these data demonstrated that the presence of *HY5* masked the effects of the *nf-yc triple* mutant on hypocotyl elongation, especially in cWL conditions. These results are not trivially explained by cross regulation between NF-YC and HY5 as their transcription levels are only altered in their own mutant backgrounds (Figure S2.2). Finally, because some commercial white light sources contain contaminating UV radiation, we additionally examined hypocotyl lengths of cWL-grown plants grown under Mylar to filter out UV light. In accordance with previous work, *hy5* mutants were longer in the absence of UV (Huang et al. 2012); however, no difference was detected in the *nf-yc triple* mutant, and while not statistically significantly different, the minor

difference in *nf-yc triple hy5* mutants can be completely accounted for by the loss of *HY5* (Figure S2.4).

NF-YC3, 4, and 9 and HY5 have both shared and independent regulatory targets

To further dissect the genetic relationship between NF-YC and HY5, we compared the transcriptome profiles of seven day old, cWL grown *nf-yc triple*, *hy5*, and *nf-yc triple hy5* mutant seedlings using RNA Sequencing (RNA-Seq, NCBI GEO accession GSE81837).

When compared to wild type, *hy5* mutants had 1368 up-regulated and 941 down-regulated genes, whereas analysis of differentially expressed genes in the *nf-yc triple* mutant showed a smaller set of 645 up-regulated genes and 493 down-regulated genes (at least 1.5 fold, adjusted $p < 0.05$, Table S2.1). Direct comparison of the *hy5* and *nf-yc triple* down-regulated genes showed substantial overlap, with approximately 40% of the *nf-yc triple* down-regulated genes being contained in the *hy5* data set (Figure 2.5A). Gene-ontology (GO) analysis for genes down-regulated in both the *nf-yc triple* and *hy5* mutants identified enrichment in many categories involved in photomorphogenesis and early seedling development, including response to light stimulus and pigment biosynthetic processes (Table S2.2). Comparison between up-regulated gene sets yielded similar results with ~50% shared between the *nf-yc triple* and *hy5* (Figure 2.5B). GO

enrichment analyses of genes up-regulated in both *nf-yc triple* and *hy5* yielded categories in cellular stress responses and cellular responses to hormones, including ethylene, salicylic acid, and jasmonic acid (Table S2.3).

To further investigate the regulatory relationship between HY5 and the NF-YCs, we analyzed genes that were either significantly up-regulated or down-regulated in *nf-yc triple hy5* mutants (Table S2.1). These data sets were then sub-divided into four groups based on the level of differential gene expression in the *nf-yc triple hy5* mutant relative to both *nf-yc triple* and *hy5*: **Group I)** Genes differentially expressed more in the *nf-yc triple hy5* mutant than both *nf-yc triple* and *hy5*; **Groups II-III)** Genes differentially expressed more in the *nf-yc triple hy5* mutant compared only to the *nf-yc triple* (**II**) or *hy5* (**III**); and **Group IV)** Genes not differentially expressed compared to either the *nf-yc triple* or *hy5* (i.e., still differentially expressed in the quadruple mutant relative to wild type, but no change from *nf-yc triple* and *hy5* (Figure 2.5C-D)). GO enrichment analyses of these four groups represent putative biological processes that NF-YCs and HY5 regulate cooperatively (genes more differentially expressed in *nf-yc triple hy5* than parental lines) and independently (genes not differentially expressed in *nf-yc triple hy5* relative to *nf-yc triple* and/or *hy5* (Table S2.2 and Table S2.3)).

Analysis of genes significantly more down-regulated in the *nf-yc triple hy5* mutant (relative to its parental genotypes) identified several over-represented categories, including flavonoid biosynthesis and polyol metabolic processes (Table S2.2). Among genes up-regulated to a similar level in the *nf-yc triple hy5, nf-yc triple,* and *hy5* data sets, and consistent with the synergistic hypocotyl phenotype of the *nf-yc triple hy5* mutant, there was a significant enrichment for genes involved in cell wall organization, cell wall biogenesis, and cell wall macromolecule metabolic processes (Table S2.3). Taken together, these data identify putative targets and biological processes regulated both cooperatively and independently by NF-YCs and HY5, solidifying the existence of a complex functional relationship.

Longer hypocotyls in nf-yc triple hy5 mutants are largely a function of greatly increased cell elongation

Previous research established that the elongated hypocotyls in *hy5* mutants are directly related to increased epidermal cell length (Oyama et al. 1997); therefore, we additionally examined individual files of epidermal cells along the hypocotyls of *nf-yc triple hy5* mutants for total cell number and mean cell length (Figure 2.6A-C). The mean length of individual epidermal cells in *hy5* (82μm) was ~90% greater than Col-0 (43μm), while cells in the *nf-yc triple* mutant measured only ~15% longer than Col-0. Reflecting the synergistic hypocotyl elongation phenotypes of *nf-yc triple hy5* mutants, the epidermal cells of the quadruple mutant (158μm) were ~270% longer than

those measured in Col-0. Total epidermal cells in the quadruple mutant were also increased >60% compared to Col-0. Therefore, the very long hypocotyls of *nf-yc triple hy5* mutants can be explained by a combination of modestly increased cell count and highly increased cell elongation.

NF-YC hypocotyl elongation phenotypes in monochromatic light are additive with HY5 and do not completely overlap with those of HFR1 and LAF1

HY5 regulates photomorphogenesis regardless of wavelength, whereas HFR1 and LAF1 are more specific to FR light responses (Ballesteros et al. 2001; Fairchild et al. 2000; Fankhauser and Chory 2000; Soh et al. 2000; Jang et al. 2005). To better compare the spectrum of *nf-yc* mutant defects to these other transcription factors, we first examined both the *nf-yc triple* and *nf-yc triple hy5* lines in cB, cFR, and cR over a gradient of light intensities (Figure 2.7A-C). Under all but the lowest cB fluence rates, the *nf-yc triple hy5* mutants had significantly longer hypocotyls than all other lines (Figure 2.7A). Considering our previous observation that *nf-yc triple* light perception defects were only apparent in SD conditions (Figure 2.1B), it was somewhat surprising to find that *nf-yc triple hy5* defects in cB were most pronounced at the highest light intensities (Figure 2.7A).

The *nf-yc triple hy5* mutants were significantly longer than their *nf-yc triple* and *hy5* parental lines under all cFR conditions (Figure 2.7B). One possible

cause for defects in FR light perception could be differential expression of *HY5*, *HFR1*, or *LAF1* in *nf-yc* mutants – i.e., NF-Y complexes could control the expression of these genes. Figure S2.2 shows that *HY5* is not differentially expressed in an *nf-yc triple* background in cWL and we additionally examined the expression of *HY5*, *HFR1*, and *LAF1* in cFR grown plants. Consistent with previous reports (Jang et al. 2013; Jang et al. 2007), modest differences in *HY5*, *HFR1, and LAF1* were either insignificant or not reproducible in repeated expression analyses in the various mutant backgrounds (Fig 7D). We conclude that *nf-yc* mutant phenotypes are not likely related to simple changes in the expression of these well-known regulators of light perception. Further, as discussed below, *nf-yc triple* mutants appear to have a different spectrum of light defects than either *hfr1* or *laf1* mutants.

Under low fluence rate cR and cFR, the *nf-yc triple* mutant alone had significantly longer hypocotyls than Col-0, which is similar to *hy5* (Figure 2.7B-C). However, *hfr1* and *laf1* are only reported to have defects in cFR (both) and cB light (*hfr1*). To confirm these previous reports with our experimental conditions, we directly compared *hfr1* and *laf1* to *nf-y* mutants under low fluence rate cR (Figure 2.7E-F). As previously reported, *hfr1* and *laf1* appeared identical to wild type Col-0 plants, whereas the *nf-yc triple* mutants were consistently ~40% longer than Col-0 and similar to *hy5* mutants. We additionally compared the *nf-yc triple* to *hfr1* and *laf1* in SD

conditions (Figure 2.7G-H). As expected, the *hfr1* and *laf1* mutants appeared phenotypically identical to Col-0, while the *nf-yc triple*, *hy5*, and *nf-yc triple hy5* mutants were all significantly longer. Collectively, our data suggests that NF-YCs regulate hypocotyl elongation via an independent pathway(s) from HY5, and at least partially independent of HFR1 and LAF1, with broad roles in light perception at variable fluence rates.

Loss of NF-YC function partially suppresses cop1 mutant phenotypes

In the dark, HY5, HFR1, and LAF1 are all targeted for degradation by the proteasome in a COP1-dependent manner (Ang and Deng 1994; Jang et al. 2005; Seo et al. 2003). COP1 mutants have short hypocotyls and other photomorphogenic phenotypes even when grown in the dark, and these phenotypes are partially suppressed in *hy5 cop1, laf1 cop1,* and *hfr1 cop1* double mutants (Ang and Deng 1994; Kim et al. 2002; Yang et al. 2005). Therefore, we examined if the *nf-yc triple* mutation could also suppress the short phenotype of dark-grown *cop1-4* mutants. Similar to *cop1 hy5*, an *nf-yc triple cop1* mutant had ~80% longer hypocotyls than the *cop1* single mutant when grown in constant darkness (Figure 2.8A). Because *cop1-4* mutants are known to have reduced rosette diameters (dwarf phenotype) and early flowering (Liu et al. 2008), we further characterized these phenotypes in *nf-yc triple cop1* mutants. For rosette diameter, the *nf-yc triple* mutant was once again able to partially suppress *cop1* (Figure 2.8B).

One possibility is that this suppression is simply a function of *nf-yc triple* mutants being late flowering – i.e., because they are later flowering, the rosettes have time to achieve a greater diameter prior to the phase change to reproductive growth. However, a control cross between *cop1* and an even later flowering *constans* mutant (the alternatively named *co-sail* or *co-9* allele (Laubinger et al. 2006)) had no impact on rosette diameter when crossed to *cop1-4*. This suggests that *nf-yc* loss of function alleles genuinely suppress the small *cop1* rosette diameter phenotype and that this NF-YC function is genetically separable from its role in flowering time.

Finally, we also tested whether the *nf-yc triple* mutant could suppress the early flowering phenotype of *cop1* (Figure 2.8C). The *nf-yc triple cop1* mutant plants flowered moderately, but significantly, later than Col-0, intermediate to the early flowering *cop1-4* mutant and late flowering *nf-yc triple* mutant. This result is not surprising as an important role for COP1 in flowering is to suppress CO function via protein degradation (Liu et al. 2008; Jang et al. 2008). Because CO and NF-Y function together to regulate photoperiod-dependent flowering (Cao et al. 2014; Kumimoto et al. 2013; Wenkel et al. 2006; Ben-Naim et al. 2006), the basis of early flowering in *cop1* is largely a function of CO protein (and potentially NF-Y, see below) hyper-accumulation (Jang et al. 2008; Liu et al. 2008). Measurements of *FLOWERING LOCUS T* (*FT*) expression – the regulatory target of CO and

NF-Y function in flowering time – correlated with expectations from the flowering time measurements (Figure 2.8D).

Like HY5, HFR1, and LAF1, NF-YA proteins are degraded in the dark

NF-YC regulation of light perception appears to share many parallels with HY5, HFR1, and LAF1, including the suppression of *cop1* mutant phenotypes (Yang et al. 2005; Ang and Deng 1994; Kim et al. 2002). As with these other photomorphogenic transcription factors, it is tempting to speculate that NF-YC proteins might be targets of COP1-mediated proteasome degradation in the dark. However, this does not appear to be the case as native antibodies to both NF-YC3 and NF-YC4 show modest fluctuations, but largely stable accumulation throughout both short day and long day cycles (Figure S2.5– recall also that expression of any *one* NF-YC from a native promoter rescues the mutant phenotype, Figure S2.3). Nevertheless, NF-YC proteins function within the context of a heterotrimeric complex and reduction of the NF-YA or NF-YB components could also disrupt activity.

In this regard, overexpression of most NF-YAs leads to small, dwarf phenotypes that are not unlike those observed for *cop1* mutants (Siriwardana et al. 2014; Leyva-Gonzalez et al. 2012). In fact, when we examined *NF-YA* overexpressing plants (35S promoter driven), they were

found to have significantly shortened hypocotyls in both cD and cR conditions (Figure 2.9A). While shortened hypocotyls in cD is a classic constitutive photomorphogenic phenotype, expressing *p35S:NF-YA* hypocotyl lengths in cR as a percentage of cD growth additionally showed that most of these plants were specifically defective in red light perception (Figure 2.9A).

While it is unknown which of the 10 Arabidopsis NF-YAs is natively involved in hypocotyl elongation, two recent publications suggested that NF-YA2 may be found in complex with NF-YC3, 4, and 9 (Hou et al. 2014; Siriwardana et al. 2016). Therefore, using qPCR, we examined the expression of *NF-YA2* in 24hr cWL or after 24-48hrs of cD and found that expression was strongly down-regulated in cD (Figure 2.9B). At the same time we compared *NF-YA2* expression to a subset of other *NF-YA* genes - *NF-YA1*, *7*, *9*, and *10*. *NF-YA10* is the most closely related paralog to *NF-YA2* (encoding 63% identical full length proteins, (Siefers et al. 2009)) and it showed the same pattern of down-regulation in cD. However, the less related *NF-YA1* and *9* genes (42% identical to each other, but only 23-24% to NF-YA2) remained stably expressed in cD, while *NF-YA7* was actually up-regulated. Thus, expression of the *NF-YA* gene family in response to cD is quite variable, and suggests potential for light regulated accumulation and depletion.

To determine if NF-YA proteins might be targets of degradation in the dark, we examined the accumulation of NF-YA2 and NF-YA7 expressed from constitutive 35S promoters. We chose to use a constitutive promoter to differentiate between changes in protein accumulation due to reduced gene expression (see Figure 2.9B) versus active protein degradation processes. Strongly suggesting an active degradation process, NF-YA2 protein accumulation was strongly reduced in cD conditions, even when expression was driven from the 35S promoter (Figure 2.9C). This was in stark contrast to NF-YA7, which had stable protein accumulation in the dark. To determine if the proteasome was involved in the process, we additionally performed cell-free protein degradation assays (as previously described) and determined that NF-YA2 was rapidly degraded (Figure 2.9D). However, the addition of the proteasome inhibitor MG132 strongly reduced the apparent degradation of NF-YA2 protein. We note that NF-YA7 also degraded in an MG132 dependent manner, suggesting that it can also be targeted by the proteasome for degradation, even if darkness may not be the driving force (Figure 2.9B-C). Collectively, these data suggest that NF-YAs can also regulate light perception and are targeted for proteasome mediated degradation, perhaps controlling the overall stability of the NF-Y complexes necessary to suppress hypocotyl elongation in the light.

Discussion

The physical and genetic relationships between the NF-YCs and HY5 are similar to those reported for several other photomorphogenic transcription factors, such as HFR1 and LAF1, where both are able to physically interact with HY5, but have clearly additive light perception defects when combined in higher order mutants (Kim et al. 2002; Ballesteros et al. 2001; Jang et al. 2013; Jang et al. 2008). While we observed *nf-yc triple* mutant phenotypes primarily in low-intensity monochromatic light, the most pronounced *nf-yc triple hy5* phenotypes (relative to both the *nf-yc triple* and *hy5* mutants) were seen in high-intensity monochromatic light. Because HY5 protein accumulation and activity are light-intensity dependent (Osterlund et al. 2000), one possible explanation for this relationship, supported by our RNA-Seq data, is that HY5 and the NF-Y complexes share a subset of regulatory functions. When *nf-yc triple* mutants are grown in low light, HY5 function becomes essential for maintaining normal photomorphogenesis; however, under these same conditions, HY5 is also increasingly targeted for degradation and the *nf-y*-related photomorphogenic defect becomes apparent. Alternatively, in saturating, high light conditions, the relative NF-Y contribution to photomorphogenesis is masked due to high accumulation and activity of HY5 and other photomorphogenic transcription factors.

Are NF-YC by HY5 interactions important in light signaling?

Unexpectedly, while we were able to detect a physical interaction between NF-YC9 and HY5 through FRET-FLIM analyses, we were not able to detect an interaction between NF-YC3 or NF-YC4 and HY5. This is surprising because the histone fold domains of NF-YC3, 4, and 9 are nearly identical (in fact , those of NF-YC3 and NF-YC9 are identical (Siefers et al. 2009)); however, the amino- and carboxy-terminal regions are more divergent and could be involved in the NF-YC by HY5 physical interaction. Because of the extreme spatial constraints required for FRET to occur, it is not valid to conclude from these experiments that NF-YC3 and NF-YC4 cannot interact with HY5 (Bastiaens and Squire 1999; Müller et al. 2013), and Y2H analyses did previously show a positive interaction between NF-YC4 and HY5 . The question also remains whether or not the ability of NF-YCs to physically interact with HY5 is of biological importance relative to their specific functions in light signaling. Both HY5 and the NF-YCs are also regulators of ABA signaling (Chen and Xiong 2008; Kumimoto et al. 2013; Siriwardana et al. 2014) and it is possible that a physical interaction between them is only related to this or another undefined pathway. This possibility is supported by the additive, and even synergistic, mutant phenotypes of the *nf-yc triple hy5* plants – i.e., if these proteins are physically interacting in a linear pathway or at a common hub in light signaling, how does the quadruple *nf-yc3 nf-yc4 nf-yc9 hy5* mutation result in these synergistic phenotypes? Alternatively, arguing for the relevance of physical interactions in light signaling, there is clearly a significant amount

of overlap in the putative regulatory targets of NF-YC3, 4, 9 and HY5. Similarly, it was also previously suggested that subsets of photomorphogenic responses might be regulated by combinations of both overlapping (where physical interactions were relevant) and non-overlapping functions between HY5, HFR1, and LAF1 (Jang et al. 2013). In future experiments, it will be informative to examine the stability of NF-YA proteins in the presence or absence of these other transcriptional regulators as it was previously shown that HFR1 and LAF1 were co-dependent for their protein stability (i.e., they required each other to avoid proteasome-mediated degradation; (Jang et al. 2007)). Ultimately, deciphering these putative cooperative versus individual roles remains an exciting challenge for future research.

Is there a separate pathway for NF-YC function in light perception?

While we have demonstrated that NF-YC3, 4, and 9 function at least partially independently of HY5 in light perception, one pressing question is whether the NF-Y complex is functioning through other known light-responsive transcription factors, such as HFR1 or LAF1. Directly addressing this hypothesis would require the creation of *nf-yc triple hfr1* and *nf-yc triple laf1* quadruple mutants; however, because *HFR1* is linked to *NF-YC9*, traditional crossing techniques would be prohibitively difficult. Therefore targeting of these loci in the *nf-yc triple* mutant with CRISPR-Cas9 is

currently underway. To support that the NF-YCs are functioning separately, or at least differently, from HFR1 or LAF1, we identified strong phenotypes in the *nf-yc triple* mutant in SD- and low cR-grown seedlings, where *hfr1* and *laf1* showed no mutant phenotype. Additionally, we found no differential expression of *HFR1* or *LAF1* in cFR-grown *nf-yc triple* mutants. These data suggest that the function of NF-YCs in light perception is at least partially separable from the function of HFR1 and LAF1.

Similar to the NF-YCs, both STH2/BBX21 and STH3/BBX22 function as photomorphogenesis-activating transcription factors over a broad range of light conditions and are also able to physically interact with HY5 (Datta et al. 2007; Datta et al. 2008). It is possible that the NF-YCs are functioning in an STH2/STH3-dependent manner; however, phenotypes of *sth2 sth3* double mutants and *sth2 sth3 hy5* triple mutants suggest that this might not be the case. In contrast to the genetic relationship between the *nf-yc triple* and *hy5*, the hypocotyls of *sth2 sth3 hy5* triple mutants were not longer than *hy5* in cR. Further, we observed the most severe *nf-yc triple* phenotypes in low-intensity light while the *sth2 sth3* mutant phenotypes were only observed in high-intensity light (Datta et al. 2008). Nevertheless, these observations do not preclude genetic interactions for subsets of shared functions, similar to what we have already suggested with HY5.

When examining the protein domains of STH2/BBX21 and STH3/BBX22, it is tempting to speculate that there may be indirect physical interactions with the NF-YC proteins as part of a larger light perception complex. STH2 and STH3 have B-box domains, thus their alternate BBX21 and BBX22 designations (Khanna et al. 2009), and these domains are necessary for direct physical interactions with HY5 (Datta et al. 2007; Datta et al. 2008). BBX proteins also often have so-called CO, CO-LIKE, and TIMING OF CAB (CCT) domains (Khanna et al. 2009; Gangappa and Botto 2014). For example, CO (BBX1) is a BBX-CCT protein and mutations in either of these domains impacts its ability to regulate flowering time (Robson et al. 2001). It is well-established that NF-YC3, 4, and 9 can all physically interact with CO and the CCT domain is both necessary and sufficient for this interaction (Wenkel et al. 2006; Kumimoto et al. 2010; Cao et al. 2014). While STH2 and STH3 do not have a CCT domain, recent evidence demonstrated that BBX proteins can heterodimerize with other BBX proteins (Qi et al. 2012). Therefore, it is possible that NF-Y complexes may interact with BBX-CCT proteins via the CCT domain and recruit other non-CCT containing BBX proteins, such as STH2 and 3, to these complexes.

Are NF-Y complexes activators of photomorphogenesis or skotomorphogenesis? Or both?

In contrast to our genetic evidence showing that NF-Ys act as suppressors of hypocotyl elongation, NF-YB9/LEC1 appears to have the opposite role.

This idea comes from recent evidence demonstrating that inducible overexpression of *NF-YB9/LEC1* also resulted in elongated hypocotyls, suggesting that NF-YB9/LEC1 might actually function as an enhancer of hypocotyl elongation (Junker et al. 2012; Baumlein and Junker 2012). Consistent with this finding, embryonic hypocotyls are shortened in *lec1* mutants (West et al. 1994). Further, recent data shows that the hypocotyls of both light and dark grown *lec1* mutants are significantly shorter than wild type plants (Huang et al. 2015). Interestingly, overexpression of a repressor of photomorphogenesis - PHYTOCHROME-INTERACTING FACTOR 4 - results in elongated hypocotyls, but this phenotype is partially dependent on the presence of *LEC1* (Huang et al. 2015). These results raise a few interesting questions: why would NF-YB9 act opposite to the NF-YA and NF-YC members of the complex (as reported here) and what mechanism would allow this result? Considering these questions, it is important to remember that each NF-Y subunit – A, B, and C – is part of a 10 member family. Thus, many unique NF-Y complexes could theoretically form and, depending on their composition, actually act to competitively suppress or enhance a given process. In this scenario, some members of a given NF-Y family might enter a complex, but render it inactive, while other members of the same family would have the opposite effect.

Our previous research on ABA-mediated seed germination provides some precedence for the above idea. We demonstrated that members of the

same NF-Y family can act in opposing manners, either enhancing or delaying germination when their expression is altered (both NF-YA and NF-YC examples exist; (Siriwardana et al. 2014; Kumimoto et al. 2013)). Similarly, some BBX proteins also show these opposing functionalities. For example, BBX24 and 25 are hypothesized to interfere with BBX22 function by entering into non-functional complexes with HY5 (Gangappa et al. 2013). Fitting this scenario nicely, NF-YB9/LEC1, and its closest relative NF-YB6/LEC1-LIKE, are quite unique and very different from the other eight NF-YB proteins in Arabidopsis. This includes 16 amino acid differences in their highly conserved histone fold domains that are completely unique to only this pair (Kwong et al. 2003). However, an alternative hypothesis must be considered related to the most recent *lec1* data (Huang et al. 2015): *lec1* shortened hypocotyls may not be developmental patterns related to loss of skotomorphogenesis or post embryonic in nature, but, instead, are lasting patterns laid down during embryogenesis. This possibility is supported by both the modest magnitude of the effects and the finding that *lec1* plants are short in all conditions (dark and light). This is not the case for the *nf-yc* mutants reported here as they are indistinguishable from wild-type plants in the dark and elongated in the light, clearly defining them as positive regulators of photomorphogenesis.

NF-Y complexes share the hallmarks of photomorphogenic transcription factors

While the functional relationship between the NF-YCs and HY5 is similar to that observed with many other photomorphogenic transcription factors, the NF-YCs do not appear to be transcriptionally or translationally regulated in a manner consistent with light-responsive proteins; however, because the NF-YCs act in the larger context of an NF-Y trimer, the physical properties and regulatory components of the functional unit can be spread across multiple proteins. We showed that several NF-YA subunits with photomorphogenic phenotypes are regulated by light, and that NF-YA2 is targeted for degradation by the proteasome. Regulation of the NF-YA subunit establishes NF-Y complexes as possessing all of the properties expected of a photomorphogenic transcription factors, including DNA-binding capacity, the ability to physically interact with other photomorphogenic factors, and a light-regulated mechanism to modulate function and abundance. While specific NF-YA subunits have not been identified to natively regulate the inhibition of hypocotyl elongation, a recent publication showed that over-expression of NF-YA2 led to earlier flowering (Hou et al. 2014). This suggests that NF-YA2 could be integrated into an NF-Y complex containing NF-YC3, 4, and/or 9, as each is redundantly involved in photoperiod-dependent flowering (Kumimoto et al. 2010); finally, an NF-YA2/NF-YB2,3/NF-YC9 trimer has been identified through yeast three-hybrid analyses, and further verified through two-way interaction assays (including BiFC and co-IP, (Hou et al. 2014)). The identity of photomorphogenic NF-YB proteins remains unknown and it will be

interesting to determine which, if any, non-LEC-type NF-YB will be positive regulators of photomorphogenesis.

Conclusion

The data presented here firmly establishes NF-Y complexes as positive regulators of photomorphogenesis, significantly extending recent findings (Junker et al. 2012; Huang et al. 2015). Future research on the potential regulation of NF-YA proteins by the proteasome and the identity of photomorphogenic NF-YA and NF-YB will improve our current understanding. Although only discussed at a cursory level here, research on NF-Y roles in flowering time demonstrate important interactions with the BBX protein CONSTANS and suggest that NF-Y by BBX interactions may be generalizable (Wenkel et al. 2006; Kumimoto et al. 2008; Kumimoto et al. 2010; Cao et al. 2014; Yamawaki et al. 2011). Given the numerous roles for BBX proteins in light perception (Bai et al. 2014; Gangappa et al. 2013; Xu et al. 2014a; Datta et al. 2007; Datta et al. 2006; Datta et al. 2008), we predict that future studies will uncover BBX by NF-Y interactions that are essential for light perception. This would be an exciting finding, significantly extending the regulatory reach and capacity of the four interacting families of proteins.

Methods

Growth Conditions and Plant Lines

All plants were of the Col-0 ecotype and were grown at 22°C. Prior to starting germination on plates or soil, seeds were cold-stratified in a 4°C walk-in cooler in the dark for 2-3 days. Plants grown in continuous white light were grown in a Conviron ATC13 growth chambers or a custom walk-in growth chamber. Plants in single wavelength light experiments were grown in a Percival E30-LED growth chamber after initial exposures to 4 hours of white light to induce germination. Plants used in flowering-time experiments, rosette diameter measurements, and GUS staining were grown in a previously-described soil mixture (Kumimoto et al. 2010). All other plants were grown on 0.8-2% agar plates supplemented with Gamborg B-5 Basal Medium (PhytoTechnology Laboratories, product #G398). *Nicotiana benthamiana* plants were grown under long-day conditions (16h light/8h dark) at 22℃ in a Conviron ATC13 growth chamber. For UV experiments, plants were grown in cWL for 5 days under a Mylar filter (Professional Plastics, catalog #A736990500) or mock-filter.

GUS Staining, Rosette Diameter Measurements, and Flowering Time Experiments

GUS staining was performed as previously described on 5 day old soil-grown seedlings, and images were taken on a Leica dissecting stereoscope. Flowering time was measured as the total number of rosette

and cauline leaves present shortly after bolting, and all genotypes exhibited similar developmental rates. To quantify rosette diameter, plants were photographed from above at the time of bolting, and the Feret's diameter was measured, anchored at the tip of the longest rosette leaf.

Hypocotyl Length and Cell Length Measurements

To measure hypocotyl elongation, seeds were sown onto B5 supplemented plates with 2% agar and cold-stratified at 4°C for 2 days. Before transfer to specific light conditions, all plates were set at room temperature in continuous white light for 4 hours. Plates were grown vertically for the duration of the experiments. Germination rates for the *nf-y* mutants under study were previously shown to be the same in B5 media and confirmed for the experiments reported here (Kumimoto et al. 2013). To facilitate proper measurement, all plants were straightened on the plate before taking pictures. Pictures were processed, and all individual hypocotyls were traced and measured in FIJI (Schindelin et al. 2012).

For individual cell length and total cell number measurements of single files of hypocotyl cells, seedlings were grown on plates as described above for 5 days in cWL. Seedlings were fixed in an FAA solution and dehydrated through sequential 30-minute incubations in 90% and 100% ethanol (Zeng et al. 2007). Fixed and dehydrated seedlings were individually mounted in the clearing agent methyl salicylate (Zeng et al. 2007), and immediately

taken for measurement on a Nikon Eclipse NI-U compound microscope. Using Differential Interference Contrast (DIC) optics, individual cell files were identified and measured manually through the NIS-Elements BR software, and pictures were taken at the hypocotyl-cotyledon junction for every seedling measured.

qRT-PCR Analyses

Total RNA was isolated from 7-day-old seedlings grown under continuous white light conditions and from 5-day-old seedlings grown under continuous far red light conditions using the Omega Biotek E.Z.N.A Plant RNA Kit (catalog #R6827-01), and was DNase treated on-column with Omega Biotek's RNase-free DNase set (catalog #E1091). First-strand cDNA synthesis was carried out with Invitrogen's SuperScript III Reverse Transcriptase (catalog #18080-044) and supplied oligo dT primers. qRT-PCR was performed on a Bio-Rad CFX Connect Real-Time PCR Detection System (http://www.bio-rad.com/), using Thermo Scientific's Maxima SYBR Green/ROX qPCR Master Mix (catalog #K0222). Each genotype was assayed with three independent biological replicates, consisting of approximately 100mg of starting tissue each. White light grown seedlings were normalized to *At2g32170*, while far red light grown seedlings were normalized to *At3g18780* and *At1g49240*. Statistical analysis and comparisons between samples was performed in the Bio-Rad CFX

Manager Software (http://www.bio-rad.com/) through use of the $2^{(-\Delta\Delta CT)}$ method.

Transient Transformation of *Nicotiana benthamiana* Leaves

The leaves of 4- to 6-week old *N. benthamiana* were co-infiltrated with *Agrobacterium tumefaciens* GV3101 strains harboring either a YFP-fused protein or an mCer3-fused protein, in addition to the *Agrobacterium* strain C58C1 harboring the viral silencing suppressor helper complex pCH32 (Hamilton 1997). Before infiltration into *Nicotiana* leaves, *Agrobacterium* cultures grown overnight were treated with 200uM acetosyringone in a modified induction buffer for 4 hours (Mangano et al. 2014). This induced culture was re-suspended in 10mM MES 10mM MgSO4 and directly infiltrated into young leaves. All downstream analyses were conducted 2-4 days after initial infiltration.

FRET-FLIM and FRAP Analyses

FLIM data was acquired through time-correlated single photon counting (TCSPC) on a Lecia TCS SP8 confocal laser scanning microscope using an HC PL APO 40x/1.10 water immersion objective. Fluorescent protein excitation was achieved through use of a titanium-sapphire multiphoton laser (Chameleon, Coherent) operating at 120 femtosecond pulses of 858nm infrared light. Fluorescence emissions were detected by non-descanned hybrid detectors (HyDs). Fluorescence lifetimes of entire nuclei

were fit to a single-exponential model through the SymphoTime 64 (www.picoquant.com) software, and comparison of the fluorescence lifetimes before and after FRAP was used to detect FRET. For FRAP analyses, YFP photobleaching was accomplished with a high-intensity Argon laser line at 514nm for 15 seconds, followed by recovery imaging of both mCer3 (excited at 458nm) and YFP (excited at 514nm) every second for 5 seconds. Descanned HyDs were used to detect mCer3 emission from 459nm to 512nm, and a Photomultiplier Tube (PMT) was used to detect YFP emission from 512 to 562nm, with a 458/514 notch filter in place. This process was performed a total of 3 times for each nucleus, and both mCer3 and YFP intensities were calculated relative to initial fluorescence intensity. FRAP was conducted as an internal control during FLIM measurements, allowing us to assess the level of YFP photobleaching and ensure that relatively little mCer3 was inadvertently photobleached.

Protein Work

Total protein was extracted from 14-day-old plants by grinding in lysis buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, pH 8.0, 1% Triton X-100, 1% SDS with fresh 5 mM DTT, and 100 μM MG132). NF-YA-CFP/HA was probed with high affinity anti-HA primary antibody (cat#11 867 423 001; Roche) and goat anti-rat secondary antibody (cat#SC-2032; Santa Cruz Biotechnology). NF-YC3 and NF-YC4 were detected by previously described native antibodies (Kumimoto et al. 2010). The Bio-Rad

ChemiDoc XRS imaging system was used for visualizing the protein blot after incubations with ECL plus reagent (cat#RPN2132; GE Healthcare). Equivalent loading and transfer efficiency were determined by staining the protein blot with Ponceau S (cat#P3504; Sigma-Aldrich).

RNA Sequencing and Analysis

Seedlings were grown for seven days on B5 media in continuous white light. Total RNA was isolated using the E.Z.N.A. Plant RNA Kit from (Omega Biotek, Cat#R6827). To ensure low levels of contaminating ribosomal RNA, two rounds of poly-A mRNA purification were performed using the μMACS mRNA Isolation Kit (Miltenyi Biotech, Cat#130-090-276). Indexed RNA-Seq libraries were prepared from 100 ng of poly-A RNA starting material using the NEXTflex Illumina qRNA-Seq Library Prep Kit (Bioo Scientific, Cat#5130). Sequencing of 150 bp paired end reads was performed on an Illumina HiSeq 2500 in rapid output mode at the Texas A&M Agrilife Research Facility (College Station, TX). Sample de-multiplexing was performed using CASAVA software v1.8.2 and bcl2fastq was performed using conversion software v1.8.4.

Resulting sequences were trimmed and quality checked using the pipeline detailed at the iPlant Collaborative Discovery Environment (http://www.iplantcollaborative.org). Sequences were mapped to the TAIR 10 representative gene models set using Burrows-Wheeler Aligner (Li and

Durbin 2010, 2009) within iPlant. Differential gene expression was determined using the Bioconductor package edgeR (Robinson et al. 2010). Gene Ontology over-representation analyses were performed in AmiGO 2 version 2.3.2 (Ashburner et al. 2000; Gene Ontology 2015). Raw sequencing data and the final differentially expressed gene lists were deposited with NCBI's Gene Expression Omnibus, accession number GSE81837.

Image Processing and Figure Construction

All image processing and figure construction were performed in either FIJI, Photoshop (www.adobe.com), or Prism (www.graphpad.com).

Accession Numbers

Mutant lines used in this study, including references for their original derivation and description in the literature, are reported in S4 table (Alonso et al. 2003; Sessions et al. 2002; Chen and Xiong 2008; Kumimoto et al. 2010; McNellis et al. 1994; Ballesteros et al. 2001; Sessa et al. 2005; Balasubramanian et al. 2006). AGI identifiers for all genes reported are also described in S4 Table.

Acknowledgements

We thank Dr. Ben Smith (University of Oklahoma) for assistance with FLIM-FRET measurements and Dr. Min Ni (University of Minnesota) for critical

reading of the manuscript. The *cop1-4* mutant allele and *cop1-4 co-9* cross were kindly provided by George Coupland (Max Planck Institute).

Figures and Tables

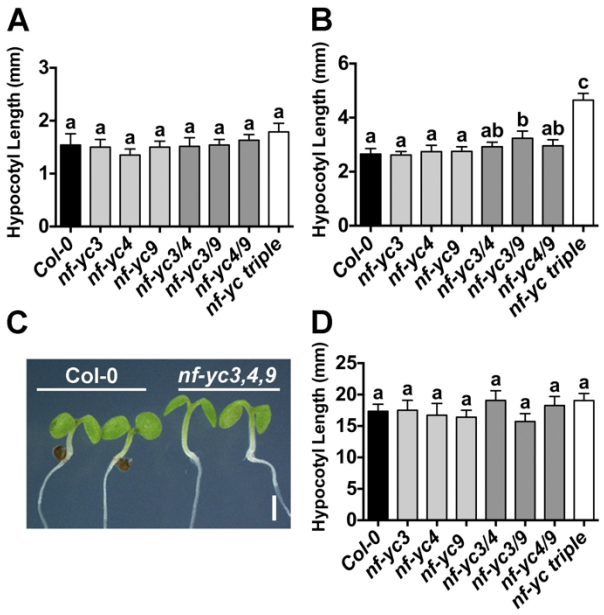

Figure 2.1. NF-YC3, 4, and 9 contribute redundantly to suppression of hypocotyl elongation in white light. Hypocotyl lengths are shown for plants grown for five days on B5 media in **A)** cWL, **B-C)** SD, or **D)** cD conditions. No differences were detected at earlier time points in cD-grown plants. Statistically significant differences (or lack thereof) are represented by lettering above bars (error bars 95% confidence interval). Statistical differences were determined by ANOVA ($P<0.01$) and subsequent multiple comparisons by either Tukey's (cWL) or Dunnett's (SD, cD) procedures. Scale bar in C) represents 2mm.

.

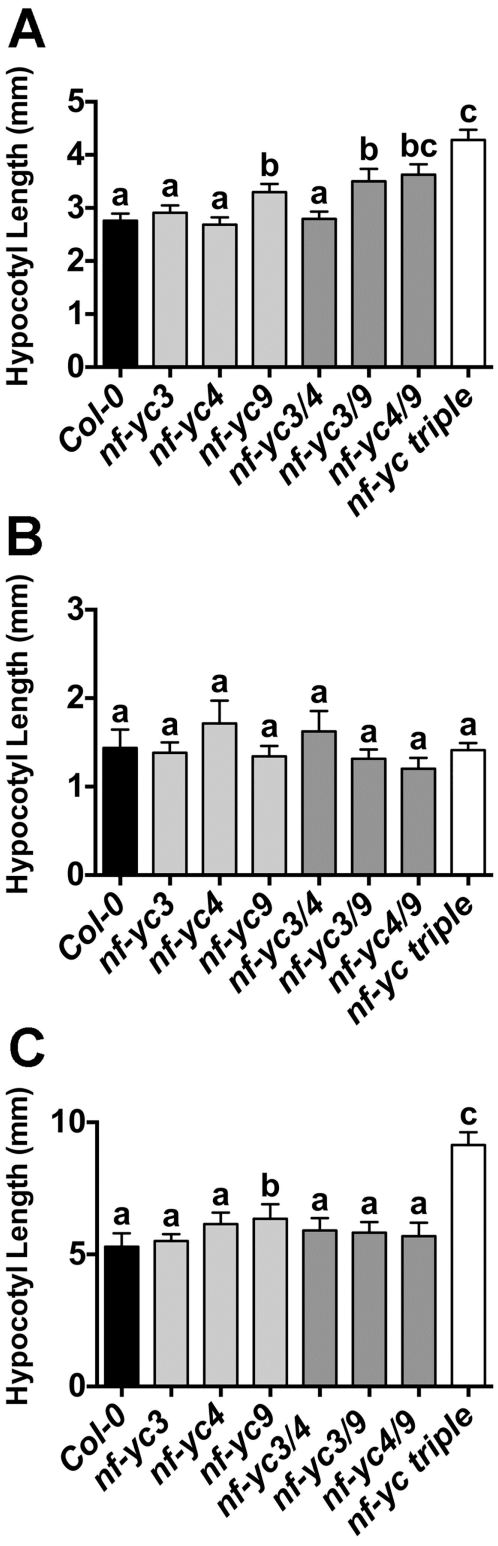

Figure 2.2. NF-YC3, 4, and 9 are necessary for suppression of hypocotyl elongation in both cB and cR light. Hypocotyl lengths are shown for five day old plants grown on B5 media in A) cB (38μmol m^{-2} s^{-1}), B) cFR (5μmol m^{-2} s^{-1}), and C) cR (6μmol m^{-2} s^{-1}). Statistically significant differences between groups (or lack thereof) are represented by lettering above bars (error bars represent 95% confidence intervals). Statistical differences were determined by standard ANOVA ($p<0.01$) when variances were not significantly different (cFR and cR) and Kruskal-Wallis ANOVA (non-parametric test, $p<0.05$) when variances were unequal (cB). Subsequent multiple comparisons were performed by either Tukey's (cFR, cR) or Dunn's (cB) procedures, respectively.

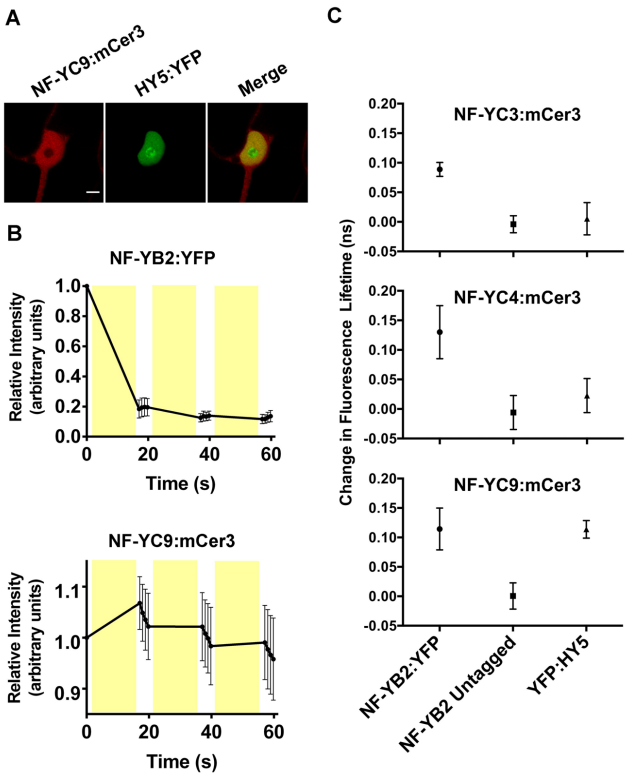

Figure 2.3. FRET-FLIM analysis shows a strong NF-YC9 by HY5 physical interaction. FRET experiments were conducted in tobacco leaves through transient 35S-driven overexpression of NF-YCs tagged with mCer3 and HY5 or NF-YB2 tagged with YFP. **A)** Nuclei expressing both mCer3 and YFP constructs were assayed for FRET through both FRAP and FLIM. **B)** A FRAP curve representative of a positive FRET result between two known interacting proteins, NF-YB2 and NF-YC9. Fluorescence intensity was calculated relative to the pre-photobleached intensity of each fluorescent protein. Yellow bars represent the timing of photobleaching events. **C)** FLIM was employed to detect FRET through lifetime measurements before and after acceptor photobleaching (FRAP) within the same nucleus. Each point is an independent combination of mCer3- and YFP-tagged proteins, and represents the shift in fluorescent lifetime elicited by acceptor photobleaching. Scale bar in A) represents 5 μm. Error bars in B-C) represent 95% CI with an n ≥ 3.

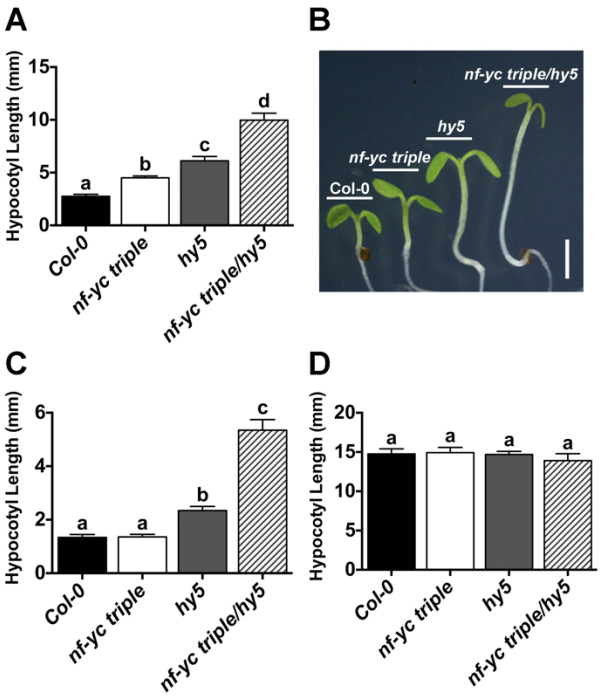

Figure 2.4. Light perception is synergistically defective in *nf-yc triple hy5* mutants. Hypocotyl lengths are shown for five day old plants grown on B5 media in **A-B)** SD, **C)** cWL, and **D)** cD. Statistically significant differences were determined and described in Figure 2.2. Scale bar in B) represents 2mm.

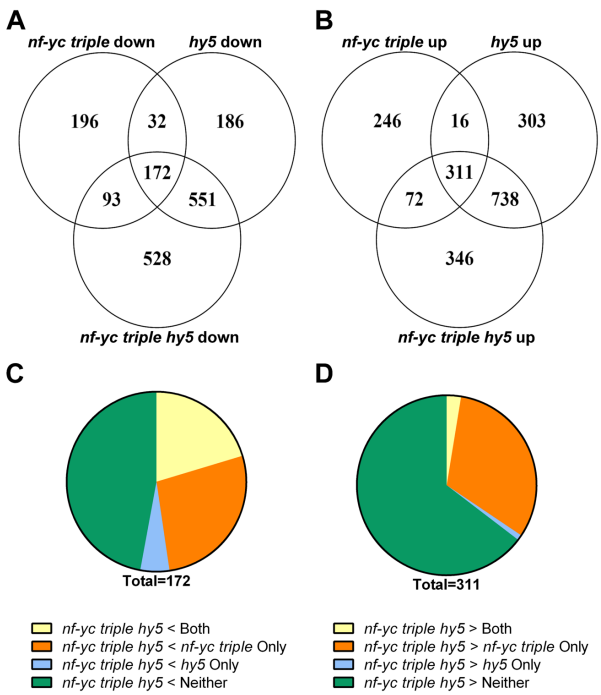

Figure 2.5. RNA-Seq analysis identified both shared and independently regulated targets for NF-YCs and HY5. Overlap between genes significantly **A)** down-regulated or **B)** up-regulated at least 1.5-fold in the *nf-yc triple*, *hy5*, and *nf-yc triple hy5* backgrounds, relative to Col-0. Genes significantly differentially expressed in all three genotypes were then broken into regulatory groups according to the level of **C)** down-regulation or **D)** up-regulation in the *nf-yc triple hy5* mutant relative to both *nf-yc triple* and *hy5*.

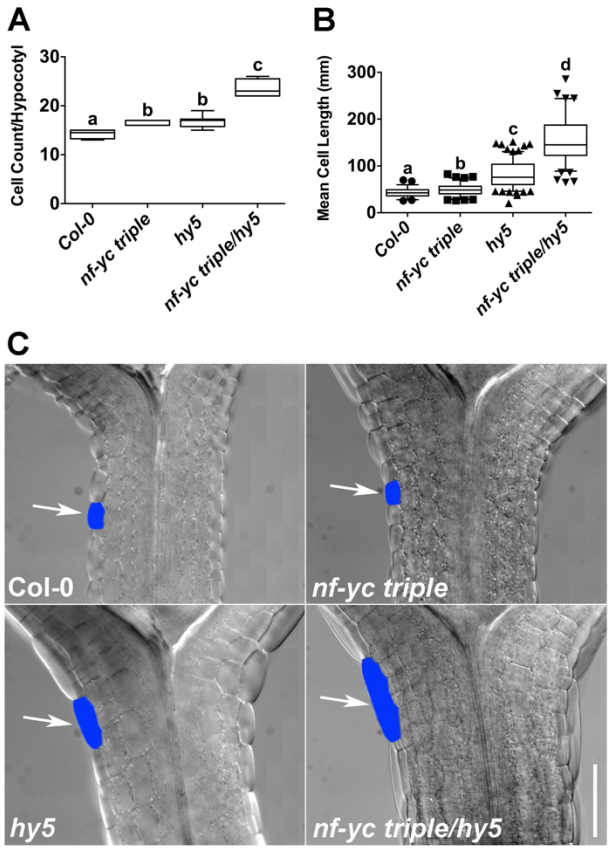

Figure 2.6. Synergistically longer hypocotyls in *nf-yc triple hy5* mutants are a function of moderately more cells and greatly increased cell elongation. Single linear files of hypocotyl epidermal cells were both **A)** counted and **B)** measured for mean length. Plant were grown in cWL and measurements were taken on five day old plants. **C)** Example hypocotyls for each genotype near the cotyledon junction – blue color marks a representative single cell in each genotype. Arrows point to typical epidermal cells for each genotype. Scale bar = 100μm.

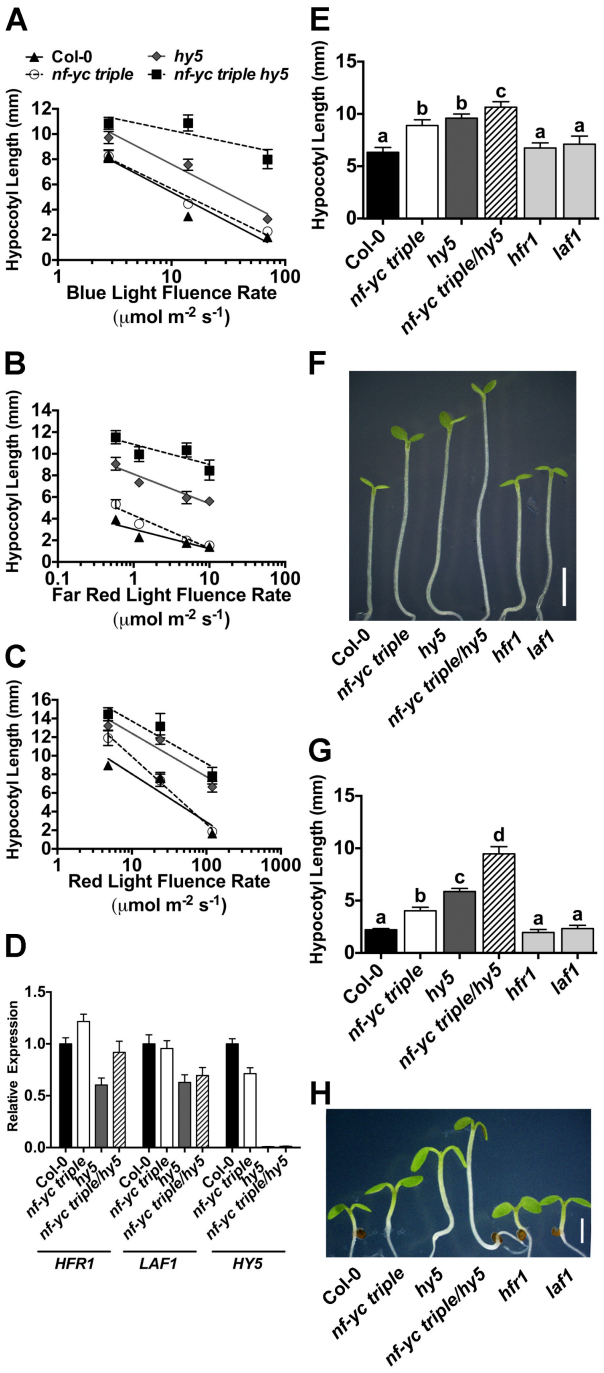

Figure 2.7. nf-yc mutants have broad light perception defects that are at least partly independent of *HY5* and do not completely overlap with those of *HFR1* and *LAF1*. Fluence rate curves for hypocotyl lengths are shown for five day old plants in **A)** cB, **B)** cFR, and **C)** cR light conditions (see symbols key in A). **D)** qPCR of *HFR1*, *LAF1*,and *HY5* in key genetic backgrounds. **E-F)** Quantification and images of typical phenotypes for mutants grown in low fluence rate cR (4.8µmol m^{-2} s^{-1}). Scale bar in F = 2mm. **G-H)** Quantification and images of typical phenotypes for mutants grown in SD, white light conditions. Scale bar in H = 2mm. Statistically significant differences were determined as described in Figure 2.2.

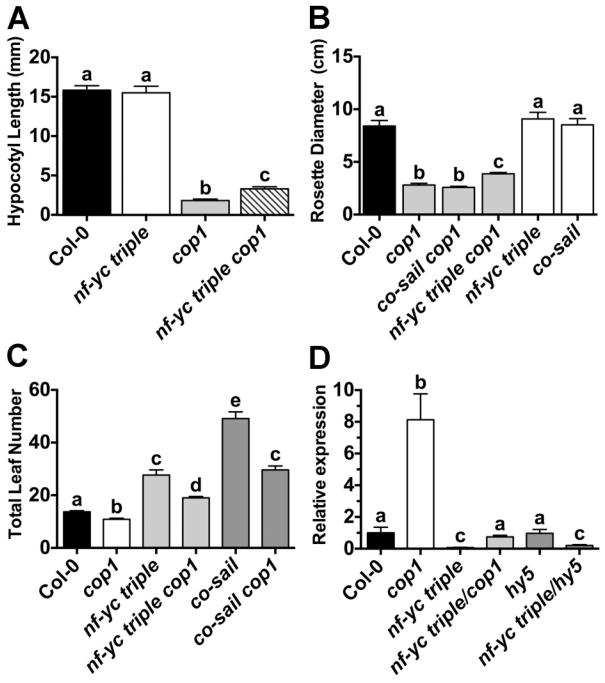

Figure 2.8. Multiple *cop1-4* mutant phenotypes are partially dependent on *NF-YC* genes. Partial suppression of cop1 mutant phenotypes are quantitated for **A)** dark-grown seedling hypocotyl elongation, **B)** rosette diameter, **C)** flowering time, and **D)** relative *FT* expression levels. Statistics and labeling as in Figure 2.2, except *FT* expression statistics which were calculated using qBase software (Biogazelle).

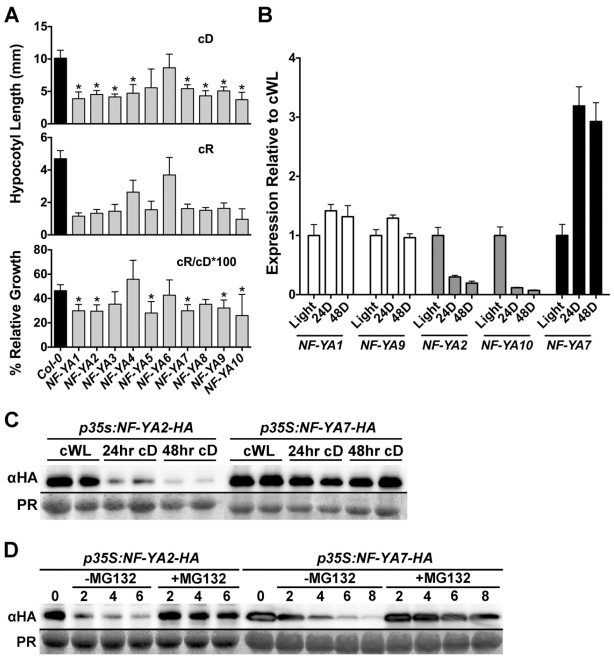

Figure 2.9. NF-YAs represent the light/dark cycle regulated components of NF-Y complexes. A) *NF-YA* overexpression causes constitutive photomorphogenesis phenotypes as well as cR light perception defects. To account for the constitutive photomorphogenic phenotypes of *NF-YA* overexpressing plants, relative percent growth in cR was determined by dividing cR hypocotyl length by cD hypocotyl length and multiplying by 100. **NOTE**: Because of their *cop1*-like, dwarf phenotypes, seed numbers are limited for NF-YA overexpressing plant. Thus, we chose to focus on cR light perception because of the relatively strong light defects measured for the *nf-yc triple* (Figure 2.2C). Statistically significant differences were determined as described in Figure 2.2. Error bars represent 95% confidence interval. **B)** Expression of select *NF-YA* genes after 24-48hrs darkness relative to cWL. Error bars represent SEM. **C)** Protein accumulation of 35S promoter-driven *NF-YA2* and *NF-YA7* constructs in cWL compared to 24-48hrs cD. **D)** Cell free protein degradation assays for NF-YA2 and NF-YA7, with and without the proteasome specific inhibitor MG132.

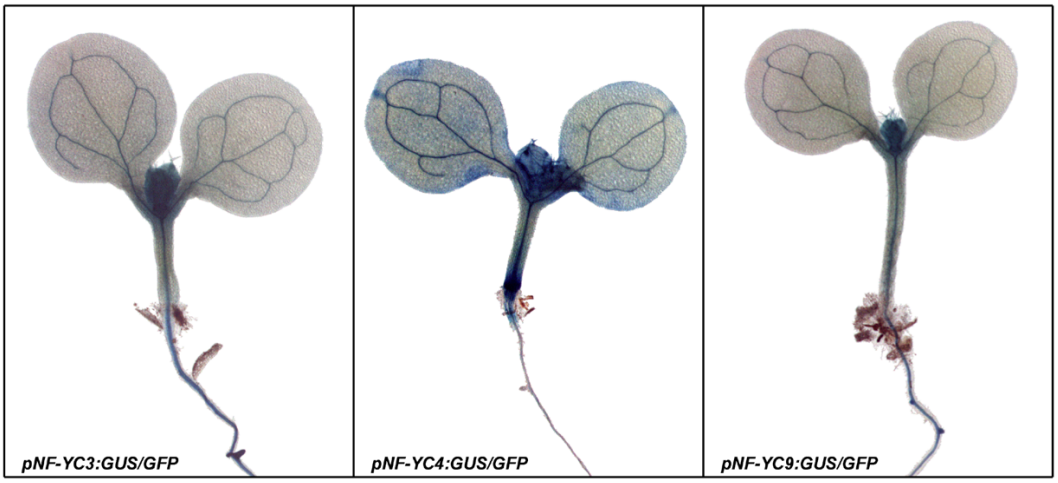

Figure S2.1. NF-YC3, 4, and 9 are expressed during early seedling development. Promoter-GUS fusions for NF-YC3 (left), NF-YC4 (middle), and NF-YC9 (right) were used to analyze expression patterns in 5-day old plants.

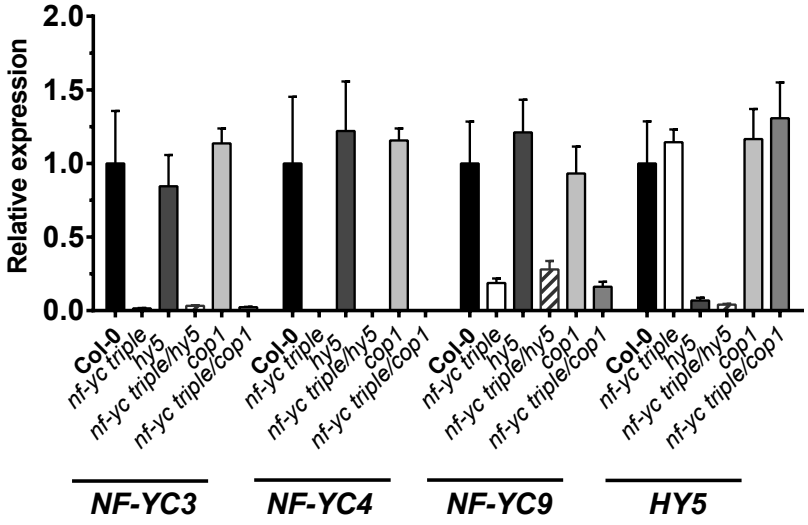

Figure S2.2. Validation of knock-out and knock-down alleles in higher-order mutants. Expression levels of NF-YC3, NF-YC4, NF-YC9, and HY5 were observed in key genotypes used in this study. Note that NF-YC9 is weakly expressed from the *nf-yc9-1* allele, as previously reported (Kumimoto et al. 2010). Error bars represent SEM.

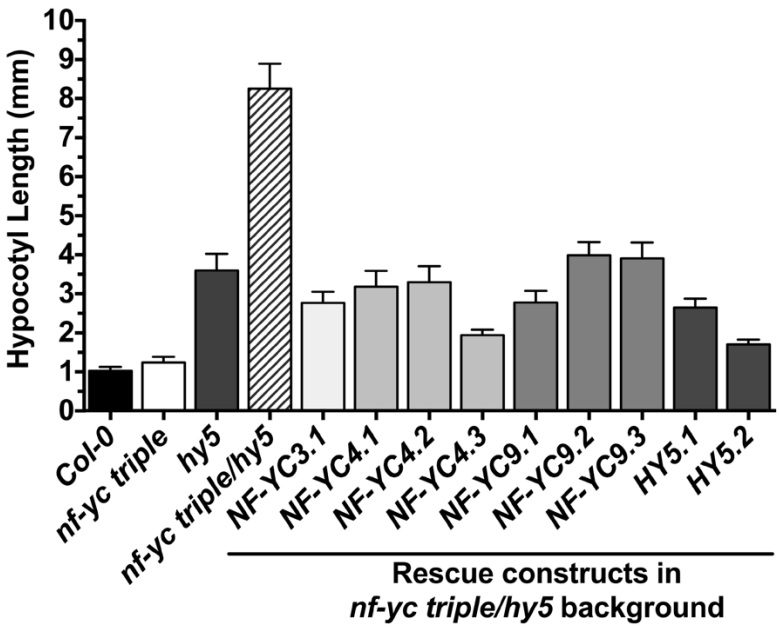

Figure S2.3. Complementation of the *nf-yc triple hy5* quadruple mutant long hypocotyl phenotype. Native-promoter driven NF-YC3, NF-YC4, and NF-YC9, as well as 35S promoter driven HY5, are able to complement the *nf-yc triple hy5* mutant hypocotyl phenotype in 5 day old, cWL-grown plants. Error bars represent 95% confidence intervals.

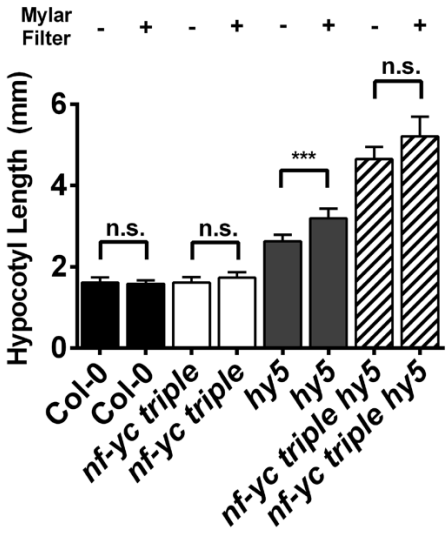

Figure S2.4. The *nf-yc triple* mutant does not have a hypocotyl elongation response to filtering out UV light. Hypocotyl elongation was assayed in 5 day old, cWL-grown seedlings in the presence or absence of a mylar filter. Error bars represent 95% confidence intervals. Significant differences (and lack thereof) were detected through unpaired t tests (Col-0, *nf-yc triple*, *hy5*) and an unpaired t test with Welch's correction for unequal variances (*nf-yc triple hy5*). ***, $p < 0.01$; n.s., not significantly different.

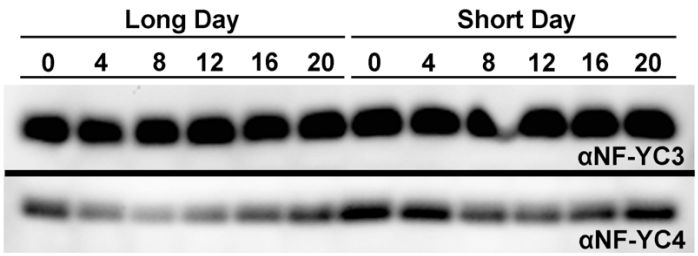

Figure S2.5. NF-YC3 and NF-YC4 accumulate at relatively steady levels over a 24 hour cycle. Protein samples were extracted every four hours from both long day (16hr light, 8hr dark) and short day (8hr light, 16hr dark) grown plants. Proteins were detected with previously described native antibodies (Kumimoto et al. 2010).

List of Tables

Table S2.1. Lists of all differentially expressed genes identified through RNA-Seq in nf-yc triple, hy5, and nf-yc triple hy5 mutants.

This table is a complex, tab-delimited Excel sheet that is not readily formatted for direct inclusion in a text-based document. This file can be freely accessed from the publisher's website, and can be directly downloaded from the following URL:

https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006333

Table S2.2. GO Enrichment of down-regulated gene sets and regulatory groupings identified through RNA-Seq.

This table is a complex, tab-delimited Excel sheet that is not readily formatted for direct inclusion in a text-based document. This file can be freely accessed from the publisher's website, and can be directly downloaded from the following URL:

https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006333

Table S2.3. GO Enrichment of up-regulated gene sets and regulatory groupings identified through RNA-Seq.

This table is a complex, tab-delimited Excel sheet that is not readily formatted for direct inclusion in a text-based document. This file can be freely accessed from the publisher's website, and can be directly downloaded from the following URL:

https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006333

Table S2.4. Allele designations, descriptions and references for all mutant lines used in this study.

This table is a complex, tab-delimited Excel sheet that is not readily formatted for direct inclusion in a text-based document. This file can be freely accessed from the publisher's website, and can be directly downloaded from the following URL:

https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006333

Chapter 3: DIMR, a Yeast-Based Synthetic Reporter System for Probing Oligomeric Transcription Factor DNA Binding

Zachary A. Myers, Swadhin Swain, Shannan Bialek, Samuel Keltner, and Ben F. Holt III

This chapter is in the final stages of data collection and preparation for submission to *Nature Methods*

Abstract

Transcription factors (TFs) are fundamental components of biological regulation, facilitating the basal and differential gene expression necessary for life. TFs exert transcriptional regulation through interactions with both DNA and other TFs, ultimately influencing the action of RNA polymerase at a genomic locus. Current approaches are proficient at identification of binding site requirements for individual TFs, but few methods have been adapted to study oligomeric TF complexes. Further, many approaches that have been turned toward understanding DNA binding of TF complexes, such as electrophoretic mobility shift assays, require protein purification steps that can be burdensome or scope-limiting when considering more exhaustive experimental design. In order to address these shortfalls and to facilitate a more streamlined approach to understanding DNA binding by TF complexes, we developed the DIMR (**D**ynamic, **I**nterdependent TF binding **M**olecular **R**eporter) system, a modular, yeast-based synthetic transcriptional activity reporter. As a proof of concept, we focused on the NUCLEAR FACTOR-Y (NF-Y) family of obligate heterotrimeric TFs in *Arabidopsis thaliana*. The DIMR system was able to reproduce the strict DNA-binding requirements of an experimentally validated NF-$Y^{A2/B2/C3}$ complex with high fidelity, including recapitulation of previously characterized mutations in subunits that either break NF-Y complex interactions or are directly involved in DNA binding. The DIMR system is a

novel, powerful, and easy-to-use approach to address questions regarding the binding oligomeric TFs to DNA.

Introduction

Transcription factors (TFs) are a fundamental component of biological control, facilitating differential gene expression in response to environmental stimuli and making complex life possible. The mechanism through which differential gene expression manifests is inherently complex, requiring coordination of stimulus perception, transcriptional and translational responses, and alteration of relevant protein activity (Crick 1970). The transcription-level integration of stimuli is of particular interest, as this process supports context-specific recruitment of RNA polymerase to a specific genomic locus (Diamond et al. 1990). This integration is accomplished through the independent, competitive, and cooperative functions of TFs and the effects of these relationships on DNA-binding and RNA polymerase recruitment (Amoutzias et al. 2008; Lickwar et al. 2012). A better understanding of the mechanisms through which transcriptional activity are modulated, particularly how TF interactions influence regulatory capacity, could facilitate the design of more precise and predictable molecular tools to address long-standing issues in the fields of agriculture, health care, and manufacturing, among others (Khalil and Collins 2010).

Current approaches to identify TF-DNA interactions are effective at *de novo* identification of binding site preferences of individual TFs *in vitro* and at identifying global binding patterns of individual TFs *in vivo*. In many cases, a combination of approaches can be used for a more complete

understanding of TF targeting. In particular, many researchers have found great success in combining approaches that effectively identify binding specificity (such as Protein Binding Microarrays (PBMs), 1-Hybrid-based library screening, or DNA Affinity Purification (DAP-seq)) and lower-resolution, global-scale binding site identification (such as Chromatin Immunoprecipitation (ChIP-seq) or Assay for Transposase-Accessible Chromatin (ATAC-seq)).

While I will not focus on the general strengths of each approach or combination of approaches (see reviews, (Mahony and Pugh 2015; Jayaram et al. 2016)), one aspect of TF function that remains less explored is the impact that TF oligomerization exerts on DNA binding (Jolma et al. 2015). Most approaches, including those introduced above, have focused on understanding some aspect of DNA-binding of an assumed individual TF; however, a much more complex interplay of TF function exists than can be readily and methodically addressed with current techniques. Transcriptional regulation through multi-component TF complexes is a common theme, but our understanding of how these complexes identify and interact with specific genomic loci remains less explored (Jolma et al. 2015; Morgunova and Taipale 2017). For example, a recent study identified dramatic shifts in cognate binding sequences of bZIP homodimers compared to heterodimers (Rodriguez-Martinez et al. 2017); however, the experimental approach used in this study required purification of a large

suite of individual transcription factors. An approach enabling rapid and straight-forward testing of complex TF units, with the ability to make targeted mutations in any component of the TF complex or its potential binding site, would create a new lens through which we could explore more nuanced aspects of TF-DNA interactions.

Traditional definitions of transcription factors often highlight the presence of both a DNA-binding and a transcriptional-regulation domain; however, many TFs function as obligate protein complexes, where individual components contain partial DNA-binding or transcriptional-regulation domains that are only fully-reconstituted within the context of a functional complex. One example of such an arrangement can be found in the NUCLEAR FACTOR-Y (NF-Y) transcription factors, a complex that functions as an obligate hetero-trimer of three distinct subunits, NF-YA, NF-YB, and NF-YC, and physically interacts with the DNA sequence *CCAAT*. Although this complex was initially identified over 30 years ago (Dorn et al. 1987; Jones et al. 1985), questions remain regarding the molecular forces driving specificity of the NF-Y complex to *CCAAT*. Several key observations and studies have led to persistent interest in the NF-Y, particularly in plant lineages, including: (1) the ubiquitous nature of the *CCAAT* box and, subsequently, NF-Y regulated genes, (2) the expansion of the NF-Y subunit families in plants compared to animals, where animals encode 1-2 members of each subunit, while higher plants often encode 10+ of each subunit

(Siefers et al. 2009), and (3) the identification and molecular characterization of the first non-canonical NF-Y complex in plants, which replaces an NF-YA subunit for a plant-specific CCT (CONSTANS, CONSTANS-LIKE, and TOC1) domain-containing protein, leading to altered DNA binding properties (Gnesutta et al. 2017a).

Published crystal structures of various NF-Y complexes, particularly of the full *Homo sapiens* NF-Y complex bound to DNA (Nardini et al. 2013), have raised increasingly complex questions regarding the specificity of the NF-Y complex for a particular *CCAAT* box. While all three NF-Y subunits have been shown to make physical contact with DNA, it appears that only NF-YA makes sequence-specific contact, inserting directly into the minor groove of the *CCAAT* box, while NF-YB and NF-YC make contacts with the DNA backbone, appropriately positioning both the DNA and NF-YA subunit for tight physical interaction (Nardini et al. 2013). Despite this understanding, the NF-Y complex appears to show more selectivity than has been experimentally derived. For example, most *CCAAT* boxes in human cell lines are not consistently bound by NF-Y complexes (Encode Project Consortium et al. 2007; Zambelli and Pavesi 2017), meaning that the mere presence of a *CCAAT* box is not predictive of NF-Y binding. Factors such as chromatin accessibility and repressive epigenetic modifications regularly limit the binding landscape of a given TF (John et al. 2011); however, the NF-Y are thought to function as 'pioneer' TFs that are able to bind less-

accessible DNA and promote further TF complex formation (Tao et al. 2017; Oldfield et al. 2014; Donati et al. 2008). Alternatively, while the only strict requirement on NF-Y binding is the presence of the pentanucleotide *CCAAT*, the flanking nucleotides could serve as a fine-tuning mechanism for complex binding and stability. Ultimately, the source of this *CCAAT* box selectivity is not well understood and is made significantly more complex when considering the expanded family size and combinatorial complexity of plant NF-Y subunits.

In particular, we hypothesized that different NF-YB/NF-YC dimers might contribute to *CCAAT* box selectivity through interactions with the nucleotides flanking the *CCAAT* pentamer; however, comparing all possible NF-YB/NF-YC combinations with even a single NF-YA component would be a significant undertaking, with 100 possible combinations in *Arabidopsis thaliana* (Petroni et al. 2012) and 208 possible combinations in *Oryza sativa* (rice, (Hwang et al. 2016)). *In vitro* approaches requiring protein purification, such as Electrophoretic Mobility Shift Assay (EMSA), become time- and cost-prohibitive at this scale. More importantly, many proteins are recalcitrant to protein purification, and can only be isolated if appropriately truncated or co-expressed with other factors. Despite these technical limitations, EMSA analysis is well-suited for these types of investigations, and any approach intended to address similar questions would need to match or complement its capabilities. A reporter system with a large

dynamic range and high sensitivity would allow scientists to address mutations that have much more modest effects on TF-DNA interactions, and could facilitate examination of more precise or biologically-relevant questions.

To address these issues and to better facilitate these types of research, we designed the DIMR (**D**ynamic, **I**nterdependent TF binding **M**olecular **R**eporter) system (Figure 3.1), a modular yeast-based transcriptional activity reporter system developed through repeated application of the synthetic biology approach of design-build-test-learn. The DIMR system allows for dose-dependent induction of a suite of transcription factors (effectors) and subsequent monitoring of transcriptional regulation. We validated our approach by testing the DNA binding capabilities of the Arabidopsis NF-$Y^{A2/B2/C3}$ complex and found that the DIMR system was able to faithfully recapitulate both wild-type *CCAAT* box binding and previously-described mutations impacting DNA binding and complex formation.

Results

DIMR system components, design philosophy, and functional description

The DIMR system is composed of three modules: (1) an **Activator Module**, encoding an inducible artificial transcription factor (ATF), (2) an **Effector Module**, encoding a single ATF-driven cassette containing viral 2A 'cleavage' sites between effector components, and (3) a **Reporter Module**, containing a dual-luciferase reporter composed of a constitutively transcribed *Renilla* luciferase and a conditionally transcribed Firefly luciferase (Figure 3.1A). Each module is carried on a yeast shuttle vector with different auxotrophic selection, and each has been designed to be as modular as possible, with restriction enzyme sites strategically placed to facilitate swapping individual effector or reporter components with minimal laboratory effort. As currently implemented, point mutations in existing effectors can be easily accomplished through commercially available site-directed mutagenesis kits, novel effector modules can be synthesized *de novo* at affordable rates, and permutations of binding sites can be cloned in as little as 3 days at a relatively low cost and with very minimal active time, resulting in a streamlined, quick, and hands-off assay to investigate TF-DNA interactions (Figure 3.1B). Finally, because of the synthetic biology-based approach taken during the design and implementation of the system, each

component of the modules can easily be further iterated upon for improved function or to accomplish different goals.

The Activator module (Figure S3.1) utilizes the chimeric Z_4EV artificial transcription factor, which contains three functional domains: (1) an engineered zinc finger with specificity to a DNA sequence not found in the genome of *Saccharomyces cerevisiae,* (2) an estrogen receptor that modulates activity and localization, and (3) a VP16 activation domain (McIsaac et al. 2013). Z_4EV is constitutively expressed under the *ACT1* promoter (Flagfeldt et al. 2009) but remains inactive and restricted to the cytosol. Activation of Z_4EV with the hormone β-estradiol initiates translocation to the nucleus and activation of the Effector cassette. This activation was previously shown through transcriptome profiling to induce remarkably few unintended effects, either through the action of β-estradiol itself or through off-target binding of Z_4EV (McIsaac et al. 2013).

The Effector module (Figure S3.2) is designed to emulate a polycistronic message through the incorporation of viral 2A 'self-cleaving' sequences. Our initial designs included cassettes expressing both 2 and 3 genes, with a translationally-fused flexible linker and unique epitope tag on each (Sabourin et al. 2007). The linkers between components A – B and B – C also encode previously-validated T2A peptide sequences (Beekwilder et al. 2014), facilitating translation of individual proteins from a single mRNA.

Restriction enzyme recognition sites were embedded into the coding sequences of the T2A linkers through the introduction of silent mutations. This allows easy swapping of effectors through either (1) the inclusion of the appropriate flanking linker sequence through gene synthesis (recommended, along with codon optimization (Kotula and Curtis 1991; Gustafsson et al. 2004)) or (2) PCR amplification to produce the appropriate over-hangs.

The Reporter module (Figure S3.3) encodes a dual luciferase reporter system, including a *Renilla* luciferase variant constitutively expressed under the *ACT1* promoter, and a conditionally expressed Firefly luciferase variant. Design of the Firefly luciferase promoter drew heavily on previously characterized yeast promoters bound by yeast NF-Y complexes. Specifically, we chose native, experimentally validated promoters of *Saccharomyces cerevisiae* whose activation required the presence of a specific *CCAAT* box (implying direct regulation by the NF-Y), and replaced the required *CCAAT* box binding site with the various permutations described below. Cloning of binding site permutations is quick and simple, requiring only a pair of appropriately designed, 5'-phosphorylated oligonucleotides and simple restriction enzyme digestion and ligation reactions.

Validation of effector induction and cleavage

As a starting point, we tested our ability to induce accumulation of Arabidopsis NF-$Y^{A2/B2/C3}$ effectors through the activation of Z_4EV. After a 6-hour induction period, we were able to visualize accumulation of NF-YA2:HA (Figure 3.2A), NF-YB2:MYC (Figure 3.2B), and NF-YC3:FLAG (Figure 3.2C). Differences were seen in the level of protein accumulation across samples, and several instances of failed T2A-mediated cleavage were clear; however, we were able to clearly and consistently observe individual NF-Y subunits in repeated experiments.

Further experiments to explore the inducibility of the effector cassette identified detectable levels of individual effector proteins with around 5 nM β-estradiol (Figure S3.4), with no obvious differences in accumulation between the different NF-Y subunits. A short time-course of effector accumulation at 10 nM β-estradiol identified fairly stable accumulation of individual effector proteins at 3 and 6 hours, though we often observed a peak accumulation at 6 hours and a drop-off in signal when extending to 9 hours (Figure S3.5). From these and other preliminary data, we established standard low- and high-induction ranges of ~1 nM and ~10 nM β -estradiol, respectively. While incomplete cleavage of individual effectors remains to be addressed, the induction scheme of the DIMR system is effective and robust.

Validation of reporter activity following effector induction

Our initial system design and testing were built upon a previously-characterized NF-Y regulated *CCAAT* box from the *FLOWERING LOCUS T (FT)* promoter in Arabidopsis. This *CCAAT* box, as well as 20 flanking base pairs on either side, was positioned immediately upstream of the minimal NF-Y regulated yeast promoter of *CITRATE SYNTHASE1 (CIT1)*. Upon induction of the NF-$Y^{A2/B2/C3}$ effector cassette, we observed a significant, dose-dependent increase in relative luminescence (Figure 3.3A). Normalization of the data into a fold change value, relative to mock induction, found an average ~3.5-fold increase in high induction conditions (10 nM, Figure 3.3B). Despite the successful recapitulation of NF-Y DNA binding in our initial designs, we were not able to statistically distinguish between low- and high-induction conditions, and the maximum fold changes observed were lower than desired. A larger dynamic range of reporter activation would facilitate addressing more nuanced questions, such as the DNA binding impact of individual effector mutations or changes in cognate binding sites.

Refinement of binding site design

From these initial tests, we next sought to refine the DIMR system for increased reporter dynamic range. We focused our system tuning on two approaches: reducing mock-level signal and increasing maximum activation. To decrease mock-level signal, we designed and tested different NF-Y regulated minimal promoter architectures to identify reporters with

lower basal level activation. We focused on promoters for the genes *CIT1* and *ASPARAGINE SYNTHETASE1 (ASN1)*, driven by the above-described *FT* CCAAT box footprint (Figure 3.4A). Luciferase activity levels were lower in mock-treated samples with the *ASN1*-based promoter compared to the originally-tested *CIT1*-based promoter (Figure 3.4B). Notably, this reduction in mock-level activation translated to a larger fold change of ~6x in *ASN1*-based reporters when comparing mock and high induction conditions.

To increase our maximum signal, we additionally tested the effects of different numbers of binding sites and the spacing between them. Multimerization of available binding sites can result in an increase in observed transcriptional activity (Khalil et al. 2012), but this larger DNA footprint can be more difficult to accommodate in cloning. To abrogate this, we also tested the effects of reducing the length of each individual binding site footprint. First, we attempted to cut the *CCAAT* box footprint roughly in half by including only 10 base pairs of flanking sequences in single, double, and triple binding site configurations (Figure 3.4C), and found that multimerization of the binding site increased total system activation. Finally, we combined the original, larger *CCAAT* box footprint into single, double, and quadruple binding site configurations (Figure 3.4D). In this set, the quadruple binding site configuration stands out at a much-improved ~15-fold increase in reporter activity from mock. The larger DNA footprint of the quadruple binding site approach required a slightly modified cloning

approach, where two pairs of annealed, phosphorylated oligos were simultaneously ligated into the Reporter module; however, we were able to preserve the flexibility of the DIMR system and avoided the need for further gene synthesis for individual binding site permutations.

Validation and assessment of refined binding site setup

With an increased signal level upon system activation, we began testing the activation requirements of the DIMR system. First, we examined dose-dependent system activation levels of the NF-$Y^{A2/B2/C3}$ complex on the quadruple *CCAAT* box over a wide range of β-estradiol induction levels (Figure 3.5A). This induction gradient was analyzed through 4-parameter logistic regression (4PLR), with an R^2 value of 0.942. No system activation was observed at induction levels below 1 nM, while activation peaked between 10 and 100 nM β-estradiol. The dose-dependent activity of the system was most pronounced between 1 nM and 10 nM, with a half maximal effective concentration (EC_{50}) of ~2.2 nM and total span of ~10.5-fold change. 4PLR is a widely used metric in chemistry, biochemistry, and computational biology to describe cooperative action, and in this case, statistically supports the idea that the individual NF-Y subunits are working interdependently in an induction-dependent manner to bind the *CCAAT* box.

To further solidify the correlation between effector induction and system activation, we tested the impacts of loss of individual DIMR modules, previously described mutations in NF-Y effectors, and changes in the CCAAT box (Figure 3.5B). Replacing either the Z_4EV activator cassette or the NF-Y effector cassette with empty vectors resulted in a complete loss of system activation upon induction. Importantly, we found that mutation of the *CCAAT* box to *CCAGC* eliminated system activation. These three pieces of data collectively suggest that the observed induction responses are accomplished through the action of the NF-Y complex on the *CCAAT* box driving the firefly luciferase reporter.

We also performed further system tests with the *NF-YB2 E65R* point mutation, a previously characterized mutation shown to prevent interaction of the NF-YB/NF-YC dimer with NF-YA (Sinha et al. 1996; Siefers et al. 2009). As expected, this point mutation abolished system activation (Figure 3.5B). While many highly-conserved residues in each of the NF-Y subunits have been previously identified and characterized to completely break DNA binding (Sinha et al. 1996; Kim et al. 1996; Sinha et al. 1995), we wanted to examine novel mutations that could help reveal the evolutionary constraints acting on the NF-Y complex. To this end, we examined the effects of one particular point mutation at a residue thought to be directly involved in *CCAAT* box binding, *NF-YA2 H183A*. With the native histidine residue replaced with an alanine at this location, we still observed a

moderate level of system activation (Figure 3.5B). This alanine-replacement approach is traditionally used to replace residues that contribute to sequence-specificity with an unobtrusive and relatively-inert alanine that is unlikely to mediate sequence-specific interactions (Luscombe et al. 2001).

Because we had previously observed instances of failed 2A site function in the form of fused effector components, we directly tested whether these fused effectors were functional by creating point mutations that break the existing 2A sites. We found that fusion of the NF-YA and NF-YB components did lower maximum reporter activity compared to a 2A-cleaved trimeric system or a fusion between NF-YB and NF-YC, but each instance of broken 2A cleavage was still able to significantly activate the reporter system (Figure S3.6). While not all transcription factor complexes will be able to effectively form and regulate DNA while fused together, this should be tested for each group of transcription factors being tested.

Finally, we tested whether yeast NF-Y orthologs might be able to compete with or take the place of our effector module-encoded Arabidopsis NF-Y proteins. Sequentially removing one NF-Y subunit from the otherwise-complete effector module induced no system activation (Figure S3.7), suggesting that native NF-Y orthologs are not responsible for any significant activation of the reporter. While interactions between NF-Y orthologs of different species has been demonstrated (Calvenzani et al. 2012), we

hypothesized that the level of induction through the activator and effector modules would quickly saturate any unintended interactions. Further, not all yeast NF-Y orthologs are thought to be constitutive expressed, with at least one ortholog only expressed in response to non-fermentable carbon sources (Bourgarel et al. 1999).

Validation of CCAAT-based NF-Y binding

To further explore the affinity of the NF-Y complex to the *CCAAT* box, we tested permutations of the *CCAAT* box with variation at the 3' end, corresponding to all possible combinations of *CCA**NN*** (Figure 3.6). We focused on variation at these locations because the crystal structure of the DNA-bound human NF-Y complex identified many more direct interactions with NF-YA and the first three bases of the *CCAAT* box than the fourth and fifth bases. Beyond this, our earlier exploration of the *NF-YA2 H183A* mutation suggests possible differences in the way human and plant NF-Y complexes bind DNA, an idea further supported by key differences in the otherwise-conserved linker region of *NF-YA2*. This linker region was proposed to be important for precise positioning of flanking α-helices for proper NF-Y complex stabilization and DNA binding, and *NF-YA2* is the only family member in Arabidopsis with this unusually long linker region.

Unsurprisingly, the highest-activated binding site of the *CCA**NN*** suite was *CCA**AT*** at ~14-fold over mock, with no other binding site permutation

activating greater than 2.2-fold over mock (Figure 3.6A). Among these non-CCAAT binding sites, we saw a small increase in reporter activity in *CCA**N**T*, but not in *CCAA**N*** (Figure 3.6B). A targeted comparison of these two binding site permutations identified significant increases relative to mock in all *CCA**N**T* binding sites, but none other than *CCAAT* in the *CCAA**N*** variants (Figure 3.6C). While further investigation is necessary, this observation runs counter to what is currently understood about DNA binding by the NF-Y complex in humans, as a direct physical interaction between NF-YA has been identified at the fourth position, but not the fifth position, of the CCAAT box.

Discussion

Using DIMR to investigate TF complex-DNA interactions

In spite of the existence and importance of multi-component transcription factor assemblies, relatively few molecular or biochemical approaches have been devised to explore the intricacies of DNA binding by these multimeric complexes. Current and widely adopted methods are not well-suited for many questions, particularly those that could benefit from extensive mutagenesis-based analyses, such as alanine scanning mutagenesis to probe the relative contributions of DNA binding amino acids within a given TF. After fine-tuning and optimizing our induction and binding site schemes, we ultimately produced a sensitive, straightforward, and versatile reporter system for assaying transcription factor complex DNA interactions. As one example of this type of study, we examined the effect of permutations of the *CCAAT* box and identified a slightly higher level of system activation when altering bases at position 4 than position 5 (*CCA**N**T* vs. *CCAA**N***). Our system bypasses the need for protein purification or deep sequencing, two common problems with other approaches. The proposed system can easily be scaled to simultaneously include dozens of effector or binding site permutations with relative simplicity, and it mostly requires basic lab equipment. The injector-fitted, luminescence-based plate reader is likely the only equipment not commonly found in a standard molecular lab, and these

types of equipment are often housed in core facilities available to most universities or institutes.

Incomplete cleavage at T2A sites in effector module

Our initial designs sought to leverage viral 2A peptide sequences to generate equimolar amounts of individual effectors through a cleavage mechanism during translation, an important goal when considering obligate oligomeric TF complexes. We use the term cleavage throughout; however, strictly speaking, the end result is an inability to form the peptide bond between two adjacent amino acids, not true cleavage of an existing peptide bond (Kim et al. 2011). Because eukaryotes do not utilize the polycistronic mRNA approach so common in bacteria, researchers have turned to 2A sites and other approaches to emulate this effect. While not a perfect solution to this problem, we found that viral T2A sites worked sufficiently well for our needs here. It is important to note, however, that incomplete cleavage of the effector cassette does alter the interpretation of system activation. Therefore, with the shortcomings of this design, we can test whether a transcription factor complex can bind *a particular consensus sequence*, but we cannot directly address whether a particular complex *can or cannot form*. While this distinction is fairly minor in many cases, particularly when working with previously-described complexes, it is an important limitation of the system that should be addressed in future iterations. Unfortunately, no completely effective eukaryotic-based system

has been described to recreate the polycistronic mRNA system so widely employed in bacteria (Blount et al. 2012). However, new 2A peptide sequences are still being identified and have been seen to vary in efficiency from organism to organism (Luke et al. 2010), raising the possibility that a 2A site more efficient in yeast remains to be identified.

To circumvent this issue, we have begun designing effector cassettes with individual components driven by separate inducible promoters. These separate promoters each contain the Z_4EV -bound consensus sequence, but utilize different minimal promoters to reduce recombination within yeast (Broach et al. 1982). Whether this modification will lead to relatively-equal protein accumulation is unknown; however, it should at least provide a framework to begin further modifying, tweaking, and improving the system for our specific needs.

Considerations and improvements on the DIMR system

While much of the parameter refinement presented above should help inform novel design considerations, some applications will require more significant alterations. In particular, issues are likely to arise where effector complexes do not possess intrinsic transcriptional activation potential or when a binding site is bound and autoactivated by endogenous yeast proteins. In cases where complexes completely lack activation potential, an activation tag might be fused to one or more effectors (Knop et al. 1999);

however, whether this approach could overcome active transcriptional repression is unclear. Similarly, instances of reporter autoactivation might be addressed through fusion of one or more effector components to a transcriptional repressor domain (Edmondson et al. 1996) and measuring inducible *reductions* of reporter activity. While we have not yet worked through either of these concerns, the modular design of the system should allow for relatively straight-forward modification and testing.

Teasing apart the details of NF-Y DNA binding: a case for DIMR

Despite consistent investigations into the mechanistic function of the NF-Y over the past 30 years in animals, many aspects of complex formation and DNA binding remain relatively unexplored in plant NF-Y orthologs. First, a significant amount of plant NF-Y research, including most DNA-binding assays, has been conducted using only conserved domains of each NF-Y subunit. In fact, the survey presented above represents one of the first comprehensive examinations of DNA binding of a complete, full-length plant NF-Y trimeric complex. While the core domains of each NF-Y subunit show significant conservation across all eukaryotic lineages, the flanking termini show remarkable divergence from one another. The biological significance of these flanking regions has remained elusive in most cases, though several protein-protein interactions with NF-Y complexes are thought to be mediated through NF-Y terminal domains (Cao et al. 2011a). A better

understanding of the functions of the less-conserved termini is critical for plant NF-Y research in particular, as the vast majority of differences between individual members of expanded NF-Y gene families are found in these flanking regions. The DIMR system is capable of facilitating a systematic domain-swap approach of the flanking regions of NF-YB or NF-YC paralogs, and could potentially uncover changes in NF-Y complex affinity or specificity achieved through modulation of the specific NF-Y subunits within a functional complex.

Like *NF-YB* and *NF-YC*, the conserved domains of *NF-YA* family members are remarkably similar to one another. One critical exception to this observation is a 4-5 amino acid elongated linker sequence encoded in the *NF-YA2* subunit. This linker is positioned between two highly-conserved alpha helices that are each responsible for NF-Y complex formation or DNA binding of the mature complex, and in addition to influencing the positions of the flanking alpha helices, this linker also makes physical contact in several places with the sugar-phosphate backbone of the *CCAAT* box (Nardini et al. 2013; Romier et al. 2003). While experimental evidence is necessary to understand the impact and significance of this atypical linker, it is possible that NF-Y complexes containing NF-YA2 have slightly different DNA binding profiles or NF-Y complex dissociation constants. It should be noted, however, that NF-YA2 is likely the best-characterized plant NF-YA subunit, as it has been a major focus of research through its regulation of

photoperiodic flowering. While the majority of our observations regarding NF-YA2 DNA binding and complex formation have supported the crystal structure of the DNA-bound human NF-Y complex, many important aspects of NF-Y form and function have not been thoroughly investigated and developed.

In fact, many of the most exciting and significant advances in NF-Y research have occurred in the last two years through the identification and characterization of the first non-canonical NF-Y complex. This complex utilizes the NF-YB2/NF-YC3 dimer to stabilize interactions between CONSTANS (CO) and its consensus binding site. This complex, termed NF-CO, recognizes a *CCACA* motif found in the proximal promoter of the *FLOWERING LOCUS T (FT)* gene and is critical for proper photoperiodic floral induction (Gnesutta et al. 2017a). The interaction between CO and the NF-YB/NF-YC dimer is mediated through a conserved, plant-specific CCT (***C****ONSTANS*, ***C****ONSTANS-LIKE*, and ***T****OC1*) domain that is found in over 40 genes in Arabidopsis (Griffiths et al. 2003; Farré and Liu 2013), raising the possibility of a significantly-expanded pool of potential NF-Y or NF-CCT complexes. Excitingly, several families of CCT domain containing proteins have been extensively studied and shown to be of critical importance for photoperiodic flowering (Griffiths et al. 2003), circadian clock entrainment and maintenance (Mizuno 2004), and seedling environmental responses (Reyes et al. 2004). Unfortunately, testing hypotheses concerned with these

non-canonical NF-CCT complexes is far from straightforward, with over a decade passing between the initial suggestions that CO might function through NF-YB/NF-YC to a complete set of work describing the NF-CO complex. Systematic testing of the ability of different NF-CCT complex to form and bind DNA is a massive undertaking, with over 4,000 possible NF-CCT complexes to test. Further, while the two are clearly related, the CCT domain shows key differences from the DNA-binding domain of NF-YA, and variation within CCT members occurs at residues predicted to be important for DNA binding in aligned NF-YA sequences. A systematic mutational analysis focusing on the differences between important residues of aligned NF-YA and CCT members could create a map of critical DNA-binding residues, and inform the search for cognate binding sequences of various NF-CCT complexes.

The DIMR system is well-suited to address these types of questions on structure and function, as supported by our analysis of the *NF-YA2 H183A* point mutation and our comparison of NF-$Y^{A2/B2/C3}$ system activation in *CCA**N**T* and *CCAA**N*** permutations. The *NF-YA2* histidine residue at position 183 is thought to make sequence-specific contact with the *CCAAT* box (Gnesutta et al. 2017a), but reduced DIMR system activation is still observed in the *H183A* point mutation. Importantly, when aligning NF-YA and CCT DNA binding domains, this residue diverges between, but not among, many clades and sub-clades. Considering this pattern of

divergence, it is intriguing to consider that this particular residue might contribute to DNA binding specificity of different NF-CCT complexes. The identification of an NF-YA mutation at this residue that retains some activity could be relevant to this observation in an evolutionary context, as a sub-optimal variant is thought to facilitate further functionalization by providing a novel evolutionary trajectory (i.e., without such a permissive intermediate, purifying selection cannot be overcome to reach a different high-fitness state (Poelwijk et al. 2007; Anderson et al. 2015)). Further, the relative tolerance of wild-type system activation at *CCA**N**T* compared to *CCAA**N*** is interesting to consider in light of the evolutionary constraints acting on the functionalization of NF-Y and NF-CCT complexes. This permissiveness could provide space for evolution to act, allowing suboptimal interactions to lead to novel functions of the complex. Direct testing of these types of evolutionarily significant hypotheses is technically challenging, requiring more exhaustive mutagenic approaches such as alanine scanning mutagenesis and subsequent functional validation (Cunningham and Wells 1989). The DIMR system presented here was designed to bridge the gap in tools necessary to more directly address these types of more nuanced questions.

Methods

Yeast strain selection

The tests presented here all utilized the BY4735 (ATCC 200897) strain (*MATα ade2Δ::hisG his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0*), a derivation of the S288C laboratory strain. BY4735 was obtained and is available from the American Type Culture Collection (https://www.atcc.org/).

Yeast growth, transformation, and induction conditions

Yeast strains were uniformly grown at 30C. Wild type strains were grown with constant agitation in YAPD liquid medium and transformed through the traditional lithium-acetate based approach, as previously described. Transformants were selected after 2-3 days growth on synthetic medium lacking Leucine, Tryptophan, and Uracil ($L^- W^- Ura^-$).

For DIMR system induction, colonies were first grown to saturation in liquid $L^- W^- Ura^-$ media. Saturated cultures were then diluted 1:1 to a total sample volume of 500μL with fresh liquid $L^- W^- Ura^-$ media containing either ethanol (mock treatments) or one of a range of β-estradiol concentrations. These 500μL induction samples were cultured on deep, 2mL 96-well plates for 6 hours before sample collection for dual luciferase reporter testing. Plates were covered with Breathe Easy strips during induction.

Unless directly stated otherwise, each system test was conducted after 6 hours of induction, with treatments corresponding to mock induction (ethanol only), low induction (1-1.5nM β-estradiol), and high induction (10nM β-estradiol).

Module design and construction

The individual modules were each incorporated and carried on a different yeast shuttle vector – *pRS314* and *pRS315* for the effector and activator constructs, respectively, and a modified *pRSII316* removing an *NcoI* recognition site within the *URA3* coding sequence through site directed mutagenesis (described below) for the reporter construct. Initial module designs were synthesized by Biomatik®, while individual effector drop-in components were synthesized by Genewiz®.

The activator module encodes the Z_4EV artificial transcription factor, driven by the constitutive *ACT1* promoter (Flagfeldt et al. 2009) and flanked by the *CYC1b* terminator (Curran et al. 2013). The effector module uses the Z_4EV-bound promoter to drive expression of the 'polycistronic' effector mRNA. Each component has been translationally fused to a 5x-glycine flexible linker (Sabourin et al. 2007) and a unique epitope tag (component A: HA, component B: MYC, component C: FLAG). Viral T2A sites were incorporated after the HA and MYC tags to facilitate cleavage into individual

effector components. The entire cassette is flanked by the CYC1b terminator. The reporter module situates the two luciferase variants in a tail-to-tail fashion. The constitutively-active Renilla luciferase is expressed under the ADH1 promoter (McIsaac et al. 2013) and is flanked by the ADH1 terminator (Curran et al. 2013), while the conditionally-expressed Firefly luciferase is driven by a modified yeast minimal promoter (*ASN1* and *CIT1* presented here, (Sundseth et al. 1997; Nevoigt et al. 2006)) that includes previously described NF-Y bound CCAAT box sequences (Cao et al. 2014; Gnesutta et al. 2017a; Siriwardana et al. 2016) and is flanked by the CYC1b terminator.

Binding site and promoter architecture cloning, site directed mutagenesis

Drop-in Binding Site (DIBS) cloning was accomplished through annealing of 5' phosphorylated oligos and subsequent ligation into the reporter module. DIBS primers were designed and ordered as pairs of complementary oligos that were then annealed together by boiling for 5 minutes in annealing buffer (10mM Tris HCl, pH 8.0, 1mM EDTA, 50mM NaCl) and slowly cooling to room temperature. The reporter module was restriction enzyme digested (SacI/BamHI, New England Biolabs) and dephosphorylated (Quick Dephosphorylation Kit, New England Biolabs), then purified and concentrated through the Zymo DNA Clean and

Concentrator kit (PN). Ligations were set up with 3:1 ratios of insert:backbone free ends with ~50-150ng of purified backbone.

Site directed mutagenesis was conducted on effector entry clones through New England Biolabs Q5® Site Directed Mutagenesis Kit, following manufacturer's instructions. Primers for mutagenesis were designed with guidance of the NEBaseChanger™ tool, and mutagenesis was verified through restriction enzyme digestion (where appropriate) and Sanger sequencing through the University of Oklahoma's Biology Core Molecular Lab.

Western blots

Total yeast proteins were extracted using Y-PER™ Yeast Protein Extraction Reagents (Thermo Scientific, Cat no: 78991) following manufacturer's protocol. The protein concentrations were measured using Pierce™ BCA Protein Assay Kit (Thermo Scientific, Cat no: 23225) following manufacturer's protocol, and was quantified on a Synergy HTX Multi-Mode Reader (BioTeK, USA) at 562 nm wavelength. The protein concentrations were calculated through generation of a BSA standard curve. A total of 3 μg of protein was loaded to Mini-PROTEAN TGX Stain-Free Gels (10% gel, BIO-RAD*)*. Proteins were transferred to standard PVDF membranes through an OWL semi-dry transfer apparatus, and the presence of NF-YA2:HA, NF-YB2:MYC, or NF-YC3:FLAG was probed with high affinity anti-

HA primary antibodies (Roche, catalog no. 11 867 423 001), anti-MYC (Abcam, catalog no. ab9106), and anti-FLAG (Abcam, catalog no. F3165) followed by rabbit anti-rat (Abcam, catalog no. ab6734), goat anti-rabbit (Abcam, catalog no. ab205718), and goat anti-mouse (Abcam, catalog no. ab6789) HRP-conjugated secondary antibodies, respectively. A Bio-Rad ChemiDoc XRS imaging system was used for visualizing the protein blot after incubations with ECL plus reagent (GE Healthcare, catalog no. RPN2132).

Dual luciferase assays

Dual Luciferase assays were performed on the BioTek® Synergy™ HTX multi-mode plate reader fitted with dual reagent injectors, using the Illumination™ Firefly & Renilla Luciferase Enhanced Assay Kit (Goldbio®, cat# I-920) per manufacturer instructions, with the following alterations: (1) we used only 5µL of each culture, (2) samples were not spun down and/or washed, and (3) we used half-volume injections of both luciferase buffers. Our initial system tuning followed the manufacturer instructions more strictly, and we could determine either no functional difference or minor improvements between manufacturer instructions and our modified instructions (data not shown).

Statistical approaches

Statistics were calculated through Graphpad Prism. One-way ANOVA on relative luminescence values was used for comparisons of DIMR system activation to mock levels. When denoting significance in fold change-based metrics, the relative luminescence data was used for statistical analysis.

Figure and model construction

Individual graphs were generated through Graphpad Prism, while full figures were composed in Adobe Photoshop CC2018. The model in Figure 3.1A was constructed in Inkscape, while the flowchart in Figure 3.1B was constructed through draw.io v10.4.5 (https://draw.io/).

Figures and Tables

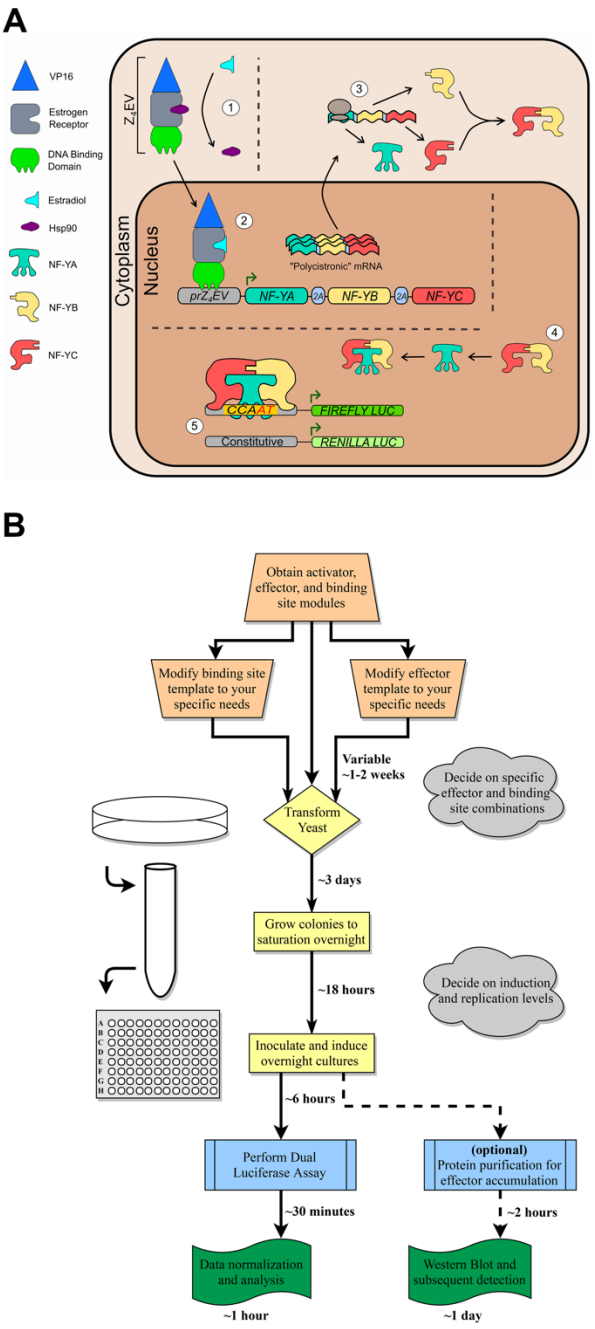

Figure 3.1. Model illustrating fundamental concepts and practical application of the DIMR system. (A) Circled numbers follow the flow of DIMR system activation: (1) β - estradiol activates the Z_4EV artificial transcription factor by competing with and replacing Hsp90, allowing for nuclear accumulation of Z_4EV; (2) Z_4EV induces transcription of the effector cassette; (3) translation of the effector cassette mRNA produces individual effector components through 2A-mediated translational cleavage; (4) effectors translocate to the nucleus and form functional complexes; and (5) transcriptional activation is measured as the ratio of conditionally-expressed firefly luciferase activity and constitutively-expressed Renilla luciferase activities from the reporter module. **(B)** Practical application of the DIMR system, including approximate timelines for design and implementation of an individual experiment.

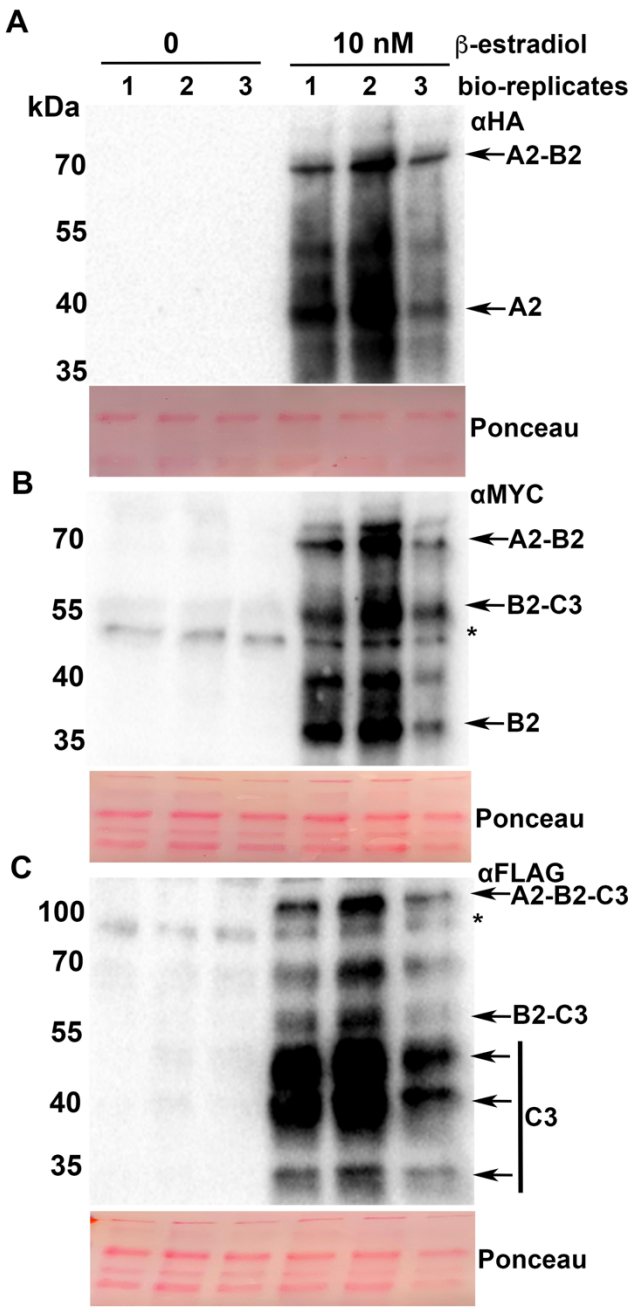

Figure 3.2. Induction and cleavage of NF-YA2, NF-YB2 and NF-YC3 effectors in yeast. Western blot analysis of **(A)** NF-YA2:HA, **(B)** NF-YB2:MYC, and **(C)** NF-YC3:FLAG accumulation in response to 10 nM β-estradiol induction for 6 hours. Three independent biological replicates, corresponding to lanes 1, 2, and 3, were tested for protein accumulation using anti-HA, anti-Myc and anti-Flag antibodies. Ponceau S staining of the PVDF membrane prior to transfer was used to test loading control. The experiment was repeated three times with similar result.

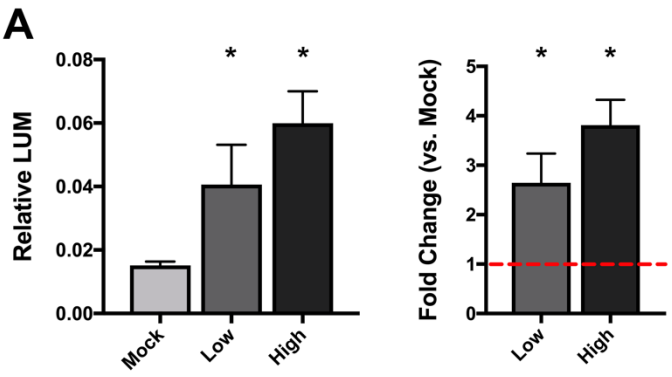

Figure 3.3. DIMR system test of NF-YA2/B2/C3 on FT CCAAT box before system optimization. (A) Relative luminescence of mock, low, and high concentration β-estradiol (0, 1uM, and 10uM, respectively) treated samples, and fold change levels (compared to mock) of low and high treated samples. Each condition includes at least 6 biological replicates. Asterisks indicate a significant increase between mock and treated samples, as determined through two-way ANOVA ($p < 0.01$). Error bars represent 95% confidence intervals.

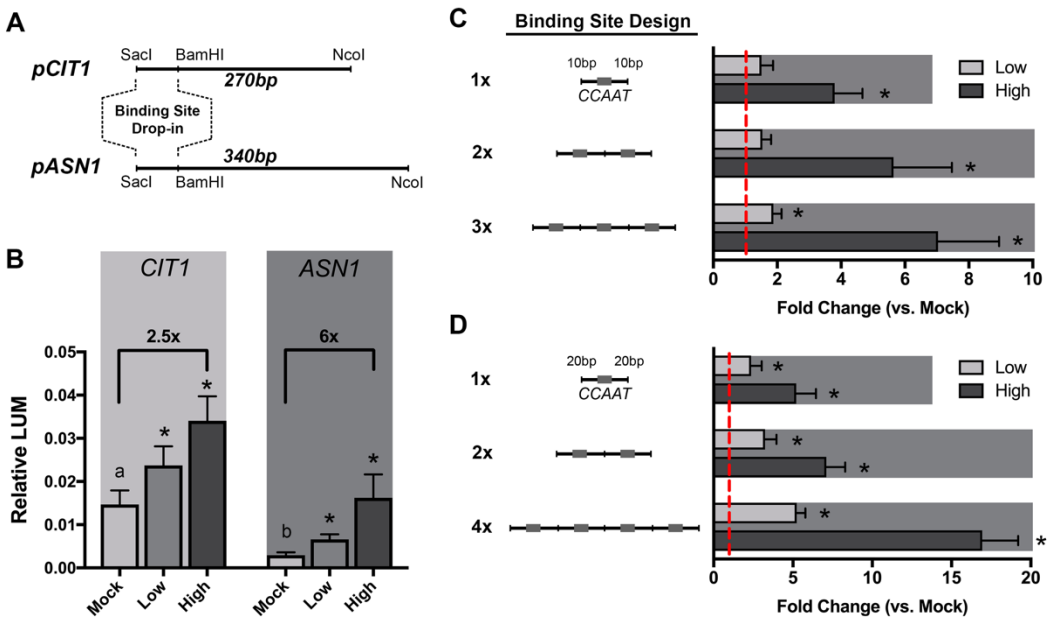

Figure 3.4. Parameter refinement of the DIMR system to optimize reporter dynamic range. (A) Model of two tested promoter architectures, derived from the *CIT1* and *ASN1* native yeast promoters. Binding site permutations were cloned into the reporter module through flanking SacI / BamHI restriction sites. **(B)** Effects of different promoter architectures of *CIT1* and *ASN1* on DIMR system activation of NF-Y$^{A2/B2/C3}$ on *FT CCAAT* box. **(C, D)** Effects of binding site multimerization and altered spacing between binding sites. In **(C)**, each binding site footprint was 25bp, while in **(D)**, each footprint was 45bp. Both sets were testing binding of NF-Y$^{A2/B2/C3}$ on *FT CCAAT* box permutations in the *ASN1*-based promoter architecture. Error bars represent 95% confidence intervals. Asterisks above individual bars indicate statistical significance over matched mock-treated samples, as determined through two-way ANOVA ($p < 0.01$). Asterisks above brackets connecting two bars indicate significance between the two conditions, determined similarly. Letters above mock-induced bars indicate significance groups between the different promoter architectures, also determined through two-way ANOVA ($p < 0.01$).

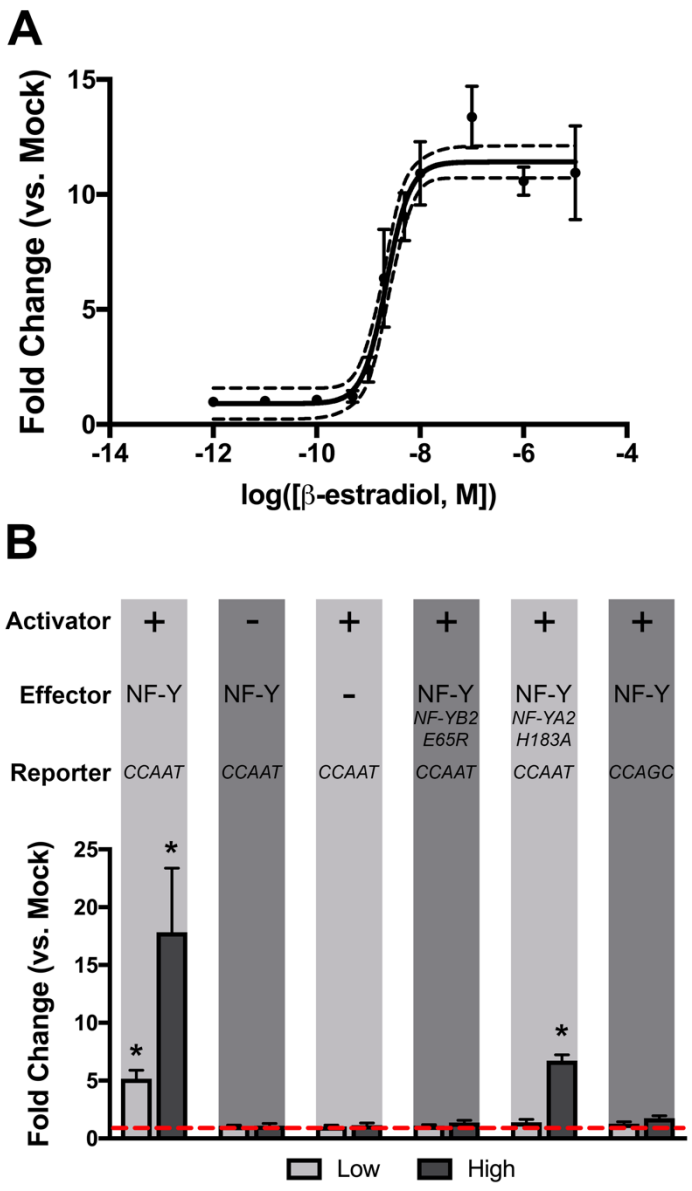

Figure 3.5. Functional assessment of DIMR system activation. (A) Dose response curve of NF-$Y^{A2/B2/C3}$ on the 4x *FT CCAAT* box in the ASN1 promoter architecture. The solid line fits the dataset to a 4-parameter logistic regression, with 95% confidence intervals plotted with dashed lines (R^2 = 0.942). **(B)** DIMR system requirements for NF-Y/CCAAT-mediated activation. Module components are generally described above each condition: +, present; -, absent; NF-Y, NF-$Y^{A2/B2/C3}$. Error bars represent 95% confidence intervals. Asterisks above individual bars indicate statistical significance over matched mock-treated samples, as determined through two-way ANOVA ($p < 0.01$); all bars lacking asterisks were found to be nonsignificant compared to matched mock-treated samples.

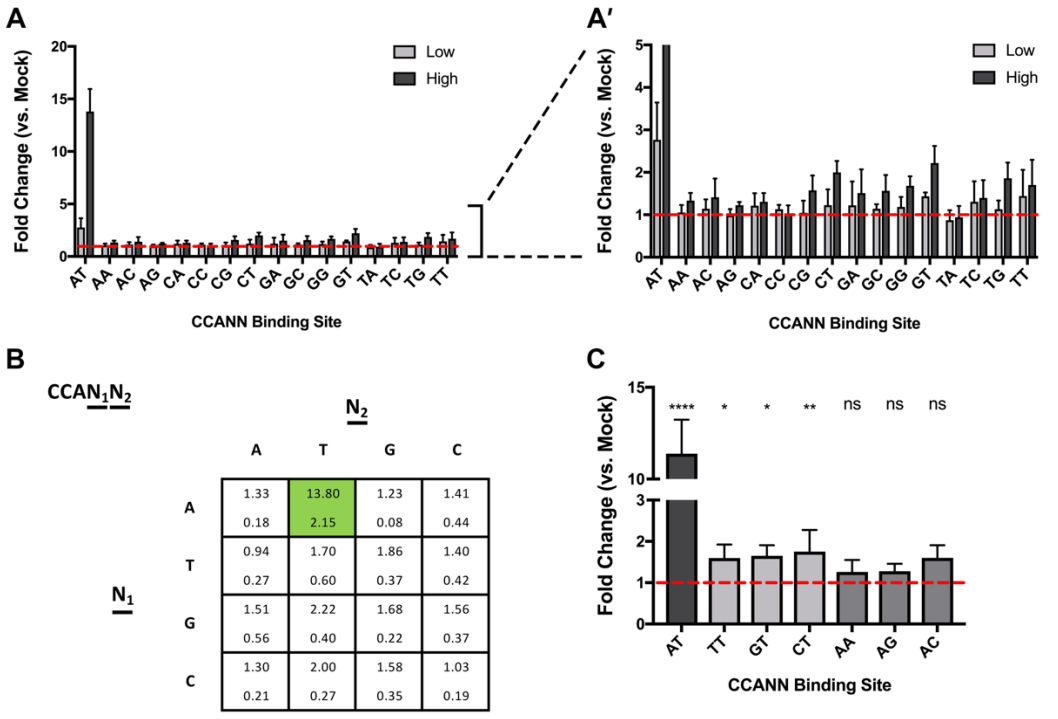

Figure 3.6. DIMR-based validation and investigation of NF-YA2/B2/C3 CCAAT binding. (A) System activation of NF-Y$^{A2/B2/C3}$ on the *FT CCAAT* box, with all possible binding site permutations corresponding to *CCA**NN***. Error bars represent 95% confidence intervals. Asterisks above individual bars indicate statistical significance over matched mock-treated samples, as determined through two-way ANOVA ($p < 0.01$); all bars lacking asterisks were found to be nonsignificant compared to matched mock-treated samples. **(B)** Graphical representation of the data presented in panel A. Each cell corresponds to a specific permutation of *CCA**NN***, and contains the average fold change (top) and standard deviation (bottom). (C) A targeted comparison of CCANT and CCAAN binding site permutations. Asterisks indicate significance compared to mock induction as determined by two-way ANOVA with Sidak's multiple comparison adjustment (****, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$).

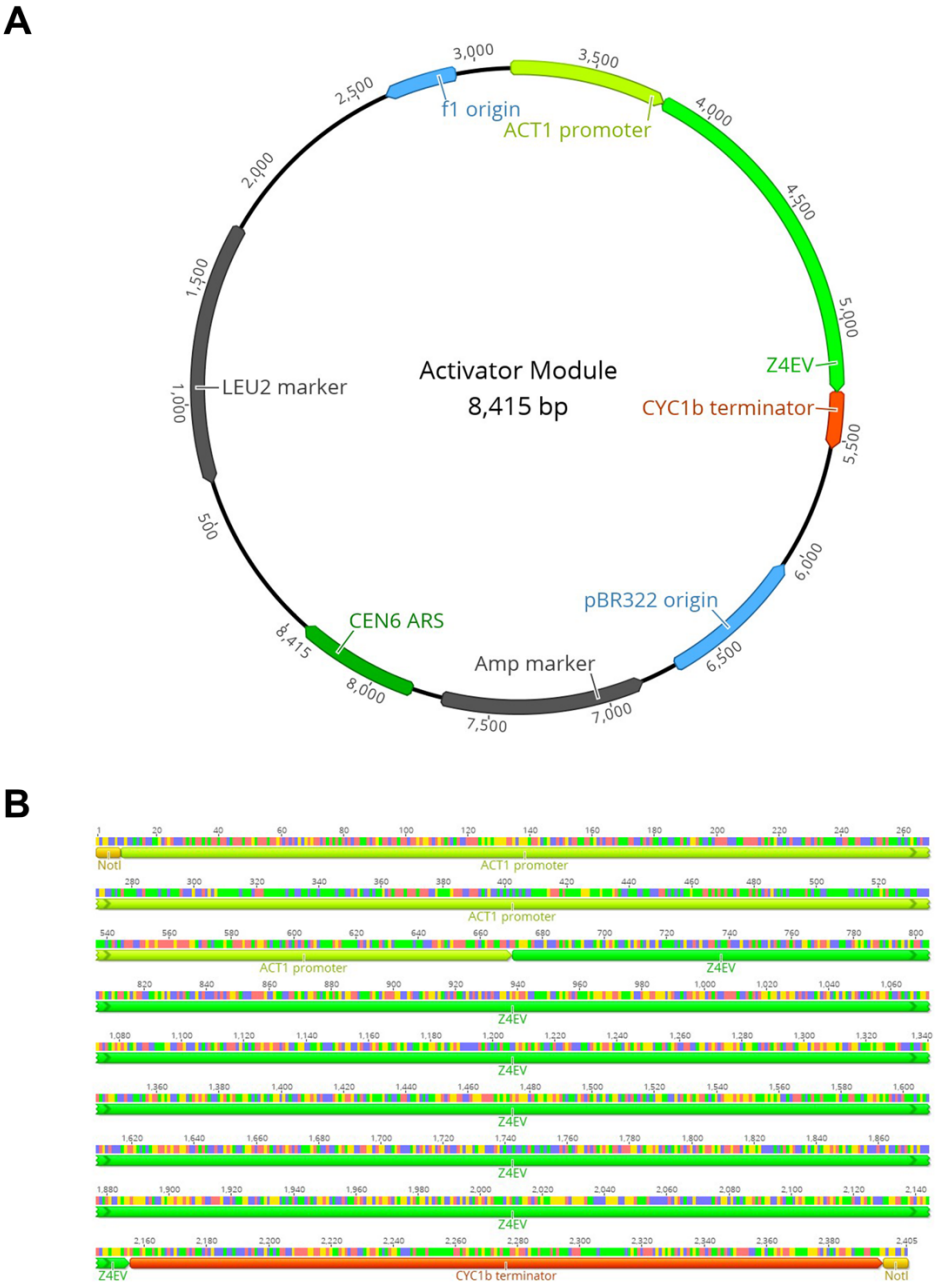

Figure S3.1. Main features of the Activator module. (A) Plasmid map of the DIMR Activator module showing important features and their orientation. (B) Zoomed-in view of the relevant Activator components.

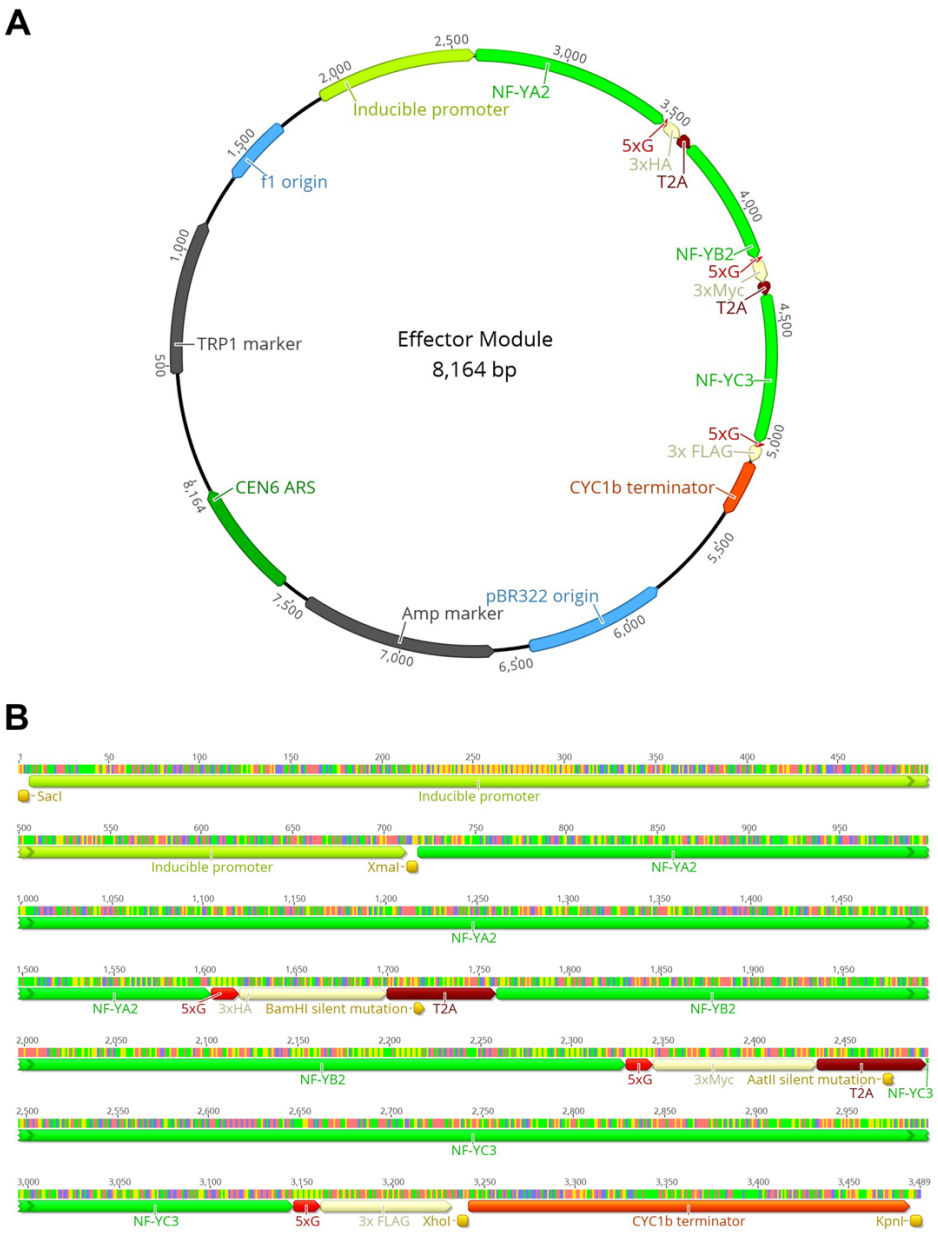

Figure S3.2. Main features of the Effector module. (A) Plasmid map of the DIMR Effector module showing important features and their orientation. (B) Zoomed-in view of the relevant Effector components, including important restriction enzymes used for swapping new components into or out of the system.

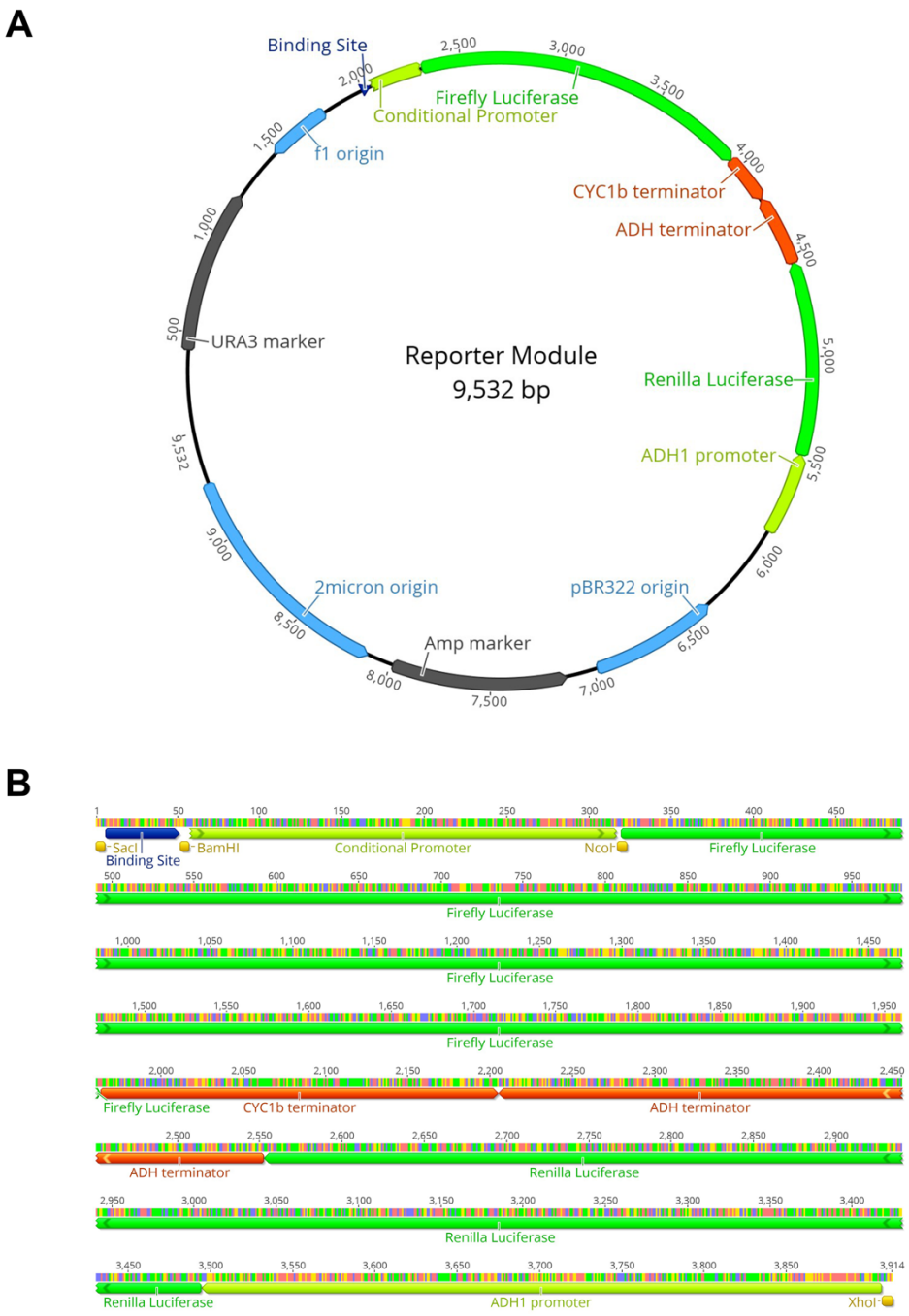

Figure S3.3. Main features of the Reporter module. (A) Plasmid map of the DIMR Reporter module showing important features and their orientation. (B) Zoomed-in view of the relevant Reporter components, including important restriction enzymes used for Drop-In Binding Site (DIBS) cloning.

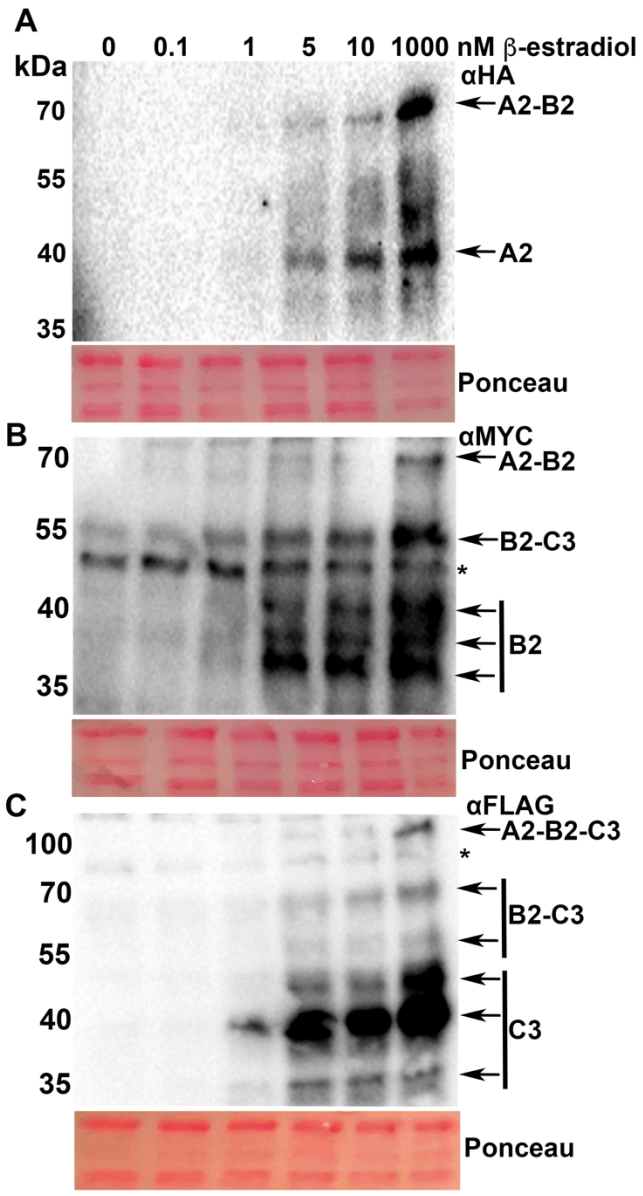

Figure S3.4. Induction gradient of NF-YA2, NF-YB2 and NF-YC3 effectors in yeast. Western blot analysis of **(A)** NF-YA2:HA, **(B)** NF-YB2:MYC, and **(C)** NF-YC3:FLAG accumulation in response to varying levels of β-estradiol induction for 6 hours. A single biological replicate was tested for protein accumulation using anti-HA, anti-Myc and anti-Flag antibodies. Ponceau S staining of the PVDF membrane prior to transfer was used to test loading control. The experiment was repeated twice with similar results.

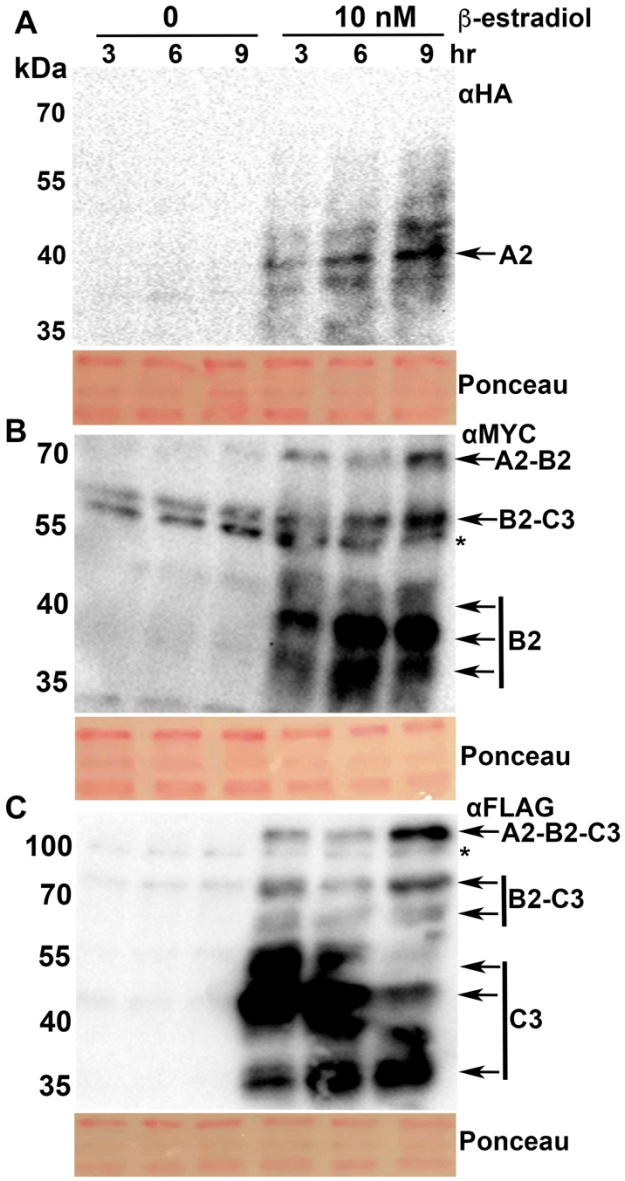

Figure S3.5. Induction time-course of NF-YA2, NF-YB2 and NF-YC3 effectors in yeast. Western blot analysis of **(A)** NF-YA2:HA, **(B)** NF-YB2:MYC, and **(C)** NF-YC3:FLAG accumulation in response to 10nMf β-estradiol induction for 3, 6, and 9 hours. A single biological replicate was tested for protein accumulation using anti-HA, anti-Myc and anti-Flag antibodies. Ponceau S staining of the PVDF membrane prior to transfer was used to test loading control. The experiment was repeated twice with similar results.

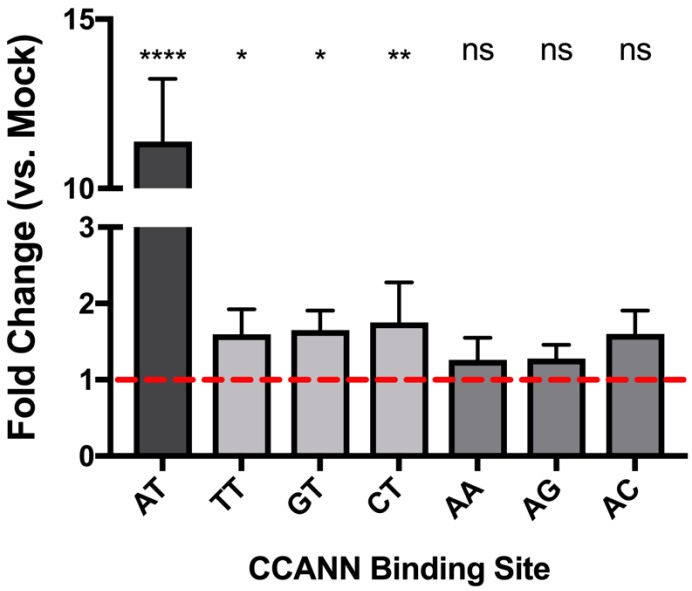

Figure S3.6. Activation of the DIMR system in effector cassettes with broken T2A sites. Point mutations were generated in the T2A sites between NF-YA / NF-YB, between NF-YB / NF-YC, and in both T2A sites simultaneously. T2A sites were broken through a proline to alanine mutation near the end of each T2A site, as previously described (Beekwilder et al. 2014). Error bars represent 95% confidence intervals. Asterisks above individual bars indicate statistical significance over matched mock-treated samples, as determined through two-way ANOVA ($p < 0.01$). Brackets with asterisks indicate significance between induced samples. ns, not significantly different.

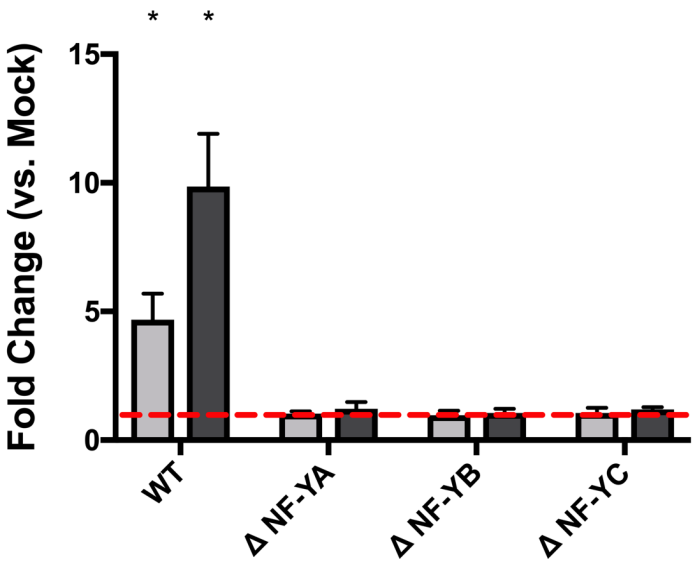

Figure S3.7. Activation of the DIMR system in the absence of individual NF-Y effectors. NF-$^{YA2/B2/C3}$ Effectors with empty components for NF-YA (ΔNF-YA), NF-YB (ΔNF-YB), or NF-YC (ΔNF-YC), were tested for ability to activate the 4x *FT CCAAT* box in the *ASN1*-based promoter. Error bars represent 95% confidence intervals. Asterisks above individual bars indicate statistical significance over matched mock-treated samples, as determined through two-way ANOVA ($p < 0.01$); all bars lacking asterisks were found to be nonsignificant compared to matched mock-treated samples.

Chapter 4: NF-YC and HY5 Show Complex Genetic Relationships in Regulation of Primary Root Elongation in *Arabidopsis thaliana*

Zachary A. Myers, Reid Selby, and Ben F. Holt III

Abstract

Because of the limited energy contained in a seed, plants must fine-tune their initial seedling growth and development very closely with environmental conditions. In order to support the transition to a photosynthetic lifestyle, seedlings must seek out the water and nutrients necessary for life from their immediate surroundings. We found indications that several *NUCLEAR FACTOR-Y, SUBUNIT C (NF-YC)* genes involved in light-mediated hypocotyl inhibition might also be involved in root elongation and hormone responses during this same developmental window. Further, the NF-YC genes share interesting physical and genetic relationships with *ELONGATED HYPCOTYL 5 (HY5)*, a key regulator of many developmental processes in seedlings. To better understand the role of the NF-YC in regulating primary root elongation, we examined mutant lines of *hy5* and *nf-yc3, 4, 9* (*nf-yc triple*), as well as the higher-order *nf-yc triple hy5* quadruple mutant, grown in a variety of light conditions. We identified several distinct genetic relationships in phenotypes of *nf-yc triple hy5* mutants that ultimately suggest a combination of both independent and cooperative transcriptional regulation by NF-YC and HY5. While direct regulatory targets have thus far remained elusive, the genetic relationships between NF-YC and HY5 clearly point to a complex interplay of convergent and divergent regulatory function during early seedling development.

Introduction

Life is built on the ability to read and appropriately respond to both internal and external signals. This response takes the form of differential gene expression, and is mediated by a class of proteins called transcription factors (TFs). TFs fine-tune the location and frequency of transcription, the process of creating an RNA intermediate from the molecular blueprint of DNA. The core processes of gene transcription are accomplished through the catalytic activity of RNA polymerase II and its associated factors, which are able to simultaneously unwind and read a template DNA strand and transcribe it into messenger RNA (mRNA). While the mechanistic function of TFs is varied, the consequence of their action is manifested through a change in the ability of the RNA polymerase complex to appropriately bind DNA and begin transcription.

In order to probe the genetic relationship between TFs of interest, scientists leverage alleles with variant TF function and observe the effect on a phenotype of interest (Wright 1932). These variant alleles often encode loss-of-function versions of a given TF. By comparing the effects of individual mutations to the effects of multiple mutations, hypotheses can be developed regarding the relationship between the individual mutations and genes. The ways in which different TFs cooperate or compete to regulate transcription is extremely varied, but some generalities can be drawn.

TF regulation that converges on a particular biological process through different signal transductions typically shows an additive phenotype in combined loss-of-function mutant backgrounds. While this does not preclude TFs regulating the same or similar sets of genes, it does suggest that they do so independent of one another. Most other genetic relationships are more nuanced, including multiple types and manifestations of epistasis, a genetic relationship where one mutant phenotype is dependent on the presence of a second mutation (Whitlock et al. 1995). Magnitude epistasis is an instance where a higher-order mutant is significantly more phenotypically different than could be expected through comparison of individual mutants, and is often described as a 'synergistic' phenotypic response. Sign epistasis and reciprocal sign epistasis are cases where higher-order mutant phenotypes flip directional effect (i.e., change sign) relative to either one or both single mutants, respectively (Weinreich et al. 2005). These phenotypic observations are often difficult to disentangle, particularly when considering regulation at a molecular level.

Research indicates that TFs work much more interdependently than initially understood, especially within families of related proteins (Hai and Curran 1991; Benbrook and Jones 1994). Broad cooperativity was recently identified amongst a pool of bZIP TFs (Rodriguez-Martinez et al. 2017) and had significant impacts on binding site specificity of each heterodimer pair, suggesting a complex interplay between independent (homodimer) and

interdependent (heterodimer) transcriptional regulation (Rodriguez-Martinez et al. 2017). This pattern of cooperative TF function within a family is extensively more complex in plants, the focus of this study, where most TF families evolved to encode many more members than those of other eukaryotes (Shiu et al. 2005).

To better contextualize the impact of these genetic relationships, we studied the interactions between mutations in the genes *ELONGATED HYPOCOTYL 5* (*HY5*) and *NUCLEAR FACTOR-YC* (*NF-YC*) in the model plant *Arabdopsis thaliana* (Arabidopsis). HY5 is a bZIP-type TF that has seen extensive study into its roles in seedling development (Ang and Deng 1994; Lee et al. 2007), coordination of root architecture (Oyama et al. 1997; Sibout et al. 2006), and hormone signaling and responses (Cluis et al. 2004; Vandenbussche et al. 2007; Chen and Xiong 2008; Li and He 2015). NF-Y is a heterotrimeric TF complex composed of three subunits (NF-YA, NF-YB, and NF-YC) that binds the pentanucleotide *CCAAT (Swain et al. 2017; Siefers et al. 2009)*, and is an important regulator of many biological processes, including plant-specific roles in flowering time (Kumimoto et al. 2008; Kumimoto et al. 2010; Gnesutta et al. 2017a; Siriwardana et al. 2016), photomorphogenesis (Myers et al. 2016; Tang et al. 2017), and hormone responses (Siriwardana et al. 2014; Kumimoto et al. 2013). Previous studies have identified both physical and genetic interactions between NF-YC and HY5 (Kumimoto et al. 2013; Myers et al. 2016), the majority of which have

been explored through their overlapping regulation of early seedling development.

Plants fine-tune their initial seedling growth and development very closely with environmental conditions (Boyer 1982). Because of the finite resources available in the endosperm of a freshly-germinated seedling, quick and proper establishment of photosynthetic capacity is critical for viability (Kircher and Schopfer 2012). In order to support the transition to a photosynthetic lifestyle, seedlings must seek out the water and nutrients necessary for life from their immediate surroundings. This nutrient acquisition is mediated through the plant's root system, where primary root elongation, lateral root initiation, and lateral root elongation are altered in response to nutrient availability (López-Bucio et al. 2003), water accessibility (Skirycz and Inzé 2010), photosynthetic activity (Kircher and Schopfer 2012), and hormone signaling (Blilou et al. 2005a; Fu and Harberd 2003). Because of the sheer number of contributing pathways, seedling root growth and development is a highly plastic trait that is not well understood (Sultan 2000).

While we might not fully understand the signal integration coordinating root growth and architecture, the developmental schemes underpinning root biology are well established. A growing seedling root can be divided into three primary developmental zones: (1) the root apical meristem, a terminal

zone that is the site of active cell division contributing to primary root growth; (2) the elongation zone, a site of active cell elongation situated just above the root apical meristem; and (3) the maturation zone, where individual cells begin to differentiate and specialize into the various tissues of the root. Primary root elongation can be modulated by altering the rate of cell division at the root apical meristem or the extent of cell extension in the elongation zone. Several signaling events are thought to be important regulators of these processes, especially cross-talk between auxin and cytokinin signaling (Bishopp et al. 2011). Auxin is produced in the shoot and forms a local maximum at the root apical meristem through the function of several *PINOID (PIN)* polar auxin transport proteins (Gälweiler et al. 1998), and this accumulation is necessary for maintenance of meristematic function by promoting cell division and inhibiting differentiation (Blilou et al. 2005b). In contrast, cytokinin is produced in the root apical meristem and transported acropetally, promoting cell differentiation and inhibiting cell division outside of the root apical meristem (Werner et al. 2001). Interestingly, despite their opposing functions, the two signaling networks also seem to require one another for proper hormone biosynthesis, distribution, and perception (Aloni et al. 2006).

We found indications in previously published RNA sequencing data that several *NF-YC* genes might be involved in root elongation and hormone responses during seedling development (Myers et al. 2016). HY5 is known

to be an important regulator in both of these areas (Oyama et al. 1997; Cluis et al. 2004; Sibout et al. 2006; Vandenbussche et al. 2007), and as such, we questioned whether NF-YC might regulate these processes through or independent from HY5. To address this question, we examined root elongation in seedlings grown in a variety of light conditions in mutant lines of *hy5* and *nf-yc3, 4, 9* (*nf-yc triple*), as well as the higher-order *nf-yc triple hy5* quadruple mutant. Previous work identified largely independent roles of NF-YC and HY5 in regulating hypocotyl elongation (Myers et al. 2016). Despite this, the two are also thought to physically interact, and differential gene expression analyses suggested that some transcriptional targets may be co-regulated by the two (Myers et al. 2016). This analysis uncovered several interesting phenotypic and genetic relationships between *nf-yc* and *hy5* mutants in seedling primary root elongation, indicating that their ability to function independently or cooperatively is context-dependent.

Results

We analyzed a previously-published dataset (Myers et al. 2016) that examined differentially expressed genes in white light grown seedlings of Arabidopsis *nf-yc triple* mutants and found significant enrichment for genes involved in processes either directly or indirectly related to root elongation. This includes Gene Ontology (GO) categories involved in photosynthesis, light responses, and response to water in significantly down-regulated genes (Table 4.1), and GO categories involved in cell wall modification, hormone responses, and root development in significantly up-regulated genes (Table 4.2). Because of these observations, and because of the previously-identified interactions between HY5 and NF-YC, we examined root elongation in 5-day old *nf-yc triple*, *hy5*, and *nf-yc triple hy5* mutants (Figure 4.1). We found that both the *nf-yc triple* and *hy5* genotypes had significantly longer primary root lengths than wild type (Col), with *hy5* also being significantly longer than the *nf-yc triple*. When examining the combination of these two backgrounds, we identified a sign epistatic relationship between the two mutant lines, where the *nf-yc triple hy5* was significantly shorter than hy5, but not significantly different than *nf-yc triple*. The impact of this type of relationship is difficult to interpret without a mechanistic understanding of each set of TFs, but does suggest that NF-YC and HY5 function interdependently in some way to regulate root elongation during seedling development.

Wavelength-dependent root elongation defects in nf-yc and hy5 mutants

Light is an important regulator of root growth and development, but plants do not perceive white light as a single signal. Instead, various classes of photoreceptors perceive light as distinct and separable signals corresponding to blue (~450nm), red (~670nm), and far red (~700nm) wavelengths of light. Phenotypes of white light grown seedlings are often more difficult to interpret because of the interactions between these distinct pathways. To deconvolute the genetic relationship observed in white light, we also examined primary root elongation in narrow wavelength ranges of blue, red, and far red light.

In seedlings grown in constant red light, we found that the *nf-yc triple* and *hy5* mutant lines had significantly shorter primary root lengths, and that this defect was additive in the *nf-yc triple hy5* mutant, which was significantly shorter than either parental mutant (Figure 4.2). While the phenotypes of each parental mutant demonstrate a role in regulating primary root elongation, this type of additive defect suggests that their regulation is primarily independent of one another. We observed similar relationships in red light grown seedlings in both hypocotyl length (which was previously known) and cotyledon area (Figure S4.1), further strengthening the argument that NF-YC and HY5 function independent of one another in red light mediated growth and development.

In seedlings grown in either blue (Figure 4.3A) or far red (Figure 4.3B) light, we observed a much more intriguing relationship. In both conditions, the primary root length of both *nf-yc triple* and *hy5* mutant lines was significantly longer than Col; however, the *nf-yc triple hy5* quadruple mutant was significantly shorter than either Col or its two parental mutant backgrounds. This observation is a typical manifestation of reciprocal sign epistasis, and suggests that the functions of NF-YC and HY5 are intimately linked in some way. Interestingly, while other phenotypic measures in red light showed the same additive genetic relationship, such as hypocotyl length and cotyledon area, the same is not the case in either blue or far red light mediated phenotypes. Previous work identified additive hypocotyl phenotypes in blue light, and magnitude epistasis-based (synergistic) relationships in hypocotyl phenotypes in far red light.

Auxin inhibition rescues hypocotyl phenotypes in nf-yc and hy5, but not nf-yc triple hy5

Finally, because we had identified GO enrichment for processes involved in various hormone responses, we asked whether *nf-yc* and *hy5* mutants responded to auxin inhibition in similar ways. We found that the long hypocotyl phenotypes of each mutant could be suppressed through auxin inhibition, but inhibition in the *nf-yc triple hy5* did not fully suppress its long hypocotyl phenotype (Figure 4.4). This relationship was present in both

white (Figure 4.4A) and red (Figure 4.4B) light, but was more pronounced in white light growth conditions. Many defects in light signaling or perception ultimately funnel down to defects in auxin biosynthesis or distribution (Jensen et al. 1998b; Takahashi et al. 2012), and as such, it is interesting that the auxin-inhibited *nf-yc triple hy5* mutant is still significantly longer than Col. One critical question raised by this observation is whether the auxin inhibition observed here is saturating, and a more thorough examination at higher concentrations of 2,4-D could address this. Aspects of seedling growth and development independent of auxin are perturbed in this higher-order mutant, despite the observation that the individual mutants could be reverted to a wild type hypocotyl length with the same dose of auxin inhibition. This suggests that some aspect of the interdependent functions of NF-YC and HY5 act separately from auxin, though any exact mechanism remains elusive.

Discussion

We identified several different genetic relationships in phenotypes of *nf-yc triple hy5* mutants that ultimately suggest a complex interplay of independent and cooperative transcriptional regulation by NF-YC and HY5. Primary root elongation in white light was enhanced in the *nf-yc triple* and *hy5* mutants but not in the *nf-yc triple hy5* higher-order mutant. In red light, the reduced primary root elongation observed in *nf-yc triple* and *hy5* was additive in the *nf-yc triple hy5* mutant. Finally, in blue and far red light, the *nf-yc triple* and *hy5* mutants had significantly longer primary roots than wild type, while the *nf-yc triple hy5* mutant had significantly shorter primary roots than wild type.

HY5 has been thought to regulate root development and associated processes since its first identification in 1990, where *hy5* loss of function mutants were shown to have altered primary root growth patterns (Oyama et al. 1997) and increased numbers of lateral roots (Chen and Xiong 2011). More recently, HY5 was shown to function as a mobile TF signal important for coordination of root elongation and nitrate uptake in response to photosynthetic activity (Chen et al. 2016). The uncovered mechanism follows *HY5* transcription and translation in photosynthetically-active leaves and its subsequent translocation to the root, where it physically interacts with the promoter of *NITRATE TRANSPORTER 2.1* (NRT2.1) (Chen et al. 2016). This same study found that *hy5* mutants grown in white light had

reduced primary root elongation rates, which is not consistent with our observations. Several important distinctions in experimental approaches could explain these differences, including key differences in growth media compositions and data collection. In particular, the nitrate concentration between the two mediums is significantly different: Whereas we conducted our experiments on Gamborg's B5 basal medium (1900mg KNO_3 / L), the previous root elongation rate experiments were conducted on ½ Murashige and Skoog medium (950mg KNO_3 / L). Reduced available nitrate levels are known to inhibit primary root elongation during early seedling development (Zhang and Forde 2000), and *hy5* mutants show reduced nitrate influx rates (Chen et al. 2016); however, a more direct set of experiments would be necessary to establish a true link between these observations.

In contrast to HY5, less is known about the roles and mechanisms of NF-Y-mediated transcriptional regulation during early seedling development; however, roles in related processes have been described in many cases. Several NF-YA genes are important for root architecture coordination (i.e., balancing primary/secondary root growth) and are post-transcriptionally regulated through the action of several microRNA169 isoforms (Hanemian et al. 2016; Sorin et al. 2014). Ectopic overexpression of NF-YB2 in Arabidopsis led to increased primary root lengths; however, the biological relevance of this observation has yet to be established (Ballif et al. 2011). It is also likely that the atypical NF-YB6 and NF-YB9 have root phenotypes

that have yet to be explored, as they have been shown to broadly impact both embryonic and post-embryonic seedling developmental schemes (Huang et al. 2015; Baumlein and Junker 2012).

Equally little is known about direct regulation of seedling root development by the NF-YC. *NF-YC1, 3, 4,* and *9* were identified as functionally-redundant genes regulating hypocotyl elongation (Myers et al. 2016), possibly through modulation of histone acetylation levels on key promoters (Tang et al. 2017); however, the focus of these studies has been on hypocotyl-based elongation, and no specific attention has yet been directed toward understanding how these same genes might regulate root elongation.

Interestingly, both *nf-yc triple* and *hy5* had shorter primary roots than wild type in red light, but longer primary roots in blue and far red light. This relationship was not seen in hypocotyl phenotypes, where both mutants present elongated hypocotyls in nearly all light conditions *(nf-yc triple* is only long in lower intensity or shorter photoperiod white light). Reductions in primary root elongation during seedling development have been linked to reduced photosynthetic activity (Kircher and Schopfer 2012), and red light is the most photosynthetically active wavelength of light; however, blue light is also photosynthetically active. Further, while GO analysis suggests impaired photosynthetic function in *nf-yc triple* (Table 4.1) and *nf-yc triple hy5* (Myers et al. 2016), only *hy5* has been experimentally shown to have

reduced photosynthetic pigment accumulation and photosynthetic rates (Kobayashi et al. 2012; Toledo-Ortiz et al. 2014; Chen et al. 2016). While the red-light root phenotypes described here could be linked to issues with reduced photosynthetic rates, the blue and far red root phenotypes suggest other defects.

Synthesizing a mechanistic understanding of NF-YC and HY5 regulation of root elongation is significantly complicated when considering their different observed genetic relationships. The data presented above suggests that HY5 and NF-YC work independently in red light, but their functional relationship in other light conditions appears at least partially interdependent. While we have not found data to support a specific, interdependent mechanism of NF-YC and HY5 function, the types of epistasis observed between them has been studied in other contexts. Sign and reciprocal sign epistasis have long been thought to influence fitness landscapes, a concept that describes the interplay of different mutations and their effect on fitness (Wright 1932, 1988). In particular, sign and reciprocal sign epistasis facilitate a rugged fitness landscape, which more faithfully captures the existence of many peaks and valleys of increasing and decreasing fitness when mutations are considered in combination (Whitlock et al. 1995; Phillips 2008; Poelwijk et al. 2011; Kvitek and Sherlock 2011). Despite the apparent importance of epistasis in evolutionary biology, research into the more-specialized forms discussed

here have primarily been constrained to studies on speciation and reproductive isolation (Anderson et al. 2010; de Visser and Elena 2007; Presgraves 2010). One recent study in *E. coli* utilized error-prone amplification of the *tetR* and *lacI* transcription factors and tested their ability to function in a synthetic signal cascade (Nghe et al. 2018). Through this approach, the authors identified a number of mutant combinations of *tetR* and *lacI* that showed either sign or reciprocal sign epistasis, and were able to model the effect of these interactions on local fitness landscapes; however, their predictive ability was built upon the foreknowledge of the mechanistic function of their synthetic signal cascade. It is unclear how the observations presented in this study might be extended to TFs without a mechanistic description, but it is apparent that these relationships have molecular and evolutionary significance.

A more complete, mechanistic understanding of the manifestations of sign and reciprocal sign epistasis could shed light on the interactions between NF-YC and HY5. Unfortunately, these types of examples are relatively few, have not focused on transcriptional regulation, and are rarely dissected to the level of detail necessary for modeling onto other epistatic relationships. Research into the evolution of multi-drug resistance in *E. coli* has identified unexpected fitness tradeoffs between mutational drug resistance and plasmid-conjugated drug resistance. Antibiotic resistance is an energetically expensive trait (Andersson and Levin 1999) and has a clear

negative impact on fitness when expressed in the absence of selection (Andersson 2006); however, in a significant number of instances, the fitness effect of this deleterious antibiotic resistance was found to change sign (e.g., from deleterious to beneficial) when present in combination with a plasmid conferring antibiotic resistance (Silva et al. 2011). The molecular underpinnings of this relationship have not been explored, despite the significant implications this research has on the expansion of multi-drug resistance in bacteria.

Despite a lack of solid models describing these epistatic interactions at the level of transcriptional regulation, at least one possible molecular explanation exists for the observed reciprocal sign epistasis between NF-YC and HY5 in blue and far red mediated primary root elongation. As discussed above, *HY5* is a critical regulator of *NRT2.1*, a high-affinity nitrate uptake transporter. Interestingly, this same transporter is significantly down-regulated in RNA sequencing analyses in the *nf-yc triple* mutant (Myers et al. 2016). While this observation has not yet been functionally validated, one can imagine a scenario where NF-YC and HY5 enhance each other's ability to regulate *NRT2.1*. In this scenario, loss of either *NF-YC* or *HY5* will result in a reduced amount of *NRT2.1* transcription, and a subsequent reduction in the rate of nitrate uptake; however, reduced uptake rates can be abrogated by increased concentrations of exogenous nitrate (Remans et al. 2006). As described above, we are using growth media with a relatively high

nitrate level, and as such, a partial reduction in the cost of nitrate uptake machinery might actually result in a net fitness gain (i.e., because sufficient nitrate can be acquired with fewer copies of NRT2.1, removing the metabolic cost of another TF and its associated targets could save net energy). Extending this model further, loss of both *NF-YC* and *HY5* would result in an insufficient amount of *NRT2.1* production, reduced net nitrate uptake, and reduced primary root elongation. This model of NF-YC and HY5-mediated regulation of primary root elongation could serve as a starting point to better understand their independent and interdependent activities, and could be easily tested by examining the effects on primary root elongation in *nf-yc triple* and *hy5* mutants in reduced nitrate conditions.

Finally, we examined the impact of auxin inhibition on hypocotyl elongation in *nf-yc triple*, *hy5*, and *nf-yc triple hy5* mutants. Auxin is a major pathway influencing early seedling development, where it promotes cell elongation in the hypocotyl and coordinates primary root elongation with lateral root initiation and development (Overvoorde et al. 2010; Jensen et al. 1998a). Hypocotyl phenotypes of *hy5* mutants are largely thought to be tied to auxin defects (Cluis et al. 2004), and *nf-yc triple* and *nf-yc triple hy5* mutants have GO enrichment in hormone- and auxin-related categories (Table 4.1 and 4.2, (Myers et al. 2016)). We found that auxin inhibition in red light grown *nf-yc triple* mutants suppressed their elongated hypocotyl phenotype, providing some initial validation of the auxin-related defects predicted

through GO enrichment analysis. Surprisingly, hypocotyl phenotypes of *nf-yc triple hy5* mutants were not completely rescued by auxin inhibition treatment, suggesting that some cooperative function of NF-YC and HY5 operates independent of auxin. A more thorough experiment looking at dose-dependent responses to auxin inhibition over a broader range of inhibition is important to ensure that the auxin inhibition treatment presented above is saturating.

Materials and Methods

Plant growth and transgenic lines

All plants were of the Col-0 ecotype and were grown at 22°C. Prior to germination on plates, seeds were cold-stratified in a dark 4°C walk-in for 2-3 days. Plants grown in continuous white light were grown in a Conviron ATC13 growth chamber or a custom walk-in growth chamber. Plants grown in monochromatic light experiments (narrow wavelength bands centered on: blue, 450nm; red, 670nm; far red, 700nm) were grown in a Percival E30-LED growth chamber after initial exposures to 4 hours of white light to induce germination. Root elongation assays were conducted on 2% agar plates supplemented with Gamborg B5 Basal Medium (PhytoTech Laboratories), while all other assays were conducted on 0.8% agar B5 plates.

Hypocotyl measurements and auxin inhibition treatments

Auxin inhibition was achieved through the inclusion of 1μM 2,4-dichlorophenoxyacetic acid (2,4-D; Cayman Chemical, PN 10190) in the growth media. To facilitate proper measurement, all plants were straightened on the plate before taking pictures. Pictures were processed and all individual hypocotyls were traced and measured in FIJI. These raw values were then analyzed through Prism as described below.

Gene Ontology enrichment analyses

Gene Ontology (GO) enrichment was explored in a previously-described and analyzed RNASeq experiment, using published differentially expressed gene lists (REF). Arabidopsis gene identifiers for significantly up- or down-regulated genes from the *nf-yc triple* genotype, compared to wild type, were fed into the PANTHER classification system (http://www.pantherdb.org/) and examined for enrichment of specific biological processes. The tables presented here contain a selection of GO categories relevant to the project presented here, and are not intended to be exhaustive.

Stats and Figure Construction

All statistics presented above were calculated through Graphpad Prism. Figures and tables were constructed through a combination of Microsoft Office, Graphpad Prism, and Photoshop.

Figures and Tables

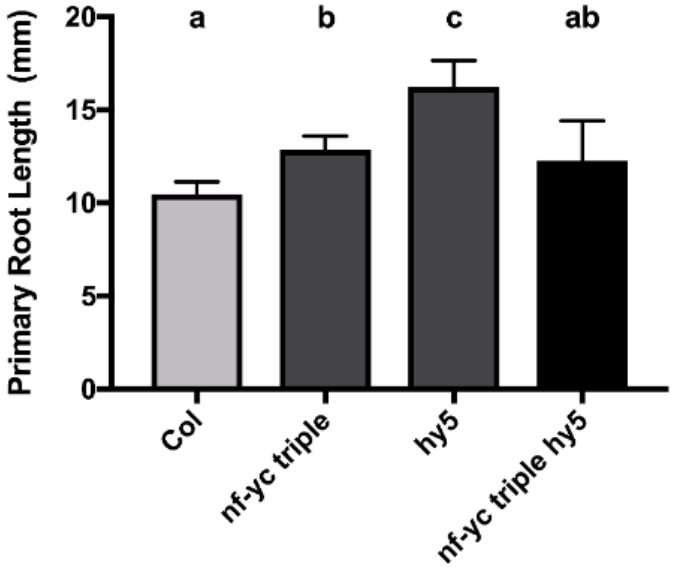

Figure 4.1. Average primary root length in white light grown seedlings. Seedlings were grown for 5 days in constant white light (~100µmol/m^2s). Each population contains at least 20 seedlings. Errors bars represent 95% confidence intervals. Letters above each bar represent statistical binning, with statistically significant differences identified through nonparametric two-way ANOVA with Dunn's correction for multiple comparisons, $p<0.05$.

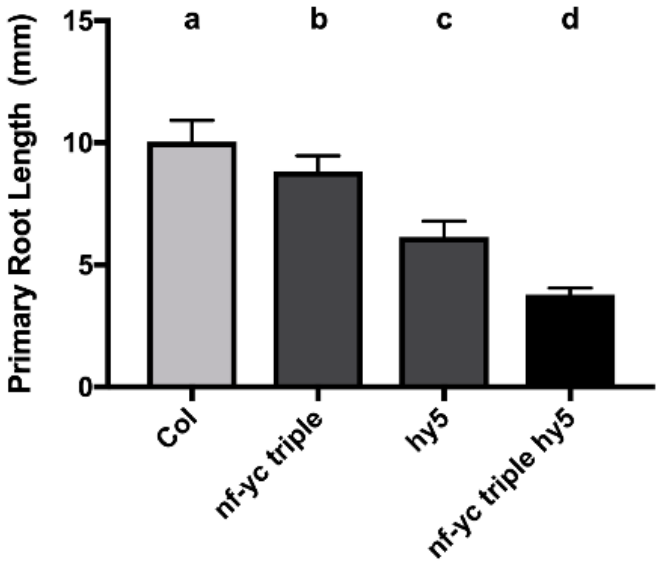

Figure 4.2. Effects of red light on primary root elongation. Seedlings were grown for 5 days in constant red light (~20µmol/m^2s). Each population contains at least 20 seedlings. Errors bars represent 95% confidence intervals. Letters above each bar represent statistical binning, with statistically significant differences identified through nonparametric two-way ANOVA with Dunn's correction for multiple comparisons, $p<0.05$.

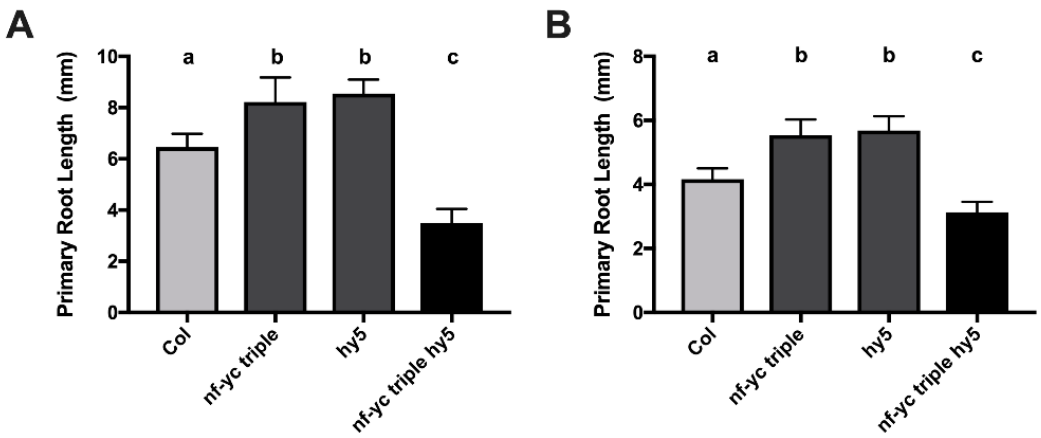

Figure 4.3. Effects of blue and far red light on primary root elongation. Seedlings were grown for 5 days in (A) constant blue (~25µmol/m^2s) or (B) constant far red (~1µmol/m^2s) light. Each population contains at least 20 seedlings. Errors bars represent 95% confidence intervals. Letters above each bar represent statistical binning, with statistically significant differences identified through nonparametric two-way ANOVA with Dunn's correction for multiple comparisons, $p<0.05$.

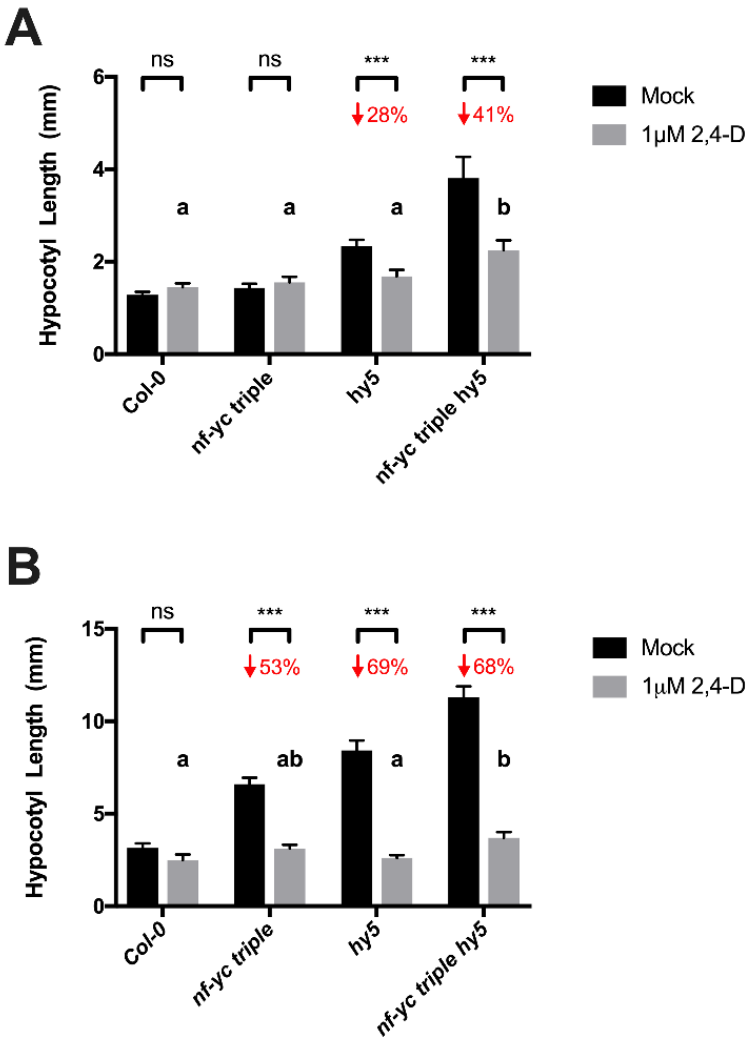

Figure 4.4. Effect of auxin inhibition on hypocotyl elongation. Seedlings were grown on either mock or 1µM 2,4-dichlorophenoxyacetic acid (2,4-D) for 5 days in (**A**) white light or (**B**) red light. before hypocotyl measurement. Each population contains at least 20 seedlings. Errors bars represent 95% confidence intervals. Letters above each bar represent statistical binning, with statistically significant differences identified through nonparametric two-way ANOVA with Dunn's correction for multiple comparisons, $p<0.05$. Brackets with asterisks indicate statistically significant differences ($p<0.01$) identified through pairwise t-tests between mock and treated samples.

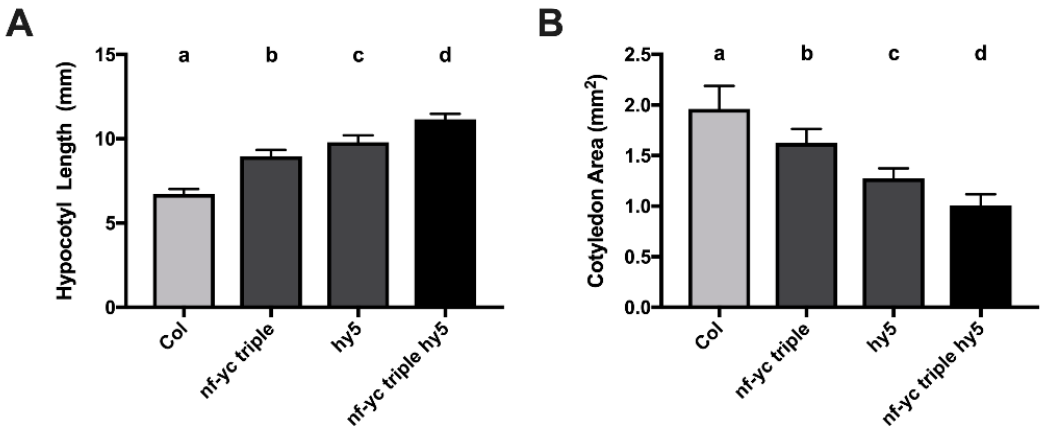

Figure S4.1. Hypocotyl length and cotyledon area averages in red light grown seedlings. Seedlings were grown for 5 days in constant red light (~20µmol/m^2s) before measuring each (**A**) hypocotyl length and (**B**) cotyledon area. Each population contains at least 20 seedlings. Errors bars represent 95% confidence intervals. Letters above each bar represent statistical binning, with statistically significant differences identified through nonparametric two-way ANOVA with Dunn's correction for multiple comparisons, $p<0.05$.

Gene Ontology Biological Process	GO ID	Fold Enrichment	P-value
Carbon fixation	**0015977**	**9.02**	**3.66E-05**
Photosynthesis, dark reaction	**0019685**	**8.59**	**1.37E-02**
Photosynthesis, light harvesting	**0009765**	**7.32**	**8.56E-10**
Photosynthesis, light reaction	**0019684**	**6.06**	**1.58E-21**
Regulation of photosynthesis	**0010109**	**5.87**	**1.58E-05**
Response to far red light	**0010218**	**4.48**	**2.35E-03**
Response to red light	**0010114**	**4.36**	**7.24E-05**
Response to blue light	**0009637**	**4.27**	**7.10E-07**
Secondary metabolite biosynthetic process	**0044550**	**3.67**	**1.56E-10**
Indole-containing compound metabolic process	**0042430**	**3.44**	**1.85E-03**
Response to cytokinin	**0009735**	**2.86**	**8.65E-07**
Response to water	**0009415**	**2.79**	**7.35E-11**

Table 4.1. GO enrichment in significantly down-regulated genes in *nf-yc triple.*

Gene Ontology Biological Process	GO ID	Fold Enrich ment	P-value
Cell wall modification involved in multidimensional cell growth	**0042547**	**10.33**	**3.82E-02**
Cellular response to ethylene stimulus	**0071369**	**6.71**	**3.11E-05**
Response to nitrogen compound	**1901698**	**5.66**	**4.34E-13**
Root morphogenesis	**0010015**	**3.99**	**3.64E-04**
Response to salicylic acid	**0009751**	**3.89**	**1.33E-02**
Developmental growth involved in morphogenesis	**0060560**	**3.42**	**1.40E-03**
Root development	**0048364**	**2.75**	**9.55E-03**
Response to hormone	**0009725**	**2.32**	**1.87E-08**

Table 4.2. GO enrichment in significantly up-regulated genes in *nf-yc triple.*

Chapter 5: Conclusion

Conclusions

The NUCLEAR FACTOR-Y (NF-Y) family of transcription factors (TFs) are important regulators of numerous biological processes across the eukaryotic lineages, from gating cell cycle progression in mammals to coordinating hormone and environmental signals in plants (Salsi et al. 2003; Petroni et al. 2012; Myers and Holt 2018). With such diverse and important roles, it is unsurprising that the NF-Y have been intensely studied for over 30 years (Dorn et al. 1987; Jones et al. 1985). While each of the three NF-Y families are evolutionarily conserved among eukaryotes, plant lineages have undergone extensive expansion (Gusmaroli et al. 2001, 2002). *Arabidopsis thaliana* (Arabidopsis), the model plant my research has focused on, encodes 10 *NF-YA*, 10 *NF-YB*, and 10 *NF-YC* genes, and as an obligate heterotrimeric TF complex, a staggering 1,000 unique NF-Y complexes could potentially form to regulate diverse biological processes (Siefers et al. 2009). This increased regulatory capacity has led to a persistent interest in plant NF-Y function, and as a primary model plant with unmatched genetic resources, Arabidopsis NF-Y research has seen extensive intellectual investment over recent years.

Significant progress has been made toward understanding NF-Y complexes, particularly through identification of biological processes they are thought to mediate. In this category, identification of plant NF-Y regulated processes has surpassed that of other organisms (reviewed in

(Swain et al. 2017)); however, the extensive expansion of plant NF-Y genes is likely the underlying force driving this disparity (i.e., it is logical that a larger TF pool will regulate more biological processes than a smaller TF pool). In Chapter 2, I contributed to our understanding of NF-Y function by identifying and describing redundant roles for three NF-YC genes in photomorphogenesis, or light-mediated development, where they are critical for proper hypocotyl elongation inhibition. I also identified that these NF-YC genes function independent of the bZIP TF *ELONGATED HYPOCOTYL5 (HY5)*, a central regulator and integrator of many plant developmental processes and abiotic responses. Interestingly, in Chapter 4, I described seedling developmental processes where the same NF-YC and HY5 proteins appear to work interdependently, supported through my observation of epistatic relationships of their associated mutants in the regulation of primary root elongation. While I did not describe a specific mechanism of NF-Y mediated regulation, I did demonstrate dark- and proteasome-dependent degradation of NF-YA2. This activity could serve to modulate when NF-Y complexes form, aligning their activity with other light-dependent regulators. A more direct investigation of the mechanics of NF-YA2 degradation, particularly a dissection of which proteasome components are involved, would significantly strengthen our understanding of this regulatory mechanism that is so widely utilized in the coordination of early seedling development.

Beyond regulation of specific biological processes, scientists have also developed a detailed understanding of the mechanism of NF-Y DNA binding, including mapping of individual amino acid and DNA nitrogenous base interactions on the *CCAAT* box (Kim et al. 1996; Sinha et al. 1996; Bi et al. 1997). In this area, our understanding of plant NF-Y DNA binding has primarily been explored through functional validation of mammalian NF-Y research. Because of this perspective, the impacts of NF-Y family expansion in plants have remained largely unaddressed, including whether different NF-Y complex compositions alter binding specificity or affinity. Animal studies have established that only NF-YA makes sequence-specific contact with the *CCAAT* box, while NF-YB and NF-YC interact with the DNA backbone, simultaneously making the DNA minor groove more accessible and providing a binding platform for NF-YA (Nardini et al. 2013; Ronchi et al. 1995). Through these interactions with the DNA backbone, specific NF-YB/NF-YC dimers might fine-tune affinity of the NF-Y complex to specific *CCAAT* boxes. A direct comparison of the effects of the sequences flanking the *CCAAT* box on NF-Y complex DNA binding could directly address this hypothesis but would be technically challenging or overly labor intensive with current approaches. To better facilitate this type of research, as well as other types discussed below, I developed DIMR (**D**ynamic, **I**nterdependent TF binding **M**olecular **R**eporter), an adaptable yeast-based synthetic reporter system capable of probing TF DNA interactions. In Chapter 3, I described the design and implementation of the DIMR system, using known

and novel mutations of the NF-Y to establish and refine its performance. In the next section I will expand on my motivation for implementing DIMR and describe several interesting lines of inquiry for which it is well-suited.

Future Directions and Perspective

Recent advances in plant NF-Y research have solidified a new, unexplored frontier of NF-Y DNA binding in the form of non-canonical NF-Y complexes, where the NF-YA subunit of the complex is replaced with a plant-specific CCT (CONSTANS (CO), CONSTANS-LIKE and TOC1) protein, to form an NF-CCT complex (Gnesutta et al. 2017a). Thus far, this non-canonical complex has only been validated for CO, NF-YB2, and NF-YC3, and this NF-CCT^{CO} complex was shown to exhibit altered DNA binding specificity (Gnesutta et al. 2017a). NF-CCT^{CO} was found to bind the consensus sequence *CCACA* in the promoter of *FLOWERING LOCUS T*, a key integrator of flowering time signals.

While it remains to be established whether NF-CCT complexes form outside of NF-CCT^{CO}, sequence similarity amongst NF-Y and CCT proteins is high at many DNA-binding residues, suggesting a similar mechanism of DNA binding. One residue stands out as a notable exception to this conservation, a histidine in NF-YA2 that is known to make direct physical contact with the fourth base the *CCA**A**T* box (Nardini et al. 2013). Identity at this particular residue has diverged between, but not among, clades and sub-clades of

CCT members. Various groups of CCT proteins, including members of the *PSEUDO RESPONSE REGULATOR (PRR,* (Mizuno 2004)*)* and *CONSTANS-LIKE (COL, (Griffiths et al. 2003))* families, have been shown to bind consensus sites containing the first three bases of the ***CCA****AT* box (Ben-Naim et al. 2006; Gendron et al. 2012). Taking these observations in conjunction, it seems possible that this divergent residue contributes to binding site specificity immediately downstream of *CCA*. Testing hypotheses associated with this observation would require screening putative NF-CCT complexes for interactions with many permutations of binding sites. While already a complex task, Arabidopsis encodes over 40 CCT domain containing proteins, meaning traditional approaches become less tenable and require significant investment.

The DIMR system is well-suited to address this type of question, and is currently being adapted to recapitulate the observed binding of NF-CCT^{CO}. A functional demonstration of NF-CCT^{CO} will lay the groundwork for testing other NF-CCT complexes and describing their binding specificity. Testing representatives of the various CCT sub-clades for binding activity when co-expressed with NF-YB2 and NF-YC3 could identify regulatory functions similar to those observed in NF-CCT^{CO}, and would require testing transcriptional regulatory capacity on a suite of binding site permutations, similar to the approach presented in Chapter 3. While the DIMR system makes this study relatively straightforward, some assumptions have been

made here that may not hold. Particularly, it is unlikely that every NF-CCT complex is mediated through NF-YB2 and NF-YC3, despite their broad native expression patterns and promiscuous NF-Y complex formation (Siefers et al. 2009; Hackenberg et al. 2012). Different NF-YB and/or NF-YC paralogs could be easily substituted into the DIMR system to address this possibility, but testing all possible combinations of NF-YB/NF-YC dimers would require a significant investment.

The value and usefulness of the DIMR system is rooted in the ability to probe TF DNA interactions in a more mechanistic and high-throughput manner than can be readily achieved otherwise. This type of molecular system has been primarily used to address fundamental questions about biology; however, the translational approaches facilitated by a more complete understanding of TF functions are significant. The DIMR system provides an avenue to better understand how TFs function, but it is also an excellent framework for initial validation and refinement of synthetic gene circuits. This type of synthetic biology gene circuit design has already begun to accelerate many translational approaches, such as drug discovery, microbial biofuel production, and secondary metabolite production (see reviewed in (Khalil and Collins 2010)).

In addition to facilitating research into mechanistic TF DNA binding, the DIMR system also presents an excellent space to explore the evolution of

TFs and their cognate binding sequences. Applying evolutionary models (in the form of mutational and substitution rates) to a particular phylogeny can predict and reconstruct ancestral sequences of related TFs (Thornton 2004; Harms and Thornton 2010; Siddiq et al. 2017), and testing these ancestral sequences against extant sequences could shed light on the forces driving their evolution. In particular, this approach would be applied to related TFs that have opposing functions, such as NF-Y regulation of abscisic acid (ABA) mediated seed germination. Subsets of NF-YA and NF-YC hypo- and hyper-morphs were found to have opposite phenotypes when grown on ABA (Kumimoto et al. 2013; Siriwardana et al. 2014). Ancestral reconstruction and subsequent DNA binding analyses could identify the function of the common ancestor of these clades (Pinney et al. 2007), and would provide a fascinating insight into the evolution of these divergent regulatory schemes. More intricate experimental designs utilizing multiple ancestrally-reconstructed NF-Y subunits could provide important insights into the mechanisms of co-evolution, where many different constraints have been proposed to influence the evolution of TF families (Hughes 1994; Force et al. 1999; Baker et al. 2013). For any of these approaches, the DIMR system would provide a tunable, straight-forward readout for TF regulatory capacity with minimal modification.

To conclude, in this dissertation, I have expanded our understanding of plant NF-YC developmental roles during early seedling development and

uncovered complex physical and genetic relationships between NF-YC and HY5. Further, I developed a yeast synthetic transcriptional reporter system with the goal of facilitating higher-throughput, mechanistic studies of NF-Y DNA binding requirements. Beyond my specific research goals, any study of TF DNA interactions could be furthered through application of the DIMR system.

References

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F (2010) cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. The Plant cell 22 (5):1425-1440. doi:tpc.110.074682 [pii]

Alam M, Tanaka T, Nakamura H, Ichikawa H, Kobayashi K, Yaeno T, Yamaoka N, Shimomoto K, Takayama K, Nishina H (2015) Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. Plant biotechnology journal 13 (1):85-96

Aloni R, Aloni E, Langhans M, Ullrich C (2006) Role of cytokinin and auxin in shaping root architecture: regulating vascular differentiation, lateral root initiation, root apical dominance and root gravitropism. Annals of botany 97 (5):883-893

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science (New York, NY) 301 (5633):653-657

Amoutzias GD, Robertson DL, Van de Peer Y, Oliver SG (2008) Choose your partners: dimerization in eukaryotic transcription factors. Trends in biochemical sciences 33 (5):220-229

Anderson DW, McKeown AN, Thornton JW (2015) Intermolecular epistasis shaped the function and evolution of an ancient transcription factor and its DNA binding sites. Elife 4:e07864

Anderson JB, Funt J, Thompson DA, Prabhu S, Socha A, Sirjusingh C, Dettman JR, Parreiras L, Guttman DS, Regev A (2010) Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. Current Biology 20 (15):1383-1388

Andersson DI (2006) The biological cost of mutational antibiotic resistance: any practical conclusions? Current opinion in microbiology 9 (5):461-465

Andersson DI, Levin BR (1999) The biological cost of antibiotic resistance. Current opinion in microbiology 2 (5):489-493

Ang LH, Deng XW (1994) Regulatory hierarchy of photomorphogenic loci: allele-specific and light-dependent interaction between the HY5 and COP1 loci. Plant Cell 6 (5):613-628

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics 25 (1):25-29

Bai S, Saito T, Honda C, Hatsuyama Y, Ito A, Moriguchi T (2014) An apple B-box protein, MdCOL11, is involved in UV-B- and temperature-induced anthocyanin biosynthesis. Planta 240 (5):1051-1062. doi:10.1007/s00425-014-2129-8

Baker CR, Hanson-Smith V, Johnson AD (2013) Following gene duplication, paralog interference constrains transcriptional circuit evolution. Science (New York, NY) 342 (6154):104-108. doi:10.1126/science.1240810

Balasubramanian S, Sureshkumar S, Lempe J, Weigel D (2006) Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLoS Genet 2 (7):e106

Ballesteros ML, Bolle C, Lois LM, Moore JM, Vielle-Calzada JP, Grossniklaus U, Chua NH (2001) LAF1, a MYB transcription activator for phytochrome A signaling. Genes & development 15 (19):2613-2625. doi:10.1101/gad.915001

Ballif J, Endo S, Kotani M, Macadam J, Wu Y (2011) Over-expression of HAP3b enhances primary root elongation in Arabidopsis. Plant Physiol Biochem 49 (6):579-583. doi:S0981-9428(11)00025-8 [pii]

Bastiaens PI, Squire A (1999) Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell. Trends Cell Biol 9 (2):48-52

Baudin M, Laloum T, Lepage A, Rípodas C, Ariel F, Frances L, Crespi M, Gamas P, Blanco FA, Zanetti ME (2015) A phylogenetically conserved group of nuclear factor-Y transcription factors interact to control nodulation in legumes. Plant physiology 169 (4):2761-2773

Baumlein H, Junker A (2012) Multifunctionality of the LEC1 transcription factor during plant development. Plant signaling & behavior 7 (12)

Beekwilder J, van Rossum HM, Koopman F, Sonntag F, Buchhaupt M, Schrader J, Hall RD, Bosch D, Pronk JT, van Maris AJ, Daran JM (2014) Polycistronic expression of a beta-carotene biosynthetic pathway in Saccharomyces cerevisiae coupled to beta-ionone production. Journal of biotechnology 192 Pt B:383-392. doi:10.1016/j.jbiotec.2013.12.016

Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E (2006) The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. Plant J 46 (3):462-476

Benbrook DM, Jones NC (1994) Different binding specificities and transactivation of variant CRE's by CREB complexes. Nucleic Acids Research 22 (8):1463-1469. doi:10.1093/nar/22.8.1463

Bi C, Ma Y, Wang XF, Zhang DP (2017) Overexpression of the transcription factor NF-YC9 confers abscisic acid hypersensitivity in Arabidopsis. Plant molecular biology 95 (4-5):425-439. doi:10.1007/s11103-017-0661-1

Bi W, Wu L, Coustry F, de Crombrugghe B, Maity SN (1997) DNA binding specificity of the CCAAT-binding factor CBF/NF-Y. The Journal of biological chemistry 272 (42):26562-26572

Bishopp A, Benková E, Helariutta Y (2011) Sending mixed messages: auxin-cytokinin crosstalk in roots. Current opinion in plant biology 14 (1):10-16

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005a) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature 433:39. doi:10.1038/nature03184

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005b) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature 433 (7021):39

Blount BA, Weenink T, Ellis T (2012) Construction of synthetic regulatory networks in yeast. FEBS letters 586 (15):2112-2121

Bourgarel D, Nguyen CC, Bolotin-Fukuhara M (1999) HAP4, the glucose-repressed regulated subunit of the HAP transcriptional complex involved in the fermentation–respiration shift, has a functional homologue in the respiratory yeast Kluyveromyces lactis. Molecular microbiology 31 (4):1205-1215

Boyer JS (1982) Plant Productivity and Environment. Science (New York, NY) 218 (4571):443. doi:10.1126/science.218.4571.443

Brambilla V, Fornara F (2016) Y flowering? Regulation and activity of CONSTANS and CCT-domain proteins in Arabidopsis and crop species. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms

Broach JR, Guarascio VR, Jayaram M (1982) Recombination within the yeast plasmid 2μ circle is site-specific. Cell 29 (1):227-234

Calvenzani V, Testoni B, Gusmaroli G, Lorenzo M, Gnesutta N, Petroni K, Mantovani R, Tonelli C (2012) Interactions and CCAAT-Binding of Arabidopsis thaliana NF-Y Subunits. PLOS ONE 7 (8):e42902. doi:10.1371/journal.pone.0042902

Cao S, Kumimoto RW, Gnesutta N, Calogero AM, Mantovani R, Holt BF, III (2014) A Distal CCAAT/NUCLEAR FACTOR Y Complex Promotes Chromatin Looping at the FLOWERING LOCUS T Promoter and Regulates the Timing of Flowering in Arabidopsis. Plant Cell 26:1009-1017. doi:10.1105/tpc.113.120352

Cao S, Kumimoto RW, Siriwardana CL, Risinger JR, Holt BF, III (2011a) Identification and Characterization of NF-Y Transcription Factor Families in the Monocot Model Plant *Brachypodium distachyon*. PLoS ONE 6 (6):e21805

Cao S, Siriwardana CL, Kumimoto RW, Holt BF, III (2011b) Construction of high quality Gateway™ entry libraries and their application to yeast two-hybrid for the monocot model plant Brachypodium distachyon. BMC Biotechnol 11 (1):53. doi:1472-6750-11-53 [pii]

Catala R, Medina J, Salinas J (2011) Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 108 (39):16475-16480. doi:10.1073/pnas.1107161108

Ceribelli M, Dolfini D, Merico D, Gatta R, Vigano AM, Pavesi G, Mantovani R (2008) The histone-like NF-Y is a bifunctional transcription factor. Mol Cell Biol 28 (6):2047-2058

Chattopadhyay S, Ang LH, Puente P, Deng XW, Wei N (1998) Arabidopsis bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. The Plant cell 10 (5):673-683

Chen H, Xiong L (2008) Role of HY5 in abscisic acid response in seeds and seedlings. Plant signaling & behavior 3 (11):986-988

Chen H, Xiong L (2011) Genetic interaction of two abscisic acid signaling regulators, HY5 and FIERY1, in mediating lateral root formation. Plant signaling & behavior 6 (1):123-125. doi:14231 [pii]

Chen X, Yao Q, Gao X, Jiang C, Harberd Nicholas P, Fu X (2016) Shoot-to-Root Mobile Transcription Factor HY5 Coordinates Plant Carbon and Nitrogen Acquisition. Current Biology 26 (5):640-646. doi:https://doi.org/10.1016/j.cub.2015.12.066

Chory J (2010) Light signal transduction: an infinite spectrum of possibilities. Plant J 61 (6):982-991. doi:10.1111/j.1365-313X.2009.04105.x

Cluis CP, Mouchel CF, Hardtke CS (2004) The Arabidopsis transcription factor HY5 integrates light and hormone signaling pathways. Plant J 38 (2):332-347. doi:10.1111/j.1365-313X.2004.02052.x

Combier JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie T, Ott T, Gamas P, Crespi M, Niebel A (2006) MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. Genes & development 20 (22):3084-3088

Crick F (1970) Central Dogma of Molecular Biology. Nature 227 (5258):561-563. doi:10.1038/227561a0

Cunningham BC, Wells JA (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science (New York, NY) 244 (4908):1081-1085

Curran KA, Karim AS, Gupta A, Alper HS (2013) Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. Metabolic engineering 19:88-97. doi:10.1016/j.ymben.2013.07.001

Datta S, Hettiarachchi C, Johansson H, Holm M (2007) SALT TOLERANCE HOMOLOG2, a B-box protein in Arabidopsis that activates transcription and positively regulates light-mediated development. The Plant cell 19 (10):3242-3255. doi:10.1105/tpc.107.054791

Datta S, Hettiarachchi GH, Deng XW, Holm M (2006) Arabidopsis CONSTANS-LIKE3 is a positive regulator of red light signaling and root growth. The Plant cell 18 (1):70-84. doi:tpc.105.038182 [pii]

Datta S, Johansson H, Hettiarachchi C, Irigoyen ML, Desai M, Rubio V, Holm M (2008) LZF1/SALT TOLERANCE HOMOLOG3, an Arabidopsis B-box protein involved in light-dependent development and gene expression, undergoes COP1-mediated ubiquitination. Plant Cell 20 (9):2324-2338. doi:10.1105/tpc.108.061747

de Visser JAG, Elena SF (2007) The evolution of sex: empirical insights into the roles of epistasis and drift. Nature Reviews Genetics 8 (2):139

Diamond MI, Miner JN, Yoshinaga SK, Yamamoto KR (1990) Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science (New York, NY) 249 (4974):1266-1272

Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic acids research 11 (5):1475-1489

Donati G, Gatta R, Dolfini D, Fossati A, Ceribelli M, Mantovani R (2008) An NF-Y-dependent switch of positive and negative histone methyl marks on CCAAT promoters. PLoS ONE 3 (4):e2066

Dorn A, Bollekens J, Staub A, Benoist C, Mathis D (1987) A multiplicity of CCAAT box-binding proteins. Cell 50 (6):863-872

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32 (5):1792-1797. doi:10.1093/nar/gkh340

Edmondson DG, Smith MM, Roth SY (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. Genes & development 10 (10):1247-1259

Encode Project Consortium T, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Stamatoyannopoulos JA, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SCJ, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Dutta A, Guigó R, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Flicek P, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermüller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung W-K, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Dermitzakis ET, Margulies EH, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei C-L, Ruan Y, Snyder M, Birney E, Struhl K, Gerstein M, Antonarakis SE, Gingeras TR, Brown JB, Flicek P, Fu Y, Keefe D, Birney E, Denoeud F, Gerstein M, Green ED, Kapranov P, Karaöz U, Myers RM, Noble WS, Reymond A, Rozowsky J, Struhl K, Siepel A, Stamatoyannopoulos JA, Taylor CM, Taylor J, Thurman RE, Tullius TD, Washietl S, Zheng D, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Collins FS, Margulies EH, Cooper GM, Asimenos G, Thomas DJ, Dewey CN, Siepel A, Birney E, Keefe D, Hou M, Taylor J, Nikolaev S, Montoya-Burgos JI, Löytynoja A, Whelan S, Pardi F, Massingham T, Brown JB, Huang H, Zhang NR, Bickel P, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Gerstein M, Antonarakis SE,

Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Pachter L, Green ED, Sidow A, Weng Z, Trinklein ND, Fu Y, Zhang ZD, Karaöz U, Barrera L, Stuart R, Zheng D, Ghosh S, Flicek P, King DC, Taylor J, Ameur A, Enroth S, Bieda MC, Koch CM, Hirsch HA, Wei C-L, Cheng J, Kim J, Bhinge AA, Giresi PG, Jiang N, Liu J, Yao F, Sung W-K, Chiu KP, Vega VB, Lee CWH, Ng P, Shahab A, Sekinger EA, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Clelland GK, Wilcox S, Dillon SC, Andrews RM, Fowler JC, Couttet P, James KD, Lefebvre GC, Bruce AW, Dovey OM, Ellis PD, Dhami P, Langford CF, Carter NP, Vetrie D, Kapranov P, Nix DA, Bell I, Patel S, Rozowsky J, Euskirchen G, Hartman S, Lian J, Wu J, Urban AE, Kraus P, Van Calcar S, Heintzman N, Hoon Kim T, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JNS, Yu Y, Birney* E, Weissman S, Ruan Y, Lieb JD, Iyer VR, Green RD, Gingeras TR, Wadelius C, Dunham I, Struhl K, Hardison RC, Gerstein M, Farnham PJ, Myers RM, Ren B, Snyder M, Thomas DJ, Rosenbloom K, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Haussler D, Kent WJ, Dermitzakis ET, Armengol L, Bird CP, Clark TG, Cooper GM, de Bakker PIW, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Thomas DJ, Woodroffe A, Batzoglou S, Davydov E, Dimas A, Eyras E, Hallgrímsdóttir IB, Hardison RC, Huppert J, Sidow A, Taylor J, Trumbower H, Zody MC, Guigó R, Mullikin JC, Abecasis GR, Estivill X, Birney E, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VVB, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799. doi:10.1038/nature05874

Fairchild CD, Schumaker MA, Quail PH (2000) HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction. Genes & development 14 (18):2377-2391

Fankhauser C, Chory J (1997) Light control of plant development. Annu Rev Cell Dev Biol 13:203-229. doi:10.1146/annurev.cellbio.13.1.203

Fankhauser C, Chory J (2000) RSF1, an Arabidopsis locus implicated in phytochrome A signaling. Plant Physiol 124 (1):39-45

Farré EM, Liu T (2013) The PRR family of transcriptional regulators reflects the complexity and evolution of plant circadian clocks. Current opinion in plant biology 16 (5):621-629

Flagfeldt DB, Siewers V, Huang L, Nielsen J (2009) Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. Yeast (Chichester, England) 26 (10):545-551. doi:10.1002/yea.1705

Fleming JD, Pavesi G, Benatti P, Imbriano C, Mantovani R, Struhl K (2013) NF-Y coassociates with FOS at promoters, enhancers, repetitive elements, and inactive

chromatin regions, and is stereo-positioned with growth-controlling transcription factors. Genome Res 23 (8):1195-1209. doi:10.1101/gr.148080.112

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151 (4):1531-1545

Frontini M, Imbriano C, diSilvio A, Bell B, Bogni A, Romier C, Moras D, Tora L, Davidson I, Mantovani R (2002) NF-Y recruitment of TFIID, multiple interactions with histone fold TAF(II)s. The Journal of biological chemistry 277 (8):5841-5848

Frontini M, Imbriano C, Manni I, Mantovani R (2004) Cell cycle regulation of NF-YC nuclear localization. Cell cycle (Georgetown, Tex 3 (2):217-222

Fu X, Harberd NP (2003) Auxin promotes Arabidopsis root growth by modulating gibberellin response. Nature 421 (6924):740-743. doi:10.1038/nature01387

Gälweiler L, Guan C, Müller A, Wisman E, Mendgen K, Yephremov A, Palme K (1998) Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular tissue. Science (New York, NY) 282 (5397):2226-2230

Gangappa SN, Botto JF (2014) The BBX family of plant transcription factors. Trends in plant science 19 (7):460-470. doi:10.1016/j.tplants.2014.01.010

Gangappa SN, Crocco CD, Johansson H, Datta S, Hettiarachchi C, Holm M, Botto JF (2013) The Arabidopsis B-BOX protein BBX25 interacts with HY5, negatively regulating BBX22 expression to suppress seedling photomorphogenesis. The Plant cell 25 (4):1243-1257. doi:10.1105/tpc.113.109751

Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA (2012) Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. Proceedings of the National Academy of Sciences 109 (8):3167-3172

Gene Ontology C (2015) Gene Ontology Consortium: going forward. Nucleic Acids Res 43 (Database issue):D1049-1056. doi:10.1093/nar/gku1179

Genoud T, Schweizer F, Tscheuschler A, Debrieux D, Casal JJ, Schafer E, Hiltbrunner A, Fankhauser C (2008) FHY1 mediates nuclear import of the light-activated phytochrome A photoreceptor. PLoS Genet 4 (8):e1000143. doi:10.1371/journal.pgen.1000143

Gnesutta N, Kumimoto RW, Swain S, Chiara M, Siriwardana C, Horner DS, Holt BF, Mantovani R (2017a) CONSTANS Imparts DNA Sequence Specificity to the Histone Fold NF-YB/NF-YC Dimer. The Plant cell 29 (6):1516-1532. doi:10.1105/tpc.16.00864

Gnesutta N, Mantovani R, Fornara F (2018) Plant Flowering: Imposing DNA Specificity on Histone-Fold Subunits. Trends in plant science. doi:10.1016/j.tplants.2017.12.005

Gnesutta N, Saad D, Chaves-Sanjuan A, Mantovani R, Nardini M (2017b) Crystal Structure of the Arabidopsis thaliana L1L/NF-YC3 Histone-fold Dimer Reveals Specificities of the LEC1 Family of NF-Y Subunits in Plants. Molecular plant 10 (4):645-648. doi:10.1016/j.molp.2016.11.006

Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis. Plant physiology 131 (4):1855-1867

Gusmaroli G, Tonelli C, Mantovani R (2001) Regulation of the CCAAT-Binding NF-Y subunits in Arabidopsis thaliana. Gene 264 (2):173-185.

Gusmaroli G, Tonelli C, Mantovani R (2002) Regulation of novel members of the Arabidopsis thaliana CCAAT-binding nuclear factor Y subunits. Gene 283 (1-2):41-48

Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends in biotechnology 22 (7):346-353. doi:10.1016/j.tibtech.2004.04.006

Hackenberg D, Wu Y, Voigt A, Adams R, Schramm P, Grimm B (2012) Studies on differential nuclear translocation mechanism and assembly of the three subunits of the Arabidopsis thaliana transcription factor NF-Y. Molecular plant 5 (4):876-888. doi:10.1093/mp/ssr107

Hai T, Curran T (1991) Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. Proceedings of the National Academy of Sciences 88 (9):3720. doi:10.1073/pnas.88.9.3720

Hamilton C (1997) A binary-BAC system for plant transformation with high molecular weight DNA. Genes & development 200 (1-2):107-116

Hanemian M, Barlet X, Sorin C, Yadeta KA, Keller H, Favery B, Simon R, Thomma BP, Hartmann C, Crespi M (2016) Arabidopsis CLAVATA1 and CLAVATA2 receptors contribute to Ralstonia solanacearum pathogenicity through a miR169-dependent pathway. New Phytologist 211 (2):502-515

Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne J-B, Reynolds DB, Yoo J (2004) Transcriptional regulatory code of a eukaryotic genome. Nature 431 (7004):99

Harms MJ, Thornton JW (2010) Analyzing protein structure and function using ancestral gene reconstruction. Current opinion in structural biology 20 (3):360-366

Hilioti Z, Ganopoulos I, Bossis I, Tsaftaris A (2014) LEC1-LIKE paralog transcription factor: how to survive extinction and fit in NF-Y protein complex. Gene 543 (2):220-233

Hiltbrunner A, Tscheuschler A, Viczian A, Kunkel T, Kircher S, Schafer E (2006) FHY1 and FHL act together to mediate nuclear accumulation of the phytochrome A photoreceptor. Plant Cell Physiol 47 (8):1023-1034. doi:10.1093/pcp/pcj087

Holm M, Ma L-G, Qu L-J, Deng X-W (2002) Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. Genes & development 16 (10):1247-1259. doi:10.1101/gad.969702

Hossain MS, Shrestha A, Zhong S, Miri M, Austin RS, Sato S, Ross L, Huebert T, Tromas A, Torres-Jerez I (2016) Lotus japonicus NF-YA1 Plays an Essential Role During Nodule Differentiation and Targets Members of the SHI/STY Gene Family. Molecular Plant-Microbe Interactions 29 (12):950-964

Hou X, Zhou J, Liu C, Liu L, Shen L, Yu H (2014) Nuclear factor Y-mediated H3K27me3 demethylation of the SOC1 locus orchestrates flowering responses of Arabidopsis. Nature communications 5:ncomms5601

Huang M, Hu Y, Liu X, Li Y, Hou X (2015) Arabidopsis LEAFY COTYLEDON1 Mediates Postembryonic Development via Interacting with PHYTOCHROME-INTERACTING FACTOR4. Plant Cell 27 (11):3099-3111. doi:10.1105/tpc.15.00750

Huang X, Ouyang X, Yang P, Lau OS, Li G, Li J, Chen H, Deng XW (2012) Arabidopsis FHY3 and HY5 positively mediate induction of COP1 transcription in response to

photomorphogenic UV-B light. The Plant cell 24 (11):4590-4606. doi:10.1105/tpc.112.103994

Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proceedings of the Royal Society of London Series B: Biological Sciences 256 (1346):119-124

Hwang Y-H, Kim S-K, Lee KC, Chung YS, Lee JH, Kim J-K (2016) Functional conservation of rice OsNF-YB/YC and Arabidopsis AtNF-YB/YC proteins in the regulation of flowering time. Plant cell reports 35 (4):857-865

Jang I-C, Yang J-Y, Seo HS, Chua N-H (2005) HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling. Genes & development 19 (5):593-602. doi:10.1101/gad.1247205

Jang IC, Henriques R, Chua NH (2013) Three transcription factors, HFR1, LAF1 and HY5, regulate largely independent signaling pathways downstream of phytochrome A. Plant Cell Physiol 54 (6):907-916. doi:10.1093/pcp/pct042

Jang IC, Yang SW, Yang JY, Chua NH (2007) Independent and interdependent functions of LAF1 and HFR1 in phytochrome A signaling. Genes & development 21 (16):2100-2111. doi:10.1101/gad.1568207

Jang S, Marchal V, Panigrahi KC, Wenkel S, Soppe W, Deng XW, Valverde F, Coupland G (2008) Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. The EMBO journal 27 (8):1277-1288

Jayaram N, Usvyat D, Martin AC (2016) Evaluating tools for transcription factor binding site prediction. BMC bioinformatics:1

Jensen PJ, Hangarter RP, Estelle M (1998a) Auxin transport is required for hypocotyl elongation in light-grown but not dark-grown Arabidopsis. Plant Physiology 116 (2):455-462

Jensen PJ, Hangarter RP, Estelle M (1998b) Auxin Transport Is Required for Hypocotyl Elongation in Light-Grown but Not Dark-Grown Arabidopsis. Plant Physiology 116 (2):455. doi:10.1104/pp.116.2.455

John S, Sabo PJ, Thurman RE, Sung M-H, Biddie SC, Johnson TA, Hager GL, Stamatoyannopoulos JA (2011) Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nature genetics 43 (3):264

Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E, Taipale J (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature 527 (7578):384

Jones KA, Yamamoto KR, Tjian R (1985) Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell 42 (2):559-572

Junker A, Monke G, Rutten T, Keilwagen J, Seifert M, Thi TM, Renou JP, Balzergue S, Viehover P, Hahnel U, Ludwig-Muller J, Altschmied L, Conrad U, Weisshaar B, Baumlein H (2012) Elongation-related functions of LEAFY COTYLEDON1 during the development of Arabidopsis thaliana. Plant J 71 (3):427-442. doi:10.1111/j.1365-313X.2012.04999.x

Kahle J, Baake M, Doenecke D, Albig W (2005) Subunits of the heterotrimeric transcription factor NF-Y are imported into the nucleus by distinct pathways involving importin beta and importin 13. Mol Cell Biol 25 (13):5339-5354

Khalil AS, Collins JJ (2010) Synthetic biology: applications come of age. Nature Reviews Genetics 11 (5):367

Khalil AS, Lu TK, Bashor CJ, Ramirez CL, Pyenson NC, Joung JK, Collins JJ (2012) A synthetic biology framework for programming eukaryotic transcription functions. Cell 150 (3):647-658

Khanna R, Kronmiller B, Maszle DR, Coupland G, Holm M, Mizuno T, Wu SH (2009) The Arabidopsis B-box zinc finger family. The Plant cell 21 (11):3416-3420. doi:tpc.109.069088 [pii]

Kim IS, Sinha S, de Crombrugghe B, Maity SN (1996) Determination of functional domains in the C subunit of the CCAAT- binding factor (CBF) necessary for formation of a CBF-DNA complex: CBF- B interacts simultaneously with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule. Mol Cell Biol 16 (8):4003-4013.

Kim JH, Lee S-R, Li L-H, Park H-J, Park J-H, Lee KY, Kim M-K, Shin BA, Choi S-Y (2011) High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PloS one 6 (4):e18556

Kim S-K, Park H-Y, Jang YH, Lee KC, Chung YS, Lee JH, Kim J-K (2016) OsNF-YC2 and OsNF-YC4 proteins inhibit flowering under long-day conditions in rice. Planta 243 (3):563-576

Kim YM, Woo JC, Song PS, Soh MS (2002) HFR1, a phytochrome A-signalling component, acts in a separate pathway from HY5, downstream of COP1 in Arabidopsis thaliana. Plant J 30 (6):711-719

Kircher S, Schopfer P (2012) Photosynthetic sucrose acts as cotyledon-derived long-distance signal to control root growth during early seedling development in Arabidopsis. Proceedings of the National Academy of Sciences 109 (28):11217. doi:10.1073/pnas.1203746109

Kliebenstein DJ, Lim JE, Landry LG, Last RL (2002) Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. Plant Physiol 130 (1):234-243. doi:10.1104/pp.005041

Knop M, Siegers K, Pereira G, Zachariae W, Winsor B, Nasmyth K, Schiebel E (1999) Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast (Chichester, England) 15 (10B):963-972

Kobayashi K, Obayashi T, Masuda T (2012) Role of the G-box element in regulation of chlorophyll biosynthesis in Arabidopsis roots. Plant signaling & behavior 7 (8):922-926

Kotula L, Curtis PJ (1991) Evaluation of foreign gene codon optimization in yeast: expression of a mouse IG kappa chain. Bio/technology (Nature Publishing Company) 9 (12):1386-1389

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular biology and evolution 33 (7):1870-1874. doi:10.1093/molbev/msw054

Kumimoto RW, Adam L, Hymus GJ, Repetti PP, Reuber TL, Marion CM, Hempel FD, Ratcliffe OJ (2008) The Nuclear Factor Y subunits NF-YB2 and NF-YB3 play additive roles in the promotion of flowering by inductive long-day photoperiods in Arabidopsis. Planta 228 (5):709-723. doi:10.1007/s00425-008-0773-6

Kumimoto RW, Siriwardana CL, Gayler KK, Risinger JR, Siefers N, Holt III BF (2013) NUCLEAR FACTOR Y Transcription Factors Have Both Opposing and Additive Roles in ABA-Mediated Seed Germination. PLoS One 8 (3):e59481. doi:10.1371/journal.pone.0059481

Kumimoto RW, Zhang Y, Siefers N, Holt BF (2010) NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in Arabidopsis thaliana. Plant J 63 (3):379-391. doi:10.1111/j.1365-313X.2010.04247.x

Kusnetsov V, Landsberger M, Meurer J, Oelmuller R (1999) The assembly of the CAAT-box binding complex at a photosynthesis gene promoter is regulated by light, cytokinin, and the stage of the plastids. J Biol Chem 274 (50):36009-36014.

Kvitek DJ, Sherlock G (2011) Reciprocal Sign Epistasis between Frequently Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. PLOS Genetics 7 (4):e1002056. doi:10.1371/journal.pgen.1002056

Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ (2003) LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. The Plant cell 15 (1):5-18

Laloum T, Baudin M, Frances L, Lepage A, Billault-Penneteau B, Cerri MR, Ariel F, Jardinaud MF, Gamas P, Carvalho-Niebel F (2014) Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. The Plant Journal 79 (5):757-768

Laloum T, De Mita S, Gamas P, Baudin M, Niebel A (2013) CCAAT-box binding transcription factors in plants: Y so many? Trends in plant science 18 (3):157-166. doi:10.1016/j.tplants.2012.07.004

Laubinger S, Marchal V, Le Gourrierec J, Wenkel S, Adrian J, Jang S, Kulajta C, Braun H, Coupland G, Hoecker U (2006) Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development (Cambridge, England) 133 (16):3213-3222

Lee D-K, Kim HI, Jang G, Chung PJ, Jeong JS, Kim YS, Bang SW, Jung H, Do Choi Y, Kim J-K (2015) The NF-YA transcription factor OsNF-YA7 confers drought stress tolerance of rice in an abscisic acid independent manner. Plant Science 241:199-210

Lee J, He K, Stolc V, Lee H, Figueroa P, Gao Y, Tongprasit W, Zhao H, Lee I, Deng XW (2007) Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. The Plant cell 19 (3):731-749. doi:10.1105/tpc.106.047688

Levine M, Manley JL (1989) Transcriptional repression of eukaryotic promoters. Cell 59 (3):405-408

Leyva-Gonzalez MA, Ibarra-Laclette E, Cruz-Ramirez A, Herrera-Estrella L (2012) Functional and transcriptome analysis reveals an acclimatization strategy for abiotic stress tolerance mediated by Arabidopsis NF-YA family members. PLoS One 7 (10):e48138. doi:10.1371/journal.pone.0048138

Li C, Distelfeld A, Comis A, Dubcovsky J (2011) Wheat flowering repressor VRN2 and promoter CO2 compete for interactions with NUCLEAR FACTOR-Y complexes. Plant J. doi:10.1111/j.1365-313X.2011.04630.x

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14):1754-1760. doi:10.1093/bioinformatics/btp324

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26 (5):589-595. doi:10.1093/bioinformatics/btp698

Li J, Li G, Gao S, Martinez C, He G, Zhou Z, Huang X, Lee J-H, Zhang H, Shen Y, Wang H, Deng XW (2010) Arabidopsis transcription factor ELONGATED HYPOCOTYL5 plays a role in the feedback regulation of phytochrome A signaling. The Plant cell 22 (11):3634-3649. doi:10.1105/tpc.110.075788

Li L, Zheng W, Zhu Y, Ye H, Tang B, Arendsee ZW, Jones D, Li R, Ortiz D, Zhao X, Du C, Nettleton D, Scott MP, Salas-Fernandez MG, Yin Y, Wurtele ES (2015) QQS orphan gene regulates carbon and nitrogen partitioning across species via NF-YC interactions. Proceedings of the National Academy of Sciences of the United States of America 112 (47):14734-14739. doi:10.1073/pnas.1514670112

Li Q-F, He J-X (2015) BZR1 interacts with HY5 to mediate brassinosteroid- and light-regulated cotyledon opening in Arabidopsis in darkness. Molecular plant. doi:10.1016/j.molp.2015.08.014

Li W, Mallano AI, Bo L, Wang T, Nisa Z, Li Y (2016) Soybean transcription factor GmNFYB1 confers abiotic stress tolerance to transgenic Arabidopsis plants. Canadian Journal of Plant Science 97 (3):501-515

Lickwar CR, Mueller F, Hanlon SE, McNally JG, Lieb JD (2012) Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature 484 (7393):251

Lin R, Ding L, Casola C, Ripoll DR, Feschotte C, Wang H (2007) Transposase-derived transcription factors regulate light signaling in Arabidopsis. Science (New York, NY) 318 (5854):1302-1305. doi:10.1126/science.1146281

Liu J-X, Howell SH (2010) bZIP28 and NF-Y transcription factors are activated by ER stress and assemble into a transcriptional complex to regulate stress response genes in Arabidopsis. The Plant cell 22 (3):782-796. doi:10.1105/tpc.109.072173

Liu L, Adrian J, Pankin A, Hu J, Dong X, von Korff M, Turck F (2014) Induced and natural variation of promoter length modulates the photoperiodic response of FLOWERING LOCUS T. Nat Commun 5:4558. doi:10.1038/ncomms5558

Liu LJ, Zhang YC, Li QH, Sang Y, Mao J, Lian HL, Wang L, Yang HQ (2008) COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. The Plant cell 20 (2):292-306. doi:10.1105/tpc.107.057281

Liu X, Hu P, Huang M, Tang Y, Li Y, Li L, Hou X (2016) The NF-YC-RGL2 module integrates GA and ABA signalling to regulate seed germination in Arabidopsis. Nat Commun 7:12768. doi:10.1038/ncomms12768

López-Bucio J, Cruz-Ramírez A, Herrera-Estrella L (2003) The role of nutrient availability in regulating root architecture. Current Opinion in Plant Biology 6 (3):280-287. doi:https://doi.org/10.1016/S1369-5266(03)00035-9

Luan M, Xu M, Lu Y, Zhang L, Fan Y, Wang L (2015) Expression of zma-miR169 miRNAs and their target ZmNF-YA genes in response to abiotic stress in maize leaves. Gene 555 (2):178-185. doi:10.1016/j.gene.2014.11.001

Luke G, Escuin H, De Felipe P, Ryan M (2010) 2A to the fore - research, technology and applications. Biotechnology & genetic engineering reviews 26:223-260

Luscombe NM, Laskowski RA, Thornton JM (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. Nucleic acids research 29 (13):2860-2874

Ma L, Li J, Qu L, Hager J, Chen Z, Zhao H, Deng XW (2001) Light control of Arabidopsis development entails coordinated regulation of genome expression and cellular pathways. The Plant cell 13 (12):2589-2607

Ma X, Zhu X, Li C, Song Y, Zhang W, Xia G, Wang M (2015) Overexpression of wheat NF-YA10 gene regulates the salinity stress response in Arabidopsis thaliana. Plant Physiology and Biochemistry 86:34-43

Mahony S, Pugh BF (2015) Protein–DNA binding in high-resolution. Critical reviews in biochemistry and molecular biology 50 (4):269-283

Maity SN, de Crombrugghe B (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription. Trends Biochem Sci 23 (5):174-178

Mangano S, Gonzalez CD, Petruccelli S (2014) Agrobacterium tumefaciens-mediated transient transformation of Arabidopsis thaliana leaves. Methods Mol Biol 1062:165-173. doi:10.1007/978-1-62703-580-4_8

Manimaran P, Reddy SV, Moin M, Reddy MR, Yugandhar P, Mohanraj S, Balachandran S, Kirti P (2017) Activation-tagging in indica rice identifies a novel transcription factor subunit, NF-YC13 associated with salt tolerance. Scientific reports 7:9341

Mantovani R (1999) The molecular biology of the CCAAT-binding factor NF-Y. Gene 239 (1):15-27

Mantovani R, Li XY, Pessara U, Hooft van Huisjduijnen R, Benoist C, Mathis D (1994) Dominant negative analogs of NF-YA. The Journal of biological chemistry 269 (32):20340-20346

Markwardt ML, Kremers G-J, Kraft CA, Ray K, Cranfill PJC, Wilson KA, Day RN, Wachter RM, Davidson MW, Rizzo MA (2011) An improved cerulean fluorescent protein with enhanced brightness and reduced reversible photoswitching. PLoS One 6 (3):e17896. doi:10.1371/journal.pone.0017896

McIsaac RS, Oakes BL, Wang X, Dummit KA, Botstein D, Noyes MB (2013) Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast. Nucleic Acids Res 41 (4):e57. doi:10.1093/nar/gks1313

McNellis TW, von Arnim AG, Araki T, Komeda Y, Misera S, Deng XW (1994) Genetic and molecular analysis of an allelic series of cop1 mutants suggests functional roles for the multiple protein domains. The Plant cell 6 (4):487-500. doi:10.1105/tpc.6.4.487

Meinke DW (1992) A Homoeotic Mutant of Arabidopsis thaliana with Leafy Cotyledons. Science (New York, NY) 258 (5088):1647-1650. doi:258/5088/1647 [pii]

Meinke DW, Franzmann LH, Nickle TC, Yeung EC (1994) Leafy Cotyledon Mutants of Arabidopsis. The Plant cell 6 (8):1049-1064. doi:10.1105/tpc.6.8.1049

Mitchell PJ, Tjian R (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science (New York, NY) 245 (4916):371-378

Mitsuda N, Ohme-Takagi M (2009) Functional analysis of transcription factors in Arabidopsis. Plant & cell physiology 50 (7):1232-1248. doi:10.1093/pcp/pcp075

Mizuno T (2004) Plant response regulators implicated in signal transduction and circadian rhythm. Current opinion in plant biology 7 (5):499-505

Morgunova E, Taipale J (2017) Structural perspective of cooperative transcription factor binding. Current Opinion in Structural Biology 47:1-8

Mu J, Tan H, Hong S, Liang Y, Zuo J (2012) Arabidopsis Transcription Factor Genes NF-YA1, 5, 6, and 9 Play Redundant Roles in Male Gametogenesis, Embryogenesis, and Seed Development. Molecular plant. doi:10.1093/mp/sss061

Müller SM, Galliardt H, Schneider J, Barisas BG, Seidel T (2013) Quantification of Förster resonance energy transfer by monitoring sensitized emission in living plant cells. Front Plant Sci 4:413. doi:10.3389/fpls.2013.00413

Myers ZA, Holt BF (2018) NUCLEAR FACTOR-Y: still complex after all these years? Current Opinion in Plant Biology 45:96-102. doi:https://doi.org/10.1016/j.pbi.2018.05.015

Myers ZA, Kumimoto RW, Siriwardana CL, Gayler KK, Risinger JR, Pezzetta D, Holt Iii BF (2016) NUCLEAR FACTOR Y, Subunit C (NF-YC) Transcription Factors Are Positive Regulators of Photomorphogenesis in Arabidopsis thaliana. PLoS Genet 12 (9):e1006333. doi:10.1371/journal.pgen.1006333

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol 20 (1):87-90. doi:10.1038/nbt0102-87

Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vonrhein C, Moras D, Romier C, Bolognesi M, Mantovani R (2013) Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell 152 (1-2):132-143. doi:10.1016/j.cell.2012.11.047

Neff MM, Fankhauser C, Chory J (2000) Light: an indicator of time and place. Genes & development 14 (3):257-271

Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G (2006) Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. Applied and environmental microbiology 72 (8):5266-5273. doi:10.1128/aem.00530-06

Nghe P, Kogenaru M, Tans SJ (2018) Sign epistasis caused by hierarchy within signalling cascades. Nature Communications 9 (1):1451. doi:10.1038/s41467-018-03644-8

Oldfield AJ, Yang P, Conway AE, Cinghu S, Freudenberg JM, Yellaboina S, Jothi R (2014) Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. Mol Cell 55 (5):708-722. doi:10.1016/j.molcel.2014.07.005

Osterlund MT, Wei N, Deng XW (2000) The roles of photoreceptor systems and the COP1-targeted destabilization of HY5 in light control of Arabidopsis seedling development. Plant Physiol 124 (4):1520-1524

Overvoorde P, Fukaki H, Beeckman T (2010) Auxin control of root development. Cold Spring Harbor perspectives in biology 2 (6):a001537-a001537. doi:10.1101/cshperspect.a001537

Oyama T, Shimura Y, Okada K (1997) The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. Genes & development 11 (22):2983-2995

Palstra R-JTS (2009) Close encounters of the 3C kind: long-range chromatin interactions and transcriptional regulation. Briefings in Functional Genomics and Proteomics 8 (4):297-309

Pelletier JM, Kwong RW, Park S, Le BH, Baden R, Cagliari A, Hashimoto M, Munoz MD, Fischer RL, Goldberg RB, Harada JJ (2017) LEC1 sequentially regulates the transcription of genes involved in diverse developmental processes during seed development. Proceedings of the National Academy of Sciences of the United States of America 114 (32):E6710-E6719. doi:10.1073/pnas.1707957114

Petroni K, Kumimoto RW, Gnesutta N, Calvenzani V, Fornari M, Tonelli C, Holt BF, 3rd, Mantovani R (2012) The promiscuous life of plant NUCLEAR FACTOR Y transcription factors. The Plant Cell 24 (12):4777-4792. doi:10.1105/tpc.112.105734

Phillips PC (2008) Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nature Reviews Genetics 9 (11):855

Pikaard CS, Scheid OM (2014) Epigenetic regulation in plants. Cold Spring Harbor perspectives in biology 6 (12):a019315

Pinney JW, Amoutzias GD, Rattray M, Robertson DL (2007) Reconstruction of ancestral protein interaction networks for the bZIP transcription factors. Proceedings of the National Academy of Sciences 104 (51):20449-20453

Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ (2007) Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445 (7126):383

Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ (2011) Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. Journal of theoretical biology 272 (1):141-144

Presgraves DC (2010) The molecular evolutionary basis of species formation. Nature Reviews Genetics 11:175. doi:10.1038/nrg2718

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The *CONSTANS* gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80:847-857

Qi Q, Gibson A, Fu X, Zheng M, Kuehn R, Wang Y, Wang Y, Navarro S, Morrell JA, Jiang D, Simmons G, Bell E, Ivleva NB, McClerren AL, Loida P, Ruff TG, Petracek ME, Preuss SB (2012) Involvement of the N-terminal B-box domain of Arabidopsis BBX32 protein in interaction with soybean BBX62 protein. The Journal of biological chemistry 287 (37):31482-31493. doi:10.1074/jbc.M112.346940

Qu B, He X, Wang J, Zhao Y, Teng W, Shao A, Zhao X, Ma W, Wang J, Li B (2015) A wheat CCAAT box-binding transcription factor increases the grain yield of wheat with less fertilizer input. Plant physiology 167 (2):411-423

Quach TN, Nguyen HT, Valliyodan B, Joshi T, Xu D, Nguyen HT (2015) Genome-wide expression analysis of soybean NF-Y genes reveals potential function in development and drought response. Molecular Genetics and Genomics 290 (3):1095-1115

Remans T, Nacry P, Pervent M, Girin T, Tillard P, Lepetit M, Gojon A (2006) A central role for the nitrate transporter NRT2. 1 in the integrated morphological and physiological responses of the root system to nitrogen limitation in Arabidopsis. Plant physiology 140 (3):909-921

Reyes JC, Muro-Pastor MI, Florencio FJ (2004) The GATA family of transcription factors in Arabidopsis and rice. Plant physiology 134 (4):1718-1732

Rizzini L, Favory J-J, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI, Ulm R (2011) Perception of UV-B by the Arabidopsis UVR8 protein. Science (New York, NY) 332 (6025):103-106. doi:10.1126/science.1200660

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26 (1):139-140. doi:10.1093/bioinformatics/btp616

Robson F, Costa MM, Hepworth SR, Vizir I, Piñeiro M, Reeves PH, Putterill J, Coupland G (2001) Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28 (6):619-631

Rodriguez-Martinez JA, Reinke AW, Bhimsaria D, Keating AE, Ansari AZ (2017) Combinatorial bZIP dimers display complex DNA-binding specificity landscapes. Elife 6:e19272

Roeder RG (1991) The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends in biochemical sciences 16:402-408

Romier C, Cocchiarella F, Mantovani R, Moras D (2003) The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. The Journal of biological chemistry 278 (2):1336-1345

Ronchi A, Bellorini M, Mongelli N, Mantovani R (1995) CCAAT-box binding protein NF-Y (CBF, CP1) recognizes the minor groove and distorts DNA. Nucleic Acids Res 23 (22):4565-4572

Sabourin M, Tuzon CT, Fisher TS, Zakian VA (2007) A flexible protein linker improves the function of epitope-tagged proteins in Saccharomyces cerevisiae. Yeast (Chichester, England) 24 (1):39-45. doi:10.1002/yea.1431

Salsi V, Caretti G, Wasner M, Reinhard W, Haugwitz U, Engeland K, Mantovani R (2003) Interactions between p300 and multiple NF-Y trimers govern cyclin B2 promoter function. The Journal of biological chemistry 278 (9):6642-6650

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science (New York, NY) 288 (5471):1613-1616

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K,

Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9 (7):676-682. doi:10.1038/nmeth.2019

Seo HS, Yang J-Y, Ishikawa M, Bolle C, Ballesteros ML, Chua N-H (2003) LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1. Nature 423 (6943):995-999. doi:10.1038/nature01696

Serfling E, Jasin M, Schaffner W (1985) Enhancers and eukaryotic gene transcription. Trends in genetics 1:224-230

Sessa G, Carabelli M, Sassi M, Ciolfi A, Possenti M, Mittempergher F, Becker J, Morelli G, Ruberti I (2005) A dynamic balance between gene activation and repression regulates the shade avoidance response in Arabidopsis. Genes & development 19 (23):2811-2815. doi:10.1101/gad.364005

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. The Plant cell 14 (12):2985-2994

Sharrock RA, Quail PH (1989) Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes & development 3 (11):1745-1757

Shi H, Ye T, Zhong B, Liu X, Jin R, Chan Z (2014) AtHAP5A modulates freezing stress resistance in Arabidopsis through binding to CCAAT motif of AtXTH21. New Phytologist 203 (2):554-567

Shi H, Zhong S, Mo X, Liu N, Nezames CD, Deng XW (2013) HFR1 sequesters PIF1 to govern the transcriptional network underlying light-initiated seed germination in Arabidopsis. The Plant cell 25 (10):3770-3784. doi:10.1105/tpc.113.117424

Shiu S-H, Shih M-C, Li W-H (2005) Transcription Factor Families Have Much Higher Expansion Rates in Plants than in Animals. Plant Physiology 139 (1):18. doi:10.1104/pp.105.065110

Sibout R, Sukumar P, Hettiarachchi C, Holm M, Muday GK, Hardtke CS (2006) Opposite root growth phenotypes of hy5 versus hy5 hyh mutants correlate with increased constitutive auxin signaling. PLoS Genet 2 (11):e202. doi:06-PLGE-RA-0345R2 [pii]

Siddiq MA, Hochberg GKA, Thornton JW (2017) Evolution of protein specificity: insights from ancestral protein reconstruction. Current Opinion in Structural Biology 47:113-122. doi:https://doi.org/10.1016/j.sbi.2017.07.003

Siefers N, Dang KK, Kumimoto RW, Bynum WE, Tayrose G, Holt BF (2009) Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. Plant Physiol 149 (2):625-641. doi:10.1104/pp.108.130591

Silva RF, Mendonça SC, Carvalho LM, Reis AM, Gordo I, Trindade S, Dionisio F (2011) Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations. PLoS genetics 7 (7):e1002181

Singh RP, Prasad HK, Sinha I, Agarwal N, Natarajan K (2011) Cap2-HAP Complex Is a Critical Transcriptional Regulator That Has Dual but Contrasting Roles in Regulation of Iron Homeostasis in Candida albicans. The Journal of biological chemistry 286 (28):25154-25170. doi:M111.233569 [pii]

Sinha S, Kim IS, Sohn KY, de Crombrugghe B, Maity SN (1996) Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. Mol Cell Biol 16 (1):328-337

Sinha S, Maity SN, Lu J, de Crombrugghe B (1995) Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. Proceedings of the National Academy of Sciences of the United States of America 92 (5):1624-1628

Siriwardana CL, Gnesutta N, Kumimoto RW, Jones DS, Myers ZA, Mantovani R, Holt BF (2016) NUCLEAR FACTOR Y, Subunit A (NF-YA) Proteins Positively Regulate Flowering and Act Through FLOWERING LOCUS T. PLoS Genet 12 (12):e1006496. doi:10.1371/journal.pgen.1006496

Siriwardana CL, Kumimoto RW, Jones DS, Holt BF (2014) Gene Family Analysis of the Arabidopsis NF-YA Transcription Factors Reveals Opposing Abscisic Acid Responses During Seed Germination. Plant Mol Biol Report 32 (5):971-986. doi:10.1007/s11105-014-0704-6

Skirycz A, Inzé D (2010) More from less: plant growth under limited water. Current Opinion in Biotechnology 21 (2):197-203. doi:https://doi.org/10.1016/j.copbio.2010.03.002

Soh MS, Kim YM, Han SJ, Song PS (2000) REP1, a basic helix-loop-helix protein, is required for a branch pathway of phytochrome A signaling in arabidopsis. The Plant cell 12 (11):2061-2074

Somers DE, Quail PH (1995) Temporal and spatial expression patterns of PHYA and PHYB genes in Arabidopsis. Plant J 7 (3):413-427

Song L, Huang S-sC, Wise A, Castanon R, Nery JR, Chen H, Watanabe M, Thomas J, Bar-Joseph Z, Ecker JR (2016) A transcription factor hierarchy defines an environmental stress response network. Science (New York, NY) 354 (6312). doi:10.1126/science.aag1550

Sorin C, Declerck M, Christ A, Blein T, Ma L, Lelandais-Briere C, Njo MF, Beeckman T, Crespi M, Hartmann C (2014) A miR169 isoform regulates specific NF-YA targets and root architecture in Arabidopsis. New Phytol 202 (4):1197-1211. doi:10.1111/nph.12735

Stacey MG, Hicks SN, von Arnim AG (1999) Discrete domains mediate the light-responsive nuclear and cytoplasmic localization of Arabidopsis COP1. The Plant cell 11 (3):349-364

Sultan SE (2000) Phenotypic plasticity for plant development, function and life history. Trends in plant science 5 (12):537-542. doi:https://doi.org/10.1016/S1360-1385(00)01797-0

Sundseth R, MacDonald G, Ting J, King AC (1997) DNA elements recognizing NF-Y and Sp1 regulate the human multidrug-resistance gene promoter. Mol Pharmacol 51 (6):963-971

Suzuki M, Wang HH, McCarty DR (2007) Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. Plant Physiol 143 (2):902-911

Swain S, Myers ZA, Siriwardana CL, Holt BF, 3rd (2017) The multifaceted roles of NUCLEAR FACTOR-Y in Arabidopsis thaliana development and stress responses. Biochimica et biophysica acta 1860 (5):636-644. doi:10.1016/j.bbagrm.2016.10.012

Takahashi K, Hayashi K-i, Kinoshita T (2012) Auxin Activates the Plasma Membrane H^+-ATPase by Phosphorylation during Hypocotyl Elongation in Arabidopsis. Plant Physiology 159 (2):632. doi:10.1104/pp.112.196428

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences of the United States of America 101 (30):11030-11035. doi:10.1073/pnas.0404206101

Tang Y, Liu X, Liu X, Li Y, Wu K, Hou X (2017) Arabidopsis NF-YCs Mediate the Light-Controlled Hypocotyl Elongation via Modulating Histone Acetylation. Molecular plant 10 (2):260-273. doi:10.1016/j.molp.2016.11.007

Tao Z, Shen L, Gu X, Wang Y, Yu H, He Y (2017) Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. Nature 551:124. doi:10.1038/nature24300

Thornton JW (2004) Resurrecting ancient genes: experimental analysis of extinct molecules. Nature Reviews Genetics 5 (5):366

Tirosh I, Barkai N, Verstrepen KJ (2009) Promoter architecture and the evolvability of gene expression. Journal of biology 8 (11):95

Tiwari SB, Shen Y, Chang HC, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, Belachew A, Repetti PP, Reuber TL, Ratcliffe OJ (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. New Phytol 187 (1):57-66. doi:NPH3251 [pii]

Toledo-Ortiz G, Johansson H, Lee KP, Bou-Torrent J, Stewart K, Steel G, Rodriguez-Concepcion M, Halliday KJ (2014) The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. PLoS Genet 10 (6):e1004416. doi:10.1371/journal.pgen.1004416

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol 59:573-594. doi:10.1146/annurev.arplant.59.032607.092755

Vandenbussche F, Habricot Y, Condiff AS, Maldiney R, Van der Straeten D, Ahmad M (2007) HY5 is a point of convergence between cryptochrome and cytokinin signalling pathways in Arabidopsis thaliana. Plant J 49 (3):428-441. doi:10.1111/j.1365-313X.2006.02973.x

Vicient CM, Bies-Etheve N, Delseny M (2000) Changes in gene expression in the leafy cotyledon1 (lec1) and fusca3 (fus3) mutants of Arabidopsis thaliana L. Journal of experimental botany 51 (347):995-1003

von Arnim AG, Deng XW (1994) Light inactivation of Arabidopsis photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning. Cell 79 (6):1035-1045

Warpeha KM, Upadhyay S, Yeh J, Adamiak J, Hawkins SI, Lapik YR, Anderson MB, Kaufman LS (2007) The GCR1, GPA1, PRN1, NF-Y signal chain mediates both blue light and abscisic acid responses in Arabidopsis. Plant Physiol 143 (4):1590-1600

Wei Q, Ma C, Xu Y, Wang T, Chen Y, Lü J, Zhang L, Jiang C-Z, Hong B, Gao J (2017) Control of chrysanthemum flowering through integration with an aging pathway. Nature communications 8:829

Weinreich DM, Watson RA, Chao L (2005) PERSPECTIVE: SIGN EPISTASIS AND GENETIC COSTRAINT ON EVOLUTIONARY TRAJECTORIES. Evolution 59 (6):1165-1174. doi:10.1111/j.0014-3820.2005.tb01768.x

Wenkel S, Turck F, Singer K, Gissot L, Gourrierec JL, Samach A, Coupland G (2006) CONSTANS and the CCAAT Box Binding Complex Share a Functionally Important Domain and Interact to Regulate Flowering of Arabidopsis. The Plant cell 18 (11):2971-2984. doi:10.1105/tpc.106.043299

Werner T, Motyka V, Strnad M, Schmülling T (2001) Regulation of plant growth by cytokinin. Proceedings of the National Academy of Sciences 98 (18):10487-10492

West M, Yee KM, Danao J, Zimmerman JL, Fischer RL, Goldberg RB, Harada JJ (1994) LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in Arabidopsis. The Plant cell 6 (12):1731-1745

Whitlock MC, Phillips PC, Moore FB-G, Tonsor SJ (1995) Multiple fitness peaks and epistasis. Annual review of ecology and systematics 26 (1):601-629

Wouters FS, Bastiaens PI (1999) Fluorescence lifetime imaging of receptor tyrosine kinase activity in cells. Curr Biol 9 (19):1127-1130

Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution, vol 1. na,

Wright S (1988) Surfaces of selective value revisited. The American Naturalist 131 (1):115-123

Xu D, Li J, Gangappa SN, Hettiarachchi C, Lin F, Andersson MX, Jiang Y, Deng XW, Holm M (2014a) Convergence of Light and ABA signaling on the ABI5 promoter. PLoS Genet 10 (2):e1004197. doi:10.1371/journal.pgen.1004197

Xu L, Lin Z, Tao Q, Liang M, Zhao G, Yin X, Fu R (2014b) Multiple NUCLEAR FACTOR Y transcription factors respond to abiotic stress in Brassica napus L. PloS one 9 (10):e111354

Xuanyuan G, Lu C, Zhang R, Jiang J (2017) Overexpression of StNF-YB3. 1 reduces photosynthetic capacity and tuber production, and promotes ABA-mediated stomatal closure in potato (Solanum tuberosum L.). Plant Science 261:50-59

Yadav D, Shavrukov Y, Bazanova N, Chirkova L, Borisjuk N, Kovalchuk N, Ismagul A, Parent B, Langridge P, Hrmova M (2015) Constitutive overexpression of the TaNF-YB4 gene in transgenic wheat significantly improves grain yield. Journal of experimental botany 66 (21):6635-6650

Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T (2009) Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. Plant J 58 (5):843-856. doi:TPJ3817 [pii]

Yamawaki S, Yamashino T, Nakamichi N, Nakanishi H, Mizuno T (2011) Light-responsive double B-box containing transcription factors are conserved in Physcomitrella patens. Biosci Biotechnol Biochem 75 (10):2037-2041. doi:10.1271/bbb.110359

Yang J, Xie Z, Glover BJ (2005) Asymmetric evolution of duplicate genes encoding the CCAAT-binding factor NF-Y in plant genomes. New Phytol 165 (2):623-631

Yang M, Zhao Y, Shi S, Du X, Gu J, Xiao K (2017) Wheat nuclear factor Y (NF-Y) B subfamily gene TaNF-YB3; l confers critical drought tolerance through modulation of the ABA-associated signaling pathway. Plant Cell, Tissue and Organ Culture (PCTOC) 128 (1):97-111

Yang SW, Jang I-C, Henriques R, Chua N-H (2009) FAR-RED ELONGATED HYPOCOTYL1 and FHY1-LIKE associate with the Arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome A signals for inhibition of hypocotyl elongation. The Plant cell 21 (5):1341-1359. doi:10.1105/tpc.109.067215

Yoshida H, Okada T, Haze K, Yanagi H, Yura T, Negishi M, Mori K (2001) Endoplasmic reticulum stress-induced formation of transcription factor complex ERSF including NF-Y (CBF) and activating transcription factors 6alpha and 6beta that activates the mammalian unfolded protein response. Mol Cell Biol 21 (4):1239-1248. doi:10.1128/MCB.21.4.1239-1248.2001

Zambelli F, Pavesi G (2017) Genome wide features, distribution and correlations of NF-Y binding sites. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1860 (5):581-589. doi:https://doi.org/10.1016/j.bbagrm.2016.10.007

Zanetti ME, Blanco FA, Beker MP, Battaglia M, Aguilar OM (2011) A C subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean-Rhizobium etli symbiosis. The Plant cell 22 (12):4142-4157. doi:tpc.110.079137 [pii]

Zanetti ME, Ripodas C, Niebel A (2017) Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. Biochimica et biophysica acta 1860 (5):645-654. doi:10.1016/j.bbagrm.2016.11.007

Zeng YX, Hu CY, Lu YG, Li JQ, Liu XD (2007) Diversity of abnormal embryo sacs in indica/japonica hybrids in rice demonstrated by confocal microscopy of ovaries. Plant Breeding 126 (6):574-580. doi:10.1111/j.1439-0523.2007.01380.x

Zhang H, Forde BG (2000) Regulation of Arabidopsis root development by nitrate availability. Journal of experimental botany 51 (342):51-59

Zhang H, He H, Wang X, Wang X, Yang X, Li L, Deng XW (2011) Genome-wide mapping of the HY5-mediated gene networks in Arabidopsis that involve both transcriptional and post-transcriptional regulation. Plant J 65 (3):346-358. doi:10.1111/j.1365-313X.2010.04426.x

Zhang T, Zhang D, Liu Y, Luo C, Zhou Y, Zhang L (2015) Overexpression of a NF-YB3 transcription factor from Picea wilsonii confers tolerance to salinity and drought stress in transformed Arabidopsis thaliana. Plant Physiology and Biochemistry 94:153-164

Zhou P, Song M, Yang Q, Su L, Hou P, Guo L, Zheng X, Xi Y, Meng F, Xiao Y, Yang L, Yang J (2014) Both PHYTOCHROME RAPIDLY REGULATED1 (PAR1) and PAR2 promote seedling photomorphogenesis in multiple light signaling pathways. Plant Physiol 164 (2):841-852. doi:10.1104/pp.113.227231

Zhou Q, Hare PD, Yang SW, Zeidler M, Huang LF, Chua NH (2005) FHL is required for full phytochrome A signaling and shares overlapping functions with FHY1. Plant J 43 (3):356-370. doi:10.1111/j.1365-313X.2005.02453.x

Zoltowski BD, Imaizumi T (2014) Structure and Function of the ZTL/FKF1/LKP2 Group Proteins in Arabidopsis. Enzymes 35:213-239. doi:10.1016/B978-0-12-801922-1.00009-9