Title
Using Optogenetics to Study Dynamic Signal Encoding and Decoding in S. cerevisiae

Author
Chen, Susan Yushan

Publication Date
2019

Peer reviewed|Thesis/dissertation

Using Optogenetics to Study Dynamic Signal Encoding and Decoding in S. cerevisiae

by
Susan Y Chen

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Hana El-Samad
D971DFD8316E4DC...

Hana El-Samad
Chair

DocuSigned by:

Wendell Lim

DocuSigned by:
Matt Thomson
51879996F5FA44B...

Matt Thomson

Committee Members

Acknowledgments

My journey through graduate school would not have been possible without my family, friends, and mentors who have believed in me and supported me.

My undergraduate advisor, John Dueber, had fostered and mentored my interest in science. On a regular basis, John and I discussed project ideas. He mentored me directly, and he led by example. One Saturday afternoon, when John saw me prep seemingly endless samples for GCMS for a publication, he started pipetting with me. I cannot thank John enough for believing in me and being a tremendous mentor.

In my time at Berkeley, I've been so lucky to have a wonderful network of friends, Gabe, Anna, Madhvi, Jenn, Tim, Nav, Justin, and Su. Collectively, they've contributed to my appreciation of scientific research for its creativity, problem solving, and impact. I want to specifically thank Gabe, Anna, Madhvi, and Jenn who have made doing science in Stanley and Hildebrand so fun together – cloning up a storm with Gabe and doing figure skating jumps on the lab floor at 12am with Anna. Even beyond Berkeley, our community flourished in the form of a cross-country journal club (JAGS).

At UCSF, I feel so grateful to have built a wonderful network of friends and colleagues who have made graduate school both academically stimulating and survivable. Thank you Han, Alex, Jess, Anne, Aishwarya, and Arunan. I want to specifically shout out my wonderful friend and roommate, Han. Han has been there for me through and through the good times and the bad. I am also indebted to Han for all the delicious meals she has cooked me and, more importantly, the deep life lessons she has taught me.

Thank you also to the El-Samad lab. I'm so luck to be part of a group of dedicated, smart, socially conscious, and genuinely nice scientists. In my early days in the lab, Jacob, Ignacio, and Raj taught me a lot about systems biology and were my scientific mentors. I also thank Michael, Andrew, Lindsey, Joao, and Alain for being wonderful teammates and friends. Lastly, to my classmates, Patrick, Graham, and Ben, you've been there the whole time and my UCSF experience would be incomplete without you all.

I thank my Thesis committee, Hana El-Samad, Wendell Lim, and Matt Thomson for advising me on my thesis project. I thank Hana for teaching me to be a critical scientist and for providing the financial means for me to explore my scientific interests. I thank Wendell for his "go for the jugular" research instincts. I thank Matt for his guidance and advice, especially the words "Take ownership, this is your work, your job."

My UCSF community includes people outside my lab and thesis committee who have been mentors to me when they have no obligation to be. Thank you Steve Altschuler, Joe DeRisi, and Chiraj who have spent time conversing with me throughout the years and truly care about my scientific and personal growth. I also thank the Altschuler-Wu lab for being such excellent lunch companions.

Ultimate frisbee has been a constant source of fun and community throughout graduate school. I am thankful especially to my team FAB and SF Pod. We chose to spend every weekend in the summer together running for hours in the hot sun, to show-up and commit to each other's growth as an ultimate player and person, and to create a positive and value-driven community. Thank you FAB, for showing me what being a teammate means, for teaching me the value and power of being together.

Throughout my time at UCSF, I've frequented the Rausch House on Sundays for many a family dinner. Thank you Chenxia, Mike, Scott, P$, and Nikki for being the best housemates a person could ask for. You have created such a generous, kind, and fun community that I have the benefit of being part of. I treasure the impromptu get togethers, trips to Costco, the search for delicious Asian food, Overcooked!2, house election nights, and parties at

Samchelle's.

Graham, thank you for being the most supportive partner. We have grown and sojourned together through graduate school; from the early days of studying for quals, to working out an RNA sequencing pipeline, to identifying value propositions, to weekly lunches. In the hardest most trying times in graduate school, you were the one who can make me laugh and helped me see the big picture. You believed in me even when I didn't believe in myself. Thank you for being my rock and pillar, for believing in me.

In the last year of graduate school, Graham's parents, Karen and Rick, have become like second parents to me. They have rooted for me and generously supported me in so many ways. They are role models to me of how to navigate a work-life balance. A most sincere and whole-hearted thank you.

Lastly, I thank my family. I thank my parents, grandma, and sister for their unwavering love and support of me. My heart swells as I think of the pride that you all have in me. I especially thank my grandmother who instilled in me early on the importance of education, moral strength, and the discipline required to pursue any effort. To my mother, I cannot thank you enough for all the sacrifices you have made for me to get a good education. In my first years in the US, you drove me to the Fremont Main Library every weekend to teach me English and math. When I wanted to seriously pursue ice skating, you poured your money and time to drive me to the ice rink at 5am every day. Mom, thank you for all the sacrifices you have made so that I can have the opportunities that you never had. I love you.

Using Optogenetics to Study Dynamic Signal Encoding and Decoding in *S. cerevisiae*

Susan Y Chen

Abstract

Transcription factors are key mediators of environmental signals. In *S. cerevisiae*, cells may change the concentration, phosphorylation state, binding partner, or temporal dynamics of transcription factors to mediate a response to upstream signals. Dynamic localization of transcription factors to and from the nucleus have been observed for a number of stress-responsive transcription factors in yeast. Two key questions arise from this observation. First, what are the upstream determinants of these nuclear localization events? And second, what are the downstream target gene interpretations?

To probe the upstream determinants of nuclear localization, the activity of PKA, an upstream regulator of the transcription factor Msn2, was perturbed using a combination of genetic mutations and optogenetics. Optogenetic activation with bacterial Photoactivatible Adenylyl Cyclase (bPAC) established a causal connection of PKA activity on Msn2 dynamic nuclear localization. Genetic deletions of PKA network components illustrated

that changes to the upstream signaling network can alter Msn2 localization dynamics. The computational analyses revealed that a negative feedback on PKA could explain the observed transient dynamics of Msn2 nuclear localization, and that different PKA network components could tune those Msn2 localization dynamics.

To probe the second question of the downstream target gene interpretations of nuclear localization, another optogenetic tool was optimized and used to directly control transcription factor nuclear localization. CLASP (Controllable Light Activated Shuttling and Plasma membrane sequestration), was engineered to enable precise, modular, and reversible control of TF localization using a combination of two optimized LOV2 optogenetic constructs that sequestered the TF to the plasma membrane in the dark and delivered the TF to the nucleus in the light. CLASP achieved minute-level resolution, reversible localization of many TF cargos, large dynamic range, and tunable target gene expression. Dynamic control of the nuclear localization of Crz1, a naturally pulsatile TF, using CLASP revealed that some Crz1 target genes respond more efficiently to pulsatile TF inputs than to continuous inputs, while others exhibited the opposite behavior. Computational modeling showed that efficient gene expression in response to short pulsing required fast promoter activation and slow inactivation, and that the opposite phenotype can ensued from a multi-stage promoter activation, where a transition in the first stage was thresholded. These data directly demonstrate differential interpretation of TF pulsing dynamics by different genes, and provide plausible models that could achieve these phenotypes.

In summary, this work explored dynamic signal encoding and decoding of transcription factor nuclear localization using optogenetics in *S. cerevisiae*. Control of PKA activity and deletion of PKA network components showed encoding of dynamic transcription factor nuclear localization by the PKA network. Arbitrary control of transcription factor nuclear localization identified differential decoding of transcription factor dynamics by target genes.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cells are exquisite autonomous machines that can execute complicated processes seamlessly in response to their environments. Take for example, the single-celled organism *S. cerevisiae*. When there is abundant glucose, *S. cerevisiae* responds by utilizing the glucose in a process known as fermentation to produce energy and alcohol as a byproduct. This cellular response to glucose, fermentation, is the crucial process behind two staples of modern society, bread and beer. Yeast responds to a plethora of other conditions in their environment such as high salt[1], heat shock[2], or the presence of a mate[3], and produce the appropriate response to survive and grow.

A central question in cell signaling is how a cell (yeast, for instance) is able to faithfully relay an external signal and mount an appropriate response. The molecular players in many of these signaling pathways have been identified, but the mechanisms by which information flows through these molecular nodes is not as well understood. A common motif in yeast signaling pathways is a signaling cascade in which an environmental signal activates a kinase (e.g. PKA[4]) or phosphatase (e.g. calcineurin[5]) that then acts on their corresponding transcription factors (e.g. PKA phosphorylation of Msn2[6] and calcineurin

de-phosphorylation of Crz1[7]) to elicit a transcriptional response. An interesting observation is that many kinase- or phosphatase-mediated transcription factors dynamically localize to the nucleus in response to environmental signals[8,9] (also see Figures 4.10, 4.11, 4.12). These nuclear localization events raise the hypothesis that changes in the dynamics of TF nuclear localization may be a mechanism to encode information from upstream signals that can then be decoded by downstream genes. This work used precise molecular tools to address this two-fold question of the origins and consequences of dynamic transcription factor nuclear localization.

The PKA/Msn2 signaling pathway was used to explore how changes in upstream PKA signaling could dictate Msn2 nuclear localization dynamics. In order to study the effects of PKA and its network components on Msn2 localization, a method of direct input to PKA is desired to establish a causal link. PKA signaling can be modulated by natural inputs[1,2,10] or small molecules[11]. Both of these methods, however, fall short to directly and reversibly control PKA since natural inputs are pleiotropic and small molecules are difficult to washout. Optogenetics, on the other hand, has the potential to provide direct and reversible control. Using a previously published bacterial Photoactivatible Adenylyl Cyclase[12] (bPAC), optogenetic activation of PKA showed that Msn2 localization exhibits mirrored dynamics to PKA activity, with a characteristic transient pulse of Msn2 localization following shut-off of light input. Deletions of three negative regulators of PKA further showed distinct alterations to the dynamics of the observed transient pulse of Msn2 localization for each deletion of PKA network components. Computational modeling of these phenotypes found that a negative feedback on PKA caused the transient Msn2 pulse, and, moreover, that negative regulators of PKA that act on different parts of the negative feedback serve to tune the observed Msn2 dynamics. These insights provide an example of how an upstream signaling network modulates the dynamics of transcription factor localization.

To draw a causal link between dynamic transcription factor localization and downstream transcriptional response, a direct method of control is required. Again, optogenetics provide the most direct and reversible means of control. Two previously published tools, LANS[13] and LOVTrap[14], were combined and optimized to create a new and robust tool, CLASP, (Controlled Light-Activated Shuttling and Plasma membrane sequestration) which sequesters transcription factors to the plasma membrane in the dark, and localizes transcription factors to the nucleus in a fast, reversible, and general manner with light. Dynamic control of a naturally pulsatile transcription factor, Crz1, showed that some target genes express more efficiently in response to short pulses, while others express less efficiently. Moreover, computational modeling illustrated that fast promoter turn ON and slow turn OFF could be responsible for the efficient response to pulsed inputs, while a thresholded and slow transition between non-transcribing promoter states could produce an inefficient response. This study provides an example of how transcription factor dynamics could differentially regulate downstream genes.

Taken together, these studies make an earnest attempt to address the central question in cell signaling – how do cells relay environmental signals and mount an appropriate response? To the forest of existing knowledge, this work adds a healthy seedling. It focuses the spotlight on transcription factor dynamics. This work explored the encoding of transcription factor dynamics by an upstream kinase signaling network. It also dissected the differential decoding of transcription factor dynamics by target genes. Moreover, the precise optogenetic approaches engineered and used in these studies highlight one arch in a key cycle in biological discovery. In this cycle, new tools lead to new discoveries, and new discoveries lead to new tools. It will be very exciting to take the biological insights gained from these studies to forward engineer new synthetic tools such as frequency responsive promoters that can be used to discover new biology or be applied to synthetic biology and metabolic engineering applications.

Chapter 2

Model-guided optogenetic study of PKA signaling in budding yeast

2.1 Summary

This study addresses the first question of the upstream determinants that control transcription factor nuclear localization dynamics by using optogenetics to study how changes in the Protein Kinase A (PKA) signaling network alters the protein dynamics of PKA and nuclear localization of Msn2.

2.2 Introduction

The second messenger cyclic-AMP (cAMP) is a ubiquitous signaling molecule whose synthesis by adenylate cyclase and degradation by phosphodiesterases (PDEs) occur in all branches of life. In eukaryotes, protein kinase A (PKA) is the most conserved cAMP-responsive protein. Binding of cAMP to the regulatory subunit of PKA frees its

catalytic units to phosphorylate hundreds of targets regulating a vast swath of metabolism and cellular physiology. cAMP often exhibits pulsatile or oscillatory dynamics. In Dictyostelium, for example, waves of cAMP coordinate colony growth and differentiation[15], and in humans, oscillating cAMP levels regulate insulin secretion in pancreatic beta cells[16].

In budding yeast, cAMP levels are regulated by extracellular glucose and a range of growth- and stress-related signals. Changes in these environmental variables alter cAMP levels, modulating activity of the PKA complex, which in turn fans out to regulate a wide range of cellular processes, estimated to involve at least one-third of the genome[10]. PKA has been shown to directly phosphorylate more than two dozen proteins, including the mitochondrial protein import machinery, P-body components, autophagy proteins, glycolysis machinery, and a large number of transcription factors[17]. In conditions of plentiful resources, cAMP levels are high, and PKA promotes rapid fermentative growth by enhancing glycolysis and ribosomal production. In stressful conditions, a decrease in cAMP levels causes a drop in PKA activity, resulting in the inhibition of ribosomal and growth-related programs and the activation of stress-responsive factors such as the transcription factor Msn2. This low-PKA-activity state is in many cases transient, as negative feedback loops embedded in the PKA network and involving the small G-protein Ras cause PKA activity to rebound even as stressful conditions continue[18].

PKA exerts much of its influence through regulation of the nuclear localization and activity of several transcription regulators, most notably the stress-responsive Msn2 and its homologue, Msn4. Decreased phosphorylation by PKA of these stress-responsive transcription factors leads to their nuclear localization[19]. Observation of the PKA-regulated subcellular localization of Msn2 in yeast cells suggests that PKA activity is highly dynamic, with rapid pulses of activity occurring in "bursts" of Msn2 localization on the minute time scale[20].

Although substantial progress has been made in probing the downstream consequences of

dynamic Msn2 pulsing[9,21,22], our understanding of how PKA generates these dynamics is incomplete, partly because upstream tools to perturb the PKA system *in vivo* are typically slow (e.g., mutations) or nonspecific (e.g., stress). A tool that provides a rapid, specific, and reversible perturbation of PKA signaling is therefore needed. In this work, we develop the recently discovered bacterial photoactivatable adenylate cyclase (bPAC)[12,23] as a quantitative perturbation tool to investigate PKA dynamics in *Saccharomyces cerevisiae*. We demonstrate that by expressing this bacterial protein in yeast cells, we can achieve high-resolution temporal control of PKA activity. Using precise optogenetic stimuli, we develop and test a model of PKA signaling and uncover important aspects of its dynamics.

2.3 Results

2.3.1 Rapid *in vivo* regulation of PKA activity by a bPAC

Previous reports demonstrated that PKA exhibits complex dynamics that occurs on fast time scales, on the order of minutes[20,24,25]. Therefore, quantitative studies of this system require perturbations with a time resolution of seconds. To achieve such rapid perturbation, we expressed in budding yeast cells a recently characterized bPAC, originally from the soil bacterium *Beggiatoa*[12,26], that contains a light-sensitive BLUF (blue light receptor using FAD) domain linked to an adenylate cyclase domain. This construct allows for regulation of cAMP levels in living cells by illumination with blue light (<500 nm), which transiently alters the conformation of the light-sensitive BLUF domain, rendering the associated cyclase domain competent to catalyze the conversion of ATP to cAMP (Figure 2.1A). To take advantage of this genetically encoded tool, we assembled a custom-built system that allows us to coordinate the imaging of budding yeast in a fluorescence microscope with illumination by a blue-light LED. This system is capable of

providing graded illumination for 1–40 µW/$mm^2$2. Although substantially brighter than ambient light (∼4 µW/$mm^2$2), these blue-light intensities do not trigger a stress response in budding yeast (Supplementary Figure 2.5). Furthermore, bPAC has a fast spontaneous reversion to the inactive state[12] (<30 s; Supplementary Figure 2.6), enabling of bPAC activity at high time resolution regulation through control of light intensity. Using this infrastructure, we were able to apply controlled amounts of blue light to cells expressing the bPAC construct and quantify the resulting PKA activity in real time using a localization-based fluorescent Msn2 reporter (Supplementary Figure 2.6). Msn2 nuclear localization is modulated by PKA via phosphorylation of PKA motifs (RRxS) in the Msn2 nuclear import (nuclear localization sequence [NLS]) and export sequences[24], making the nuclear enrichment of Msn2 an accurate measure of PKA activity[9].

To test the ability of bPAC to regulate PKA *in vivo*, we built a yeast strain harboring both a hormone-inducible dominant-negative allele of Ras2S24N[10,27] and a constitutively expressed bPAC. We used the translocation of mCherry-tagged Msn2 in and out of the nucleus as a measure of PKA activity[9] (quantification of nuclear localization is discussed in Materials and Methods). Expression of Ras2S24N inhibits cAMP production by sequestering the Ras guanine exchange factor CDC25, preventing the formation of Ras-GTP and hence activation of the endogenous adenylate cyclase. As a result, when this allele is induced, we expect cAMP levels to drop, PKA to become inactive, and Msn2-mCherry to be localized to the nucleus. Consistent with this expectation, we observed complete localization of Msn2 to the nuclei of cells within 10 min of the induction of Ras2S24N expression (Figure 2.1B and Supplementary Movie S1). Activation of bPAC should provide an orthogonal non–Ras-dependent source of cAMP, which would activate PKA and cause Msn2-mCherry to be exported from the nucleus. Indeed, upon illumination with blue light, there was a rapid and synchronous exit of Msn2 from the nucleus. These data are consistent with a sharp increase in PKA activity upon illumination induced by cAMP production from the activated bPAC and show that bPAC can produce sufficient

cAMP to compensate for the loss of Ras-dependent endogenous adenylate cyclase activity.

2.3.2 The dynamics of PKA signaling induced by optogenetic stimulation is explained by a computational model

Using the precise perturbation afforded by the bPAC construct, we next explored how PKA dynamics is shaped by its signaling network. To do so, we exposed wild-type cells expressing bPAC to a 3-min pulse of light (Figure 2.2A). As discussed earlier, in resting cells, Msn2 localization is predominantly cytoplasmic due to the PKA phosphorylation of the NLS, with occasional stochastic nuclear localization in individual cells. Consequently, blue light–mediated activation of bPAC and the subsequent increase in cAMP had only the slight and expected effect of inducing uniform Msn2 cytoplasmic localization. Of interest, however, soon after the blue light was turned off, we observed a sharp transient (lasting $\sim$5 min) increase in Msn2 nuclear localization above the initial localization state (Figure 2.2B and Supplementary Movie S2). This pulse of Msn2 nuclear localization after bPAC shutoff is surprising since we expected Msn2 to return monotonically to its prestimulus steady state. The observed behavior is a hallmark of an underdamped system, whose signature is a transient overshoot (transient PKA inactivation) before return to steady state (fully active PKA) upon change in input, as compared with an overdamped system, which exhibits monotonic return to this steady state. Such unexpected behavior can be generated by a negative feedback loop with a delay, a hypothesis that we wanted to further pursue.

Feedback loops are believed to decorate the architecture of the PKA signaling network. To investigate the possibility that feedback is responsible for the underdamped behavior of the PKA system after bPAC-mediated rapid alteration of cAMP levels, we built two mathematical models of the PKA network. The first is a negative feedback model consisting of PKA, its core regulatory and activation components (Cdc25, Ras, and

cAMP), the exogenous optogenetic stimulation, and Msn2 nuclear shuttling kinetics. The negative feedback in this system has been identified genetically and is generally considered to be implemented through regulation of the Ras1/2 GTPase proteins[28,29]. The second model is an identical open loop variant that lacks the negative feedback component, as the dependence of Cdc25 levels on PKA activity is removed (Supplementary Note S1). With this open-loop model, we wanted to test whether the Msn2 pulse of activity could be simply generated by the interaction of Msn2 with PKA without the need to invoke feedback or whether it is necessarily generated by negative feedback in PKA signaling.

In contrast to prior work[25,30], we did not seek to build a model that reflects all known biochemical interactions surrounding PKA signaling. Instead, by capturing only the known essential interactions in the PKA system, we aimed to build a simple model that could recapitulate the pulse of Msn2 translocation after a bPAC pulse. Because the parameters for the interactions that are captured in the model have not been measured independently, we simulated both open- and closed-loop models for >107 log-uniformly sampled parameter sets. We then selected the 5.5×10^4 parameter sets that minimized the mean square error between the model output and the wild-type experimental data of Figure 2.2B and optimized these sets by the Nelder–Mead algorithm to improve the fit[31]. We found 1.4×10^4 parameter sets for the negative feedback model that generated a good fit to the data (Figure 2.2B). At the same time, we could not identify any parameters for the open-loop model that could recapitulate qualitatively or quantitatively the transient pulse of Msn2 (Supplementary Figure 2.7). Instead, for all parameters sampled, the open-loop model produced an overdamped Msn2 nuclear residence and depletion profile that decreased upon bPAC activation and monotonically increased to prestimulus steady state upon bPAC shutoff.

Examining the parameter sets for the feedback model that could recapitulate the Msn2 pulse data revealed a common feature they all shared: a feedback-induced delay for the

activation of the endogenous adenylate cyclase after bPAC shutoff.

In simple terms, bPAC activation injects a large concentration of cAMP into the system with two consequences. First, the endogenous cAMP production is repressed by the PKA feedback on Ras1/2 activity. In addition, as cAMP concentration increases, the PDE-mediated degradation of cAMP also increases due to mass action effects. When bPAC is shut off, this cAMP degradation initially continues at a high rate. This degradation is not immediately counteracted by the endogenous production of cAMP, as Ras1/2 activity is delayed by the engaged PKA feedback. As a result of such imbalance between degradation and production dynamics, cAMP levels transiently drop below bPAC preinduction levels (Figure 2.2B). The corresponding drop in PKA activity consequently generates a pulse of Msn2 nuclear translocation. Msn2 localization reaches its cytoplasmic steady state when the degradation and production of cAMP equilibrate. Supporting the idea that the pulse of Msn2 nuclear localization is a result of feedback-induced delay, we observed in the model that the pulse occurs precisely when one of the components in the feedback loop traversing Ras1/2 to PKA is rate limiting (Supplementary Figure 2.8). Moreover, the lack of a cAMP undershoot after bPAC shutoff for the open-loop model is in agreement with the idea that the feedback-induced delay of cAMP production produces the cAMP undershoot and subsequent Msn2 pulse. Taken together, our data suggest that the observed pulse after bPAC shutoff is likely to be a structural feature of the negative feedback surrounding PKA.

2.3.3 Quantitative features of Msn2 nuclear pulse after bPAC shutoff depend on the components of the PKA signaling network

An implication of feedback in generating the transient Msn2 pulse after bPAC shutoff is that the attributes of this pulse, such as its peak height and time to peak, can be

predictably modified by perturbations of different components that impinge on this feedback. To explore this idea, we built bPAC strains in which one of the following three negative regulators of PKA was deleted: Ira2, a GTPase-activating protein (GAP) that inhibits Ras1/2 and hence decreases PKA activity; and low- and high-affinity PDEs Pde1 and Pde2, respectively, which also decrease PKA activity by degrading cAMP. We then delivered a pulse of blue light input to these strains and measured the time dynamics of Msn2 nuclear localization. Deletion of Ira2, Pde1, and Pde2 generated qualitative and quantitative differences in the system's response to the bPAC input. Compared to WT, Δira2 showed no transient Msn2 pulse upon bPAC shutoff, whereas both Δpde1 and Δpde2 displayed a delayed and attenuated Msn2 pulse (Figure 2.3 and Supplementary Figure 2.9).

To capitalize on the model as a guide for our intuition in understanding these differences, we used the parameter sets of the feedback model that fit the wild-type (WT) data as seeds to perform a second stage of fitting to experimental data from all four strains (WT and three mutants; Supplementary Note S1). Of the original data set, only $\sim$300 parameter sets from the original ensemble produced a good fit to all experimental data, suggesting that the mutant data effectively constrained the model parameters. For any one mutant, simulated trajectories of nonfitted variables (e.g., active Ras) were quantitatively different based on the particular parameter set used. However, these trajectories had common qualitative features that provided plausible explanations for the particular Msn2 phenotype observed in the various mutants.

For Δira2, two nonexclusive parameter categories were at the root of the suppression of the Msn2 translocation pulse (Figure 2.3A and Supplementary Figure 2.10). In the first category, accumulation of Ras in the model due to the absence of the negative regulator IRA2 caused a large increase in cAMP concentration both at steady state and during bPAC application (Figure 2.3A, left). After bPAC shutoff, the degradation of cAMP still proceeds at a fast rate and may cause cAMP levels to transiently drop below their

(elevated) steady-state value. However, for the particular parameter sets corresponding to this mutant, even the lowest cAMP level achieved is still sufficient to keep PKA activity above the threshold that needs to be traversed in order for Msn2 to translocate to the nucleus (Supplementary Figure 2.11). In the second category of parameters also reproducing the Δira2 data, the lack of an Msn2 pulse resulted from altered feedback. Here the pulse of Msn2 nuclear translocation was abrogated by the fact that, due to higher Ras activity in this mutant, cAMP degradation was more quickly balanced by endogenous cAMP production after bPAC shutoff (Figure 2.3A, right).

The model also clarified the phenotypes of Δpde1 and Δpde2, which both show a delayed and attenuated Msn2 nuclear translocation pulse. In both mutants, degradation of cAMP is reduced, resulting in rapid accumulation of cAMP to levels exceeding those of the wild type upon bPAC stimulation. After bPAC shutoff, impaired degradation of this excess cAMP keeps PKA activity high for an extended period of time, therefore maintaining Msn2 cytoplasmic localization during that time and explaining the delay. Levels of cAMP eventually decline below a level capable of keeping PKA active, and a pulse of Msn2 nuclear translocation ensues. The presence of the pulse is still a result of the feedback-induced delay in endogenous cAMP production, but its attenuated magnitude is a manifestation of the combined effect of a slower cAMP degradation (slower rise time for the pulse) and a greater ability of cAMP production to balance this degradation (Figure 2.3B). The different quantitative phenotypes of Δpde1 versus Δpde2 are ascribed by the model to different affinity of these enzymes to cAMP, resulting in a differential effect of their deletion on cAMP degradation rate. This differential binding of Pde1 and Pde2 to cAMP is well documented experimentally[18,32]. In addition, the model indicates that at steady state, cAMP levels for these mutants is similar to that of the wild type, a nonintuitive result of the feedback ensuring that decreased cAMP degradation is homeostatically counteracted by decreased Ras.

The observation that the dynamics of the Msn2 nuclear pulse is profoundly affected by perturbations of the PKA network, supported by our modeling results, lends further support to the notion that this pulse is predominantly generated by the PKA signaling network itself.

2.3.4 bPAC enables frequency- and amplitude-modulated control of the PKA signaling network

To further explore the feedback model with the constrained parameters, we subjected it to light pulses of increasing amplitudes and durations, inputs for which the model was not fit (Figure 2.4A). With increasing light duration at a fixed intensity, the model predicted a corresponding graded increase in peak nuclear localization of Msn2 (Figure 2.4B). A similar conclusion holds for increasing light intensity for a constant duration (Figures 4B and Supplementary Figure 2.12). This is because in this model, the cAMP response is dependent on the integral of the total light pulse within this intensity/duration regime (Figure 2.4B, bottom). Hence both light stimulus amplitude and duration can achieve a continuous range of values for the peak of the Msn2 pulse. Experiments performed with bPAC pulses of different intensities and durations agree with this result (Figure 2.4A), supporting the basic model structure.

Finally, we asked whether the model can recapitulate the filtering properties of the PKA system, which can be determined by assessing the system's output to different bPAC input pulse trains of different frequencies (Toettcher et al., 2013 blue right-pointing triangle). Such a frequency response analysis provides a quantitative picture of the dominant time scales of PKA signaling. We therefore used the optogenetic bPAC input to apply light pulses at six distinct frequencies spanning 0.83 to 0.11 min-1, with five repeats of each frequency (except for 0.11 min-1, which only had two repeats). Experimental data revealed

a low-pass filter with a cutoff frequency of 0.33 min-1 and quality factor of 0.88 (Figure 2.4C and Supplementary Movie S3), suggesting that the response time scale of the PKA system is $< 1/0.33 \approx 3$ min and that inputs with substantially higher frequency content are significantly attenuated.

We then stimulated the computational model with a bPAC input consisting of a train of pulses with decreasing frequencies and compared these simulations to the experimental results (Figure 2.4D, top). We repeated this exercise for all of the parameter sets that recapitulated the data in previous sections. The maximum magnitude of Msn2 nuclear localization varied among parameter sets. However, all parameter sets consistently predicted that the system has the characteristics of a second-order low-pass filter with a median cutoff frequency of 0.14 min-1 with interquartile range of 0.12–0.15 min-1 and a quality factor of 0.86 with interquartile range of 0.77–0.93 (see Supplementary Note S1 for the conventions used in the transfer function and Supplementary Figure 2.13 for the phase response). The cutoff frequency arises because at high pulsing frequencies, shutoff of bPAC is now closely followed by another bPAC activation pulse, such that cAMP concentration now cannot fall below the preinduction level before the next light pulse is applied (Figure 2.4D, bottom).

The quantitative difference between the computational and experimental cutoff frequency is not surprising. Ultimately, our model was built to be the most parsimonious representation that captures the phenomenology and overall characteristics of the system and trained on a limited set of data. Specifically, since the model was trained on 3-min light pulses, it cannot predict accurately the rates of cAMP production/degradation necessary to fit shorter light pulses and hence cannot quantitatively fit data obtained at high input frequencies. Therefore, while this model reveals the main dynamic characteristics of the PKA signaling network, featuring delayed feedback, more detailed representations and experimentation are needed to accurately capture the true quantitative complexity of the PKA system.

2.4 Discussion

Growth-regulatory signaling pathways such as the PKA pathway link perception of the environment with cell proliferation. To reliably relay the environmental state, these pathways often show rapid dynamics. Rapid dynamics of PKA was previously documented in response to environmental perturbations[9,25]. Fast and minimally pleiotropic perturbation tools are therefore required to explore these dynamics and evaluate their functional roles. Optogenetic perturbations are increasingly identified as powerful tools to carry out these studies. Although optogenetic technology has seen widespread use in neuroscience, these tools were only recently introduced in the study of a wider breadth of molecular biology. In this work, we capitalize on a naturally occurring bPAC to study PKA signaling in budding yeast. We show that when coupled with real-time reporters and computer-controlled illumination, bPAC constitutes a powerful general tool for administering precise and specific perturbations to this system to probe its quantitative properties. Using this tool, we were able to study the system-level characteristics of the PKA signaling pathway.

Specifically, our studies using bPAC indicated that strong negative feedback channeled through Ras1/2 in the PKA signaling network in budding yeast causes a delay in Ras1/2-mediated endogenous production of cAMP, and we pinpointed its relevant timescale. Using these data, we were able to build a quantitative computational model that generated a rigorous predictive understanding of the role of feedback in generating PKA activity dynamics, which may be important in regulating downstream signaling and gene expression.

Using a combination of precise optogenetic perturbation and quantitative modeling, this study enabled identification of dynamic properties of the PKA signaling network that would otherwise be difficult to dissect using slower and more pleiotropic methods such as

knockouts and overexpression. More generally, our analyses revealed two salient principles. First, the presence of feedback in a system generates nonintuitive dynamic effects upon perturbation of components both within (such as Ira2) and outside the feedback loop (such as Pde1/2). Second, different regulatory components with the same overall qualitative description (e.g., both Pde1/2 and Ira2 negatively affect PKA activity) can generate vastly different qualitative and quantitative phenotypes.

Overall it is tempting to hypothesize that these regulators could be used as gateways for different environmental inputs into this system. In this scheme, different inputs can affect and perturb different regulators. Because perturbation of such regulators induces distinct dynamical phenotypes, this scheme could be the basis for encoding the identity of inputs into different dynamical patterns of the pathway. Dynamic encoding of inputs has been proposed as a strategy for implementing response specificity in a signaling pathway that propagates multiple environmental inputs[9]. Although current research has focused on identifying specific molecular implementations of such a strategy, in the future, it will be interesting to investigate the dynamically malleable platform of feedback as a possible basis for generating dynamic encoding in signaling pathways.

2.5 Figures

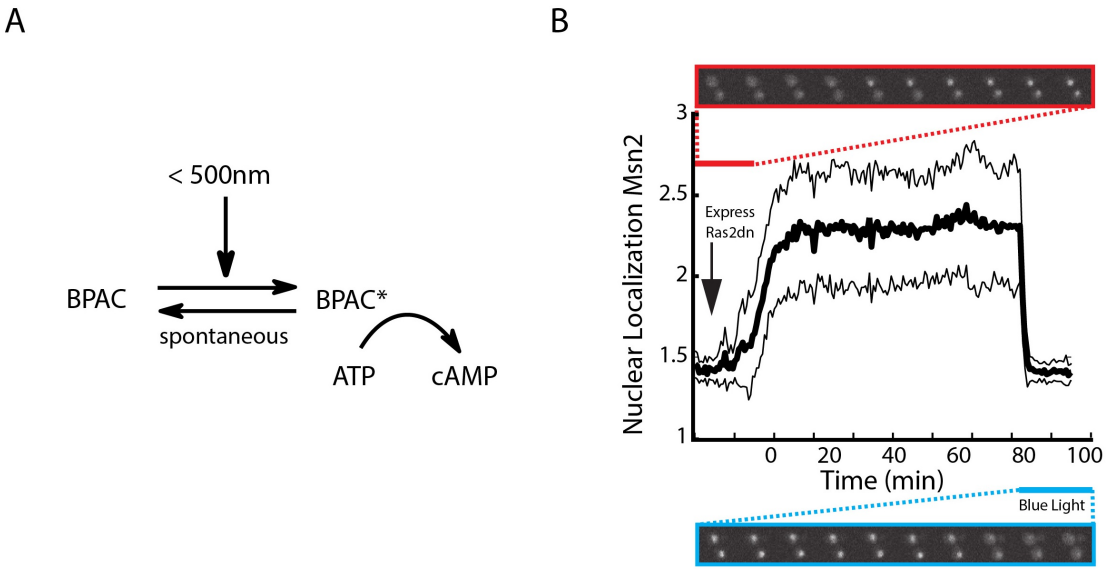


Figure 2.1: **Expression of a bPAC in budding yeast allows for real-time light-gated control of PKA signaling** Expression of a bPAC in budding yeast allows for real-time light-gated control of PKA signaling. A) The bPAC protein switches from an inactive to an active conformation in response to light. B) Nuclear localization of Msn2 after inhibition of the PKA pathway by expression of a dominant-negative Ras2 allele (S24N; red inset) and subsequent blue light activation of bPAC expressed in the same strain (blue inset). Activation of the dominant-negative allele results in rapid nuclear localization of Msn2 into the nucleus, and this is synchronously reversed by bPAC activation. Nuclear localization is defined as the ratio of nuclear to cytoplasmic Msn2 (see also Materials and Methods). Time trace (thick black line) shows average of 65 cells with SDs (thin black lines).

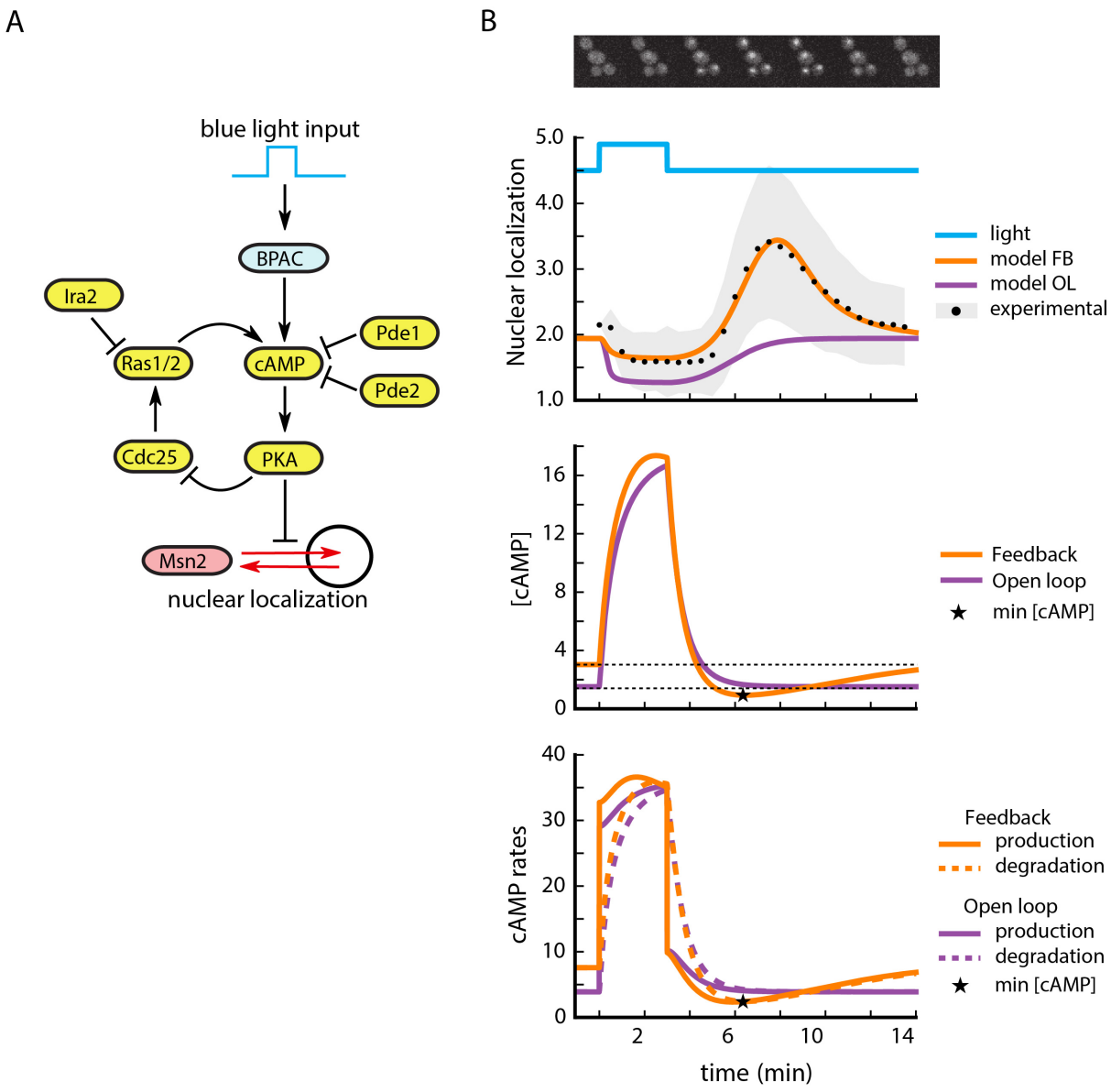

A
blue light input
BPAC
Ira2
Ras1/2
cAMP
Pde1
Pde2
Cdc25
PKA
Msn2
nuclear localization
B
Nuclear localization
5.0
4.0
3.0
2.0
1.0
light
model FB
model OL
experimental
[cAMP]
16
12
8
4
0
Feedback
Open loop
min [cAMP]
cAMP rates
40
30
20
10
0
Feedback
production
degradation
Open loop
production
degradation
min [cAMP]
2
6
10
14
time (min)

Figure 2.2: **A computational model of the cAMP-PKA circuit explains the response to bPAC stimulation** A) Diagram of the components involved in the model and their interactions. Included are production of cAMP through bPAC stimulation and endogenous cyclase activity, Msn2 nuclear localization, autoregulatory negative feedback on PKA through Cdc25 and Ras2, and the effect of Pde1 and Pde2 on cAMP degradation. B) Top, nuclear localization of Msn2 as a function of time (black dots are the mean from a population of 379 cells; shaded gray error bars indicate SD) after a single pulse of blue light (40 µW/$mm^2$2). The orange trace is the output of the computational model containing a negative feedback loop for a single representative parameter set. The purple trace is a representative output of a model not containing the feedback. Middle, concentration of cAMP as a function of time for the model with (orange) or without negative feedback (purple). The negative feedback model predicts a cAMP undershoot (minimum value denoted by a star), whereas the open loop model (purple) monotonically approaches the steady-state value (black dotted). The model undershoot is more pronounced than the experimentally observed undershoot because of the detection limit of the experiment. Bottom, cAMP production and degradation rates as a function of time. The undershoot in cAMP concentration is generated by a delayed production of cAMP. The cAMP minimum (star) is reached when the rate of cAMP production is balanced by the rate of degradation.

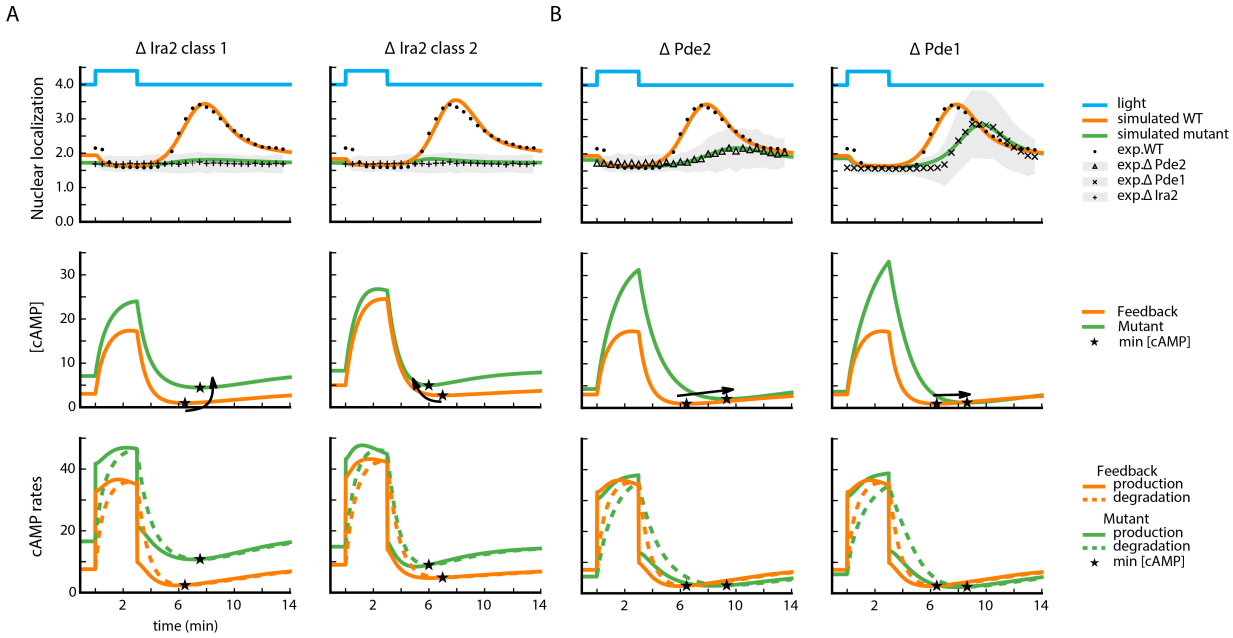


Figure 2.3: **Computational modeling is used to dissect the behavior of the Δira2, Δpde1, and Δpde2 mutants** A) Model predicted nuclear localization of Msn2 (top), concentration of cAMP (middle), and cAMP production/degradation rates (bottom) as a function of time for Δira2. The plots are shown for two representative parameter sets, one of each of the class 1 and class 2 parameter regimes that could explain this mutant. Top, experimental nuclear localization in response to 40-µW/$mm^2$2 blue light input; symbols with error bars in gray. Arrows indicate the change in minimum cAMP concentration (star) due to the deletion. B) Same plots as in A, for the Δpde1 and Δpde2 mutants.

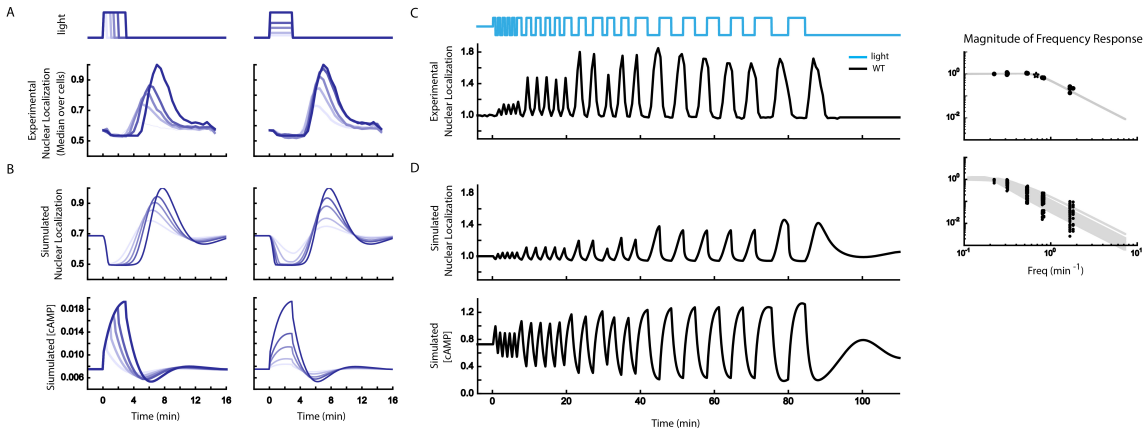


Figure 2.4: **High-precision control of cAMP using bPAC allows for systematic characterization of the PKA system using pulse width–, amplitude-, or frequency-modulated input signals** A) Peak nuclear localization of Msn2-mCherry increases with blue light duration (left) and amplitude (right) in cells expressing bPAC. Experimental WT data collected at 30-s intervals for light durations of 0.5, 1, 1.5, 2, and 3 min with amplitude fixed at 40 µW/mm^2 (left) and light amplitudes of 8, 10, 16, 20, and 40 µW/$mm^2$2 with duration fixed at 3 min (right). (See also Supplementary Figure 2.12.) Plots show median values over ∼900 cells normalized to the maximum observed median value for comparison of responses across different light doses. B) Top, the model predicts that nuclear localization shows a trend similar to the experimental data of A when light amplitude or duration increases (see Supplementary Figure 2.11). Data normalized by maximum value. Bottom row, maximum simulated cAMP concentration increases with both light duration and amplitude. C) Left, mean Msn2-mCherry nuclear localization after blue light illumination at different pulse frequencies. Right, peak mean Msn2 nuclear localization at each frequency (black dots) with fit to second-order transfer function (gray lines). D) Left, model prediction of the frequency sweep experiment in C for a representative set of parameters (parameter set 17 in Supplementary Data Set S3). Right, peak Msn2 nuclear localization at each frequency for representative parameter sets (black dots), with fits to second-order transfer function (gray lines).

2.6 Materials and Methods

2.6.1 Yeast strains and constructs

All yeast strains used for these experiments are derived from W303A-1 in which the ade2 marker was reverted to ADE2+ to reduce the autofluorescence. Msn2-mCherry was integrated at the trp1 locus, and the bPAC construct was integrated into the leu2 locus. Overexpression construct for RAS2(S24N) was integrated into the TRP1 locus of a LEU2+ MATalpha strain that contained an estradiol-inducible construct. All strains were constructed using standard yeast protocols and LioAc/PEG transformation.

The RAS2(S24N) expression vector was constructed by amplification of the RAS2 gene from yeast genomic DNA and cloned downstream of a prGAL1 promoter in a TRP1-marked single integration plasmid. Site-directed mutagenesis was performed to create the dominant-negative allele of RAS2(S24N) using a standard QuikChange protocol and the pfuTurbo enzyme mix (Stratagene, La Jolla, CA). The bPAC gene was synthesized with yeast-optimized codons by Integrated DNA Technologies (Coralville, IA) and cloned downstream of a prNOP7 promoter in the LEU2-marked single integration vector.

2.6.2 Microscopy, image acquisition, and analysis

Cells expressing Msn2-mCherry or related constructs were plated in SD complete medium onto concanavalin A–coated 96-well glass-bottom plates, allowed to settle, and then washed twice with fresh medium and 100 µl of fresh medium added.

Samples were imaged on a Nikon Ti inverted scope with arc-lamp illumination using red fluorescent protein (560/40 nm excitation, 630/75 nm emission; Chroma, Bellows Falls, VT) and yellow fluorescent protein (510/10 nm excitation, 542/27 nm emission; Semrock,

Rochester, NY) filters. Blue light illumination was provided by a 465-nm LED driven by a USB-controlled power source (MIGHTex, Pleasanton, CA) mounted on the bright-field condenser. Imaging and illumination were controlled and coordinated by custom Matlab (MathWorks, Natick, MA) software interfaced with the µmanager software suite (Edelstein et al., 2010 blue right-pointing triangle).

Images were processed and analyzed with ImageJ and custom-built Matlab scripts. Nuclear localization was computed by dividing the average intensity of the brightest 10% of pixels in the cell by the median intensity of the cell.

2.6.3 Deterministic model

An ordinary differential equation (ODE) model of the PKA regulatory network, consisting of mostly Michaelis–Menten interactions, was constructed with five state variables and 24 parameters. Latin hypercube sampling and Nelder–Mead optimization of parameters[31] were done to obtain model fits. Additional details of the computational methods are described in Supplementary Note S1.

2.7 Supplementary Material

2.7.1 Note 1

This note provides an overview of our computational methods.

Negative feedback model The main interactions in the PKA regulatory pathways are depicted in Figure 2.3A. We model these interactions by

$$
\begin{aligned}
t\,\mathrm{PKAa} &= p_9\,\mathrm{cAMP}(\mathrm{PKAt}-\mathrm{PKAa}) && - p_{10}\,\mathrm{PKAa} && (2.1a)\\
t\,\mathrm{cAMP} &= p_{23}u(t) + p_{11}\,\mathrm{RASa} + p_{10}\,\mathrm{PKAa} - \frac{p_{12}\,\mathrm{cAMP}}{\mathrm{cAMP}+p_{13}} - \frac{p_{14}\,\mathrm{cAMP}}{\mathrm{cAMP}+p_{15}p_{13}} && \\
&\quad - p_9\,\mathrm{cAMP}(\mathrm{PKAt}-\mathrm{PKAa}) && && (2.1b)\\
t\,\mathrm{RASa} &= p_{16}\frac{(\mathrm{RASt}-\mathrm{RASa})\,\mathrm{CDC25}}{\mathrm{RASt}-\mathrm{RASa}+p_{17}} && - p_{18}\frac{\mathrm{RASa}}{\mathrm{RASa}+p_{19}} && (2.1c)\\
t\,\mathrm{CDC25} &= p_{22} && - p_{20}\frac{\mathrm{PKAa}\,\mathrm{CDC25}}{\mathrm{CDC25}+p_{21}} && (2.1d)\\
t\,\mathrm{MSNn} &= p_2\frac{\mathrm{MSNt}-\mathrm{MSNn}}{\mathrm{MSNt}-\mathrm{MSNn}+p_3} && - p_4\frac{\mathrm{PKAa}\,\mathrm{MSNn}}{\mathrm{MSNn}+p_5} && (2.1e)\\
\mathrm{output} &= p_1\frac{\mathrm{MSNn}+p_0}{\mathrm{MSNt}-\mathrm{MSNn}+p_0}. && && (2.1f)
\end{aligned}
$$

In these equations, PKAa and RASa represent protein activities, cAMP and CDC25 represent the concentration of the small molecule and protein respectively, MSNn is the concentration of nuclear Msn2, and $u(t)$ represents blue light intensity. The model is characterized by 24 parameters, including the 21 shown above as well as three parameters representing total protein concentrations: $p_6 = \mathrm{MSNt}, p_7 = \mathrm{RASt}, p_8 = \mathrm{PKAt}$.

Most of activation/inactivation and production/degradation terms are modeled by Michaelis-Menten kinetics, including the kinetics of Msn2 shuttling in/out of the nucleus. The interaction of PKA with cAMP, however, is modeled as a binding reaction. Inactive

PKA is known to be a heterotetramer composed of a dimer of catalytic subunits and a dimer of regulatory subunits. cAMP is thought to activate the catalytic subunits by binding to and releasing the regulatory subunits. Here we assume that 2 cAMP molecules bind simultaneously and non-cooperatively to 2 regulatory subunits to release 2 catalytic subunits. Hence there is a 1:1 relationship between active catalytic subunits and cAMP molecules. Further, cAMP production is explicitly modeled as having two components dependent on Ras and blue light u.

Finally the output represents nuclear localization of Msn2, defined as the ratio of nuclear Msn2 to cytoplasmic Msn2, and is modeled as a rational function of Msn2n as in equation (0.1f). Here p_1 allows the output to be scaled version of Msn2 while p_0 accounts for background intensity.

The equations were numerically solved by the LSODE solver for stiff differential equations via the python package odespy.

Open loop model The open loop model replaces equation (0.1d) with

$$t\,\mathrm{CDC25} \;=\; p_{22} - p_{20}\frac{\mathrm{CDC25}}{\mathrm{CDC25} + p_{21}} \tag{0.1d*}$$

Fitting Fits of experimental data to the feedback model (Figures 2 and 3; Supplementary Figure 2.11) and the open loop model (Figure 2.2; Supplementary Figure 2.3) were obtained by the three step procedure:

1. Sample parameters log uniformly with the requirements:
   (a) Rates/concentrations are sampled between 10^{-4} and 10^{4}.
   (b) p_6 through p_8 are sampled between 10^{0} and 10^{4}.

(c) The input and output parameters: p_1 is sampled between 10^{-3} to 10^1; p_1 and p_{23} are sampled between 10^0 to 10^2.

(d) $10^{.5} < p_{15} < 10^4$. This is to ensure that the Pde1 and Pde2 terms can be distinguished and have the correct relative affinities.

For both models, $>10^7$ sets were sampled.

2. Use the result of step 1 as seed values for Nelder-Mead optimization to fit to the experimental wild type data. The objective function is the mean square error weighted by the standard deviations at each time point.

3. Use the result of step 2 as seed values for Nelder-Mead optimization to fit to all 4 experimental strains. The mutants are modeled by

   (a) ΔIra2: $p_{18} \leftarrow p_{18}/4$

   (b) ΔPde2: $p_{12} \leftarrow 0$

   (c) ΔPde1: $p_{14} \leftarrow 0$

   The objective function is the mean square error weighted by standard deviations, averaged over the 4 data sets, plus a penalty to encourage desired behavior. The penalty is a heuristic that takes a weighted average of three terms,

   (a) a term dependent on cAMP, to prevent the minimum cAMP concentration from reaching zero,

   (b) a term dependent on cAMP$t|_{t=3^-}$, to encourage the system to equilibrate by 3 minutes, and

   (c) a function of the average standard deviation of all species concentration between 30 and 35 minutes, to penalize models that do not reach steady state 30 min after bPAC induction.

The equations were numerically solved by the LSODE solver for stiff differential equations via the python package odespy.

Any fits outside the parameter ranges established in step 1 were discarded.

From this procedure, we found approximately 300 parameter sets for the feedback model that fit all the experimental data.

Frequency sweep The response of the wildtype system to an input with varying period was simulated with the 300 best parameters sets obtained from our three-step fitting procedure described above. The input to the system was identical to that used in experiments, as shown in Figure 2.4. The transfer function was found by identifying the peak Msn2 nuclear localization (measured with respect to steady state Msn2 nuclear localization) and then plotting the peaks as a function of frequency. This computationally derived transfer function was then fit to the standard transfer function for first and second order linear systems,

$$H_1(f) = \frac{C}{\sqrt{1 + (f/f_{c1})^2}} \tag{2.2a}$$

$$H_2(f) = \frac{C}{\sqrt{(1 - (f/f_{c2})^2)^2 + (f/f_{c2}Q)^2}}, \tag{2.2b}$$

where f is frequency and C, f_c, and Q are fitted parameters. The second order transfer function provided a better fit and produced the values for median cutoff frequency with IQR reported in the main text. Fitting was performed by the curvefit function of the scipy package.

The simulated phase response was determined by computing the lag of the output peak with respect to the input peak, as a proportion of the input period. This was then scaled appropriately and plotted against the input frequency.

2.8 SI Figures

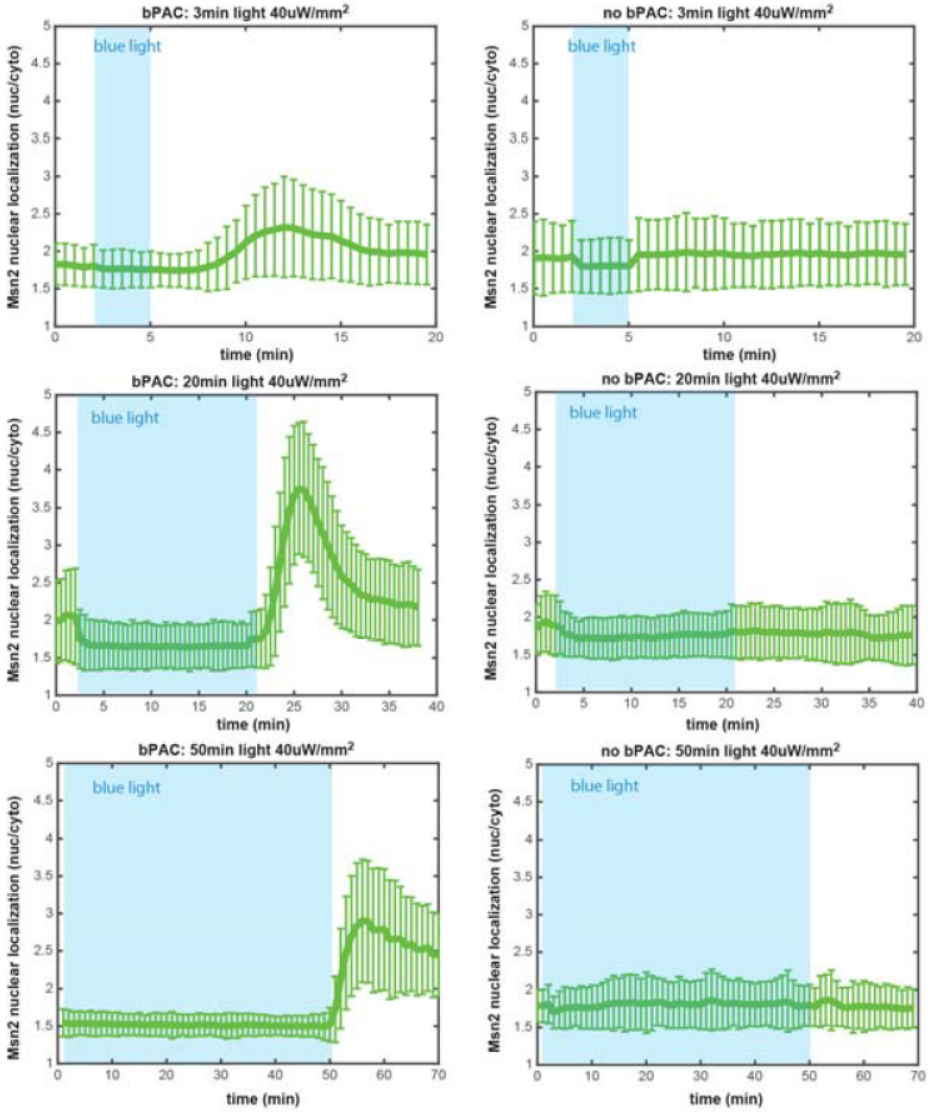


Figure 2.5: **Extended blue light and fluorescent imaging exposure do not induce nuclear localization of Msn2** Time traces of Msn2-mCherry nuclear localization are plotted in green for cells with bPAC (left 3 panels) and cells without bPAC (right 3 panels). The light blue shading indicates the time over which light is applied in the experiment. Blue light exposure span three different durations – 3 minutes, 20 minutes, and 50 minutes – at 40uW/$mm^2$2 constant amplitude.

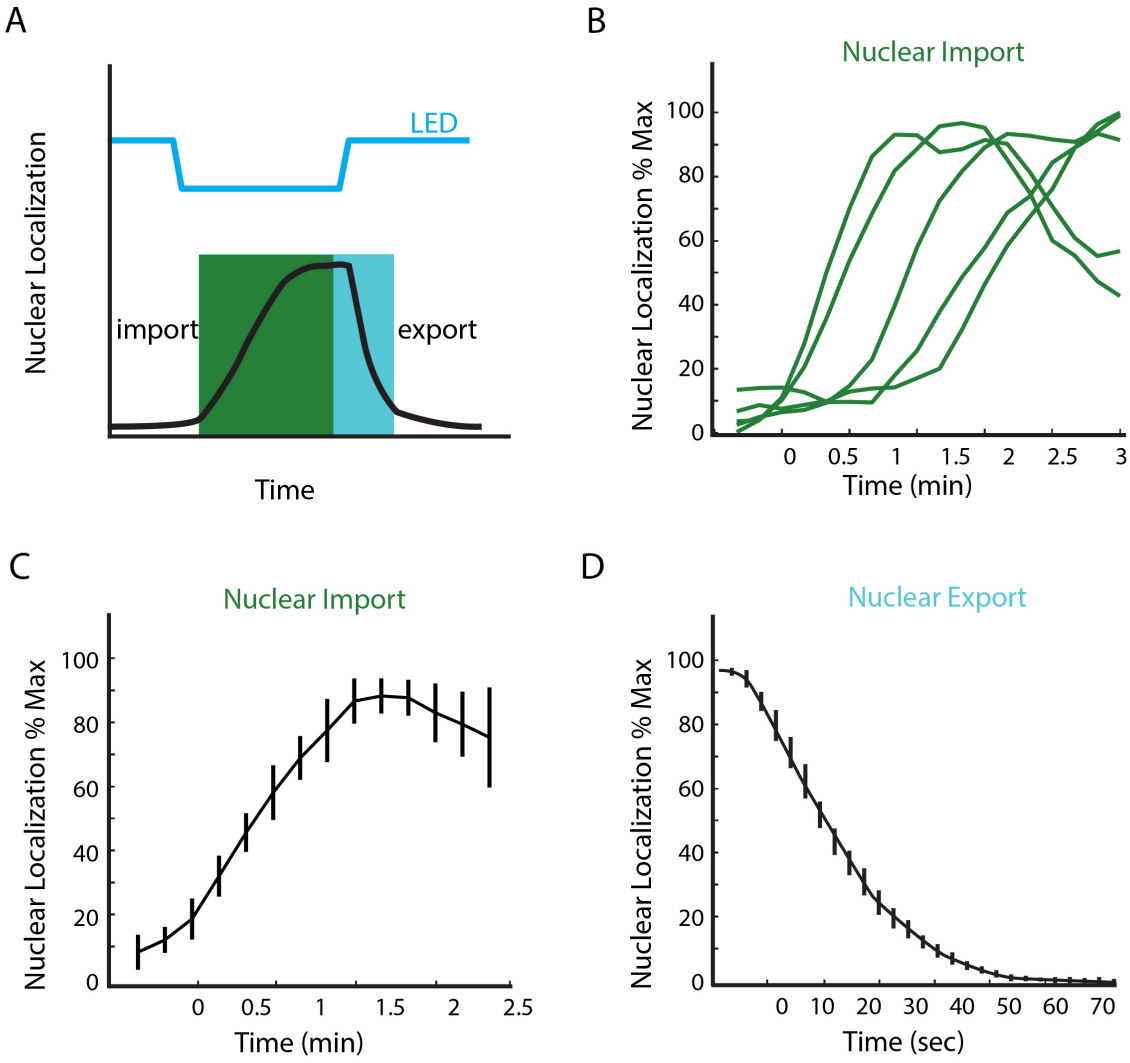


Figure 2.6: **Real-time control of PKA activity with bPAC allows for quantification of Msn2 nuclear import and export rates** A) Experimental blue light illumination protocol and expected Msn2 translocation pulse. The green shading represents the range considered as Msn2 nuclear import B, C) and the light blue shading represents the range considered as Msn2 nuclear export (D). B,C) Single cell traces of Msn2-mCherry nuclear import following shutoff of light pulse. When computationally synchronized as in C), these traces can be used to estimate an average nuclear import rate of T1/2 =30s. Error bars represent the standard deviation of three independent experiments. D) Nuclear localization of Msn2-mCherry following reactivation of blue light. Nuclear export is estimated to have T1/2 of 20s. Nuclear export was more synchronous than nuclear import, and traces can therefore be immediately used for these export rate calculations.

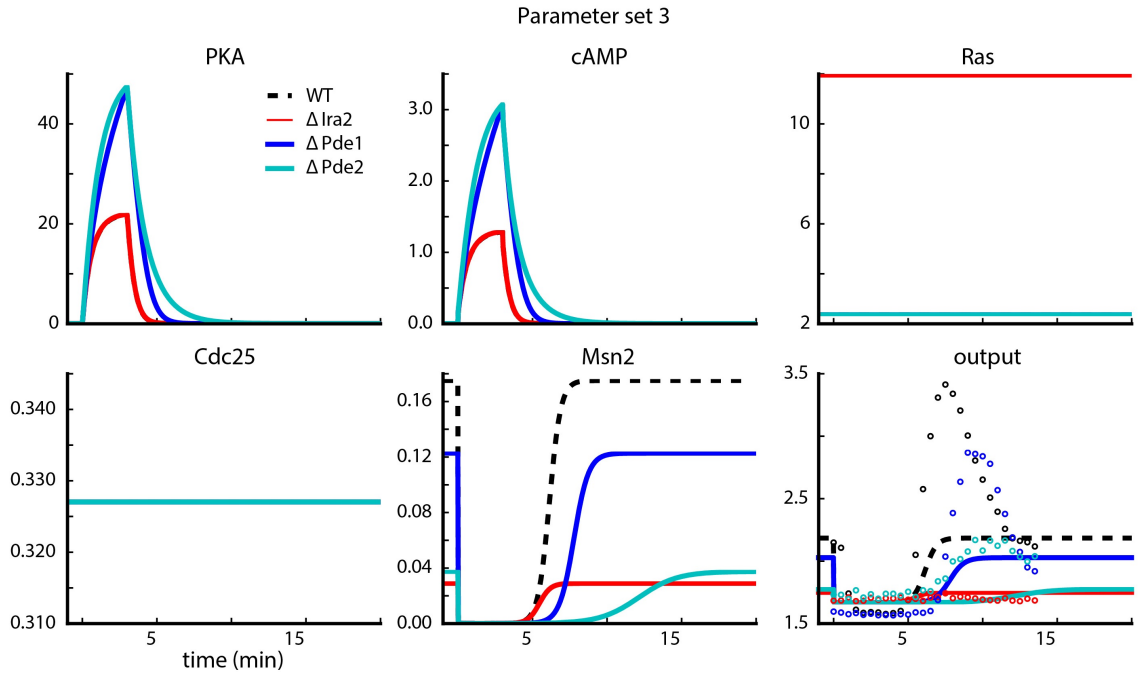


Figure 2.7: **An open loop model does not recapitulate the transient pulse of Msn2 nuclear localization in response to a 3 minute transient blue light input starting at time t = 2 minutes** The open loop model was constructed by removing the Cdc25-PKA interaction and fitting to WT data. State variables (arbitrary units) for a representative fit are shown (solid lines). The different solid line represent different strain background. Experimental data of Msn2 nuclear localization (output) is shown in circles. Other fitted parameter sets are available in Supplementary Dataset 2.

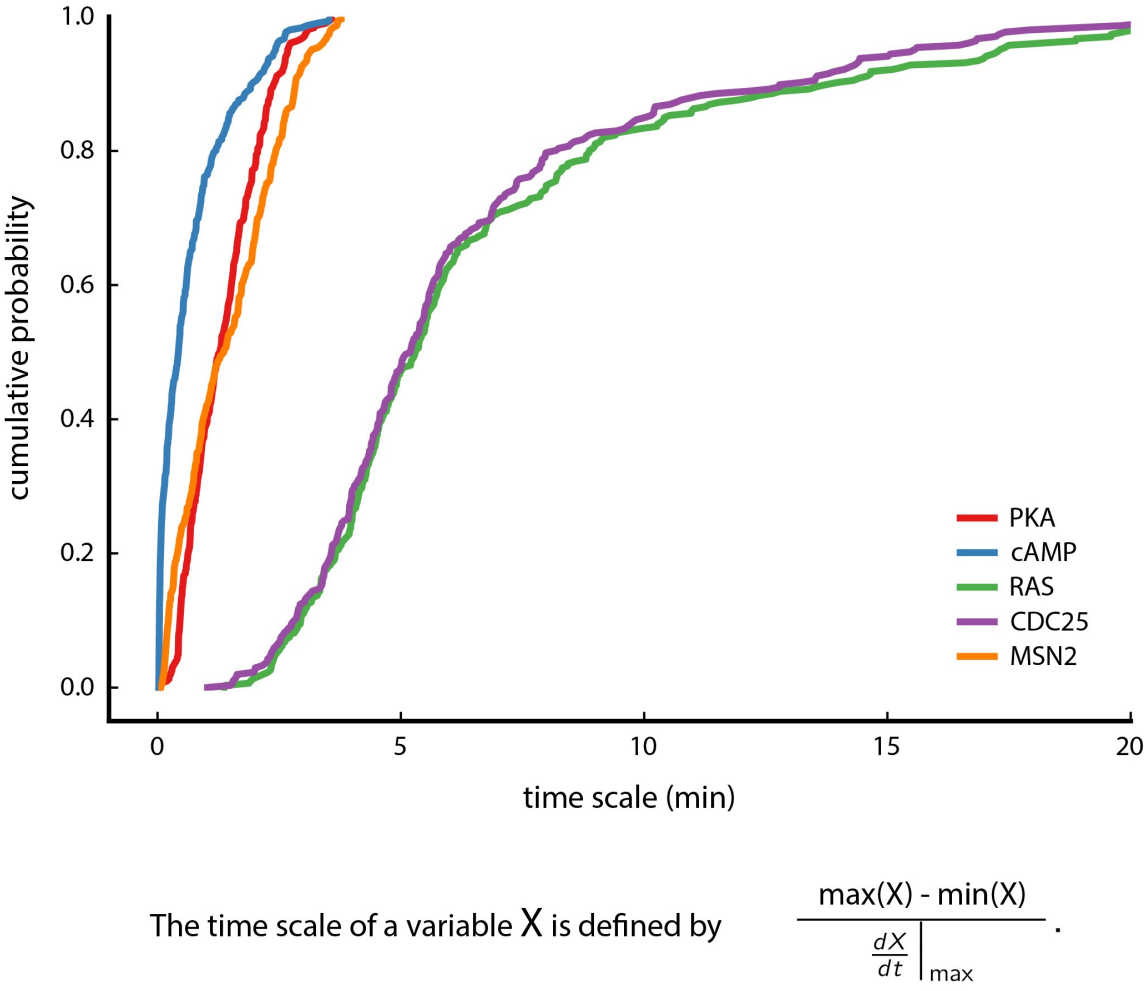
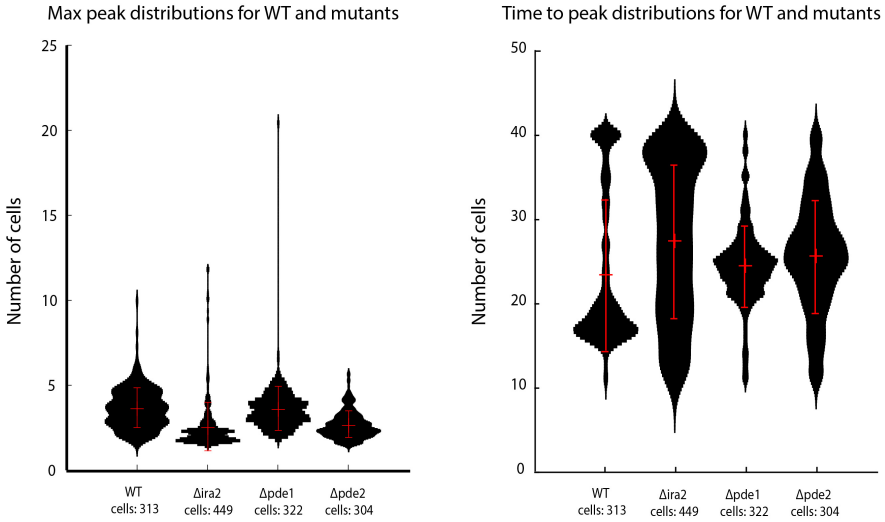


Figure 2.8: **The negative feedback model exhibits a separation of time scales, with PKA, cAMP, and nuclear Msn2 responding faster on average than Cdc25 and Ras** Cumulative distribution functions for the time scale of each variable in the model for each of the parameter sets of Supplementary Dataset 3 obtained by fitting to WT and mutant Msn2 nuclear localization. For each fitted parameter set, each protein's time scale was computed by the formula shown (Segel, 1984).

A



B

	WT	Δira2	Δpde1	Δpde2
Max peak (nuclear localization)	3.55	2.16*	3.43	2.44*
Time to peak (min)	10.5*	14*	12*	12.75*

* Two sample Kolmogorov-Smirnov test, p-value < 0.05

Figure 2.9: **Distributions of time trace features** (maximum peak and the time to peak of Msn2 nuclear localization) show that WT and mutant populations are distinct. A) Single cell distributions of the time trace features, maximum peak and time to peak, are plotted for WT, Δira1, Δira2, Δpde1 and Δpde2 samples with median and the interquartile range shown (red). B) The medians of single cell distributions for the max peak heights and time to peak are shown. The variability in the single cell data is attributed to a combination of natural biological variability to input signal and also technical noise. The two-sample Kolmogorov-Smirnov test was used to determine the significance of the difference in distributions of the WT and the deletions mutants. Asterisk signifies statistically significant ($p<0.05$) differences as compared to WT.

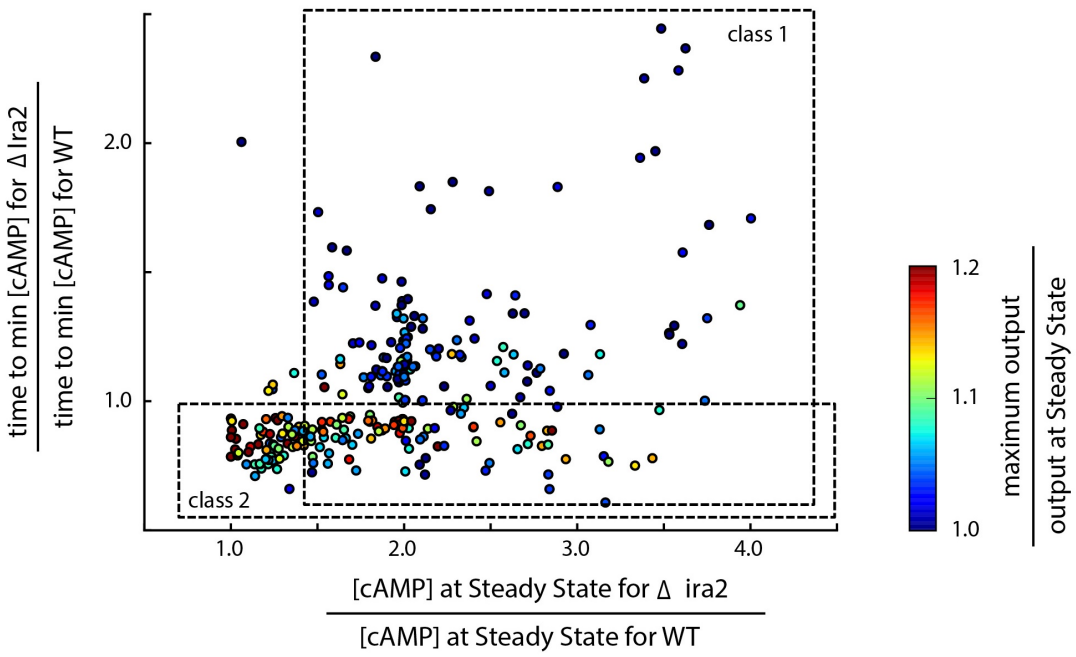


Figure 2.10: **Fitted parameter sets from Supplementary Dataset 3 exhibit two non-exclusive features that explain the observed** Δira2 phenotype. Class 1 is characterized by an increase in cAMP concentration at steady state and during bPAC application whereas class 2 is characterized by a fast attainment of the cAMP minimum. Both classes show a reduced Msn2 pulse for Δira2.

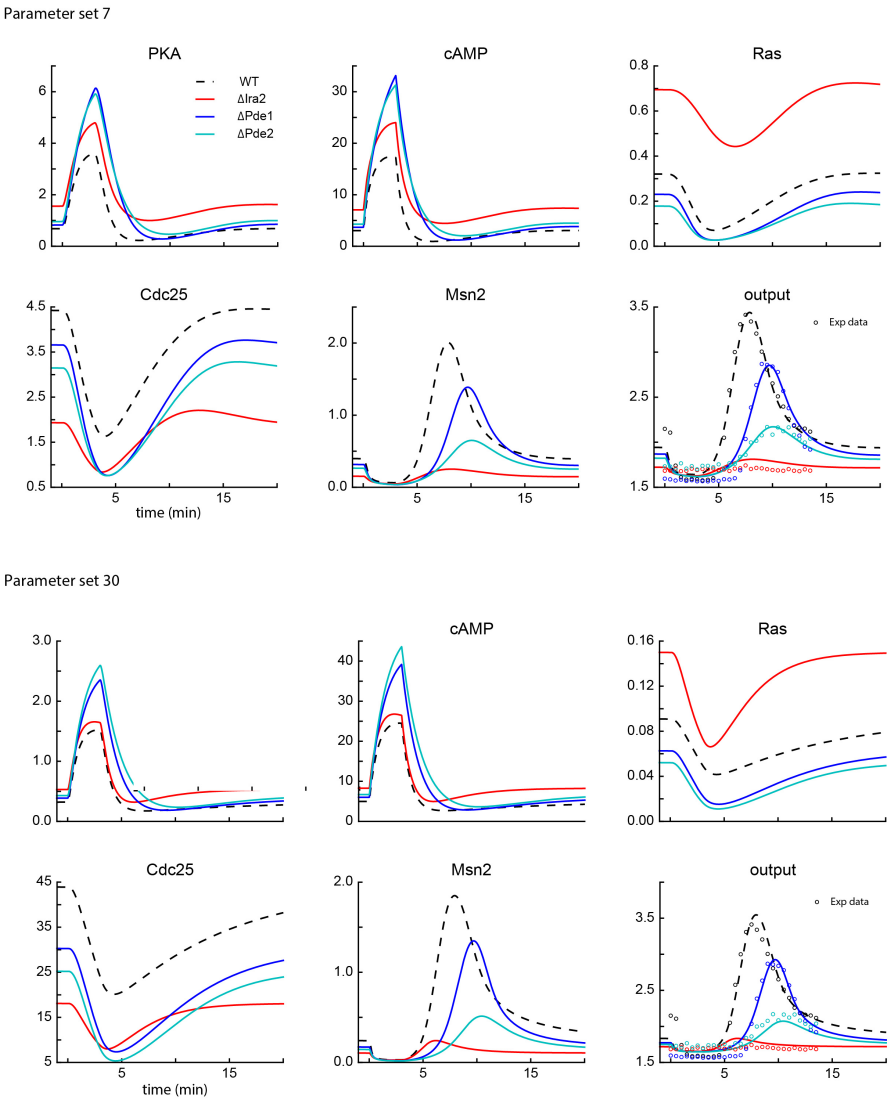


Figure 2.11: **Simulated concentrations (arbitrary units) over time from the negative feedback model in response to a transient 3 minute pulse of blue light starting at time t = 2 minutes** Two parameter sets from Supplementary Dataset 3, obtained by fitting the negative feedback model to WT and mutant Msn2 nuclear localization measured during and after blue light illumination. The two chosen parameter sets are the same as those that generated the data for model plots of Figures 2.3 and 2.4.

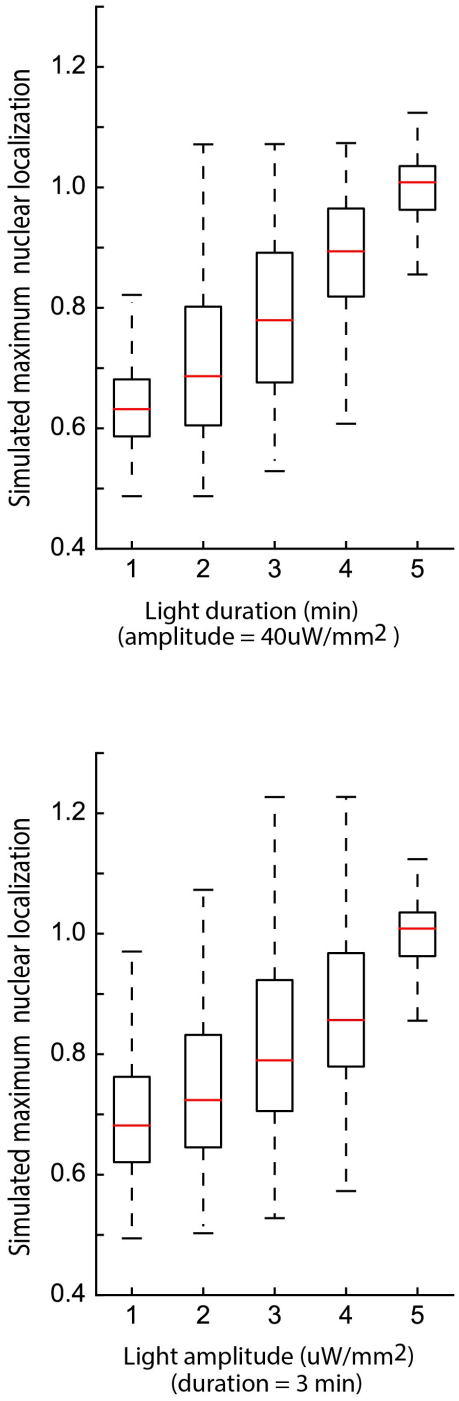


Figure 2.12: **Msn2 peak nuclear localization simulated for the parameter sets of Supplementary Dataset 3 increases with light duration (top) and light amplitude (bottom)** Data normalized so that the median of peak Msn2 in response to a blue light pulse of 3 minutes and 40uW/$mm^2$2 is 1. The distributions are across the different parameter sets of Supplementary Dataset 3; boxes indicate quartiles with median in red.

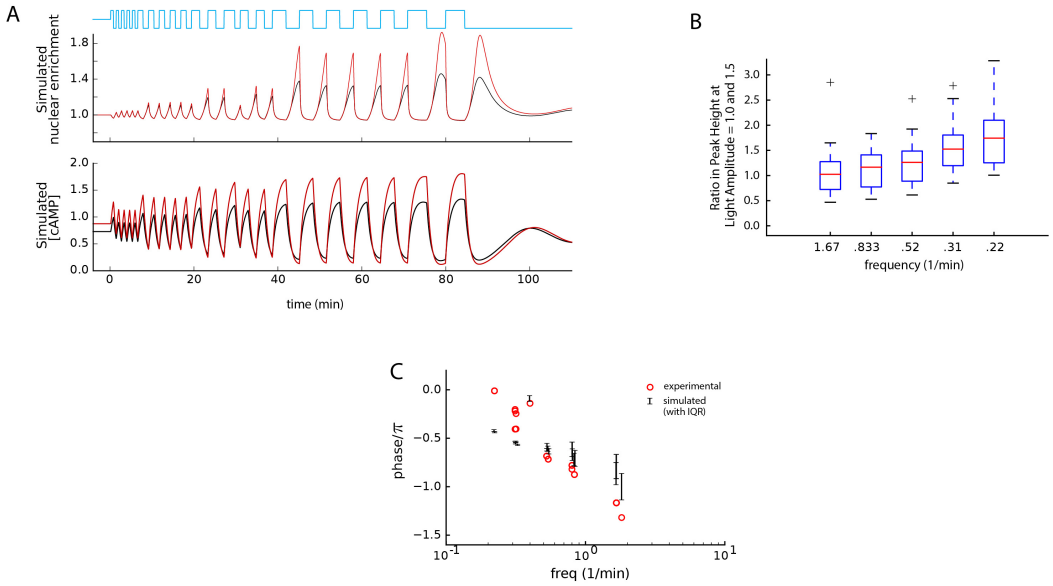


Figure 2.13: **Using bPAC for frequency sweep of Msn2 nuclear localization** A) Simulated Msn2 nuclear enrichment and cAMP concentration in response to a light input with varying frequencies, when light amplitude is increased by a factor of 1.5 (red) relative to baseline (black). If bPAC input saturates, then the model underestimates cAMP production at high frequencies. The effect of this underestimation can be assessed by increasing the light input and observing the change at high frequencies. B) The distribution of the change in peak height, as a function of input frequency, when the light input amplitude increases by a factor of 1.5. For the three highest frequencies, where the underestimation of cAMP production is most likely, the median change is close to one. C) Phase response for simulated and experimental data shows general agreement.

Chapter 3

Optogenetic control reveals differential promoter interpretation of transcription factor nuclear translocation dynamics

3.1 Summary

The idea of using optogenetics to robustly control nuclear localization of transcription factors had been explored using LANS and other natural and small molecule perturbations. The natural and small molecule perturbations were not modular and difficult to handle. And the first generation of optogenetic localization using LANS did not yield appreciable gene expression. The next course of action was to either consider a different type of molecular tool or to improve upon the current optogenetic tool. This chapter details the engineering efforts to produce an optogenetic tool that can not only move proteins to and

from the nucleus but also induce gene expression to allow study of promoter responses to dynamic transcription factor localization.

3.2 Introduction

Transcription factors (TFs) are key mediators in the transmission of information from the internal and external environment of the cell to its genome. Understanding how TFs encode information about the environment in order to coordinate transcriptional programs remains one of the most pressing problems in molecular and systems biology. Many studies have explored how modulation of TF concentration, TF post-translational modification, and combinatorial TF control can yield differential gene regulation[33–35], therefore explaining many important aspects of TF function and their information encoding capacity. These mechanisms, however, may not fully account for the complexity of signal multiplexing that is carried out by TFs. As a result, it has been proposed that TFs might also encode information in their spatio-temporal dynamics, using for example different patterns of nuclear shuttling for different inputs.

A number of studies have attempted to elucidate this TF dynamic encoding hypothesis by eliciting different TF dynamic patterns using various environmental inputs and assessing the consequences[36–45,45,46]. For example, it was shown that p53 exhibits fixed concentration pulses in response to gamma radiation, but implements only one amplitude- and duration-dependent continuous pulse in response to UV[36]. These two pulsing regimes have different physiological outcomes, with the former leading to cell cycle arrest and the latter leading to cell death[43]. Other studies programmed different TF nuclear translocation patterns by gaining control of a signaling node upstream of the TF. A prominent example of this approach is the modulation of Msn2 dynamics by regulating protein kinase A (PKA). Inhibition of an analog-sensitive PKA by a small-molecule resulted in Msn2

translocation to the nucleus[9,11,21,22,47]. With this method, it was shown that genes in the Msn2 regulon can be differentially modulated by the amplitude, duration, and frequency of Msn2 nuclear translocation pulses.

In the budding yeast *Saccharomyces cerevisiae*, there are approximately 200 known TFs, two-thirds of which are constitutively localized to the nucleus; the remaining one-third are located in the cytoplasm during exponential growth in complete media[48]. At least nine of these basally cytoplasmic TFs transiently localize into the nucleus in response to various stress conditions[8]. Furthermore, different environmental conditions elicit a range of pulsing characteristics for these TFs that differ in their duration, amplitude, and frequency[8] (Supplementary Figure 3.6), suggesting that reversible TF nuclear localization may encode regulatory information. This information is then decoded by downstream target genes in order to produce an appropriate response[49].

Control of TF localization through modulation of upstream regulators with small molecules or chemicals has been an essential method to put forward such a hypothesis of TF dynamic encoding[9,20–22,43,47,50,51]. However, this method produces pleiotropic effects that can be hard to untangle. For example, PKA controls many transcriptional regulators in addition to Msn2. As a result, modulating its activity with a small molecule may yield gene expression changes that are not solely caused by Msn2 translocation dynamics but are instead the result of combinatorial gene regulation by other PKA-responsive TFs such as Msn4[25,50] and Dot6[52].

Therefore, to causally and quantitatively probe the relationship between TF nuclear localization dynamics and transcriptional activity, a method by which TFs can be specifically, quickly, and reversibly localized to the nucleus is needed. Specificity is necessary to allow direct regulation of TF nuclear localization without pleiotropic effects, while speed and reversibility are necessary to recapitulate the minutes-level resolution with which TFs translocate into and out of the nucleus in response to environmental inputs.

Ideally, this method would also work modularly with many TF cargos, including TFs that are basally nuclear.

Here, we present CLASP, an optimized optogenetic tool that can exert precise, modular, and reversible control of transcription factor localization. CLASP uses two LOV2 light-responsive domains derived from *Avena sativa* to sequester a cargo at the plasma membrane in the dark and target it to the nucleus in response to blue light. We demonstrate how CLASP can be used as a general strategy to control many TF cargos without any further optimization. We exploit the characteristics of CLASP to control the localization of Crz1, a pulsatile TF, and show unambiguously that its target promoters have different abilities to differentially interpret pulsatile dynamics. Using data-backed computational modeling, we explore the principles by which they can do so. Our studies reveal that a more efficient response to short pulsed inputs can be achieved by a simple two-state promoter model with fast activation and slow shut-off. By contrast, to achieve more efficient gene expression from continuous inputs than from pulsed inputs, a more complicated model with at least two activation steps or thresholding needs to be invoked. Furthermore, to achieve this property in combination with a graded dose response, a promoter model needs to minimally combine two activation steps and thresholding, with a dependence of both activation steps on the TF. These results, made possible through the productive use of CLASP in iteration with computational modeling, paves the way for more thorough understanding of the general principles by which gene promoters can interpret TF dynamics.

3.3 Results

3.3.1 Construction and optimization of CLASP, a dual-LOV2 optogenetic strategy for control of nuclear shuttling

Some optogenetic tools that rapidly translocate protein cargos to the nucleus have been developed[13,53,54]. A number of these tools utilized LOV2, a light-responsive protein often isolated from *A. sativa*, to uncage a nuclear localization sequence (NLS) in response to blue light. Uncaging of this NLS caused the translocation of the optogenetic molecule to the nucleus along with any appended protein cargo. Light Activated Nuclear Shuttle (LANS) is an example of this strategy[13] (Figure 3.1A). While effective for some applications, the architecture of this class of optogenetic tools may cause leaky nuclear localization for some basally cytoplasmic TFs. An example of this phenotype is Msn2, which, in many cells, exhibited constitutive nuclear localization when fused to LANS even in the absence of light stimulation (Supplementary Figure 3.7A). Moreover, tools such as LANS cannot be used to regulate localization of basally nuclear TFs, as there is no mechanism for preventing the endogenous nuclear localization of these TF cargos.

A different optogenetic tool, LOVTRAP, a LOV2-based tool for protein sequestration, could be used for rapid translocation of cargo with less leaky basal localization. LOVTRAP is composed of a LOV2 fused to the mitochondria and Zdk1, a small peptide that is fused to the protein cargo. The interaction of LOV2 and Zdk1 in the dark sequesters the cargo to the surface of the mitochondria[14] (Figure 3.1A). However, LOVTRAP alone does not contain targeting information, and hence cannot direct the cargo to the nucleus on demand. Therefore, to enable both robust and targeted optogenetic control of many different cargos, we sought to use LOVTRAP in concert with LANS. The idea of combining optogenetic sequestration and nuclear localization was previously investigated[54,55]. However, the

resulting tools either required complex dual color stimulation[54], thereby limiting the number of fluorescent proteins that could be used in a cell, or did not demonstrate modularity for different cargos[55]. These tools also lacked optimization for use in yeast.

To construct a modular and specific tool for yeast protein nuclear localization, we first tackled optimization of the published LANS and LOVTRAP constructs. Fluorescently-tagged (mCherry) LANS[13] displayed only a moderate increase (3.4%) in nuclear/cytoplasmic enrichment in response to blue light (Figure 3.1B, upper left panel). This increase was much weaker than that seen for transcription factors in response to stress inputs (Supplementary Figure 3.6, 20-50% increase). Additionally, the published LOVTRAP tool used a TOM20 mitochondrial targeting tag that caused a strong growth defect in yeast at high expression levels (Figure 3.1B, lower panel). LOVTRAP sequestration had previously been shown to perform best when the mitochondria-bound LOV2 trap was expressed in excess of the Zdk1; as a result, these high expression levels were necessary for trapping many protein cargos, causing the growth defect to be an issue.

To improve LANS localization properties, we replaced the published LANS NLS with a small library of yeast NLS peptides[56] (Supplementary Table 3.1). We then screened blue light induced nuclear localization of mCherry-LANS constructs that had any one of these different NLS sequences. We identified a number of NLS sequences that showed an improvement in nuclear/cytoplasmic enrichment in response to blue light (Supplementary Figure 3.7B), including an NLS that increased the fold change by eight-fold. We chose this NLS sequence to move forward as a yeast enhanced LANS (yeLANS) (Figure 3.1B). Next, to rectify the growth defect associated with LOVTRAP sequestration to the mitochondria, we swapped the mitochondrial TOM20 tag with a plasma membrane Hs-RGS2 tag[57] to create pm-LOVTRAP. This modification rescued the growth defect of LOVTRAP even at high expression levels (Supplementary Figure 3.7C).

Finally, we combined yeLANS and pm-LOVTRAP to form CLASP (Controllable Light

Activated Shuttling and Plasma membrane sequestration), a construct comprised of two AsLOV2 domains. The first AsLOV2 domain is fused to the plasma membrane and sequesters a Zdk1 fused to the N-terminus of the cargo (for example, a TF). The second AsLOV2 domain is fused to the C-terminus of the cargo. This AsLOV2 domain is preceded by a nuclear export sequence (NES) and has a nuclear localization sequence (NLS) embedded in the J helix. Blue light causes a conformational change in both AsLOV2 domains, yielding the simultaneous unlocking of cargo and its targeting to the nucleus (Figure 3.1A). Strains harboring CLASP did not experience any growth defect (Supplementary Figure 3.7D).

We first tested CLASP with a red fluorescent protein (mScarlet) as a cargo. Confocal microscopy showed that mScarlet-CLASP was successfully sequestered at the membrane in the dark and translocates to the nucleus in response to blue light. Furthermore, widefield microscopy showed that nuclear localization could be maintained stably for at least 80 minutes (Figure 3.1C). Varying the duration of the light input demonstrated that CLASP could also track shorter light inputs (Supplementary Figure 3.7E-G). On average, mScarlet-CLASP nuclear localization extended four minutes longer than the duration of the input light pulse, illustrating its rapid shut-off time (Figure 3.1D, Supplementary Figure 3.7E). The maximum nuclear/cytoplasmic enrichment achieved by mScarlet-CLASP was also graded as a function of light amplitude; when subjected to one minute pulses of increasing amplitude (64-1024 a.u.), enrichment increased commensurately for a wide range and saturated after 256 a.u. of light (Figure 3.1D, Supplementary Table 3.2).

Finally, to test the ability of CLASP to respond to repeated light pulses and probe its dependence on their period, we subjected the cells to one minute pulses of blue light repeated every 2-9 minutes (Figure 3.1E, left 3 panels show one minute pulses every 2, 5, or 9 minutes). These experiments revealed that mScarlet-CLASP followed these pulses faithfully until the pulses became too rapid, that is, when the next light pulse occurred

during the time required for nuclear exit (4 minutes). This effect occurred when pulses were repeated every 2 minutes, at which point nuclear localization became almost continuous at a high level. Quantification of the mean peak-to-trough difference in nuclear localization of single cell traces for different periodic light inputs showed a clear dependence on the period of the light pulse (Figure 3.1E).

Overall, our data indicate that mScarlet-CLASP could be rapidly, reversibly, and repeatedly localized to the nucleus as frequently as every five minutes and that the duration and the magnitude of this translocation could be robustly controlled.

3.3.2 CLASP achieves precise, modular control of TF nuclear translocation and activation of target genes

The usefulness of CLASP depends on its ability to successfully control translocation of TF cargos while maintaining their function. Our next step was therefore to test the ability of CLASP to quickly and reversibly control the translocation of three basally cytoplasmic transcription factors to the nucleus. We chose a synthetic transcription factor, SynTF, constructed from Cys2-His2 zinc finger domains and a VP16 activation domain[58], as well as Msn2, the principal transcription factor in the environmental stress response[59], and Pho4, the principal transcription factor in the phosphate starvation response[60]. Both Msn2 and Pho4 have been known to translocate to the nucleus in response to stress[8,60]. The three TF cargos were also tagged with a C-terminal RFP (mScarlet) for visualization.

For all three TFs, TF-CLASP achieved its maximal nuclear localization in response to light within one minute of blue light exposure. Like the mScarlet cargo, the TF cargos reversibly translocated to the nucleus as frequently as every five minutes when induced with a one minute pulse of light. Furthermore, a sustained light input produced continuous nuclear localization of the TFs, indicating that CLASP was capable of maintaining robust nuclear

localization of associated TF cargos for an extended period of time (Figure 3.2A). The maximum nuclear/cytoplasmic enrichment fold change achieved with CLASP for Msn2 as a cargo was similar to that of Msn2 with a strong osmotic shock using 0.95M Sorbitol[61] (Supplementary Figure 3.6).

To test whether nuclear localization of the TFs led to concomitant gene expression, we constructed promoter fusions expressing YFP with promoters that were responsive to SynTF (pSYNTF-YFP), Msn2 (pHSP12-YFP), and Pho4 (pPHO84-YFP). We then exposed these strains to fixed-amplitude light inputs (Supplementary Figure 3.8A) of increasing duration (0.5-2 hours) and measured YFP fluorescence via flow cytometry. For all three TFs, increasing the duration of the light input led to increased downstream reporter gene expression, illustrating that the TF was still functional despite its fusion to CLASP. Notably, SynTF-CLASP yielded more than 20-fold activation of pSYNTF-YFP with only 2 hours of light activation (Figure 3.2B). Gene expression in the dark downstream of the three TF-CLASP constructs was similar to basal expression, and was also commensurate after light induction to gene expression generated by a constitutively nuclear TF (Supplementary Figure 3.8B-D, Supplementary Text).

Next, we explored whether CLASP could control localization of transcription factors such as Gal4, which was basally nuclear. Gal4-CLASP was successfully sequestered to the plasma membrane in the dark and reversibly translocated to the nucleus in response to light. Nuclear translocation of Gal4-CLASP also activated expression from pGAL1, a Gal4-responsive promoter (Supplementary Figure 3.8E-G), indicating that CLASP was able to control TFs irrespective of their endogenous nuclear localization.

Finally, we sought to demonstrate that different TF dynamic translocation patterns generated with CLASP could yield different gene expression outputs. Several transcription factors, such as Pho4 following phosphate starvation, translocate into the nucleus in response to a stress input and reside there continuously until the response was

completed[60]. Others, including Msn2 following a 0.4% glucose input, have been known to translocate into the nucleus with episodic and repeated pulses in response to an activating input[8]. Moreover, Msn2 has also been known to translocate with sustained pulses in response to osmotic shock (Supplementary Figure 3.6B). As a result, we sought to explore the gene expression consequences of pulsing relative to continuous localization of the three CLASP-fused TFs (SynTF, Msn2 and Pho4). We delivered two light inputs that had different dynamic patterns but the same cumulative light duration of 40 minutes. In the first case, light was switched ON for 40 minutes, and in the second, light was given in 20 episodic pulses (2 minutes ON/10 minutes OFF) (Figure 3.2C). Delivery of the same cumulative light input and measurement at the end of the time course were necessary controls to compare efficiency of response of pulsed input relative to continuous inputs. YFP fluorescence was measured for both inputs after 5 hours using flow cytometry. These data showed that continuous nuclear residence of SynTF-CLASP, Msn2-CLASP, and Pho4-CLASP produced more gene expression than pulsed translocation, indicating that these promoters respond more efficiently to continuous inputs than pulsed inputs. This directly demonstrates that TF nuclear translocation dynamics could affect downstream reporter gene expression, an idea that we wanted to explore in more depth.

3.3.3 CLASP control of the Crz1 TF reveals that its target genes differ in their efficiency of response to short pulses

To further explore the modes of decoding of TF dynamics by promoters in a biologically meaningful setting, we chose to focus on Crz1, the main TF in the calcineurin-Crz1 signaling pathway that responds to calcium stress. Crz1 has been shown to exhibit two modes of pulsatile nuclear translocation in response to calcium chloride ($CaCl_2$) stress – a single long initial pulse (40-60 min) and subsequent episodic repeated pulsing (1-4 min) (Supplementary Figure 3.9A). We reasoned that continuous nuclear localization and

pulsing of Crz1 could be interpreted differently by different target genes, a behavior that could be revealed and studied by controlling its localization using CLASP.

Crz1 has been shown to undergo phosphorylation on multiple residues to activate gene expression in calcium stress[7]. Therefore, to survey the response of Crz1 target genes to dynamic inputs using CLASP, we needed to adopt a variant of Crz1 that bypassed this regulation, an endeavor that could be necessary for studying the effects of many TFs with CLASP. We therefore built a strain in which Crz1*, an alanine mutant with 19 S/T to A substitutions of Crz1, was used as a CLASP cargo. Crz1* was basally nuclear and bypassed the post-translational modification required of Crz1[7] (Supplementary Figure 3.9B). To verify that Crz1* preserved the transcriptional profile of wild type Crz1, we carried out mRNA sequencing of cell populations in which the wild type allele of Crz1 was knocked out and Crz1* was expressed from a constitutive pADH1 promoter. We compared the up-regulated genes of the Crz1* strain (where Crz1* is basally nuclear) with genes upregulated by Crz1-yeLANS under $CaCl_2$ stress. Hierarchical clustering of the resulting gene expression profile shows similarly up- and down-regulated gene sets between these two samples (Supplementary Figure 3.9C). By probing individual Crz1 target genes with fluorescent reporters, we also found that light-induced Crz1*-CLASP, but not light-induced Crz1-CLASP (Supplementary Figure 3.9D), was able to elicit appreciable gene expression. For example, Crz1*-CLASP driving pPUN1-YFP, a canonical Crz1 responsive promoter, achieved similar gene expression fold change (1.8) as pPUN1-YFP in calcium stress (1.7) (Supplementary Figure 3.9E). Importantly, Crz1*-CLASP did not cause increased gene expression in the absence of light, indicating that CLASP was able to successfully sequester the nuclearly localized Crz1* outside of the nucleus in the dark (Supplementary Figure 3.9F).

We next identified six Crz1 gene targets (Yps1, Ena1, Mep1, Put1, Cmk2, Gyp7) for follow up studies. We used the promoters of these genes, which have also been canonically used in

the literature[5,62], to build YFP-expressing promoter fusions, each in a strain with Crz1*-CLASP tagged with mCherry for visualization (Figure 3.3A). We subjected these cells to two distinct types of inputs that mimic natural Crz1 translocation: 2 minute short repeated pulses with different periods or one continuous pulse of varying duration (Figure 3.3A). We checked that extended light exposure did not cause a growth defect in the Crz1 overexpression strain (Supplementary Figure 3.9G). We then measured the nuclear enrichment of mCherry-tagged Crz1*-CLASP continuously at 30 second intervals. We also measured gene expression from all six YFP promoter fusions at 5 hours for all inputs given (Figure 3.3A). Every input (pulsatile or continuous) generated a given nuclear occupancy, which we calculated as the integral of the measured Crz1*-CLASP nuclear enrichment time traces. A given nuclear occupancy was associated with a commensurate gene expression value (measured at five hours), and these values were plotted against each other for the two input regimes for each of the six promoters. The resulting plot for all nuclear occupancy values are referred to as the Gene Output - Nuclear Occupancy plot (Output-Occupancy plot for short). Exploration of gene expression as a function of nuclear occupancy allowed a comparison on equal footing of the overall integrated responses to pulsed and continuous inputs. The Crz1-responsive promoters showed a spectrum of qualitative and quantitative behaviors in the Output-Occupancy plots (Figure 3.3A-C, Supplementary Figure 3.9H-J). For pGYP7-YFP, like the promoters shown in Figure 3.2, a pulsed input generated lower gene expression output than a continuous input of the same nuclear occupancy for all values tested, a phenotype that we termed efficient response to continuous inputs (Figure 3.3B). For pCMK2-YFP, pulsed and continuous inputs generated almost identical gene expression output. However, for pYPS1-YFP, pulsed inputs produced higher gene expression output at all Crz1*-CLASP nuclear occupancy values tested, a phenotype that we termed efficient response to pulsed inputs. These phenotypes were qualitatively reproducible despite slight quantitative day to day variability in gene expression between experiments (Supplementary Figure 3.9H-J). The difference in output between pulsed and

continuous input as a function of nuclear occupancy was quantified as the ratio of slope of the two lines in the Output-Occupancy plot, termed the slope ratio (Figure 3.3A). This metric showed that the six Crz1 responsive promoters spanned a range that is bracketed by pYPS1-YFP (slope ratio > 1) and pGYP7-YFP (slope ratio < 1), going from a more efficient response to pulsed than continuous inputs to the opposite phenotype (Figure 3.3C). These phenotypes must reflect different promoter properties since all promoter fusions generated the same YFP as the protein output. We next turned to data-backed computational modeling to systematically explore and interpret these behaviors.

3.3.4 Efficient response to short pulses by promoters can be explained by a simple model with fast promoter activation and slow shut-off

To understand the general determinants of the behavior of the Crz1-responsive promoters, we first built a simple and parsimonious model of gene expression. The model consisted of a promoter that occupied two states, OFF (p_{off}) and ON (p_{on}), with the rate constants k_{on} and k_{off} describing the transition between the two states. The rate of promoter activation depended on the concentration of nuclear TF. The activated promoter then produced YFP mRNA at rate β_1 and β_0, which represented the basal activity of the promoter, and the YFP mRNA produced the YFP protein at rate β_2. mRNA and protein degraded with rates γ_1 and γ_2, respectively (Figure 3.4A). Since the data represented a YFP promoter fusion, the protein degradation rate (γ_2), protein production rate (β_2) and mRNA degradation rate (γ_1) were assumed to be the same for all promoters. Hence, all remaining parameters that differed among the different promoters were only related to promoter properties. Moreover, γ_1 and γ_2 were set to plausible values corresponding to YFP characteristics from the literature[63,64]. The model was simulated with the same pulsatile or continuous nuclear

localization input regimes used in the experiments and the output was the YFP protein value at 5 hours, mimicking the experimental setup and data collection procedures.

To explore a wide range of parameter regimes, we randomly sampled 10000 parameters (k_{on}, k_{off}, β_1, β_2, β_0) and generated Output-Occupancy plots and the slope ratio metric for every combination. A plot of slope ratio as a function of each sampled parameter set showed that there was no relationship between β_1, β_2, or β_0 and slope ratio (Supplementary Figure 3.10A). Instead, β_1, β_2, and β_0 scaled the range of the output curves. We therefore focused on the influence of k_{on} and k_{off} on slope ratio, plotting the values of that metric as a heatmap in the k_{on} -k_{off} plane (Figure 3.4B). Broadly, there were two clear patterns. First, all parameter combinations tested produced higher gene expression output from pulsed inputs than continuous inputs for all nuclear occupancy values (slope ratio was always ≥ 1) (Figure 3.4A-E). Second, the separation between the continuous and pulsed input curves in the Output-Occupancy plots was dictated by the values of k_{on} and k_{off}, with the separation increasing with increasing k_{on}/k_{off}. At very large k_{on}/k_{off}, the pulsed output became multi-sloped (gray region in Figure 3.4B).

Since the translation and degradation rates in this model were fixed, the observed differences at the protein level were determined by the differences in the promoter activity, p_{on}, which in turn was dictated by k_{on} and k_{off}. Hence, to gain intuition about the model results, we focused on the promoter activity, p_{on}, in three regimes marked 1, 2, and 3 in Figure 3.4B, which provided snapshots of promoter activity as k_{on} increased and k_{off} decreased. In the first regime (1), k_{on} was much smaller than k_{off} so that a pulsed input did not induce full ON switching of the promoter within the duration of a short pulse, while p_{on} could reach its maximum possible value during a continuous longer input. This caused an amplitude difference in p_{on} between pulsed and continuous inputs (Figure 3.4C). However, a relatively slower k_{off} implied that the promoter stayed ON for a period of time beyond the duration of the input, and when repeated for every pulse, this residual activity

could counteract the amplitude deficiency (Figure 3.4C, hatched region under p_{on} curves). Hence, in this k_{on}/k_{off} regime, the protein output of the short pulsed input was slightly higher than that of the continuous input. In the second regime (2 in Figure 3.4B), a larger k_{on} caused the promoter to turn fully ON within one short pulse, and a small k_{off} again caused the promoter to switch OFF slowly, maximizing the gain from every pulse (Figure 3.4D). As a result, the cumulative promoter activity generated by a short pulsed input was large, and the difference in protein output generated by a pulsed compared to a continuous input of matched nuclear occupancy was also large (Figure 3.4D). In the third regime (3 in Figure 3.4B), at an extreme where $k_{on} >> k_{off}$, the promoter was ON for the whole duration of the short pulsed input (4 hours of repeated pulses of 2 minutes each, for example), with little OFF switching of the promoter during repeated pulses. On the other hand, a continuous input, with the same cumulative nuclear residence as the short pulsed input, would generate a promoter that was ON approximately only for the duration of the pulse plus the switching off time of the promoter (Figure 3.4E, region b lower panel). The change in the slope of Output-Occupancy plot occurred when the promoter did not shut completely OFF between two consecutive pulses (compare plots for point (a) and (b) in Figure 3.4E upper and lower panel). While the model identified this regime, which could be applicable to some promoters, none of the promoters we explored showed strong multi-slope Output-Occupancy curves. Specifically, pYPS1-YFP showed a clear Output-Occupancy relationship with one slope, making it likely that this promoter operated in the k_{on}/k_{off} regime in regime 2.

To test whether the model could provide a quantitative fit to the pYPS1-YFP and pCMK2-YFP data, we further fit the parameters to the Output-Occupancy data, specifying fits to be model predictions that maximized the fit through the data points within the error bars for pYPS1-YFP (15 parameter sets) and for pCMK2-YFP (489 parameter sets) (Figure 3.4F-G). We then subjected these fits to cross-validation, asking whether the parameters with the accompanying model could reproduce the dose response

of pYPS1-YFP and pCMK2-YFP. We measured these dose responses in an independent experiment using strains in which Crz1*-CLASP was expressed at different levels using a suite of constitutive promoters of different strengths[65] (pRPL18B, pADH1, and pTEF1). These strains, which also harbored either pYPS1-YFP or pCMK2-YFP, were subjected to continuous light input over four hours, leading to maximum nuclear localization in all strains. Different strains had different amounts of nuclear Crz1*-CLASP when localized with light, therefore measurement of YFP in each strain provided a different point on the dose response (Figure 3.4F, G). We subjected the model to the same treatment *in silico*, and produced computational predictions of the dose response curves for all parameter sets that fit the Output-Occupancy data for pYPS1-YFP or pCMK2-YFP (Figure 3.4F, G, H). Computational and experimental predictions were in strong qualitative agreement.

Taken together, our data show that a simple promoter model can provide a straightforward scenario in which a promoter can respond more efficiently to repeated short nuclear pulses of a TF than a continuous input. A tight iteration of modeling and experiment further revealed that this behavior is dependent on the dynamics of promoter activity produced by fast promoter activation coupled with its slow inactivation. Fast promoter activation leads to complete activation during a short pulse, and slow promoter turnoff leads to accumulation of expression over the course of repeated pulses, providing insight into how even simple promoters can readily decode dynamic inputs.

3.3.5 Efficient response to continuous inputs by promoters can be explained by a model with a thresholded transition between non-transcribing promoter states

The simple model from the previous analysis could not produce the pGYP7-YFP phenotype (Supplementary Figure 3.11A-B). When the output difference between the

pulsed and continuous inputs was small ($k_{on} << k_{off}$) in this model, the output of the pulsed input was always higher than the continuous input. This was because while decreasing k_{on} reduced the output of the pulsed input, it also reduced the dynamic range of the output in response to a continuous input to a point where k_{on} was so small that the promoter was barely activated and the much faster k_{off} quickly shut off promoter activity, resulting in a promoter that was essentially unresponsive to either inputs (Supplementary Figure 3.10B).

In order to identify a minimal model that explained the pGYP7-YFP phenotype, we explored eight elaborations of the simple promoter switching model from Figure 3.4 using a sequence of fitting and cross-validation. In this process, each model was first fit to the Output-Occupancy data in Figure 3.3; one of the eight models failed to fit. Models that fit the Output-Occupancy data were further fit to the dose response of pGYP7-YFP, which was collected in the same way as for pCMK2 and pYPS1. The pGYP7-YFP dose response was remarkably linear, and four models failed to fit it (Figure 3.5A-D). For the 3 remaining models, the dose response data served to further constrain parameter sets. For those refined parameters, we cross-validated the models on the data from an additional experiment in which we expressed Crz1*-CLASP from a stronger promoter (pTEF1 versus pADH1), and measured gene expression following a cumulative light induction of 40 minutes administered either as pulsed or continuous input. Following these rounds of fitting and cross-validation (Supplementary Figure 3.11C-R), only two of the models surveyed were able to explain all the data we collected (Figure 3.5A-D, Supplementary Figure 3.11O-P).

The two models were structurally similar – they both extended the simple two-state model to contain another promoter state, thereby necessitating transition through an unproductive promoter state (p_{off}) before the promoter can be fully activated. Therefore, in these models, the first transition occurred reversibly between promoter state p_0 and a non-transcribing state p_{off} with rate constants r_{on} and r_{off}, while a second transition stage

occurred between p_{off} and p_{on} with rate constants k_{on} and k_{off}. Both models also necessitated a linear dependence on TF in the second transition stage, whose effect was to prevent the dose response from exhibiting a thresholded behavior. Finally, the two models necessitated a thresholded interaction in the first promoter transition stage, but differed in where it was applied – in one model, r_{on} was a thresholded function of TF, while in the other model, it was r_{off} that was thresholded by TF (Figure 3.5A, Supplementary Figure 3.11O). The threshold on either r_{on} or r_{off} prevented short pulsed inputs from fully transitioning the promoter from the p_0 state, essentially creating a filter for short inputs. Detailed descriptions of all models and their exploration can be found in the Supplementary Text and Supplementary Figure 3.11.

To gain more insight into the pGYP7-YFP phenotype, we further explored the 3-state, r_{off} threshold model for many parameter values (Figure 3.5A). We sampled the model parameters by fixing r_{on} and k_{on} to values that fit the data from Figure 3.5B-D and varied r_{off} and k_{off} within a range of four logs. We then generated Output-Occupancy plots for every parameter set and computed its corresponding slope ratio metric, which we plotted in the log10(k_{on}/k_{off}) - log10(ron/roff) plane (Figure 3.5E). Overall, we found that this model can generate both higher expression with a continuous input (slope ratio < 1, black region in Figure 3.5E, top panel) and higher expression with short pulses (slope ratio > 1, colored region on Figure 3.5E, top and bottom panel).

Quantitatively, there seemed to be three parameter constraints for this promoter model to respond more efficiently to a continuous input than a pulsed one. First, the rate of transition from p_0 to p_{off} should be slow; second, r_{off} should be fast relative to r_{on}; third, k_{off} should be fast relative to k_{on}. An analysis of the 3-state r_{on} threshold model demonstrated similar requirements (Supplementary Figure 3.12A-B). When r_{on} and r_{off} were increased tenfold, there were no parameter combinations that generated higher expression for continuous inputs than short pulses (Figure 3.5E, bottom panel, Figure

3.5F-G, top panel). The difference in the protein outputs between the pulsed and continuous inputs was determined by the amplitude differences of promoter activity p_{on} (Figure 3.5F-G, middle panel), which was in turn dictated by the amplitudes of depletion from p_0 for the short pulsed and continuous inputs (Figure 3.5F-G, middle panel). A slow transition from p_0 prevented the quick, full depletion of this state before a short pulse ended, while p_0 was fully depleted for the continuous input (Figure 3.5F, middle panel). In contrast, when r_{on} and r_{off} were fast, this difference disappeared as the transition from p_0 was now able to reach the same maximal amplitude in the duration of the short input (Figure 3.5G, middle panel). Hence, the incomplete depletion of the p_0 state in the duration of the short pulsed input accounted for the difference in protein outputs between the short pulsed and continuous inputs.

The requirement that the value of r_{off} be large relative to r_{on} was motivated by the fact that r_{off} dictated how quickly the promoter state transitioned back to the initial OFF state p_0 after the end of a short pulse. When the value of r_{off} decreased relative to r_{on} (Supplementary Figure 3.12C), the depletion of p_0 could proceed to completion during a short pulse (Supplementary Figure 3.12C, middle panel), and the resulting maximum amplitudes of the active promoter state p_{on} were more comparable for a pulsed or continuous input (Supplementary Figure 3.12C, bottom panel). Lastly, as k_{off} was decreased while keeping all other parameters constant, the p_{on} to p_{off} switching also slowed, and promoter activity continued unabated between two pulses, hence maximizing the gain of promoter activity from every input pulse and causing stronger gene expression from pulses than from a continuous input (Supplementary Figure 3.12D). This was in essence the same mechanism as described in Figure 3.4. In summary, slow transition from the initial OFF state to the secondary OFF state prevented the short pulsed input from achieving a quick depletion of the initial OFF state, essentially creating a filter for short inputs. This analysis reveals that differences in dynamic changes in promoter activity can result in more efficient response to a continuous TF input than a pulsed input.

Finally, in addition to the constraints above, we found that a threshold of log10(k_{on}/k_{off}) $\sim<$ -1.5 seemed to demarcate the transition between a linear and nonlinear promoter dose response in the parameter regime probed (light gray points, Figure 3.5E, left panel), therefore imposing quantitative bounds on this promoter model to exhibit a graded dose response as seen in the data.

3.4 Discussion

In this work, we presented a general optogenetic tool that circumvents some functional caveats of previous methodologies, and used it productively to investigate in a systematic way how promoters are able to differentially respond to pulsed versus continuous transcription factor localization. Our tool, CLASP, was inspired by previous optimization efforts that attempted to make optogenetic control more precise and malleable[66–70]. However, we pushed our optimization and characterization efforts further, producing a tool that was functional for all cargos tested, had a large dynamic range and no detectable deleterious impact on the cell. We capitalized on this tool to ask a simple and profound biological question – can genes differentiate between transcription factor inputs that differ only in their dynamic patterns? CLASP allowed us to directly test and provide a definitive demonstration of this phenomenon for a number of TF-promoter pairs. We then explored mechanistic underpinnings of this behavior using the transcription factor Crz1, whose response to calcium stress is naturally pulsatile.

The precise and robust operation of CLASP allowed us to approach this investigation methodically, establishing through rounds of computational modeling and experimentation two classes of models that could explain how Crz1 promoters may respond differentially to different Crz1 pulsing inputs. For promoters that responded more efficiently to short pulsed inputs, we demonstrated that their behaviors could be simply explained by a

two-state model of the promoter (ON or OFF), transitioning between them with first order kinetics. If the ON rate, which is dependent on the nuclear TF concentration, is fast and the OFF rate is slow, then the output from short pulses is larger than for a continuous TF input of the same nuclear occupancy. By contrast, for promoters that responded more efficiently to continuous than pulsed TF inputs, a more involved model needed to be invoked. Interestingly, for pGYP7, this behavior coincided with an additional property – a linear dose response. To explain the efficient response to continuous inputs in isolation, either a thresholded step in a two-state model or a model with additional promoter states was sufficient. To satisfy a linear dose response in isolation, k_{off}/k_{on} of a simple two-state model must be relatively large. Yet, in order to satisfy both properties simultaneously, an involved model that fulfilled the requirements for each individual property and that included a dependence on the TF at each promoter stage was needed.

What possible benefits of this differential interpretation, and of the dose response linearity, might exist for the cell? Under stress, Crz1 undergoes an initial long 40-60 minute nuclear localization, followed by pulsing in the "maintenance" phase of the calcium response. It is possible that differential interpretation of these dynamic inputs by different sets of genes is used as a mechanism to temporally program the response, with cohorts of genes activating strongly in the first long pulse and then to a lesser degree with the subsequent pulsatile episode, while others do the opposite. Moreover, it has been observed that Crz1 pulses with different amplitudes in the "maintenance" phase. In this work, we established that pGYP7 was inherently linear as a function of Crz1 input for a broad range of Crz1 levels (Figure 3.5). Hence a linear dose response extend the efficient response to a continuous input over a range of Crz1 pulse amplitudes in the maintenance phase.

Our studies presented concise phenomenological promoter models which were filtered through rounds of cross validation and explained all the data collected. These models correspond to well-understood mechanisms of promoter regulation. The 2-state kinetic

model that describes the pYPS1-YFP behavior may represent a simple promoter activated by a TF. The two promoter models that describe pGYP7-YFP, in which a thresholding step in promoter transition occurs either in the ON (ron) or the OFF (roff) rates, can represent more complex biological mechanisms. The 3-state promoter, on the other hand, can correspond to closed (p_0), open but non-transcribing (p_{off}), and open and transcribing (p_{on}) states of the promoter. Additionally, the TF-thresholded transitions of r_{on} can represent transcription factor activation of chromatin remodelers[71,72], while thresholding in the r_{off} rate constant can represent transcription factor inhibition of heterochromatin formation either by physical occlusion of nucleosomes[73,74] or inhibition of deacetylation of nucleosomes[75–77]. Available nucleosome occupancy data for Crz1 target genes support the models that describe pYPS1-YFP and pGYP7-YFP by showing a negative correlation between nucleosome occupancy and responsiveness to short pulses in our data, with genes that respond efficiently to short pulses exhibiting lower nucleosome occupancy (Supplementary Figure 3.12E-F). This correlative data aligns with the idea that a more efficient response to continuous inputs requires additional promoter regulation, such as a TF-gated promoter transition between non-transcribing promoter states, compared to promoters that respond efficiently to short inputs. Still, more mechanistic studies, such as RNA FISH for observation of promoter dynamics, are needed to pinpoint the biochemical mechanisms that underlie these models.

While we used CLASP to explore fundamental aspects of promoter dynamics, many investigations that extend this work in broader directions are readily possible. For example, we showed that CLASP can sequester and translocate basally nuclear TFs (Supplementary Figure 3.8E-G), and therefore can be used with a variety of cargos to mimic dynamic TF knockouts in the presence of their activating environmental inputs, or combinatorial TF regulation using a leave-one-out strategy. More generally, CLASP is likely transplantable to other cells and organisms. In mammalian cells, NFAT, the Crz1 homolog, is known to respond to immunological stimuli with nuclear pulsing[78]. We envision using CLASP in this

context to explore whether NFAT target genes also use differential decoding to modulate different expression programs and determine general conserved principles.

Our exploration into the possible mechanisms of differential decoding could also guide engineering of frequency-responsive promoters. A more in-depth computational and experimental analysis of Crz1 target gene promoters may reveal concrete sequence motifs for building promoters that differentially decode TF dynamics. Frequency-responsive promoters can provide a useful alternative for distinct signal encoding in experimental setups where a single input is desired or necessary. These promoters can diversify available inputs for synthetic biology and metabolic engineering applications.

Finally, while our studies focused on transcriptional regulation at the promoter level, many opportunities for further decoding of TF information can be implemented through modulation and control of translation and degradation of mRNA and protein (Supplementary Figure 3.10C). It will be fascinating to study the bounds of complexity explored by endogenous genes through combinatorial tuning of all steps of gene expression to implement sophisticated dynamic decoding capabilities.

3.5 Figures

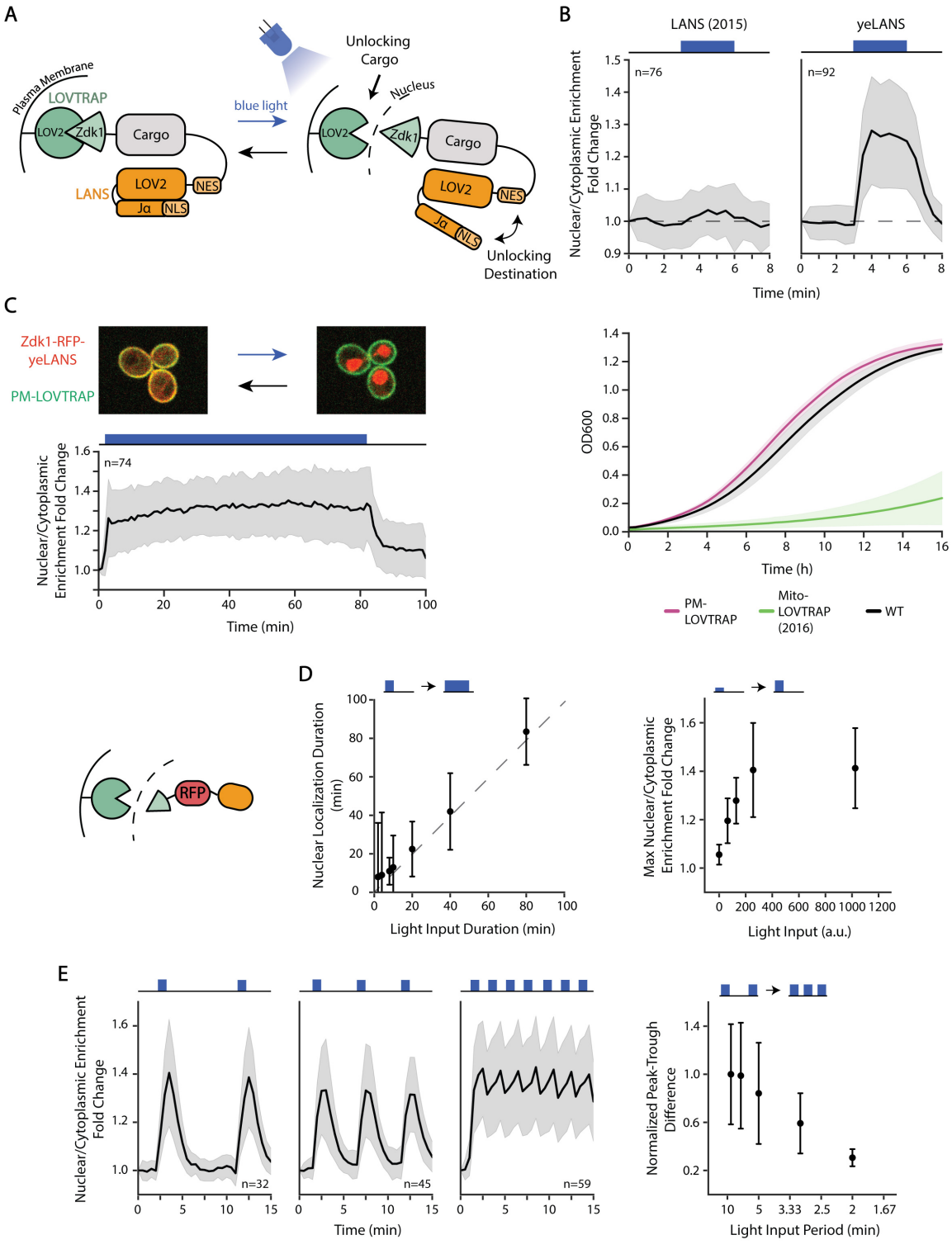

A
Plasma Membrane
LOVTRAP
LOV2
Zdk1
Cargo
LANS
LOV2
NES
Jα
NLS
blue light
Unlocking
Cargo
Nucleus
Unlocking
Destination
B
LANS (2015)
yeLANS
Nuclear/Cytoplasmic Enrichment Fold Change
n=76
n=92
Time (min)
C
Zdk1-RFP-yeLANS
PM-LOVTRAP
Nuclear/Cytoplasmic Enrichment Fold Change
n=74
Time (min)
OD600
Time (h)
PM-LOVTRAP
Mito-LOVTRAP (2016)
WT
RFP
D
Nuclear Localization Duration (min)
Light Input Duration (min)
Max Nuclear/Cytoplasmic Enrichment Fold Change
Light Input (a.u.)
E
Nuclear/Cytoplasmic Enrichment Fold Change
n=32
n=45
n=59
Time (min)
Normalized Peak-Trough Difference
Light Input Period (min)

Figure 3.1: **Design, Optimization, and Characterization of CLASP** A) Schematic illustrating CLASP mechanism. B) Optimization of LANS NLS (top panels) and LOVTRAP localization (bottom panel). Top panels show mean value of nuclear/cytoplasmic enrichment fold change for original NLS and optimized NLS (yeLANS) as a function of time when given a pulse of blue light. Nuclear/Cytoplasmic enrichment fold change is calculated relative to the nuclear/cytoplasmic enrichment at t=0. Bottom panel shows mean of OD600 in 3 growth experiments for original LOVTRAP targeted to mitochondria in addition to the optimized plasma membrane targeted LOVTRAP. C) (top panel) Confocal microscopy image showing mScarlet-CLASP localization at the plasma membrane in the dark (left) and in the nucleus (right) after 3 minutes of light exposure. (bottom panel) Quantification of mean nuclear/cytoplasmic enrichment fold change of mScarlet-CLASP as a function of time in response to a prolonged light input (80 minutes, 1024 a.u. light input amplitude). Black line represents the mean of 74 cells. D) Quantification of the response of mScarlet-CLASP to light inputs with different dynamic characteristics. Left plot shows median time to return to within 25% of basal nuclear/cytoplasmic enrichment for light pulses of different durations and constant 1024 a.u. amplitude. Median is used to minimize the effect of outliers. The dotted line is Y=X line. Right plot shows the mean response to one minute light pulses of different amplitudes. E) Nuclear/cytoplasmic enrichment fold change of mScarlet-CLASP in response to light pulsing with different periods. Left three graphs show mean enrichment fold change as a function of time in response to pulsed light inputs (1 minute light given in a 9, 5, or 2 minute period, respectively) with 1024 a.u. amplitude. Right plot quantifies median peak-to-trough difference (normalized to the median peak-to-trough difference generated by the shortest period). Median is used to minimize the effect of outliers. Error bars and shaded area, except where noted, represent standard deviation to show the spread of the data. For all panels, n represents the number of cells tracked and light input regimes are depicted on top of panels. Cartoon (left of D) represents mScarlet-CLASP. yeLANS – yeast enhanced LANS, PM-LOVTRAP – Plasma Membrane LOVTRAP, Mito-LOVTRAP – Mitochondrial LOVTRAP.

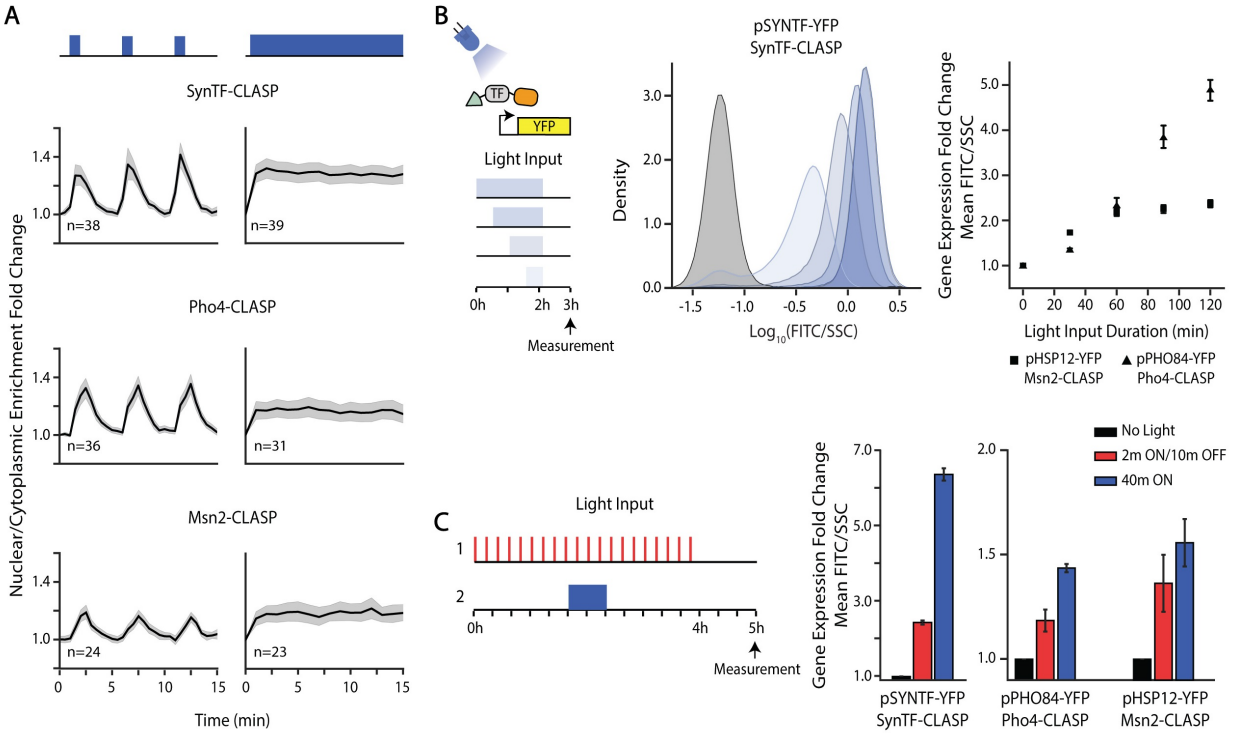


Figure 3.2: **CLASP can be used to control localization of many transcription factor cargos** A) Nuclear/cytoplasmic enrichment fold change in response to pulsed (left panels) and continuous light (right panels) for several TF-CLASP cargos. Graph shows mean of single-cell traces for transcription factors tagged with CLASP. Light is delivered for one minute at the start of each five-minute period or continuously. Shaded gray area represents 95% confidence interval and light inputs are represented in blue above graphs. B) Fluorescent reporter expression due to TF-CLASP localization. Left panel shows a schematic of the experiment – the TF is localized to the nucleus for 0.5, 1, 1.5 or 2hrs. A fluorescent reporter is measured via flow cytometry one hour after light shut-off. Center panel shows the population response of pSYNTF-YFP (promoter downstream of SynTF-CLASP) for inputs shown on the left. Darker blue shades correspond to longer light duration. Black histogram corresponds to no light. Right panel shows quantification of the YFP fold change as a function of light duration for promoters responsive to other TF-CLASP constructs following the same experimental protocol. Fluorescence readings are normalized by side scatter and then normalized to the 0m dose for each strain to show fold change. Error bars represent standard error of the mean for 9 biologically independent replicates. C) Fluorescent reporter response to pulsatile versus continuous localization of different TF-CLASP constructs. TF-CLASP constructs are given either 20 two-minute pulses of light or 1 forty-minute pulse of light, as depicted in the schematic on the left. Reporter expression is measured via flow cytometry one hour after light shut-off. Right panels show quantification of YFP fold change in response to pulsed light input, continuous light input, or no input. Error bars represent standard error of the mean for 9 biologically independent replicates. In all panels, strains are induced with a given amplitude of light (SynTF-CLASP – 1024 a.u.; Msn2-CLASP – 2048 a.u., Pho4-CLASP – 4095 a.u.).

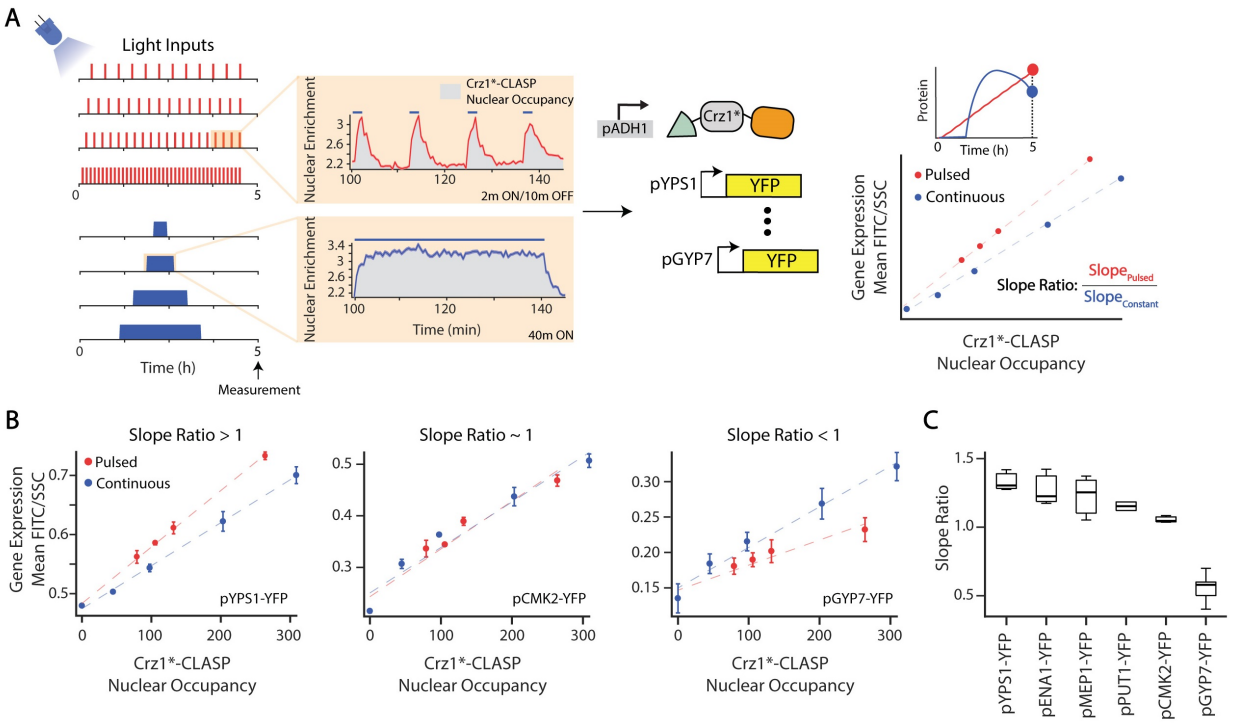


Figure 3.3: **Crz1 target genes show differing interpretation to Crz1*-CLASP short nucleoplasmic pulses** A) Schematic of experimental setup used. Two types of light inputs are given to cells expressing Crz1*-CLASP: 2 minute pulses with increasing period (20, 15, 12, and 6 minute periods) and single pulses with increasing duration (20, 40, 80, 120 minutes). Light-induced Crz1*-CLASP nuclear localization is measured with fluorescence microscopy. The mean of single cell fluorescences are plotted (solid red for pulsed input or blue line for continuous inputs), with the shaded area representing 95% confidence interval (red or blue shading). Crz1*-CLASP nuclear occupancy (x-axis in rightmost panel) is quantified as the area under the nuclear localization traces (gray shading in middle panel). Gene expression (mean FITC/SSC) is measured for 6 promoter fusions of target gene driving a fluorescent protein (YFP) at 5 hours after light input. A schematic shows gene expression values for different light input regimes are plotted as a function of nuclear occupancy, generating the Output-Occupancy plot referred to in the text. Each point in the plot is an endpoint measurement of gene expression, as highlighted by the YFP time course schematic above. Red circles represent output-occupancy for short 2 minute pulses with increasing period, and blue circles represent that for continuous single pulse with increasing durations. A best fit line (red for pulsed inputs and blue for continuous inputs) is fit through the data points for the pulsed and continuous inputs. For each Output-Occupancy plot we define the slope ratio as the ratio of the slope of the pulsed to continuous best fit lines. B) Output-Occupancy plot for three representative Crz1 target promoters pYPS1-YFP, pCMK2-YFP, and pGYP7-YFP. The error bars are standard deviation of at least 3 biological replicates. C) Slope ratios for 6 Crz1 target genes plotted in order of highest to lowest slope ratio. In all panels, Crz1*-CLASP is induced with a 512 a.u. light input.

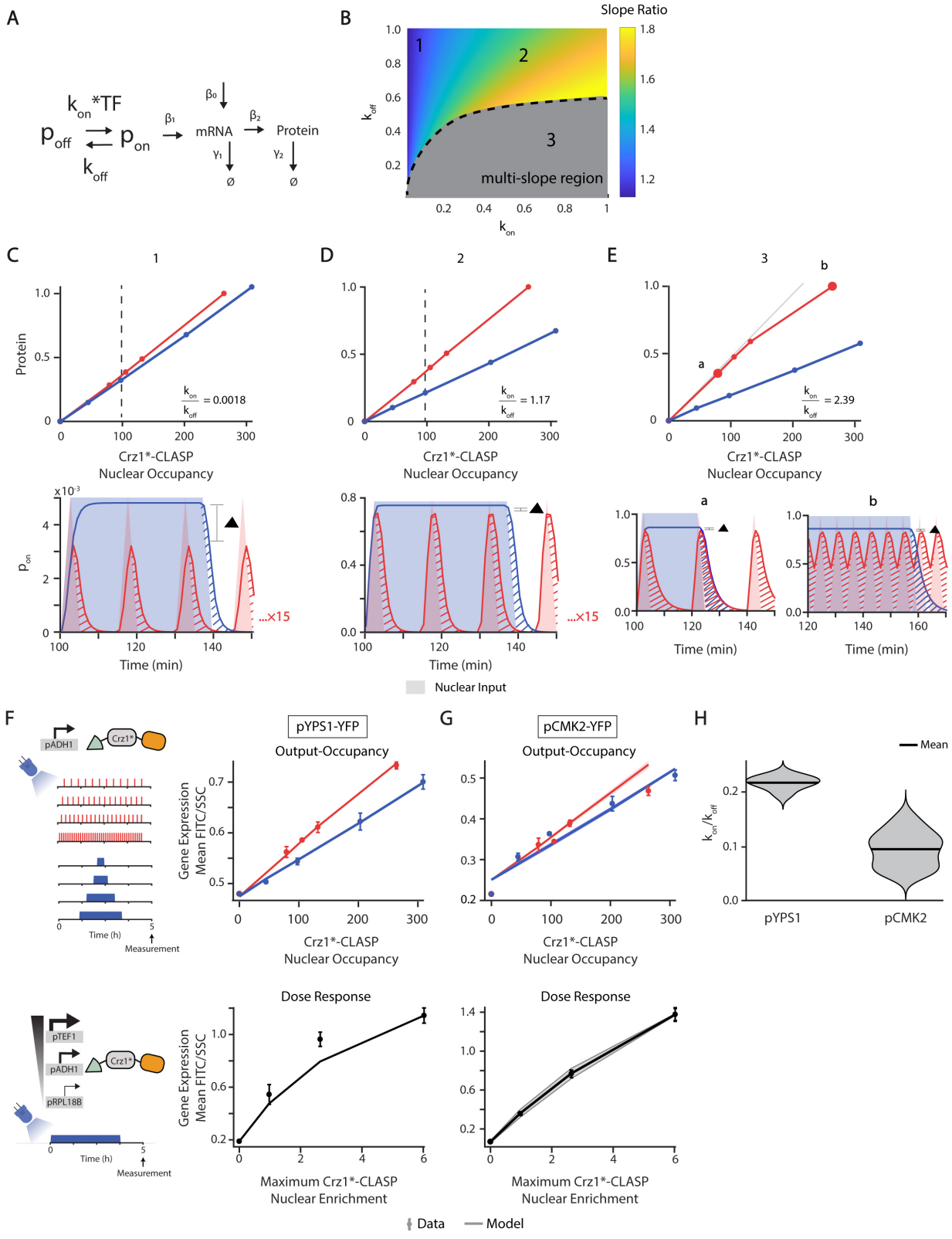

A
B
Slope Ratio
multi-slope region
C
D
E
Protein
Crz1*-CLASP
Nuclear Occupancy
Time (min)
Nuclear Input
F
pADH1
Crz1*
pTEF1
pRPL18B
Time (h)
Measurement
pYPS1-YFP
pCMK2-YFP
Output-Occupancy
Gene Expression
Mean FITC/SSC
G
H
Mean
pYPS1
pCMK2
Dose Response
Maximum Crz1*-CLASP
Nuclear Enrichment
Data
Model

Figure 3.4: **Efficient response to short pulses by promoters can be explained by a simple model with fast promoter activation and slow shut-off** A) Schematic of a two-state promoter model, where the input is Crz1*-CLASP nuclear localization (TF) and the output is fluorescent protein level (Protein). The promoter turns ON with rate constant k_{on} and OFF with rate constant k_{off}. B) Heatmap of slope ratio (defined in Figure 3.3 and main text) resulting from the model in (A) as a function of k_{on} and k_{off}, which both vary from 0.0001-1. The ratio k_{on}/k_{off} increases in a clockwise direction on the heatmap. Points 1 and 2 highlight the increasing slope ratio in the linear region. Point 3 resides in the multi-slope region (gray) of the Output-Occupancy plot, and hence is not quantified by slope ratio. In this heatmap, β_1 varies from 0.0001-10, β_0 from 0.000001-0.01 and β_2 from 0.0001-10. The parameter γ_1 is set to 0.05 and γ_2 to 0.0083. C-E) (upper panels) Output-Occupancy plots generated by the model for different parameter sets that correspond to points 1, 2 and 3 in the heatmap of panel B. The slope ratio for point 1 is 1.1 with a k_{on}/k_{off} equal to 0.0018 (C). The slope ratio for point 2 is 1.73 with a k_{on}/k_{off} equal to 1.17 (D). The k_{on}/k_{off} for point 3 is 2.39 (E). (Lower panels) Examples of time courses of promoter activity (p_{on}) for a light input that produces the equivalent of 40 minutes (dotted line in upper panel) in nuclear localization either continuously or in short pulses. Solid lines are the p_{on} pulses while shading denotes nuclear localization. The red and blue hashes represent residual promoter activity beyond the nuclear localization input. The red residual promoter activity is repeated 15 times while the blue residual activity is repeated one time. The black triangle denotes the difference between the amplitudes generated by the 2 minute pulsed and 40 minute continuous inputs. For panel (E), p_{on} is plotted for two nuclear localization values (denoted a and b in upper panel) to illustrate a regime in which p_{on} does not completely shut off between repeated pulses. F) (upper panel) Output-Occupancy plots for pYPS1-YFP. Circles are experimentally measured values. Solid lines are the mean of the 15 parameter sets that fit the data and shaded areas are the standard deviation of the mean. (lower panel) Parameters that fit the Output-Occupancy are used to predict the dose response of pYPS1-YFP (solid black line and gray shading are the mean and standard deviation, respectively, generated by the model). The gray circles are the experimentally measured dose response. G) Same as in (F) but repeated for pCMK2-YFP with results from 489 parameter sets that maximize fit to the experimental data. Dose response model predictions are for these parameter sets. H) k_{on}/k_{off} ratio of the parameter sets that fit both the Output-Occupancy and dose response data for pYPS1-YFP and pCMK2-YFP. The black line denotes mean and the gray cloud denotes the distribution of the k_{on}/k_{off} parameter values.

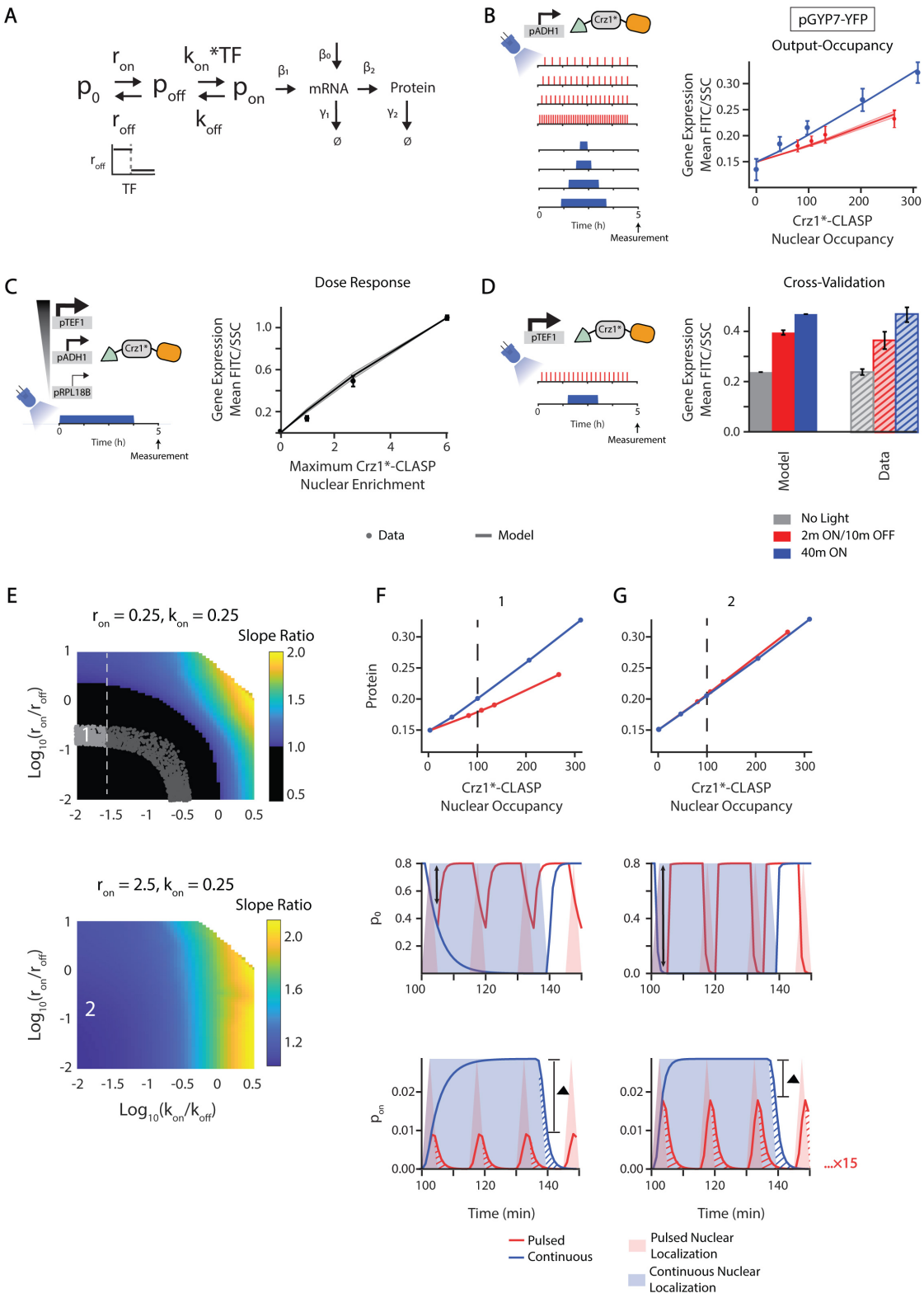

A
B
C
D
E
F
G
pGYP7-YFP
Output-Occupancy
Gene Expression Mean FITC/SSC
Crz1*-CLASP Nuclear Occupancy
Time (h)
Measurement
Dose Response
Maximum Crz1*-CLASP Nuclear Enrichment
Cross-Validation
Model
Data
No Light
2m ON/10m OFF
40m ON
Slope Ratio
Protein
Time (min)
Pulsed
Continuous
Pulsed Nuclear Localization
Continuous Nuclear Localization
...x15

Figure 3.5: **Efficient response to continuous inputs by promoters can be explained by a model with a thresholded transition between non-transcribing promoter states** A) Schematic of the three-state model where r_{off}, the inactivation rate constant from p_0 to p_{off}, is thresholded by TF concentration and where the activation from p_{off} to p_{on} is linearly dependent on TF. B) (left panel) Schematic of experimental setup. (right panel) Output-Occupancy plot for pGYP7-YFP. Circles are experimentally measured values while lines denote the mean model output for 96 parameter sets that fit the data points within the error bars, the same metric as used in Figure 3.4. The solid line denotes the mean and shaded areas denote the standard deviation of the model outputs for these parameter sets. Parameters were sampled (r_{on} from 0.1-100, r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01, threshold from 0-0.5) or set (β_2 = 0.06, γ_1 = 0.05, γ_2 = 0.0083). C) (left panel) Schematic of experimental setup. (right panel) Dose response plot for pGYP7-YFP. The parameters that fit the Output-Occupancy data were used to further fit the dose response of pGYP7-YFP using a least squared error criterion (25 parameter sets). Solid black line is the mean generated by the model. The gray circles are the experimentally measured dose response. D) (left panel) Schematic of experimental setup. (right panel) The parameters that fit the Output-Occupancy are subjected to cross-validation using an experiment where Crz1*-CLASP expression is increased (expressed from a pTEF1 promoter), and cells are exposed to either short (2 minutes ON/10 minutes OFF) or continuous input (40 minutes of light). The model generated outputs (solid gray, red, and blue bars) are plotted with the experimental data (dashed gray, red, and blue bars). The gray bars correspond to no light input. E) (top panel) Heatmap shown in the log10(k_{on}/k_{off})-log10(ron/roff) plane of slope ratio of Output-Occupancy relationship resulting from the model in (A). Parameters are sampled (r_{off} from 0.0025-25, k_{off} from 0.0025-25) or set (r_{on} =0.25, k_{on} = 0.25, β_1 = 0.0001, β_2 =0.06, γ_1 =0.05, γ_2 =0.0083, threshold= 0.5, β_0 =0.000001). Point 1 highlights a parameter set that fits the output-occupancy, dose response, and cross-validation datasets. Black region is where slope ratio < 1. Gray dotted line indicates when log10(k_{on}/k_{off}) $\sim=$ -1.5, at which point the dose response changes from linear to nonlinear with increase in the log10(k_{on}/k_{off}) value. All parameters that show a qualitative fit to Output-Occupancy data are displayed as light and dark gray dots. The light gray dots represent parameter sets where all pGYP7-YFP data are quantitatively fit. (bottom panel) Heatmap of slope ratio as in (B) with a r_{on} = 2.5, 10 times larger than than in (B). k_{on} is also set to 0.25. Parameters are sampled (r_{off} from 0.025-250, k_{off} from 0.0025-25) or set (β_2 = 0.0001, β_2 = 0.06, γ_1 = 0.05, γ_2 = 0.0083, threshold = 0.5, β_0 = 0.000001). Point 2 highlights the effect of increasing both r_{on} and r_{off} while maintaining the ratio log10(ron/roff). F-G) (upper panels) Output-Occupancy plots generated by the model for different parameter sets that correspond to points 1 and 2 in the heatmaps of panel E. The slope ratio for point 1 is 0.51 with log10(k_{on}/k_{off}) = -1.58 and log10(ron/roff) = -0.89. The slope ratio for point 2 is 1.04 with log10(k_{on}/k_{off}) = -1.58 and log10(ron/roff) = -0.89. Point 2 is chosen to highlight the effect of increasing both r_{on} and r_{off} while maintaining the ratio log10(ron/roff). (middle panels) Example of a time course of promoter state p_0 for a light input that produces the equivalent of 40 minutes (dotted line in upper panel) in nuclear localization either continuously or in short pulses. Solid lines are the p_0 pulses

Figure 3.5: while shading denotes nuclear localization. The black double arrows denote the maximum depletion of the p_0 state for the pulsed input. (lower panels) Example of a time course of promoter activity p_{on} for a light input that produces the equivalent of 40 minutes (dotted line in upper panel) in nuclear localization either continuously or in short pulses, similar to middle panels. The red and blue hashes represent residual promoter activity beyond the nuclear localization input. The red residual promoter activity is repeated 15 times while the blue residual activity is repeated one time. The black triangle denotes the difference between the amplitudes generated by the 2 minute pulsed and 40 minute continuous input.

3.6 Materials and Methods

3.6.1 Plasmid and strain construction

Hierarchical golden gate assembly was used to assemble plasmids for yeast strain construction using the method in Lee et al. BsaI, BsmBI, and NotI cut sites were removed from individual parts to facilitate downstream assembly and linearization. Parts were either generated via PCR or purchased as gBlocks from IDT. These parts were then assembled into transcriptional units (promoter-gene-terminator) on cassette plasmids. These cassettes were assembled together to form multi-gene plasmids for insertion into the yeast genome at the TRP, URA, or LEU locus. Cassettes were digested with NotI and then transformed into yeast as described in Lee S et al., 2013 or Lee ME et al., 2015.

3.6.2 Yeast strains, media, and growth conditions

The base *S. cerevisiae* strain used for experimentation was W303 or BY4741. Base strain for each engineered strain is noted in the strain list. From these base strains, knockout of endogenous transcription factors was done with a one-step replacement using a plasmid that contains 40 base pair overlaps in the 5' and 3' UTR of the transcription factor[79]. The 40 base pair overhangs flank the *Candida Albicans* HIS selectable marker.

Single colonies were picked from auxotrophic SD (6.7 g/L Bacto-yeast nitrogen base without amino acids, 2 g/L complete supplement amino acid mix, 20 g/L dextrose) agar plates. For microscopy and growth measurement studies, colonies were picked into 1 ml SDC media. For flow cytometry studies, colonies were picked into 1 ml YPD (yeast extract, peptone, 2% glucose) or SDC (6.7 g/L Bacto-yeast nitrogen base without amino acids, 2 g/L complete supplement amino acid mix, 20 g/L dextrose) media. Colonies were

grown overnight from 30°C to saturation. Prior to the start of an experiment, cells were diluted into 1-3 ml of SDC and grown for 4 hours to an OD of 0.05-0.1 prior to the start of an experiment. A TECAN Spark 10M plate reader (TECAN, Mannedorf, Switzerland) was used for growth measurements.

3.6.3 Microscopy and blue light delivery

Cells were imaged in 96-well Matriplates (MGB096-1-2-LG-L; Brooks Life Science Systems, Spokane, WA). For widefield microscopy, blue light optogenetic stimulation of samples was done using a custom built "optoPlate" as described in Bugaj et al (Bugaj et al., 2018). Individually addressable LEDs (in 96-well format) were controlled by an Arduino Micro microcontroller and programmed with different dynamic light patterns using custom Arduino scripts. Custom adapters for fitting optoPlate on to 96-well matrix plates were designed in AutoCad and 3D printed. For confocal microscopy, blue light stimulation was done using GFP laser illumination. A Nikon Ti inverted scope, with mercury arc-lamp illumination using RFP (560/40 nm excitation, 630/75 nm emission; 572/35 nm excitation, 632/60 nm emission; both manufactured by Chroma, Bellows Falls, VT) and near-infrared FP (640/30 nm excitation, 700/35 nm emission; Chroma, Bellows Falls, VT) filters, was used for widefield microscopy imaging of samples. Images were taken with an Andor EMCCD camera. Automated imaging was controlled and coordinated by custom Matlab (MathWorks, Natick, MA) software interfaced with the μmanager software suite (Edelstein et al., 2010). Confocal microscopy of samples took place on a Nikon Ti inverted scope with a Yokogawa CSU-22 spinning disk confocal scanner unit; cells were excited with laser illumination for Cy3 (561 nm, 100 mW Coherent OBIS; ET610/60nm emission filter) and Cy5 (640 nm,100 mW Coherent OBIS; ET700/75nm emission filter). Imaging was controlled with Nikon Elements 5.02 build 1266 (Nikon Instruments, Melville, NY).

3.6.4 Quantification of nuclear localization

The software tool "ilastik" was used for image segmentation to determine nuclear occupancy[80]. Time lapse images of the IRFP nuclear marker were used to identify nuclei objects. Nuclear occupancy of each nucleus was defined as the mean pixel intensity of this nucleus. Cell tracking of nuclear/cytoplasmic enrichment was done using automated yeast cell tracking software implemented in Matlab[81]. Photobleaching was corrected for cells that underwent constant illumination and frequent imaging. This correction was done by fitting an exponential decay function to each nuclear and cytoplasmic trace and then dividing each trace by its decay function. For all microscopy analysis, "nuclear/cytoplasmic enrichment" represented the mean pixel intensity of the nucleus divided by the mean pixel intensity of the cytoplasm, and "fold change" represented division by the value of the signal at t=0. Nuclear localization duration was defined as the time from light ON to the time when the nuclear/cytoplasmic enrichment has returned to 75% of the starting (t=0) value. Normalized peak-trough difference was quantified across all pulses for single cell traces, and represented the difference between the local maximum (the peak) and the local minimum (the trough) values divided by the maximum peak-trough difference in the population-averaged traces.

3.6.5 Flow cytometry

Analysis of fluorescent protein reporter expression was performed with a BD LSRII flow cytometer (BD Biosciences) equipped with a high-throughput sampler. For steady-state measurements, cultures were diluted in TE before running through the instrument. Cultures were run on the instrument 1 hour (+/- 20 min) after optical stimulation using the optoPlate, to allow for YFP maturation. YFP (Venus) fluorescence was measured using the FITC channel and RFP (mCherry/mScarlet) was measured using the PE-Texas

Red channel. For steady-state measurements, a maximum of 10,000 events were collected per sample. Fluorescence values were calculated using the height (H) measurement for the appropriate channel and normalized to cell size by dividing by side scatter (SSC-H). All analysis of flow cytometry data was performed in Python 2.7 using the package FlowCytometryTools and custom scripts.

3.6.6 Growth assays

Growth was measured using a TECAN Spark 10M plate reader (TECAN, Mannedorf, Switzerland) using 600nm excitation. Cultures were plated into Corning 3904 96-well assay plates (Corning, Corning, NY) and grown at 30°C while shaking until saturation. Data was analyzed in Python 2.7 or Matlab using custom scripts. To quantify log phase growth rate, only the OD600 measurements which were between .1 and 1 for each strain were used. A linear regression was then fit to the natural logarithm of the log phase OD600 values (as y) and time (as x). The slope from this regression was plotted as the log phase growth rate.

3.6.7 Computational modeling

Ordinary differential equation (ODE) models of gene expression focusing on promoter kinetics were constructed. For the simple kinetic model that described an efficient response to short pulses, a model was constructed with three state variables and seven parameters. Nine models were constructed and tested for efficient response to continuous pulses . These models either contained three or five state variables with up to ten parameters. Latin hypercube sampling was done to randomly sample parameters. ODE solvers 45 and 113 in Matlab were used. Least squared error and fit within the error bars of the data were metrics used to obtain model fits. Additional details of the computational methods are described in the Supplementary Text.

3.6.8 Treatment with $CaCl_2$ stress

Cells were grown at 30°C in YPD medium to saturation overnight. Cells were then diluted prior to the start of an experiment and grown for 4 hours to an OD of 0.05-0.1. For microscopy experiments, cells were plated in SDC with concanavalin A (conA) for 15 minutes to adhere them to the bottom of the glass imaging plate. Prior to imaging, the SDC was removed and replaced with a solution of SDC with 0.2M $CaCl_2$. For flow cytometry experiments, cells in SDC were diluted to OD 0.1 in a media of SDC with 0.2M $CaCl_2$ and grown in the media for the duration of the experiment. Prior to measurement, the 0.2M $CaCl_2$ media was removed by centrifugation with 3 washes in 1X TE.

3.7 Supplementary Text

3.7.1 Experimental Controls - Measuring the basal and constitutively nuclear gene expression of TFs

To further assess how TF-CLASP-induced expression compares to endogenous gene expression, we measured the level of reporter gene expression when the TFs were constitutively localized to the nucleus by C-terminally tagging them with the same NLS used in yeLANS (TF-NLS), or in their basal localization by C-terminally tagging them with only mScarlet. All TF-NLS, TF-mScarlet, and TF-CLASP constructs were expressed from pRPL18b. We compared this value to expression achieved when TF-CLASP was induced with 2 hours of blue light. SynTF-CLASP achieved 52% of pSYNTF-YFP expression produced through constitutive nuclear localization of SynTF (Supplementary Figure 3.8B). Furthermore, the mean SynTF-CLASP-induced gene expression in the dark (.09) was similar to the the mean basal gene expression in a strain in which the SynTF was only tagged with mScarlet (.07) (Supplementary Figure 3.8B). Pho4-CLASP activated pPHO84-YFP to 14% of the gene expression achieved with constitutive nuclear localization (Supplementary Figure 3.8C) while Msn2-CLASP was more efficient at inducing pHSP12-YFP gene expression than constitutive Msn2 nuclear localization (23% greater expression, Supplementary Figure 3.8D). Since Msn2 is subject to faster degradation in the nucleus (Chi et al., 2001; Durchschlag et al., 2004), transient localization with CLASP may be more efficient at inducing gene expression. For both pHSP12-YFP and pPHO84-YFP, reporter expression in the dark was lower in a strain that had either Msn2-CLASP or Pho4-CLASP than in their respective controls with either Msn2 or Pho4 when only tagged with mScarlet (28% and 88% lower, respectively). In fact, pPHO84-YFP showed basal bimodal expression in the constitutively expressed Pho4 strain, but not in the Pho4-CLASP strain (Supplementary Figure 3.8C). These data suggest that CLASP can

potently sequester TFs in the dark.

3.7.2 Modeling

Model equations and sampling details of the pYPS1-YFP and pCMK2-YFP phenotypes

Simple Two-State Promoter Model This model described the efficient response of pYPS1-YFP to short pulses. The model described a two-state promoter that activates mRNA production which then activates protein production and is depicted in Figure 3.4A. We modeled these interactions by:

$$\frac{dp_{on}}{dt} = k_{on} \cdot TF \cdot (1 - p_{on}) - k_{off} \cdot p_{on} \tag{3.1}$$

$$\frac{dmRNA}{dt} = \beta_0 + \beta_1 \cdot p_{on} - \gamma_1 \cdot mRNA \tag{3.2}$$

$$\frac{dProtein}{dt} = \beta_2 \cdot mRNA - \gamma_2 \cdot Protein \tag{3.3}$$

In these equations, p_{on} represented promoter activity while mRNA and Protein represented concentration of mRNA and protein, respectively. Here we assumed that promoter activity was conserved such that $1 = p_{on} + p_{off}$. TF represented the concentration of nuclear transcription factor.

The model was characterized by 7 parameters. Most of the activation/inactivation and production/degradation terms were modeled by first-order mass action kinetics. The parameter, β_0, was zeroth-order, to reflect basal promoter activity. We chose this simple

model form because we were interested in a parsimonious model that could explain the experimental phenotype of pYPS1-YFP. Note that the rate of promoter activation was dependent on TF concentration because Crz1 has been shown to activate genes through binding of a known promoter element, the calcineurin-dependent response element (CDRE), through its zinc finger domain[7].

The input to the model was the concentration of nuclear transcription factor (TF), while the output represented protein concentration (Protein). The equations were numerically solved by the ODE solver ode113 for nonstiff differential equations via MATLAB.

Parameters k_{on}, k_{off}, β_0,and β_1 were sampled over 4-5 orders of magnitude systematically and randomly using Latin Hypercube Sampling (LHS). k_{on} and k_{off} varied from 1e-4 to 1. β_0 was varied from 1e-6 to 1e-2. β_1 was varied from 1e-4 to 10. Parameters $\beta_2 = 0.06$,1 = 0.05, and $\gamma_2 = 0.0083$ were fixed to values according to literature[21,64].

Parameter Search and Model Fitting From the parameter sets sampled, the slope ratio (defined in Figure 3.3), a summary metric for the degree of efficiency in response to short pulses, was calculated for each parameter set. These slope ratio values were plotted as heatmaps as a function of the parameters k_{on} and k_{off} to demonstrate the effect of parameters on efficient response to short pulses (slope ratio > 1) or efficient response to continuous pulses (slope ratio < 1).

The model was used to fit the experimental data. Fits of the experimental data to the simple two-promoter state model (Figure 3.4) were obtained by the following procedure:

1. Sample 10^4 parameters randomly using LHS with the aforementioned parameter ranges

2. The model outputs are used to construct the Output-Occupancy plots and compared to the experimental Output-Occupancy plots. Fits are determined to be model

outputs that are within error of the data for >= 9 of the 10 data points. If no quantitative fit can be obtained (as in the pCMK2 case), the best 100 model outputs with the lowest mean squared error to the best fit line of the experimental data are determined to be fits.

3. Cross-validation of parameter fits to the Output-Occupancy data are then performed using the dose response. The mean squared error is used as the metric to assess fit.

From this procedure, we identified parameter sets that fit all of the experimental data for both pYPS1-YFP and pCMK2-YFP.

Model exploration and sampling details for the pGYP7-YFP phenotype: List of models

Two-state models with r_{off} and r_{on} thresholding These models involved a TF concentration gated activation, r_{on}, or inactivation, r_{off}, rate constant. The kinetic model with r_{off} thresholding consisted of equations (1)-(3), but utilized the following equation instead of (1):

$$\frac{dp_{on}}{dt} = r_{on} \cdot TF \cdot (1 - p_{on}) - r^*_{off} \cdot p_{on} \tag{1a}$$

where $r*_{off} = 0$ when TF $<$ *threshold*. Otherwise $r*_{off} = r_{off}$. *threshold* is a parameter value which denoted the threshold TF concentration at which $r*_{off}$ switches from 0 to r_{off}.

The kinetic model with r_{on} thresholding consisted of equations (1)-(3), and utilized the following equation in place of (1):

$$\frac{dp_{on}}{dt} = r^*_{on} \cdot TF \cdot (1 - p_{on}) - r_{off} \cdot p_{on} \tag{1b}$$

where $r^*_{on} = r_{on}$ when TF $\geq$ thr. Otherwise $r^*_{on} = 0$.

These two models could represent a binary interaction of the transcription factor (TF) with promoter elements, where below a TF concentration, the TF had no effect on promoter activity and above a TF concentration, the promoter turned on at its maximal rate.

The same parameter ranges were sampled in this model as in the simple kinetic model. The additional parameter *threshold* was sampled randomly from TF = 0 to 2.7, the maximum value of the TF input to the model.

Cooperative Model. Similar to the two-state thresholded models, the cooperative model described a nonlinear relationship between TF concentration and protein output. The model was represented by the equations:

$$\frac{dmRNA}{dt} = \beta_0 + \beta_1 \cdot \frac{TF^n}{TF^n + k_d^n} - \gamma_1 \cdot mRNA \tag{3.4}$$

$$\frac{dProtein}{dt} = \beta_2 \cdot mRNA - \gamma_2 \cdot Protein \tag{3.5}$$

where n is the hill coefficient, and kd = $\frac{k_{off}}{k_{on}}$.

The same parameter ranges were sampled in this model as in the simple kinetic model. The additional parameter n was sampled randomly from n = 0.5 to 4, a biologically

relevant range[21].

3-State Models. We considered five 3-state models with different relationships of TF and the rate constants for the transition between promoter states. The first such model was a **3-state model**. An additional promoter state, p_0, is added. The rate equations describing this model are:

$$\frac{dp_0}{dt} = r_{off} \cdot p_{off} - r_{on} \cdot p_0 \tag{3.6}$$

$$\frac{dp_{off}}{dt} = k_{off} \cdot p_{on} + r_{on} \cdot p_0 - (r_{off} + k_{on} \cdot TF) \cdot p_{off} \tag{3.7}$$

$$\frac{dp_{on}}{dt} = k_{on} \cdot TF \cdot p_{off} - k_{off} \cdot p_{on} \tag{3.8}$$

$$\frac{dmRNA}{dt} = \beta_0 + \beta_1 \cdot p_{on} - \gamma_1 \cdot mRNA \tag{3.9}$$

$$\frac{dProtein}{dt} = \beta_2 \cdot mRNA - \gamma_2 \cdot Protein \tag{3.10}$$

In this model, the p_0 and p_{off} could be thought of as non-transcribing promoter states that represented nucleosome occluded and open, respectively. p_{on} represented an active transcribing promoter state. The rate constants r_{off} and r_{on} described the transitions between the occluded and open promoter states. The promoter underwent a transition between p_0 and p_{off} (with rate constants, r_{on} and r_{off} respectively), but p_{off} was still not a promoter state conducive for transcription. A second transition from p_{off} to p_{on} (with rate constants, k_{on} and k_{off}, respectively) was needed to start transcription.

3-state model with r_{off} threshold but no linear TF dependence on k_{on}. This model was the same as the 3-state model with r_{off} threshold, except no linear dependence of the TF in the transition from p_{off} and p_{on}.The model was described by equations (6)-(10); the following equations replaced equations (6) and (7):

$$\frac{dp_0}{dt} = r_{off} \cdot p_{off} - r_{on} \cdot TF \cdot p_0 \tag{6d}$$

$$\frac{dp_{off}}{dt} = k_{off} \cdot p_{on} + r_{on} \cdot TF \cdot p_0 - (r_{off} + k_{on} \cdot TF) \cdot p_{off} \tag{7d}$$

3-state model with r_{on} threshold. This model contained a TF concentration threshold dependence of r_{on} between the occluded p_0 and open p_{off} promoter states, and was described by equations (6) - (10); where model equations (6) and (7) were replaced by:

$$\frac{dp_0}{dt} = r_{off} \cdot p_{off} - r^*_{on} \cdot p_0 \tag{6b}$$

$$\frac{dp_{off}}{dt} = k_{off} \cdot p_{on} + r^*_{on} \cdot p_0 - (r^*_{off} + k_{on} \cdot TF) \cdot p_{off} \tag{7b}$$

where $r^*_{on} = r_{on}$ when TF $\geq$ *threshold*. Otherwise $r^*_{on} = 0$.

For this model, the transcription factor modulated the rate of transition from p_0 to p_{off} such that when TF concentration reached the threshold concentration, *threshold*, the transition rate switches from zero to a value. Biologically, this model could represent TF interaction with chromatin acetylators and other modifiers that could promote an open chromatin conformation on the promoter.

3-state model with r_{off} threshold. Similarly, the model with a threshold dependence on the inactivating transition from the open p_{off} to occluded p_0 states is described by the model equations (6)-(10); where equations (6) and (7) are replaced by:

$$\frac{dp_0}{dt} = r^*_{off} \cdot p_{off} - r_{on} \cdot p_0 \tag{6c}$$

$$\frac{dp_{off}}{dt} = k_{off} \cdot p_{on} + r_{on} \cdot p_0 - (r^*_{off} + k_{on} \cdot TF) \cdot p_{off} \tag{7c}$$

where $r^*_{off} = 0$ when TF $\geq$ *threshold*. Otherwise $r^*_{off} = r_{off}$.

For this model, the transcription factor modulated the rate of transition from p_{off} to p_0 such that when TF concentration reached the threshold concentration, *threshold*, the transition rate switched from a value to zero. Biologically, this model could represent either physical hindrance of heterochromatin formation or TF-modulated repression of a chromatin de-acetylase. The parameters r_{on} and r_{off} were randomly sampled in the range from 10e-4 to 1.

3-state model with linear TF dependence of r_{on} and k_{on}. This model contained a linear dependence on TF for the transitions between both p_0 to p_{off} and p_{off} to p_{on}. This model was described by the model equations (6)-(10); where equations (6) and (7) were replaced by:

$$\frac{dp_0}{dt} = r^*_{off} \cdot p_{off} - r_{on} \cdot p_0 \tag{6a}$$

$$\frac{dp_{off}}{dt} = k_{off} \cdot p_{on} + r_{on} \cdot p_0 - (r^*_{off} + k_{on}) \cdot p_{off} \tag{7a}$$

Parameter Search and Model Fitting. Parameter search and model fitting were done in the same way as for modeling of pYPS1-YFP and pCMK2-YFP in the section above, except two rounds of fitting were done with the Output-Occupancy and dose response data

of pGYP7-YFP. Fits to the dose response were determined to be parameter sets whose least squared error was 0.8 standard deviations below the mean of the least squared error distribution. An experiment with a strain expressing pTEF1 driven Crz1*-CLASP exposed to short pulsed and continuous inputs, as described in the main text, was used to cross-validate the model fits.

Detailed exploration of model fits to the pGYP7-YFP data

The simple kinetic model that described the pYPS1-YFP and pCMK2-YFP phenotypes produced no parameter sets for which a pulsed input generated lower gene expression output than a continuous input (Supplementary Figure 3.11A-B). Hence, we explored model elaborations, introduced in the previous section, of the simple promoter switching model.

We first tested whether the **two-state models with either r_{off} or r_{on} thresholding** (Supplementary Figure 3.11E-F, G-H), or the **cooperative model** (Supplementary Figure 3.11C-D), could generate a promoter that responded efficiently to continuous pulses. The rationale here was that if the promoter spent some time below its threshold of activation for any input, then the effect of this TF concentration thresholding would be smaller for a continuous pulse that does this once, than for a sequence of short pulses where this would be done repeatedly. In agreement with this intuition, this suite of models was able to generate Output-Occupancy plots that mirrored the pGYP7-YFP experimental data for many parameters (Supplementary Figure 3.11C-D, E-G, G-H, left panel). Supplementary Figure 3.11G shows an illustrative example of this class of models, where many parameters sets (380) were shown to maximize fits through the data points within the error bars (Supplementary Figure 3.11G, left panel). Upon further fitting with independently-obtained dose response data for pGYP7-YFP (obtained in the same way as explained above for pYPS1-YFP and pCMK2-YFP), this model however failed to fit the

data, as did all models with only two promoter states (Supplementary Figure 3.11C-D, E-F, G-H, middle panel). The failure of these models to fit the pGYP7-YFP dose response showed a characteristic pattern – while the pGYP7-YFP dose response was linear in the TF regime we measured, the computationally predicted dose response was thresholded given the model structure we imposed (Supplementary Figure 3.11C-D, E-F, G-H, middle panel).

Next, we increased the complexity of the model by adding a second layer of promoter transitions to generate the **3-state model** (Supplementary Figure 3.11K). This model had the same linear structure as the two-state promoter model of Figure 3.4, and hence could not produce a more efficient response to continuous input over the pulsed one (Supplementary Figure 3.11L).

We also tested whether a **3-state model with r_{off} threshold but no linear TF dependence on k_{on}** could produce better fits (Supplementary Figure 3.11I). This model was indeed able to generate Output-Occupancy plots that match the pGYP7-YFP experimental data (423 parameter sets within error) (Supplementary Figure 3.11J, left panel), but with these parameters, it again produced a thresholded dose response that failed to fit that of pGYP7-YFP (Supplementary Figure 3.11J, middle panel).

Following these results, we reasoned that the introduction of a linear dependence on TF concentration in the immediate step before promoter activation could mitigate the effects of a threshold on an earlier promoter transition step, therefore producing a linear dose response. Hence, the **3-state model with r_{off} threshold** was tested (Figure 3.5A). With this addition, the model was able to generate Output-Occupancy plots that maximize fit to the experimental data for many parameters (96 parameter sets) and for a subset of those (25 parameter sets), was also able to recapitulate the pGYP7-YFP dose response (Figure 3.5B-C). The **3-state model with r_{on} threshold** was similarly able to recapitulate the data (Supplementary Figure 3.11O-P), albeit with a poorer fit for the Output-Occupancy plot. Finally, we tested the **3-state model with linear TF dependence of r_{on} and k_{on}**

(Supplementary Figure 3.11M). This model was also able to produce qualitative fits to the Output-Occupancy plot (Supplementary Figure 3.11N, left panel) and dose response data (Supplementary Figure 3.11N, middle panel).

To further test these three successful models and also further invalidate the discarded models, we subjected them to cross-validation using an independent experiment in a strain where Crz1*-CLASP expression was increased (now expressed from pTEF1 instead of pADH1). We subjected these cells to either short pulses (2 minutes ON/10 minutes OFF) or a continuous input (40 minutes of light) in a timespan of 4 hours, and measured pGYP7-YFP levels at 5 hours. These data revealed that the efficient response to continuous input was still preserved at the higher pTEF1 expression level. All discarded models (Supplementary Figure 3.11, right panels) were inconsistent with these data, predicting instead a reversal of the phenotype with an increase in the TF concentration. Notably, the **3-state model with linear TF dependence of r_{on} and k_{on}** (Supplementary Figure 3.11N, right panel) also failed this cross-validation because the dependence on TF caused the rate of transition from p_0 to p_{off} to increase with increased TF, and thus the efficient response to continuous inputs could only be produced for relatively low TF concentrations. Hence, only two minimal models were able to explain all the data we collected (Supplementary Figure 3.11O-P, Figure 3.5A-D).

3.7.3 Supplementary Methods

Delivery of stress inputs for microscopy

For each environmental perturbation, cells were grown overnight to saturation in YPD, diluted in prior to the experiment, and grown to an OD of 0.1. 200ul cells were plated with conA. Just before imaging, the SDC media was removed from the microscopy well and the appropriate environmental stress media was applied to the cells.The media for glucose

depletion consisted of 0.67% YNB w/o AA w/ ammonium sulfate, 0.79% CSM, 0.05% glucose. The media for Osmotic shock was composed of 0.67% YNB w/o AA w/ ammonium sulfate, 0.79% CSM, 2% glucose, and 0.95M sorbitol (Gasch et al., 2000). The phosphate depletion media contained 0.17% Pi-depleted YNB (without amino acids and ammonium sulfate), 0.1% ammonium sulfate, 2% glucose, 25mM sodium citrate (pH 4.7), 0.79% CSM.

RNAseq of Crz1 19A and 5A mutant

Single colonies were picked and grown to saturation in YPD at 30°C overnight. Cells were then diluted and grown for 4 hours to an OD of 0.3. Cells were harvested by centrifugation and frozen with liquid nitrogen. RNA was extracted using phenol chloroform (Sambrook and Russell, 2006). RNA quality was assessed using the Agilent RNA Pico kit. The Lexogen Quantseq 3' mRNA-Seq Library Prep Kit was used for RNA preparation. mRNA libraries were quantified using Qubit dsDNA HS Assay Kit and subject to single-end sequencing on an Illumina HiSeq 4000. Fastq files from illumina were aligned using STAR (Dobin et al., 2013). Downstream processing of read counts and differential gene expression was conducted using custom Matlab scripts.

Automated Flow cytometry

Cells were cultured as described above. Cells were inoculated and grown to an OD of 0.1 then diluted 1:10 for a total of 30mls for each reaction chamber. Cells were subjected to light input perturbations in the reaction chambers. Control of fluidics was achieved using LABView. The first 750 events of sample were discarded, and 2,000-10,000 events were collected per sample. Gene expression was measured using the FITC channel. Cytometer outputs were analyzed using custom matlab scripts. For more information on the automated flow cytometry hardware and LabView control of dynamic sample acquisition, see Harrigan et al., 2018.

3.8 SI Figures

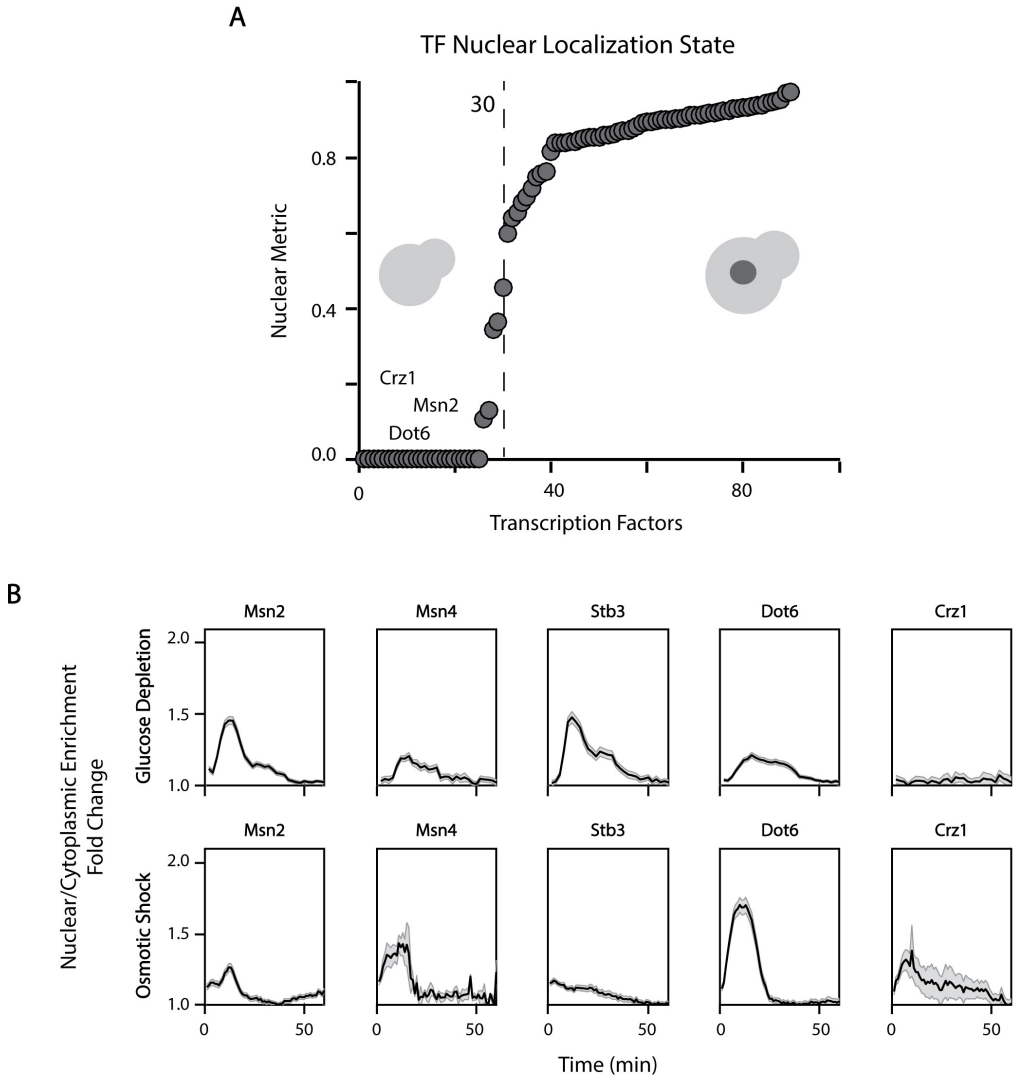


Figure 3.6: **Approximately one-third of TFs are basally cytoplasmic in log phase and a subset are shown to exhibit transient nuclear localization** A) LOC scores of available transcription factors from the CYCLoPs database are plotted[48]. The LOC score, as defined in Chong et al., 2015, is the number of cells assigned to a specific location (nucleus in this instance) over the total number of cells in any subcellular location. Increasing LOC score denotes increasing nuclear enrichment. B) Fold change of nuclear enrichment for a panel of stress-responsive transcription factors (Msn2, Msn4, Stb3, Dot6, and Crz1) are plotted as a function of time in response to environmental inputs (Glucose depletion and osmotic shock). For glucose depletion, SDC media (2% glucose) is replaced with SD media with 0.05% glucose. For osmotic shock, SDC media is replaced with SDC media with 0.95M sorbitol. Imaging begins at t = 0 after addition of environmental perturbation and samples are imaged every 30 seconds. The solid black lines represent the mean of single cell traces and the shading represents the standard error of the mean.

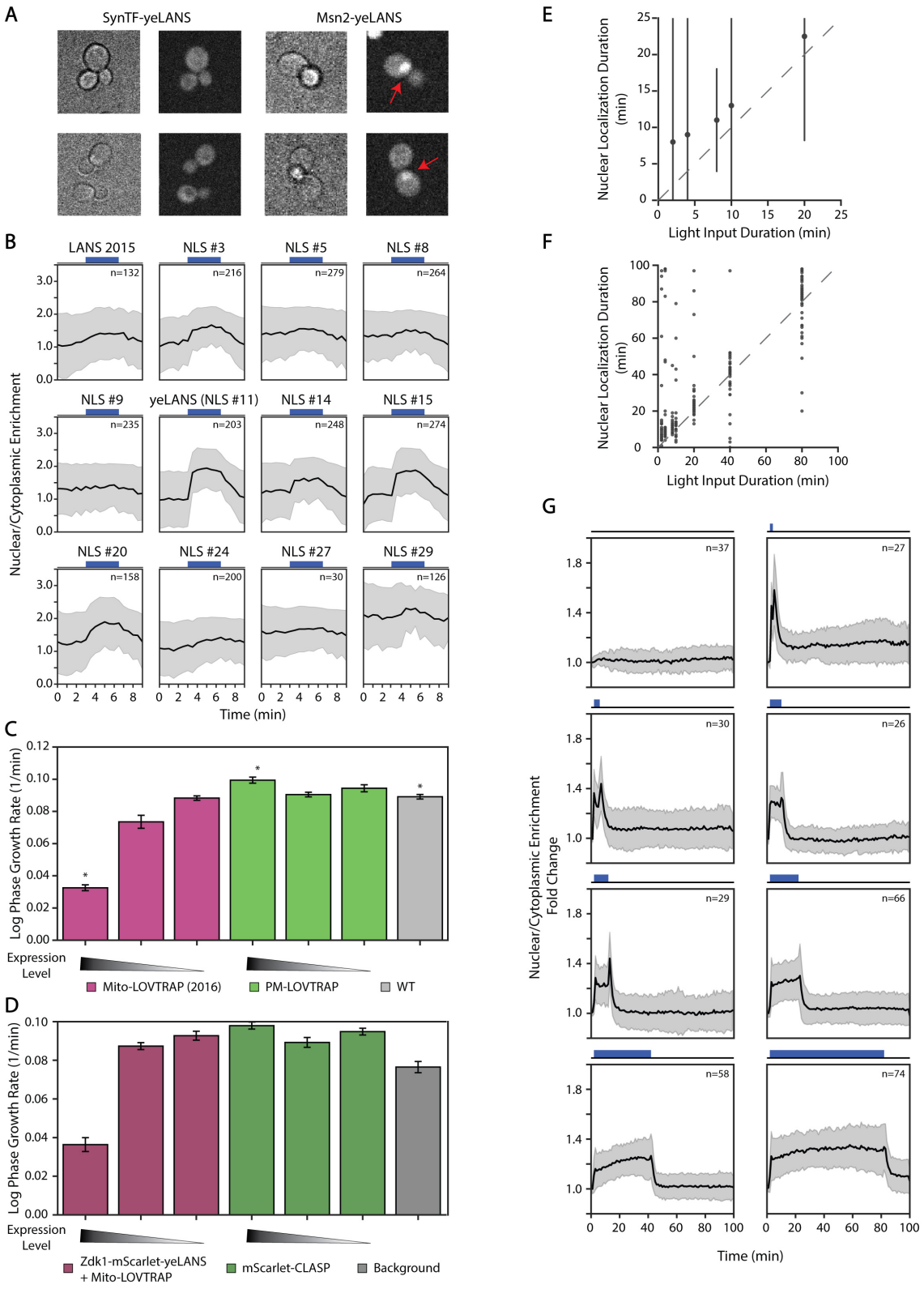

A
SynTF-yeLANS
Msn2-yeLANS
B
LANS 2015
NLS #3
NLS #5
NLS #8
NLS #9
yeLANS (NLS #11)
NLS #14
NLS #15
NLS #20
NLS #24
NLS #27
NLS #29
Nuclear/Cytoplasmic Enrichment
Time (min)
C
Log Phase Growth Rate (1/min)
Expression Level
Mito-LOVTRAP (2016)
PM-LOVTRAP
WT
D
Log Phase Growth Rate (1/min)
Expression Level
Zdk1-mScarlet-yeLANS + Mito-LOVTRAP
mScarlet-CLASP
Background
E
Nuclear Localization Duration (min)
Light Input Duration (min)
F
Nuclear Localization Duration (min)
Light Input Duration (min)
G
Nuclear/Cytoplasmic Enrichment Fold Change
Time (min)

Figure 3.7: **Optimization of LANS and LOVTRAP and CLASP characterization** A) Confocal microscopy images of yeast expressing SynTF-yeLANS and Msn2-yeLANS in the absence of blue light. Red arrows (inset) denote examples of cells that exhibit nuclear/cytoplasmic localization of Msn2-yeLANS. B) Mean nuclear/cytoplasmic enrichment (nuclear intensity divided by cytoplasm intensity) is plotted as a function of time. Shaded error represents standard deviation and light input regimes are illustrated above graphs. n refers to number of cells tracked and subplot headings (e.g., NLS 3) correspond to NLS peptides listed in Table S2. C) Comparison of Mito-LOVTRAP and PM-LOVTRAP strains. Mito-LOVTRAP and PM-LOVTRAP are expressed from pTDH3 (highest), pRPL18B (medium), and pREV1 (lowest) promoters. Strains marked with an asterisk denote those for which growth curves are plotted in Figure 3.1B of main text. Error bars represent standard error of the linear regression. D) Comparison of Zdk1-mScarlet-yeLANS + Mito-LOVTrap and CLASP. Both component (e.g. Mito-LOVTRAP and Zdk1-mScarlet-yeLANS) are expressed at the same level, using either pTDH3, pRPL18B, or pREV1 promoters. Background (control strain) denotes the WT strain with pSYNTF-YFP integrated in the LEU locus. Zdk1-mScarlet-yeLANS + Mito-LOVTRAP and mScarlet-CLASP strains also have this integration. Error bars represent standard error of the linear regression. E) A zoom in of Figure 3.1D in the main text that shows median duration of nuclear localization as a function of light input duration; the line X = Y is denoted by the dashed line. The zoomed graph illustrates that for short pulse durations, the OFF time – time that nuclear localization extends past the pulse – is not linearly related to light input duration. F) A scatterplot that shows duration of nuclear localization as a function of light input duration. Each point represents a single cell. G) Mean nuclear/cytoplasmic enrichment fold change as a function of time for mScarlet-CLASP induced with blue light. Light input regimes are illustrated above graphs (indicating 0, 2, 4, 8, 10, 20, 40, or 80 minute light input). In all plots, except where noted, error (bars or shading) represents standard deviation.

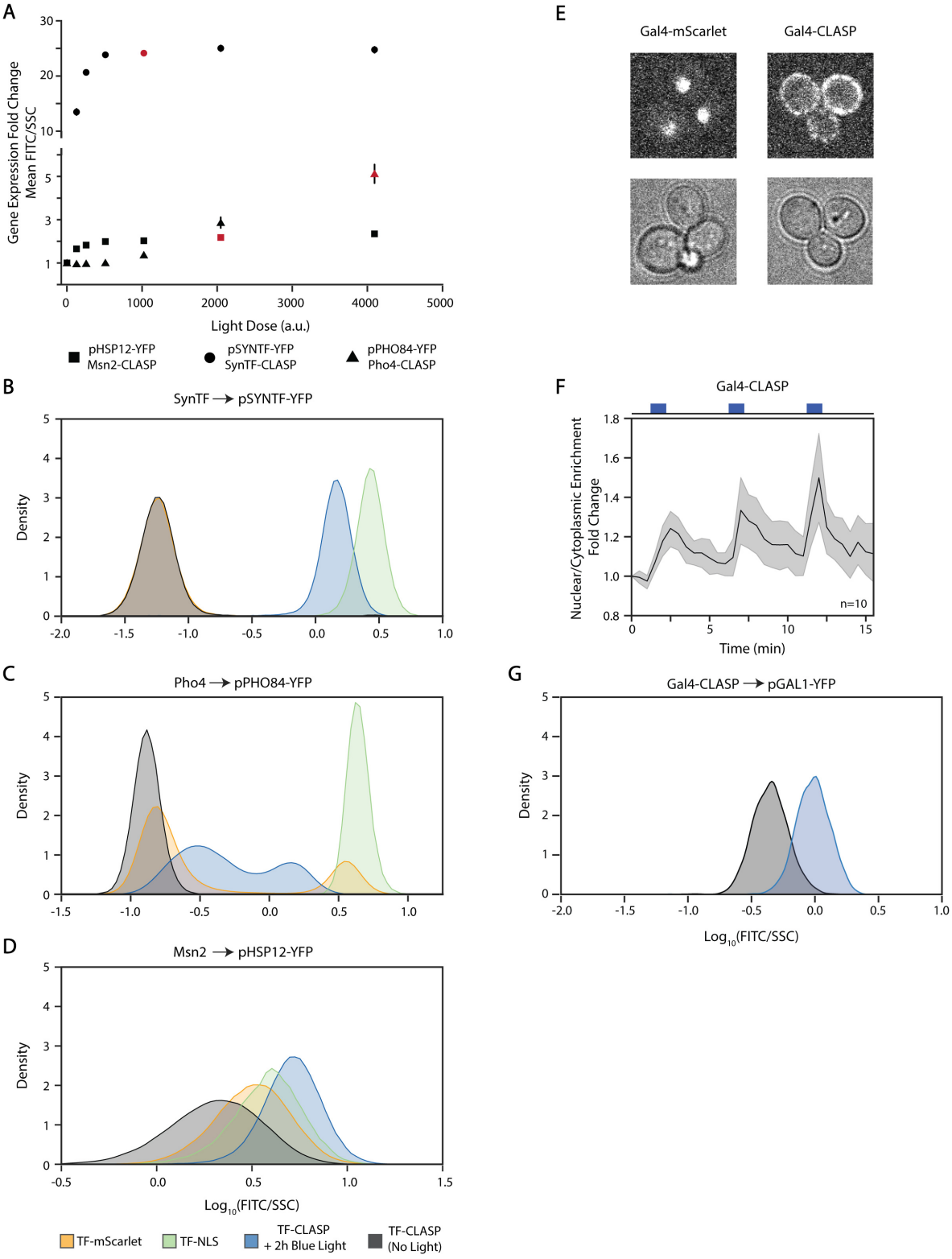

A
Gene Expression Fold Change
Mean FITC/SSC
Light Dose (a.u.)
pHSP12-YFP
Msn2-CLASP
pSYNTF-YFP
SynTF-CLASP
pPHO84-YFP
Pho4-CLASP
B
SynTF → pSYNTF-YFP
Density
C
Pho4 → pPHO84-YFP
Density
D
Msn2 → pHSP12-YFP
Density
Log10(FITC/SSC)
TF-mScarlet
TF-NLS
TF-CLASP
+ 2h Blue Light
TF-CLASP
(No Light)
E
Gal4-mScarlet
Gal4-CLASP
F
Gal4-CLASP
Nuclear/Cytoplasmic Enrichment
Fold Change
n=10
Time (min)
G
Gal4-CLASP → pGAL1-YFP
Density
Log10(FITC/SSC)

Figure 3.8: **Characterization of TF-CLASP strains** A) Mean FITC/SSC is plotted as a function of light intensity (a.u.) for strains that are exposed to different amplitudes of light for two hours (continuous input). Marked in red is the lowest light dose which yielded near-maximal expression for each strain (>90%); this dose is used in all microscopy and flow cytometry experiments for each strain. Error bars represent standard error of the mean for 9 biological replicates. B-D) Each subplot shows the probability density functions of log10(FITC/SSC) of gene expression of corresponding fluorescent promoter fusion for TF-CLASP, TF-NLS (constitutive nuclear localization), and TF-mScarlet (basal localization) strains. Histograms display expression from 9 biological replicates (data from replicates are pooled). TF-mScarlet strains are not exposed to light, for facile comparison to TF-CLASP (No Light) expression. TF-NLS strains are exposed to two hours of blue light (continuous input) to control for the effect of blue light on YFP fluorescence when comparing to TF-CLASP (Light) expression. For all panels, TF cargos are expressed from pRPL18B. E) RFP (top panels) and brightfield (bottom panels) images of mScarlet-tagged Gal4 (left panels) and mScarlet-tagged Gal4-CLASP (right panels). F) Gal4 nuclear enrichment is plotted as a function of time. Light input regime is illustrated above graph. Shaded gray area represents 95% confidence interval. n refers to the number of cells tracked. G) Gene expression of pGal1-YFP resulting from Gal4-CLASP nuclear localization following two hours of blue light input. Shown is the probability density function of log10(FITC/SSC) of pGAL1-YFP in the dark (gray) or after light exposure (blue).

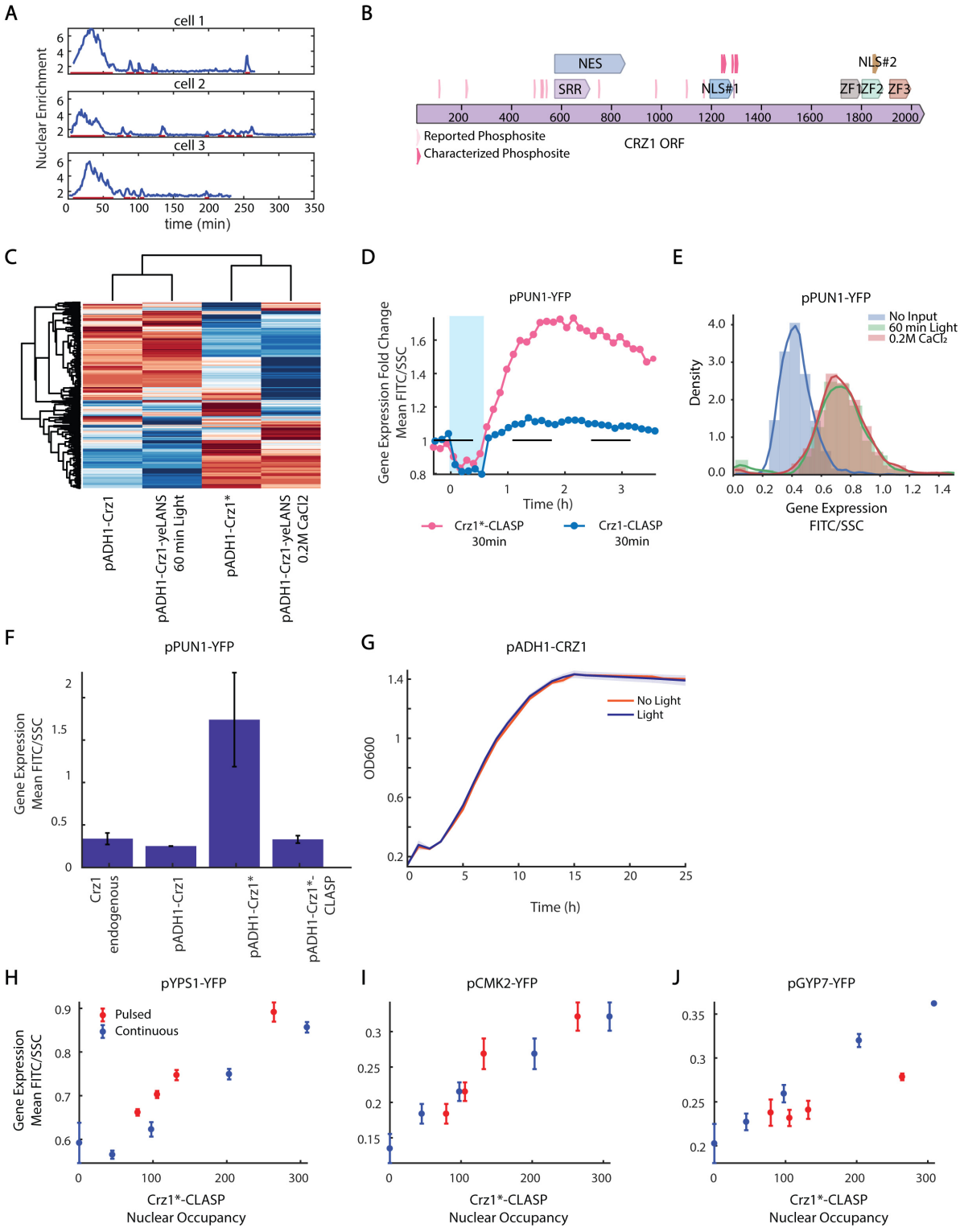

A
cell 1
cell 2
cell 3
Nuclear Enrichment
time (min)
B
NES
SRR
NLS#1
NLS#2
ZF1 ZF2 ZF3
Reported Phosphosite
Characterized Phosphosite
CRZ1 ORF
C
pADH1-Crz1
pADH1-Crz1-yeLANS 60 min Light
pADH1-Crz1*
pADH1-Crz1-yeLANS 0.2M CaCl2
D
pPUN1-YFP
Gene Expression Fold Change Mean FITC/SSC
Time (h)
Crz1*-CLASP 30min
Crz1-CLASP 30min
E
pPUN1-YFP
No Input
60 min Light
0.2M CaCl2
Density
Gene Expression FITC/SSC
F
pPUN1-YFP
Gene Expression Mean FITC/SSC
Crz1 endogenous
pADH1-Crz1
pADH1-Crz1*
pADH1-Crz1*-CLASP
G
pADH1-CRZ1
No Light
Light
OD600
Time (h)
H
pYPS1-YFP
Pulsed
Continuous
Gene Expression Mean FITC/SSC
Crz1*-CLASP Nuclear Occupancy
I
pCMK2-YFP
Crz1*-CLASP Nuclear Occupancy
J
pGYP7-YFP
Crz1*-CLASP Nuclear Occupancy

Figure 3.9: **Characterization of Crz1, Crz1-CLASP, and Crz1*-CLASP nuclear translocation and gene expression with $CaCl_2$ or blue light input** A) Single cell traces of Crz1 nuclear enrichment over time for 3 representative cells following 0.2M $CaCl_2$. The red lines indicate nuclear localization events. B) Schematic of the CRZ1 Open Reading Frame (ORF). Labeled are the Nuclear Localization Sequences (NLS1 and NLS2) and the Nuclear Exit Sequence (NES), as well as the Serine-Rich Region (SRR), which is calcium responsive. The light pink triangles denote reported S/T phosphosites, while the dark pink triangles denote reported and characterized S/T phosphosites. The 19 dark and light pink phosphosites are mutated from S/T -¿ A to construct Crz1*. C) Heatmap of gene expression for 5657 genes. Samples in each column of the heatmap are pADH1-Crz1 with no input, pADH1-Crz1-yeLANS with 60 minutes of light, pADH1-Crz1* with no input, and pADH1-Crz1-yeLANS with 0.2M $CaCl_2$ delivered at the start of the experiment. All samples are in log phase and all measurements are taken 60 minutes after delivery of input. D) Gene expression (Mean FITC/SSC) of the Crz1 reporter gene pPUN1-YFP driven by either Crz1-CLASP (blue) or Crz1*-CLASP (pink) when given 30 minutes of blue light. E) Probability density functions of gene expression of pPUN1-YFP, measured by FITC/SSC, in response to 0.2M $CaCl_2$ (which causes an initial Crz1 nuclear localization pulse of 40-60 minutes), 60 minutes of blue light exposure, and no input. Measurements are taken at 5 hours after delivery of input. F) Basal gene expression of pPUN1-YFP for different Crz1 strains: endogenous Crz1, pAdh1-Crz1 in a Crz1 KO background, pAdh1-Crz1* in a Crz1 KO background, and pAdh1-Crz1*-CLASP (without light input) in a Crz1 KO background. G) OD600, a measurement for growth, plotted as a function of time, for pAdh1-Crz1 with (blue) and without (red) light input (intensity 512 a.u.) over a period of 24 hours, indicating that light exposure does not affect population growth. Measurements are taken every hour. H) Characterization of additional Crz1*-CLASP gene expression in response to blue light, as in Figure 3.3. Output-Occupancy plot for pYPS1-YFP. I) Output-Occupancy plot for pCMK2-YFP. J) Output-Occupancy plot for pGYP7-YFP. These data are biological replicates taken from different days for the data shown in Figure 3.3.

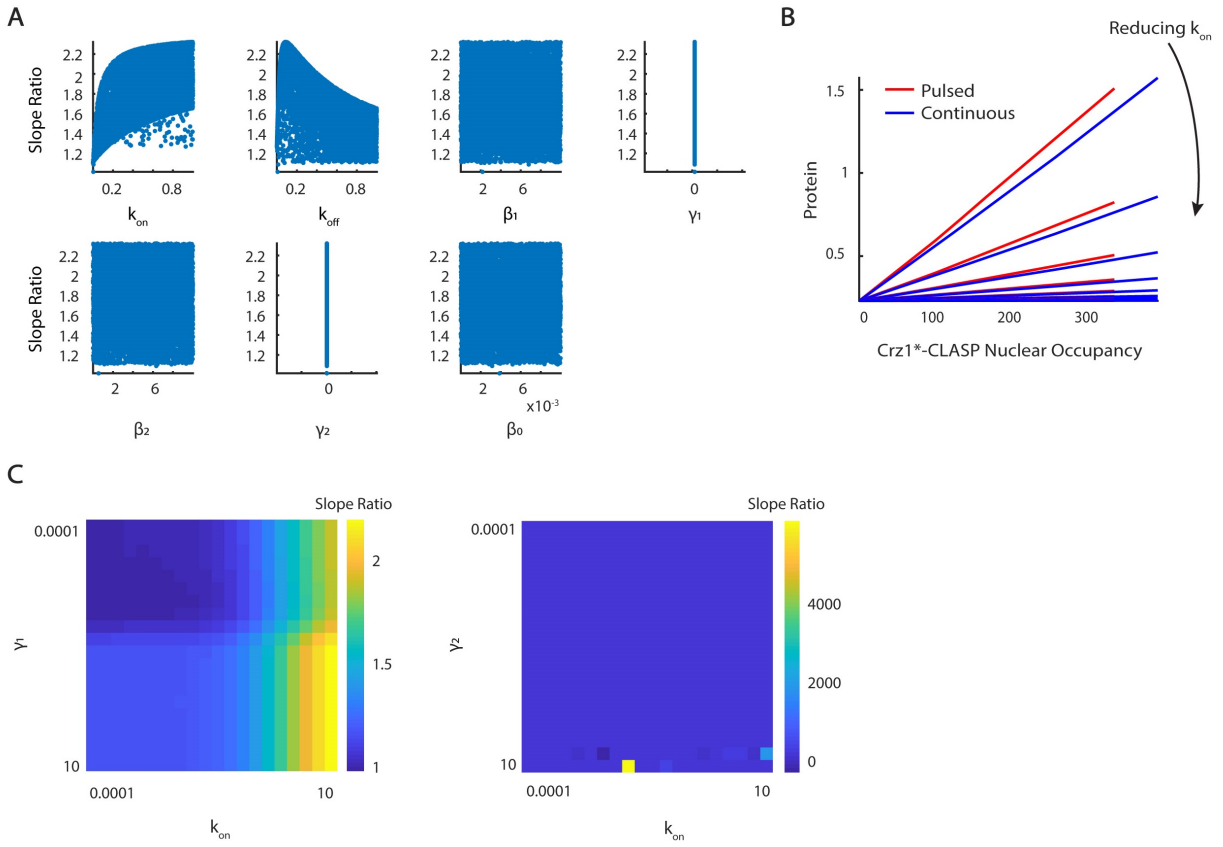


Figure 3.10: **Characterization of two-state promoter model of Figure 3.4 in main text** A) Slope ratio as a function of parameter values for each sampled parameter. Each blue circle represents a parameter set in the set of 10000 parameter sets searched. k_{on} varies from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_2 from 0.0001-10 and β_0 from 0.000001-0.01. γ_1 is set to 0.05 and γ_2 to 0.0083. B) Output-Occupancy plot for a parameter set with efficient response to short pulses. Parameter values are: $k_{on} = 1$, $k_{off} = 0.8$, $\beta_1 = 0.0001$, $\beta_2 = 0.1$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$, and $\beta_0 = 0.000001$. k_{on} is reduced 2x, 4x, 8x, 16x, and 32x. The red line represents the gene expression values from increasing pulsed inputs and the blue line represents gene expression values from increasing constant inputs. C) (left panel) Heatmap of slope ratios in the γ_1-k_{on} plane. Parameters are sampled (1 from 0.0001-10, k_{on} from 0.0001-10) or set ($k_{off} = 1$, $\beta_1 = 0.0001$, $\beta_2 = 0.1$, $\gamma_2 = 0.0083$, $\beta_0 = 0.000001$). (right panel) Heatmap of slope ratios resulting from the kinetic model in the γ_2-k_{on} plane. Parameters are sampled (2 from 0.0001-10, k_{on} from 0.0001-10) or set ($k_{on} = 0.001$, $\beta_1 = 0.0001$, $\beta_2 = 0.1$, $\gamma_1 = 0.05$, $\beta_0 = 0.000001$).

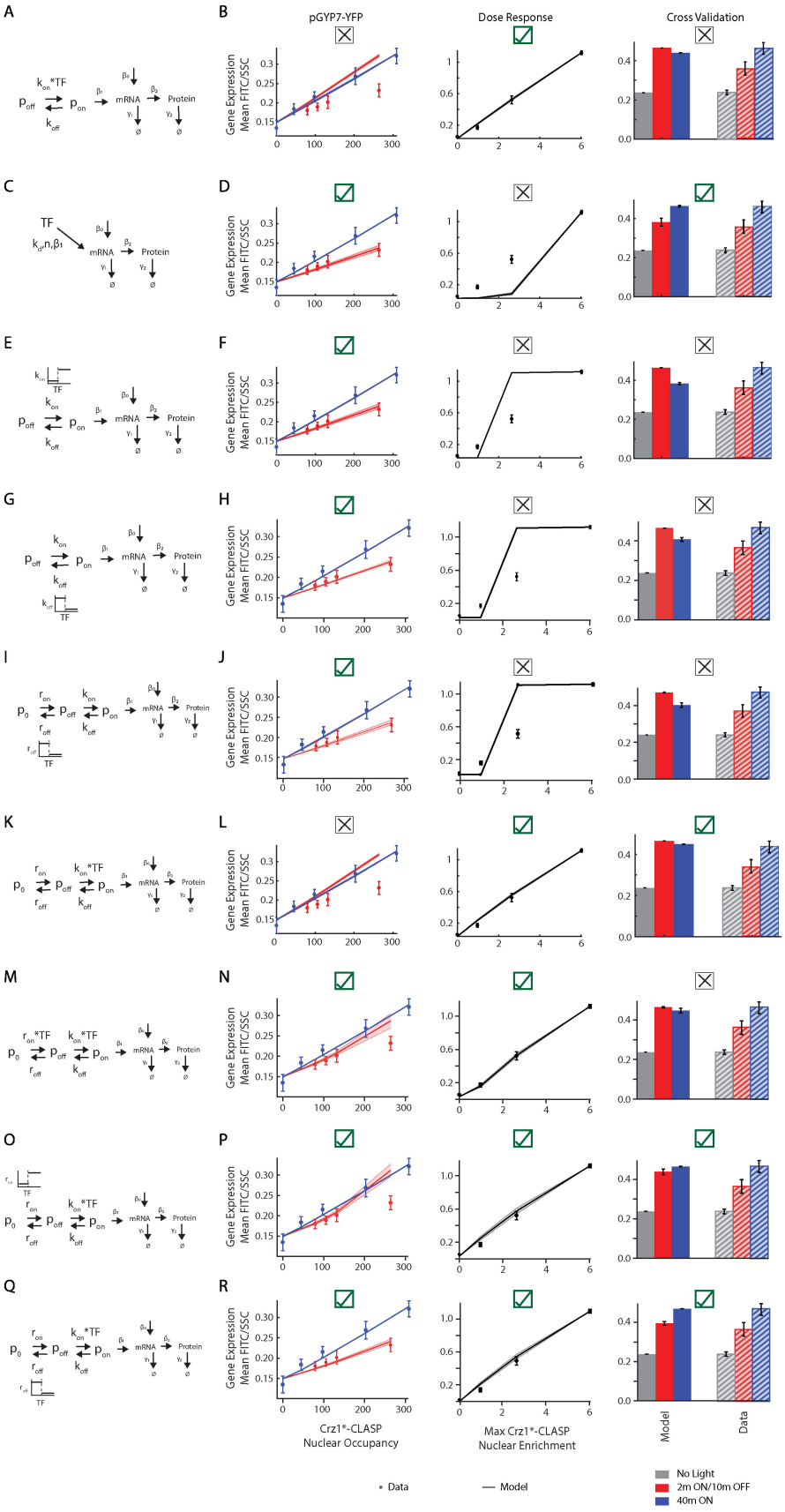

A
B
pGYP7-YFP
Dose Response
Cross Validation
Gene Expression
Mean FITC/SSC
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
Crz1*-CLASP
Nuclear Occupancy
Max Crz1*-CLASP
Nuclear Enrichment
Model
Data
Data
Model
No Light
2m ON/10m OFF
40m ON

Figure 3.11: **Exploration of various models for pGYP7-YFP data** A) Schematic of the kinetic model, where the input is Crz1*-CLASP nuclear localization (TF) and the output is fluorescent protein level (Protein). B) (left panel) Output-Occupancy plot for pGYP7-YFP. Circles are experimentally measured values while lines denote the output of the model for 200 parameter sets out of 10000 that maximize fits through data points. The solid line denotes the mean and shaded areas the standard deviation of the model outputs. Parameters are sampled (k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel) Dose response plot for pGYP7. The parameters that fit the Output-Occupancy data are used to further fit the dose response of pGYP7-YFP using a best fit to least squared error criterion. Parameter sets below the mean of the least squared error distribution are plotted (solid black line is the mean generated by the model). The gray dots are the experimentally measured dose response. (right panel) The parameters that fit the Output-Occupancy are then subject to cross-validation using an experiment where Crz1*-CLASP expression is increased (expressed from a pTEF1 promoter), and cells are exposed to either short (2 minutes ON/10 minutes OFF) or continuous input (40 minutes of light). The model generated outputs (solid red and blue bar) are plotted with the experimental data (hashed red and blue bar). The gray bars are samples not exposed to light. C) Schematic of a model with cooperativity D) (left panel) Same plots as in (B), with 2481 parameter sets for this model. Parameters are sampled (kd from 0.01-100, n from 0.5-4, β_1 from 0.0001-10, β_0 from 0.000001-0.01) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 35 parameter sets. E) Schematic of a 2-state model with thresholding on the activation constant, r_{on}. F) (left panel). Same plots as in (B), with 148 parameter sets for this model. Parameters are sampled (r_{on} from 0.1-100,r_{off} from 0.1-100, β_1 from 0.0001-10, β_0 from 0.000001-0.01) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$, threshold = 0.5). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 16 parameter sets. G) Schematic of a two-state promoter model with a thresholded promoter inactivation constant, r_{off}, where the input is Crz1*-CLASP nuclear localization (TF) and the output is fluorescent protein level (Protein). H) (left panel) Same plots as in (B), with 380 parameter sets for this model. Parameters are sampled (r_{on} from 0.0001-1, r_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01, threshold from 0-2.7) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 52 parameter sets. I) Schematic of a 3-state model with thresholding in the inactivation constant, r_{off}, between the promoter off-states, p_0 and p_{off}, and no TF dependence in the step before promoter activation. J) (left panel) Same plots as in (B), with 423 parameter sets for this model. Parameters are sampled (r_{on} from 0.1-100, r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01, threshold from 0-2.7) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 84 parameter sets. K) Schematic of a 3-state model with linear dependence of transition from the p_0 to p_{off}. L) (left panel). Same plots as in (B), with 1288 parameter sets for this model.

Figure 3.11: Parameters are sampled (r_{on} from 0.1-100,r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 16 parameter sets. M) Schematic of 3-state model with linear dependence on TF in both transitions from p_0 to p_{off} and p_{off} to p_{on}. N) (left panel) Same plots as in (B), with 1638 parameter sets for this model. Parameters are sampled (r_{on} from 0.1-100, r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted, in the same manner as in (B, middle panel, right panel) with 228 parameter. O) Schematic of the 3-state model with thresholding in the activation constant, r_{on}, between promoter off-states, p_0 and p_{off}. P) (left panel) Same plots as in (B), with 1649 parameter sets for this model. Parameters are sampled (r_{on} from 0.1-100, r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01, threshold from 0-0.5) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 455 parameter sets. Q) Schematic of the 3-state model with thresholding in the inactivation constant, r_{off}, between promoter off-states, p_0 and p_{off}. R) (left panel) Same plots as in (B), with 96 parameter sets for this model. Parameters are sampled (r_{on} from 0.1-100, r_{off} from 0.1-100, k_{on} from 0.0001-1, k_{off} from 0.0001-1, β_1 from 0.0001-10, β_0 from 0.000001-0.01, threshold from 0-0.5) or set ($\beta_2 = 0.06$, $\gamma_1 = 0.05$, $\gamma_2 = 0.0083$). (middle panel, right panel) Plotted in the same manner as in (B, middle panel, right panel) with 25 parameter sets.

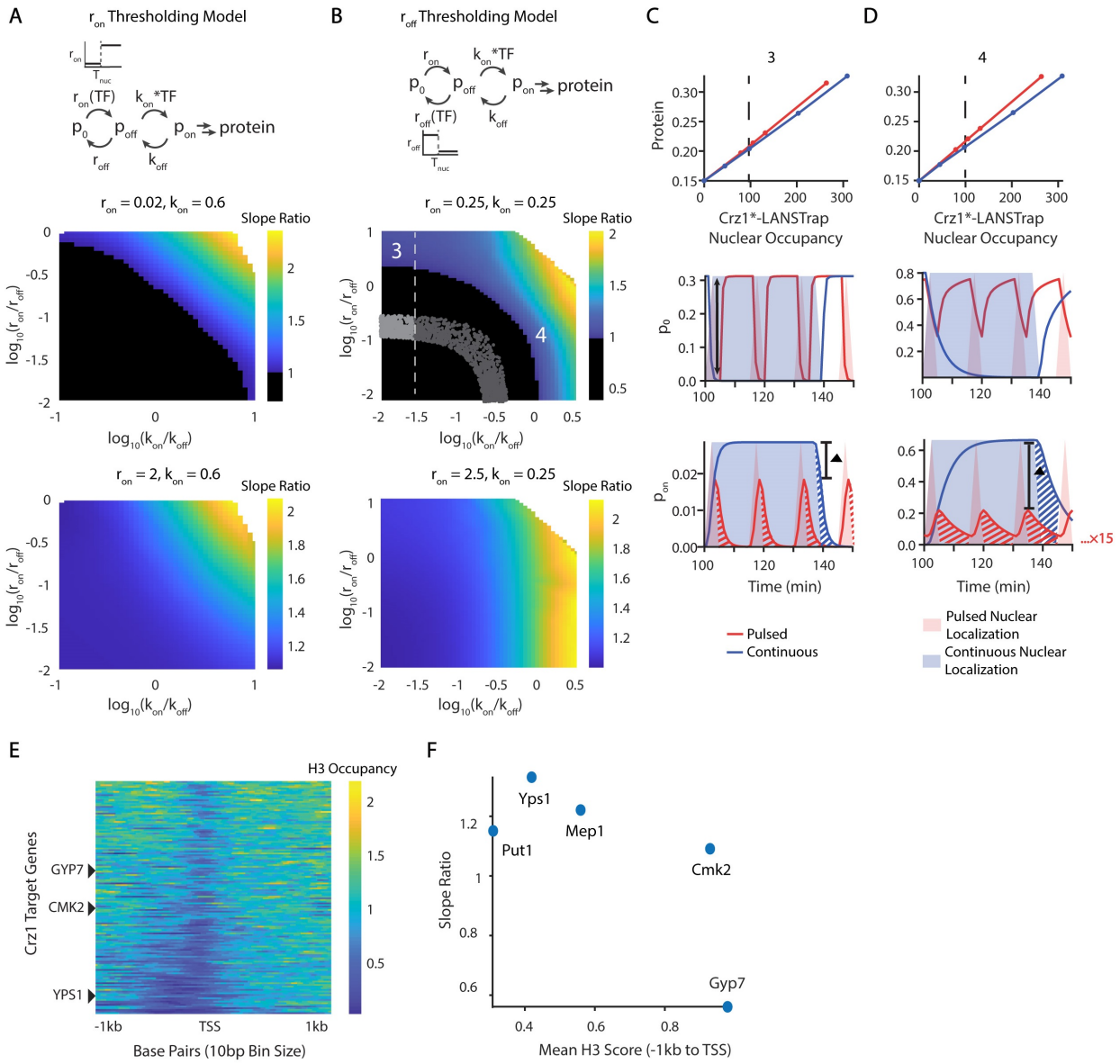

A
r_on Thresholding Model
r_on(TF)
k_on*TF
p_0
p_off
p_on
protein
r_off
k_off
T_nuc
r_on = 0.02, k_on = 0.6
r_on = 2, k_on = 0.6
Slope Ratio
log10(r_on/r_off)
log10(k_on/k_off)
B
r_off Thresholding Model
r_off(TF)
r_on = 0.25, k_on = 0.25
r_on = 2.5, k_on = 0.25
3
4
C
D
Protein
Crz1*-LANSTrap Nuclear Occupancy
p_0
p_on
Time (min)
Pulsed
Continuous
Pulsed Nuclear Localization
Continuous Nuclear Localization
...x15
E
H3 Occupancy
Crz1 Target Genes
GYP7
CMK2
YPS1
-1kb
TSS
1kb
Base Pairs (10bp Bin Size)
F
Yps1
Mep1
Put1
Cmk2
Gyp7
Slope Ratio
Mean H3 Score (-1kb to TSS)

Figure 3.12: **Exploration of the models that fit pGYP7-YFP** A) Comparison of the 3-state models with either r_{on} or r_{off} thresholding in the transition from p_0 to p_{off} (upper panel) Schematic of the 3-state model with thresholding in the activation constant, r_{on}, between promoter off-states, p_0 and p_{off}. (middle panel) Heatmap of slope ratio in the log10(k_{on}/k_{off})-log10(ron/roff) plane. r_{on} is set to 0.02 and k_{on} = 0.6. Parameters are sampled (r_{off} from 0.0002-0.02, k_{off} from 0.002-0.2) or set (β_1 = 0.0001, β_2 =0.06, γ_1 =0.05, γ_2 =0.0083, threshold= 0.5, β_0 =0.000001). (lower panel) Same heatmap as in (A, middle panel) except with r_{on} set to 2, and r_{off} ranges from 0.02-2. B) (upper panel) Schematic of the 3-state model with thresholding in the inactivation constant, r_{off}, between promoter OFF-states, p_0 and p_{off}. (middle panel) Same heatmap as in (A, middle panel) with r_{on} is set to 0.25 and k_{on} = 0.25, that is previously described in Figure 3.5E. Parameters are sampled (r_{off} from 0.0025-2.5, k_{off} from 0.0025-0.25) or set (β_1 = 0.0001, β_2 =0.06, γ_1 =0.05, γ_2 =0.0083, threshold= 0.5, β_0 =0.000001). (lower panel) Same heatmap as in (B, middle panel) except with r_{on} set to 2.5, and r_{off} ranges from 0.025-25. C-D) Additional parameter requirements of the 3-state r_{on} threshold model for fitting pGYP7-YFP. (upper panels) Output-Occupancy plots are generated by the model for different parameter sets that correspond to points 3 and 4 in the heatmap in B. The slope ratio for point 3 is 1.05 with log10(k_{on}/k_{off}) = -1.58 and log10(ron/roff) = 0.6. The slope ratio for point 4 is 1.25 with log10(k_{on}/k_{off}) = 0.1 and log10(ron/roff) = -0.89. Point 3 is chosen to highlight the effect of decreasing r_{off}, while Point 4 is chosen to highlight the effect of decreasing k_{off}. (middle panels) Example of a time course of promoter state p_0 for a light input that produces the equivalent of 40 minutes (dotted line in upper panel) in nuclear localization either continuously or in short pulses. Solid lines are the p_0 pulses while shading denotes nuclear localization. The black double arrows denote the maximum depletion of the p_0 state for the pulsed input. (lower panels) Example of a time course of promoter activity p_{on} for a light input that produces the equivalent of 40 minutes (dotted line in upper panel) in nuclear localization either continuously or in short pulses, similar to the (middle panels). The red and blue dashes represent residual promoter activity beyond the nuclear localization input. The red residual promoter activity is repeated 15 times while the blue residual activity is repeated one time. The black triangle denotes the difference between the amplitudes generated by the 2 minute pulse and 40 minute continuous input. E) Correlation of nucleosome occupancy and sensitivity to pulsing. Heatmap of H3 occupancy for the Crz1 target genes as specified by Yoshimoto 2002. H3 occupancy is defined as counts of H3 enrichment over the IgG antibody, which signals no pull down of histones. The dataset and determination of start sites are obtained from Sen et al., 2015 and Malabat et al., 2015, respectively. The software deepTools 2.0 is used to compute the H3 occupancy values. -1 and +1 kb from the transcription start site (TSS) is used. The positions of YPS1, CMK2, and GYP7 in the heatmap are denoted with black triangles. F) Slope ratios of Crz1 target genes as a function of their mean H3 nucleosome occupancy scores averaged from -1kb to the Transcription Start Site (TSS). The correlation coefficient is r^2 = 0.26.

Table 3.1: **NLS Optimization** This table contains the amino acid and peptide sequences used in the screen for NLSs in the optimization of CLASP for maximal nuclear localization dynamic range.

Number	NLS	Score (Kosugi 2009)	AA Length	Class
0	paaKRvKld	na	9	original LANS, class2
3	raaKRpRtt	10	9	class2
5	paaKRpRtt	9	9	class2
8	apaKRaRtt	8	9	class2
9	paaKRlCtt	9	9	class2
11	aaaKRswsmaf	10	11	class3
14	aaaKRswvmaf	9	11	class3
15	aaaKRswsaaf	10	11	class3
20	KRpatlandspaaKRR	9	16	bipartite
24	KRKRwendip	na	10	class1
27	psRKRKRdhyav	na	12	class1
29	tspsRKRKwdqv	na	12	class1

Table 3.2: **OptoPlate Transfer Function** The transfer function of the OptoPlate. Conversion from arbitrary units (a.u.) to intensity (milliWatts).

Light Dose	milliWattage	Std Dev
0	-0.0001499	0.0002004
0	-0.0002513	0.000252
0	-0.001819	0.000978
64	0.1011	0.008522
64	0.07921	0.007053
64	0.1279	0.01123
128	0.1944	0.02126
128	0.1962	0.01824
128	0.2633	0.02653
256	0.3961	0.0393
256	0.4012	0.0362
256	0.5107	0.04832
512	0.7849	0.07396
512	0.7776	0.09084
512	1.041	0.09508
1024	1.53	0.1346
1024	1.477	0.1491
1024	1.986	0.2349
2048	3.088	0.3326
2048	3.026	0.2563
2048	3.45	0.9
3072	4.302	0.4414
3072	4.235	0.4559
4095	5.407	0.5494
4095	5.486	0.4617
4095	6.56	0.587

Chapter 4

Vignettes: the origin stories of the bPAC and CLASP papers

4.1 Summary

The following chapter is a compilation of the unpublished data in exploring and studying the upstream determinants and downstream interpretation of dynamic transcription factor nuclear localization. These studies center on using different natural stress, genetic, chemical, and optogenetic perturbations to either naturally alter or synthetically control transcription factor nuclear localization, and quantifying the outputs of nuclear localization and gene expression. These studies also chronicle the sustained efforts of a PhD student in pursuing and answering a biological question. The work presented in this chapter (and other data not included here) formed the basis of the published bPAC and CLASP work detailed in Chapters 2 and 3.

4.2 Deletions of PKA components result in alteration of transcription factor nuclear localization in response to natural stress inputs that impinge on the PKA network

Protein kinase A (PKA) is a key mediator of nutrient and stress signaling in *S. cerevisiae*. As such PKA integrates signals from many different inputs such as glucose depletion[24], nitrogen starvation[82], heat shock[83], and hyperosmotic stress[1]. PKA is embedded in a complex of feedback loops that can precisely tune its activity[84]. It is thought that environmental stresses may impinge on different components of the PKA feedback network to activate or inactivate PKA. In addition to these natural inputs to PKA, recent work has shown that PKA can also be activated optogenetically by a bacterial photoactivatible adenylyl cyclase (bPAC)[12]. The perturbations on PKA are directly reflected in one of its target transcription factors Msn2[25] (Figure 4.1).

In order to test whether different environmental inputs do impinge on PKA to alter both PKA and Msn2 activity, Msn2 nuclear localization in response to different doses of glucose depletion and osmotic shock (KCl) are measured (Figure 4.2). There are two main conclusions. First, Msn2 nuclear localization is proportional to the duration and amplitude of stress for both inputs. Second, in response to osmotic shock at a high dose (0.5M KCl), Msn2 nuclear localization is more sustained than that for a high level of glucose deprivation (0.1% glucose). This data establishes that these two environmental inputs do alter PKA and Msn2 dynamics.

Next, components of the PKA network were knocked out to identify the mediators of glucose depletion or osmotic shock signaling in the PKA network (Figure 4.3). For glucose depletion, Δras1 and Δras2 increased Msn2 localization, while for Δpde1, Δpde2, Δira1,

and Δira2, the decreased Msn2 nuclear localization. Δira2 abolished the Msn2 localization response. This suggests that Δira2 plays a key role in mediating the glucose depletion signal. For osmotic shock, Δras1 and Δras2 behaved similarly, with little effect on Msn2 localization. Interestingly, Δira2 and Δpde2 exhibit slower localization than the WT or other mutants. This suggests that Δira2 and Δpde2, with Δpde2 having a more dominant effect, both play an important role in mediating the osmotic stress signaling through PKA.

Lastly, optogenetic production of cAMP with bPAC was used to causally relate genetic perturbations to the PKA network to changes in PKA signaling and Msn2 nuclear localization dynamics. Bcy1 is an inhibitor of PKA, and Bcy1 mutants were used to alter the PKA network. Single mutations (S-> G, E, D, R, and T) were made at the 145 position of the Bcy1 gene (Figure 4.4). The serine 145 in Bcy1 is thought to either be phosphorylated by PKA or other kinases in response to environmental inputs[85]. The different nuclear localization dynamics in response to different Bcy1 mutants demonstrate that interactions that alter the activity of a PKA inhibitor, such as Bcy1, can greatly affect Msn2 nuclear localization dynamics.

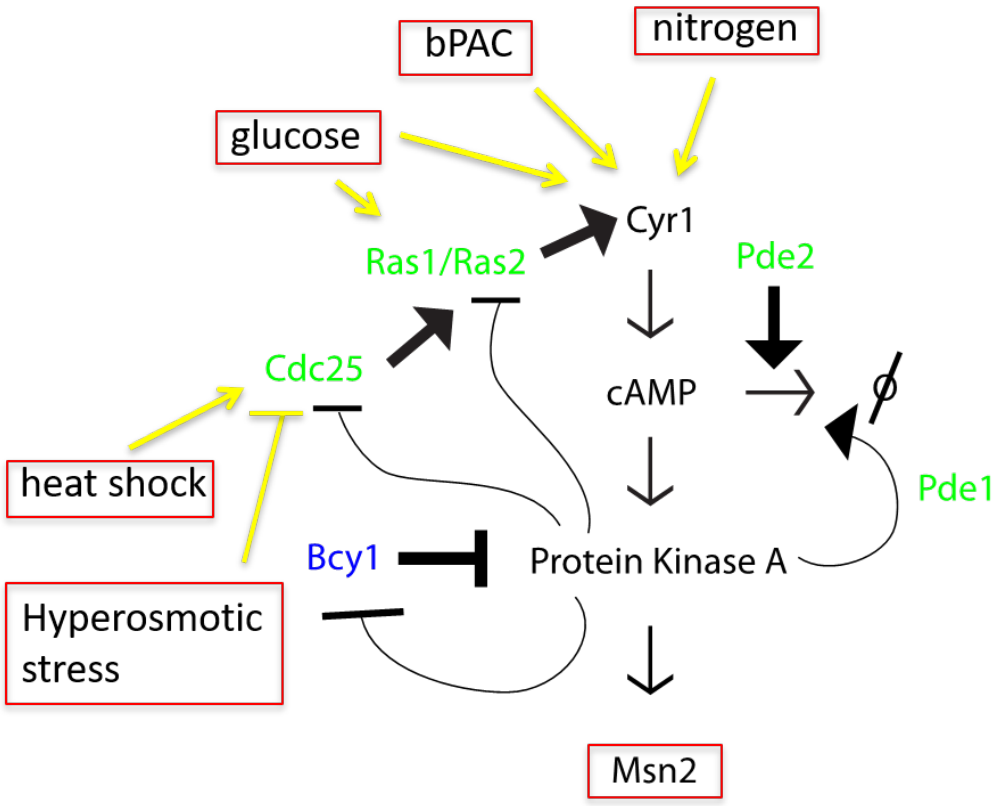


Figure 4.1: **Many natural stress inputs impinge on PKA** Schematic of the PKA signaling network and inputs that impinge on the network.

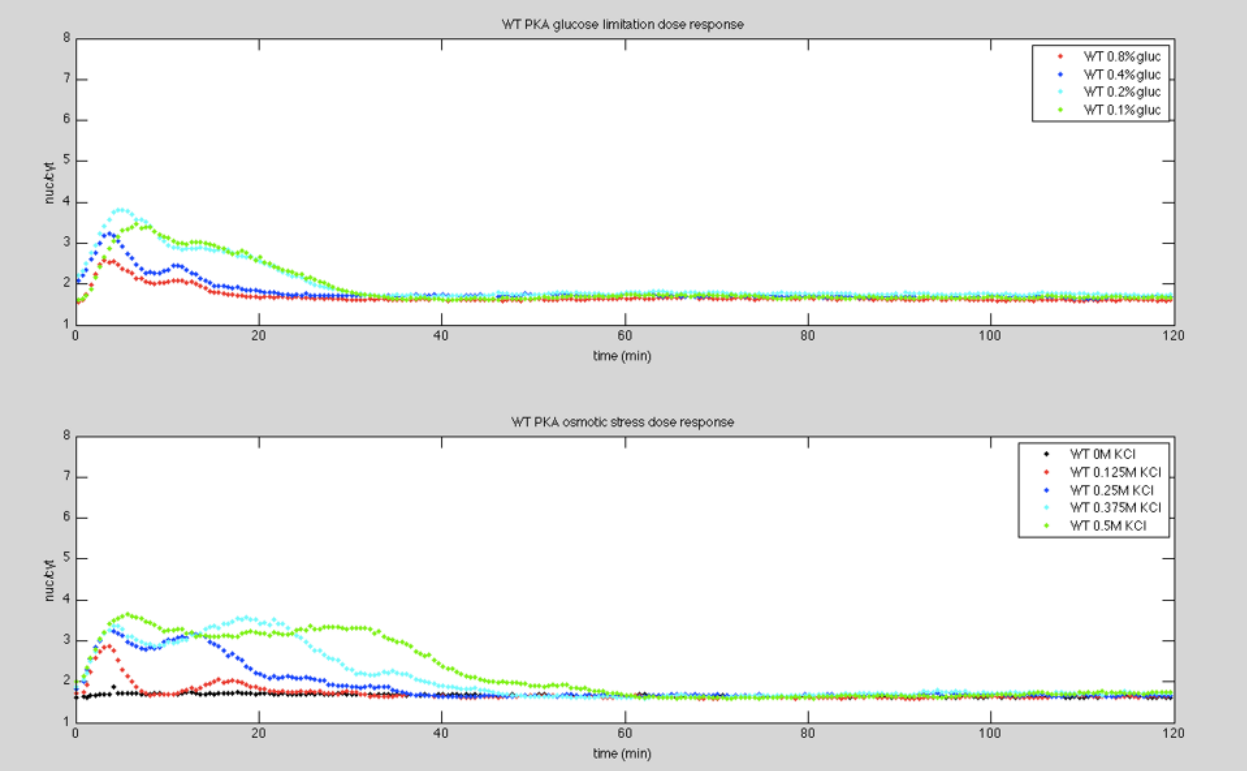


Figure 4.2: **Glucose and osmotic shock result in distinct patterns of Msn2 nuclear localization** (top panel) Time course of Msn2 nuclear localization in response to decreasing glucose availability in SDC. (bottom panel) Time course of Msn2 nuclear localization in response to increasing potassium chloride (KCl).

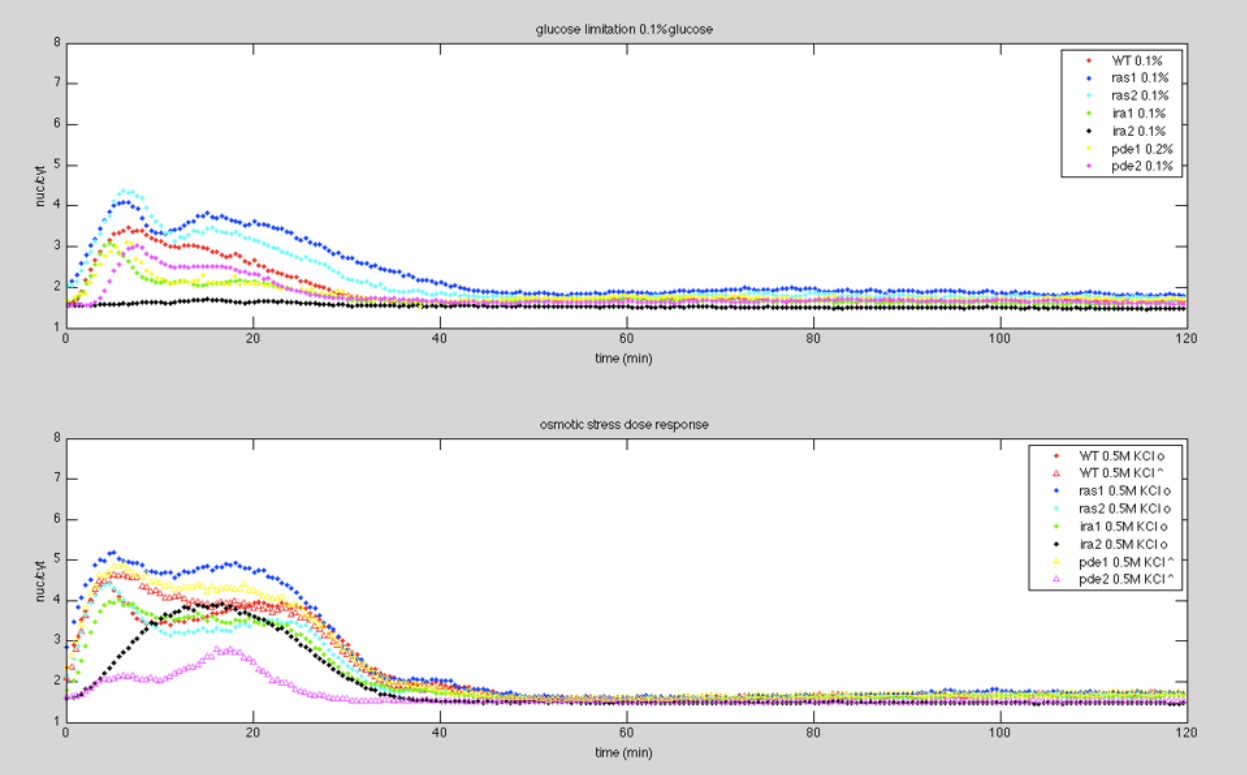


Figure 4.3: (top panel) **Deletion of different components of the PKA network result in distinct nuclear Msn2 nuclear localization in response to stresses** Time course of Msn2 nuclear localization in 0.1% glucose in a background with a single component of the PKA network knocked out. (bottom panel) Same as in the top expanel except in 0.5M KCl.

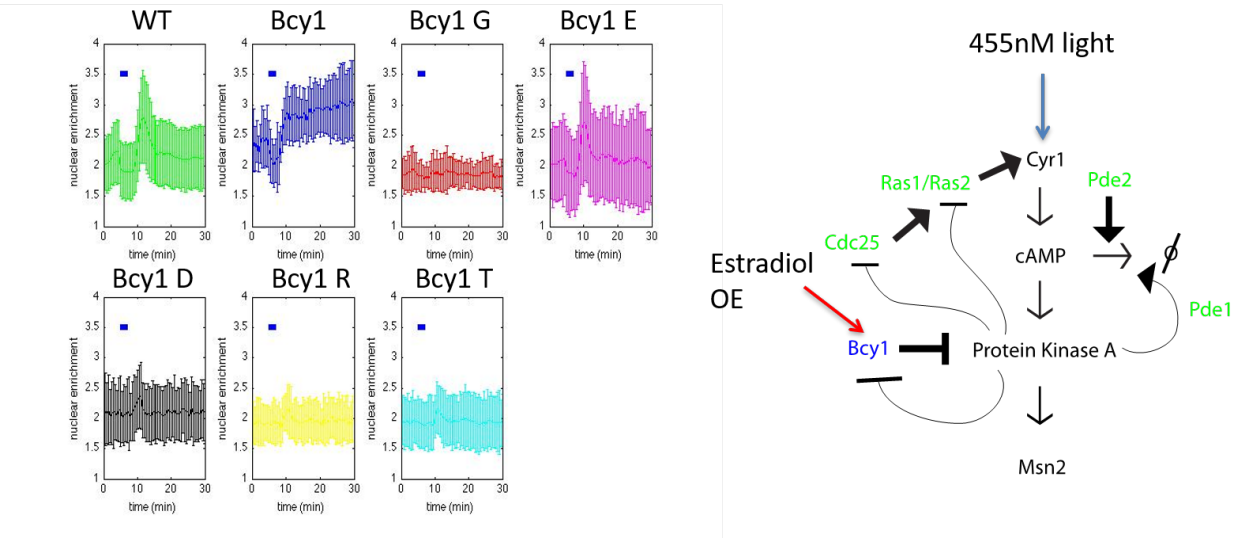


Figure 4.4: **Using optogenetics instead of natural inputs to perturb PKA in deletion backgrounds** (left panel) Time course of Msn2 nuclear localization in different mutational backgrounds of a PKA inhibitor Bcy1. The PKA network is activated using blue light with the optogenetic system bacterial photoactivated adenylyl cyclase (bPAC). (bottom panel) Schematic of the PKA network, identifying the blue light perturbation with the blue arrow and the induction of a Bcy1 mutatnt with the red arrow.

4.3 PKA phosphorylation sites on Msn2 differentially modulate Msn2 localization dynamics

To more controllably map changes in nuclear localization dynamics to Msn2 sequence changes, various Msn2 NLS and NES mutants were made and the Msn2 nuclear localization measured. Msn2 has a number of serine residues that are known to be phosphorylated by PKA (Figure 4.5) and PKA de-phosphorylation has been documented to change the localization Msn2[86] (Figure 4.6A). Disrupting the NLS and NES will provide some mechanistic insights to the variable dynamics of nuclear localization due to the intrinsic sequence. In order to observe the dynamic changes in Msn2, bPAC is used to deliver a transient input (Figure 4.6B). The single S to A mutations typically responded

with similar dynamics as the wildtype (Figure 4.7A), while the R to G mutants tended to have attenuated localization responses (Figure 4.7B). The S to A mutations in the NES have a more dramatic effect than the R to G mutants in the NES and compared to all the NLS mutations (Figure 4.8). Nuclear localization is sustained for the S to A NES mutants. Taken together, the phosphorylation state of Msn2 dictates its dynamic nuclear localization patterns.

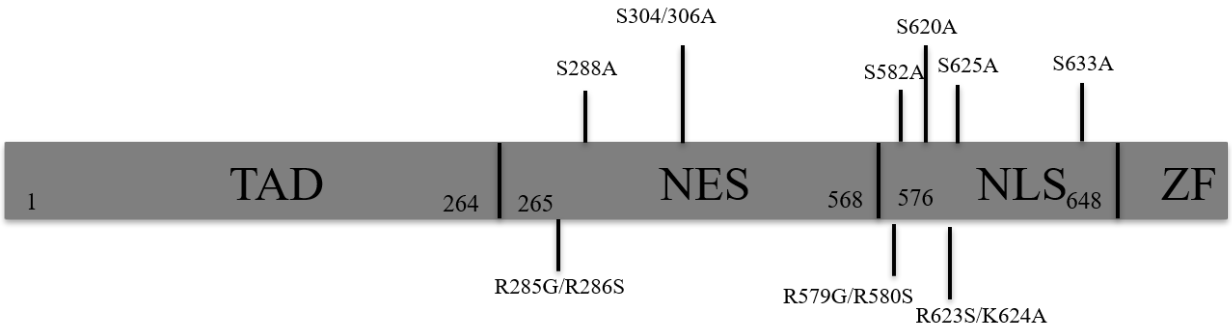


Figure 4.5: **Msn2 known phoshporylation sites** The mutants of PKA include point mutations to the sites of phosphorylation in the nuclear localization sequence (NLS) and nuclear exit sequence (NES). Other mutants have arginine point mutations also in the NLS or NES near the PKA binding sites.

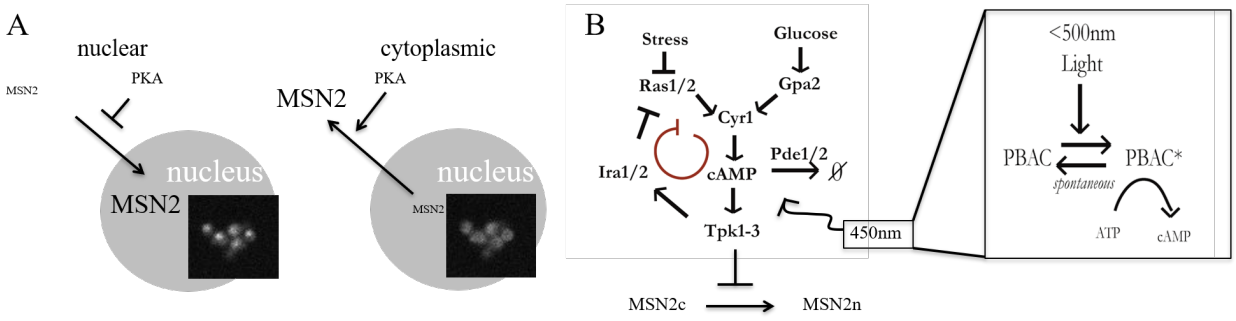


Figure 4.6: **Msn2 localization controlled by PKA** A) PKA regulation of Msn2 localization. PKA has been experimentally shown to inhibit nuclear localization and facilitate nuclear export. B) In the endogenous PKA network, PKA is activated by cAMP as a result of upstream environmental perturbations. By using bPAC, light induction artificially and controllably activates PKA, which then inhibits nuclear localization of Msn2.

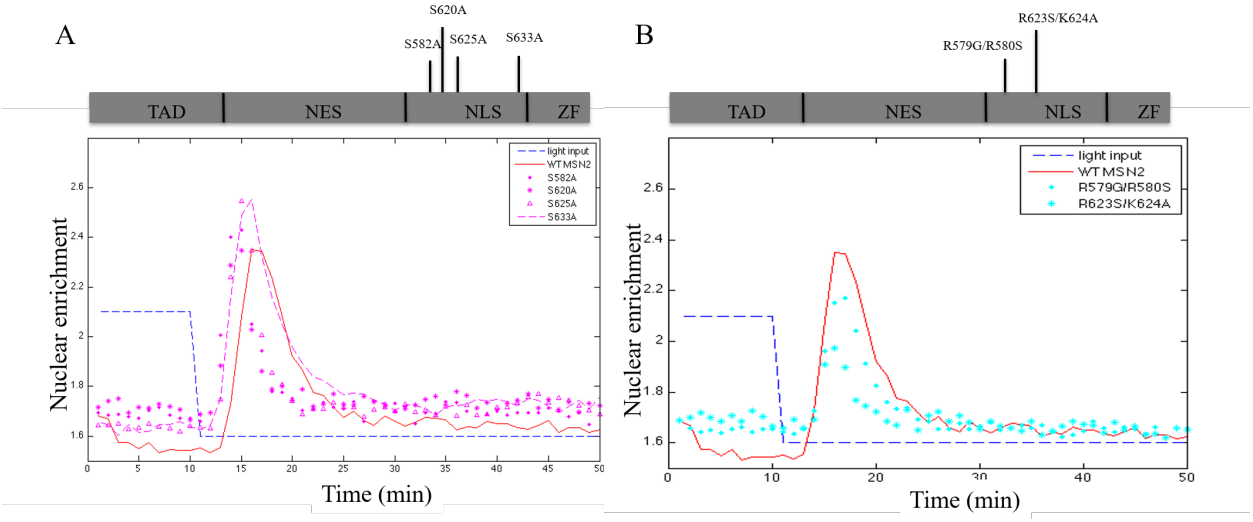


Figure 4.7: **Msn2 nuclear localization with NLS mutants** A) Nuclear enrichment of single phosphorylation knockout NLS mutants. A pulse of light at 450nm is delivered as the input for a duration of five minutes. Time traces of nuclear localization are shown. Initial light induction increases PKA activity, but subsequent nuclear localization occurs due to the negative feedback in the native PKA network. B) Nuclear enrichment of arginine NLS mutants.

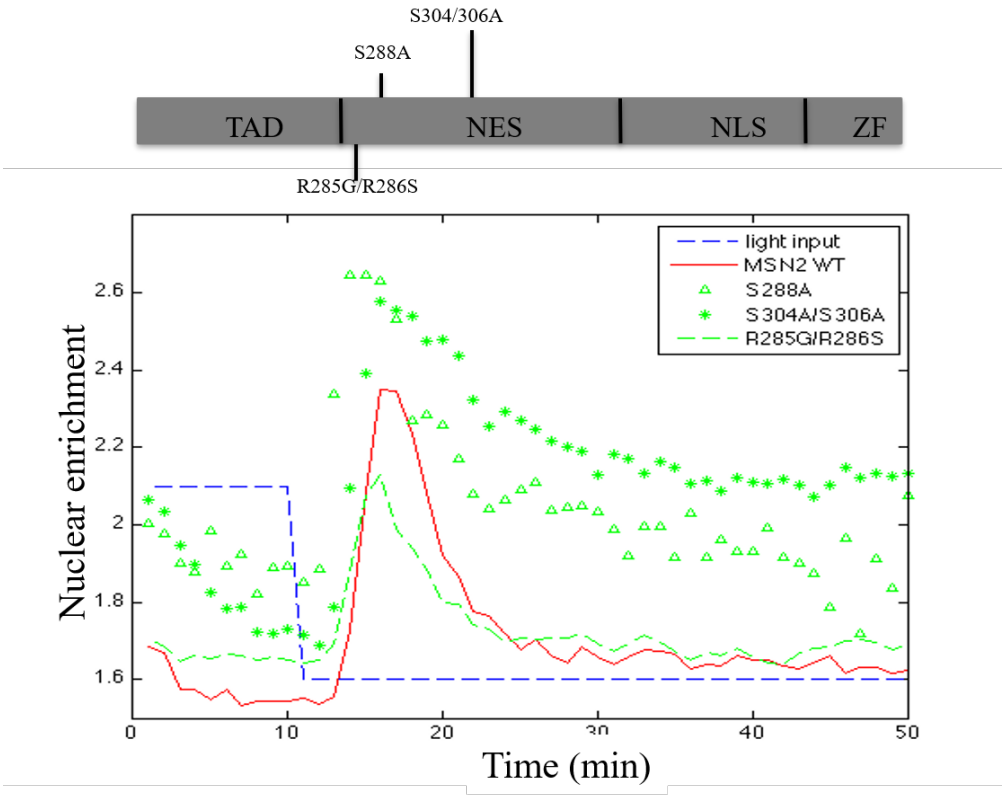


Figure 4.8: **Msn2 nuclear localization with NES mutants** Nuclear enrichment of three different NES mutants. Mutant S288A is a PKA phosphomutant. S304/S306A is another phosphomutant. And R285G/R285S is a possible PKA binding mutant.

4.4 A screen of transcription factor nuclear localization dynamics in response to environmental stress inputs

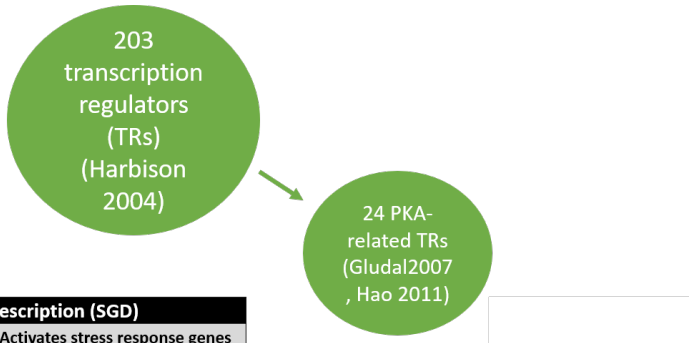
From the study of PKA regulated Msn2 nuclear localization, a question arose of whether PKA could regulate the nuclear localization events of many other transcription factors, and moreover whether different environmental inputs could regulate localization events perhaps through PKA. A search into existing databases showed that many transcription factors have either validated for putative PKA phosphorylation sites (Figure 4.9, left)[9,87]. A further search into the stress inputs that upregulate transcription factors with known PKA phosphorylation sites (obtained from PhosphoGRID) show that the transcription factors are activated by a variety of different inputs (Figure 4.9, right).

Hence, it would be interesting to observe the localization of transcription factors against these different stress input, to identify which transcription factors respond to which stress inputs. A screen of 15 transcription factors against 6 stress inputs measured the nuclear localization over time (Figure 4.10, 4.11, and 4.12). A number of transcription factors did not localize in response to any stress inputs. These are Cad1, Com2, Crf1, Gat1, Yap1, and Nrg2. Stb3 is the only transcription factor that was only responsive to one stress input. For the other transcription factors Crz1, Dot6, Maf1, Msn2, Msn4, Sko1, Rtg3, and Pho4, nuclear localization occurred in response to multiple stress inputs. A search in the

literature shows that some of these nuclear localization events had been previously observed, but this study is the first to systematically characterize the effects of stress inputs on transcription factor localization. Several localization events have not been previously observed in the literature, and this includes Msn2 response to cell wall integrity stress, Msn2 response to alkaline PH (PH8), and Msn4 response to nitrogen limitation (Table 4.1).

It can be concluded from this study that a given stress input activates multiple transcription factors. This finding is consistent with the understanding that nutrient and stress signaling involve activation of different expression regulons or the same expression regulon under combinatorial control. Perhaps more interestingly, this study revealed that a given transcription factor can mediate multiple different stress inputs. These different dynamics of localization may hence play a role in discrimination of these upstream signals. This is the study that motivated the development of CLASP.

TRs	Isoform	PKA Phosphosites
Adr1	Tpk1	S230
Azf1	Tpk1, Tpk2	S303
Cin5	Tpk1	S196
Crz1	Tpk1	S409, S410, S423, S427, S429
Dot6		S282, S322, S423, S313
Gis1		S424
Gln3		S267, S395, S684
Gzf3/Deh1	Tpk1	S92
Hsf1		S97, S458
Maf1	Tpk1	S90, S101, S177, S178, S209, S210
Mig1		S310, S367
Msn2		S582, S620, S625, S633, S288, S304
Msn4	Tpk1	S263, S279, S316, S488, S531, S532, S558
Opi1	Tpk1, Tpk3	S31, S251
Pho4	Tpk1	S5, S204, S243
Rgt1	Tpk1	S96, S146, S202, S283, S284
Rtg1		S163
Rtg3		S135, S255
Sko1	Tpk1	S380, S393, S399
Ssn2	Tpk3	S608, S1236
Stb3		S146, S234, S235, S254, S285, S286, S345, S348
Tod6		S280, S298, S318, S333
Ume6	Tpk1	S141, S645
Yap1		S503



TRs	Description (SGD)
Crz1	Activates stress response genes
Dot6	Represses ribosome biogenesis
Maf1	Negative regulator of RNA PolIII, stress activated
Msn2	Activates stress response genes
Msn4	Activates stress response genes
Pho4	Activates under phosphate limitation
Rtg3	Related to retrograde and TOR pathways
Sko1	Involved in osmotic and oxidative stress responses
Stb3	Represses ribosome biogenesis
Yap1	Required for oxidative stress tolerance

Figure 4.9: **Many stress related transcription factors have PKA phosphorylation sites** (left) A table of transcription factors with their known kinase phosphorylator and phosphosites. (right) A table of known PKA regulated transcription factors and the stresses they are upregulated by. A graphic showing the number of total transcription factors in yeast (203) and the number of PKA regulated transcription factors (24).

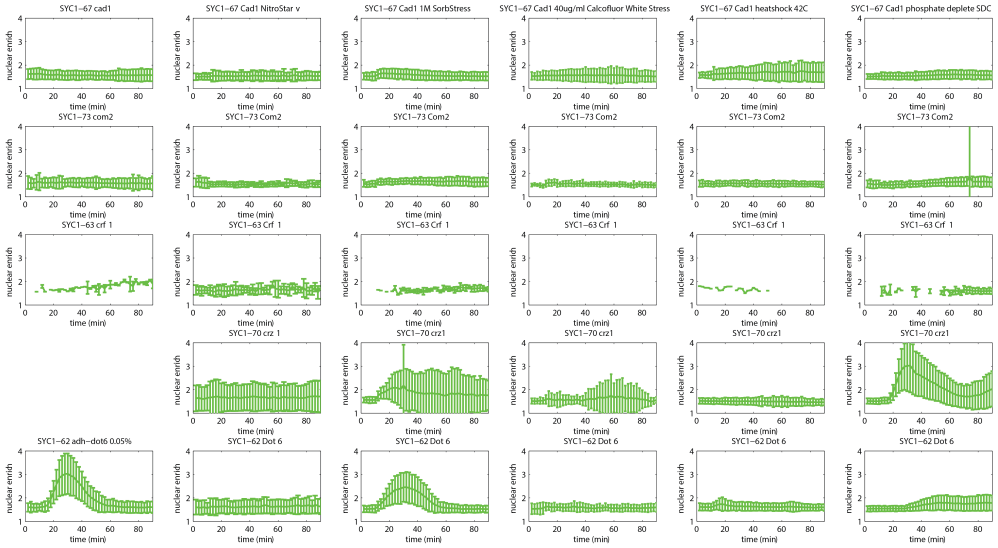


Figure 4.10: **A screen of transcription factor localization in response to stress inputs 1** The transcription factors: Cad1, Com2, Crf1, Crz1, and Dot6 are screen against 0.05% glucose depletion, nitrogen starvation, osmostic stress (sorbitol), calcofluor white, heat shock, and phoshpate depletion. The mean nuclear localization is plotted with error bars representing standard deviation. Imaging was done every 2 minutes in the RFP channel.

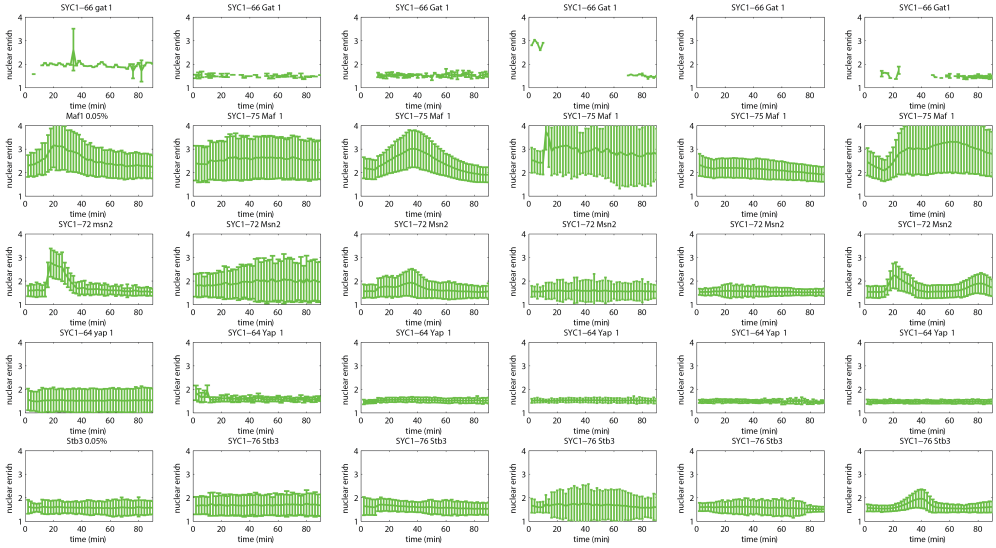


Figure 4.11: **A screen of transcription factor localization in response to stress inputs 2** The transcription factors: Gat1, Maf1, Msn2, Yap1, and Stb3 are screen against 0.05% glucose depletion, nitrogen starvation, osmostic stress (sorbitol), calcofluor white, heat shock, and phoshpate depletion. The mean nuclear localization is plotted with error bars representing standard deviation. Imaging was done every 2 minutes in the RFP channel.

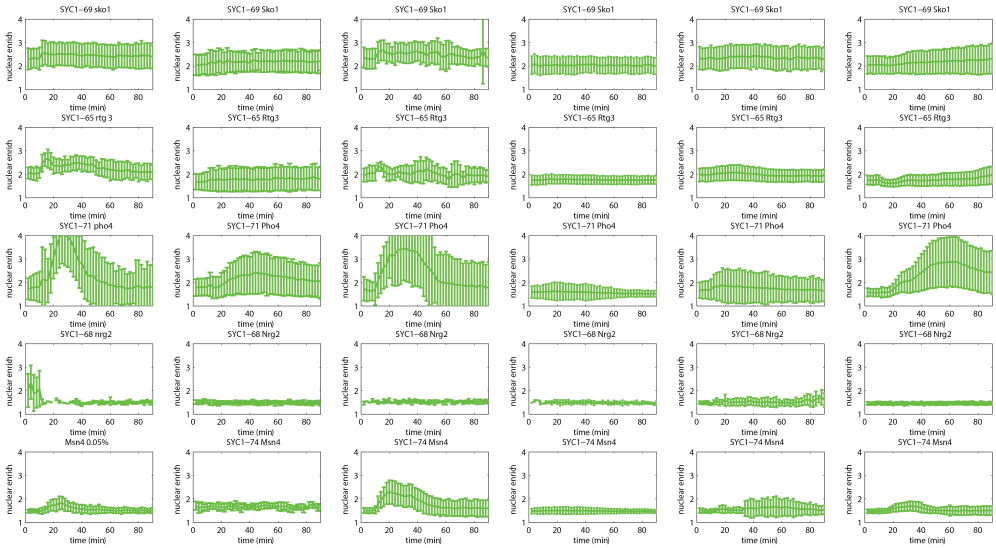


Figure 4.12: **A screen of transcription factor localization in response to stress inputs 3** The transcription factors: Sko1, Rtg3, Pho4, Nrg2, and Msn4 are screen against 0.05% glucose depletion, nitrogen starvation, osmostic stress (sorbitol), calcofluor white, heat shock, and phoshpate depletion. The mean nuclear localization is plotted with error bars representing standard deviation. Imaging was done every 2 minutes in the RFP channel.

Table 4.1: **Summary table of transcription factor nuclear localization in response to stress inputs** For several of the transcription factors that responded to all of the stress inputs, whether nuclear localization has been reported is written in red with literature references noted.

	Glucose Limitation	Osmotic Stress	Cell Wall Stress	Alkaline	Nitrogen Limitation
Pho4	Swinnern 2006 (YES)	Yale 2001 (YES)	Huang 2002 (YES)	Serrano 2002 (YES)	Swinnern 2006 (YES)
Maf1	Hallet 2014 (YES)	Boisnard 2009 (YES)	(YES)	(YES)	Hallet 2014 (YES)
Dot6	Hippman 2009 (YES)	Enjalbert 2008 (YES)	(YES)	(YES)	Stracka 2014 (YES)
Msn2	Gorner 2002 (YES)	Hao 2011 (YES)	Garcia 2009 (NO)	(YES)	Pan 1999 (NO)
Msn4	Martinez-Pastor 1996 (YES)	Martinez-Pastor 1996 (YES)	Garcia 2009 (NO)	(YES)	(NO)

4.5 Optogenetics could be used to test correlation between upstream dynamics and downstream gene expression

The previous study showed that many transcription factors localize to the nucleus in response to stress inputs. It was also known that many of these transcription factors are regulated by PKA and that PKA is a key mediator of environmental stress inputs. It's been thought that the negative feedback network in which PKA is embedded allows PKA to propagate upstream signals in dynamically distinct ways to its target transcription factors (Figure 4.13). PKA is known to de-phosphorylate Msn2, which results in Msn2 nuclear localization[86]. Localized Msn2 then binds to STRE promoter motifs to activate Msn2 target genes[6]. Msn2 exhibits transient nuclear localization in response to hyperosmotic shock and sustained nuclear localization in response to oxidative stress (Figure 4.14). In order to examine whether dynamically distinct Msn2 nuclear localization

might differentially express genes, gene expression datasets[59] in response to environmental stresses were re-analyzed by identifying Msn2 target genes, defined as genes bearing STRE promoter motifs, and comparing gene expression output between the two stress inputs (Figure 4.14). Genes common to hyperosmotic and oxidative stress that were up- or down-regulated could be clustered into two groups. One group had similar expression levels between the two stress, while in the other group, genes were up-regulated in oxidative stress but down-regulated in hyperosmotic stress (Figure 4.14). This data identifies a gene set of Msn2 target genes that are differentially expressed based on the stress input, and this differential expression is correlated with Msn2 nuclear localization dynamics.

In order to establish a causal relationship between nuclear localization dynamics and differential gene expression, synthetic control of nuclear localization was required. From previous work on bPAC, it is known that glucose addition results in a transient increase in cAMP concentration and PKA activity[88] (Figure 4.15). This transient increase in cAMP/PKA activity would translate into a mirrored transient decrease in Msn2 nuclear localization, and could be mimicked with light input using bPAC activation (Figure 4.15). Hence, RNA sequencing was performed on cells experienced similar PKA/Msn2 dynamic changes; a transient 40 minutes pulse of light and glucose addition were used as inputs. The log_2 fold change of gene expression showed that expression profiles clustered similarly for both light stimulation and glucose addition, suggesting a causal relationship between PKA dynamics and the downstream gene expression in response to glucose addition (Figure 4.16).

This study showed the ability of using optogenetics to successfully mimic upstream protein activity dynamics and draw causal conclusions about the role of dynamics in controlling gene expression. However, this method was ultimately relegated because control of Msn2 nuclear localization with bPAC was limited in that activation of bPAC caused a decrease in Msn2 nuclear localization that was not measurable with our experimental system. Hence,

it was difficult to measure the dynamics of Msn2 to causally correlate with Msn2 target gene expression. Hence, a different method to control Msn2 was pursued. This effort is detailed in the next section. Ultimately, the goal of controlling transcription factor localization was broader than just the control of Msn2, and a more general tool was necessary that de-coupled localization from any upstream kinase or natural input. Never the less, the idea of using optogenetics to control transcription factor nuclear localization was incepted from these sets of experiments.

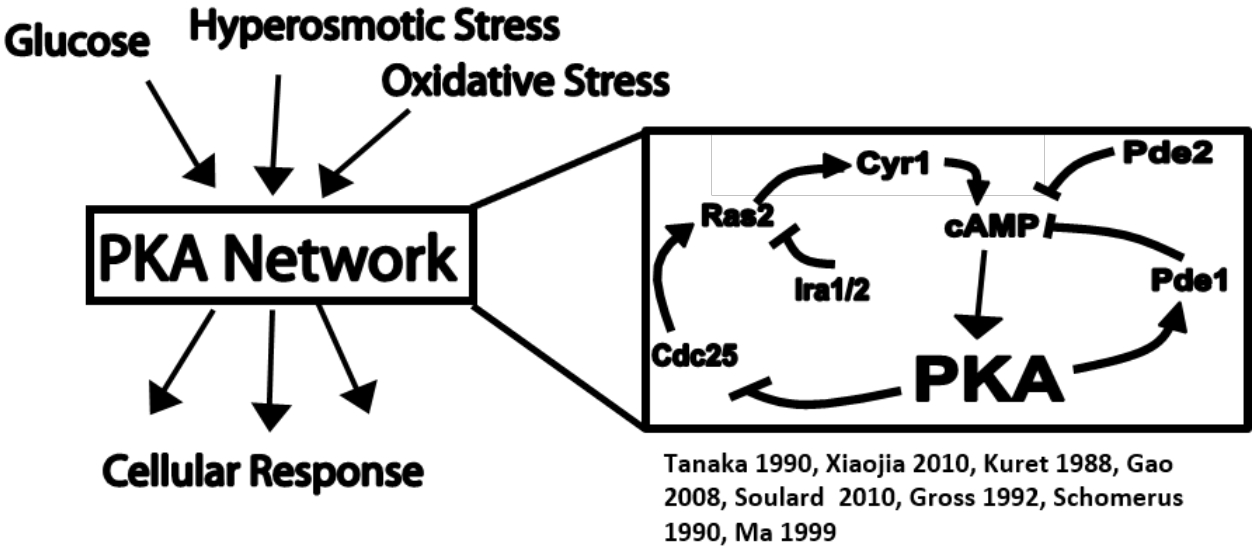


Figure 4.13: **PKA is a signaling hub** The PKA network is a signaling hub that integrates many environmental responses and elicits downstream transcriptional responses. The PKA network consists of two main negative feedback loops.

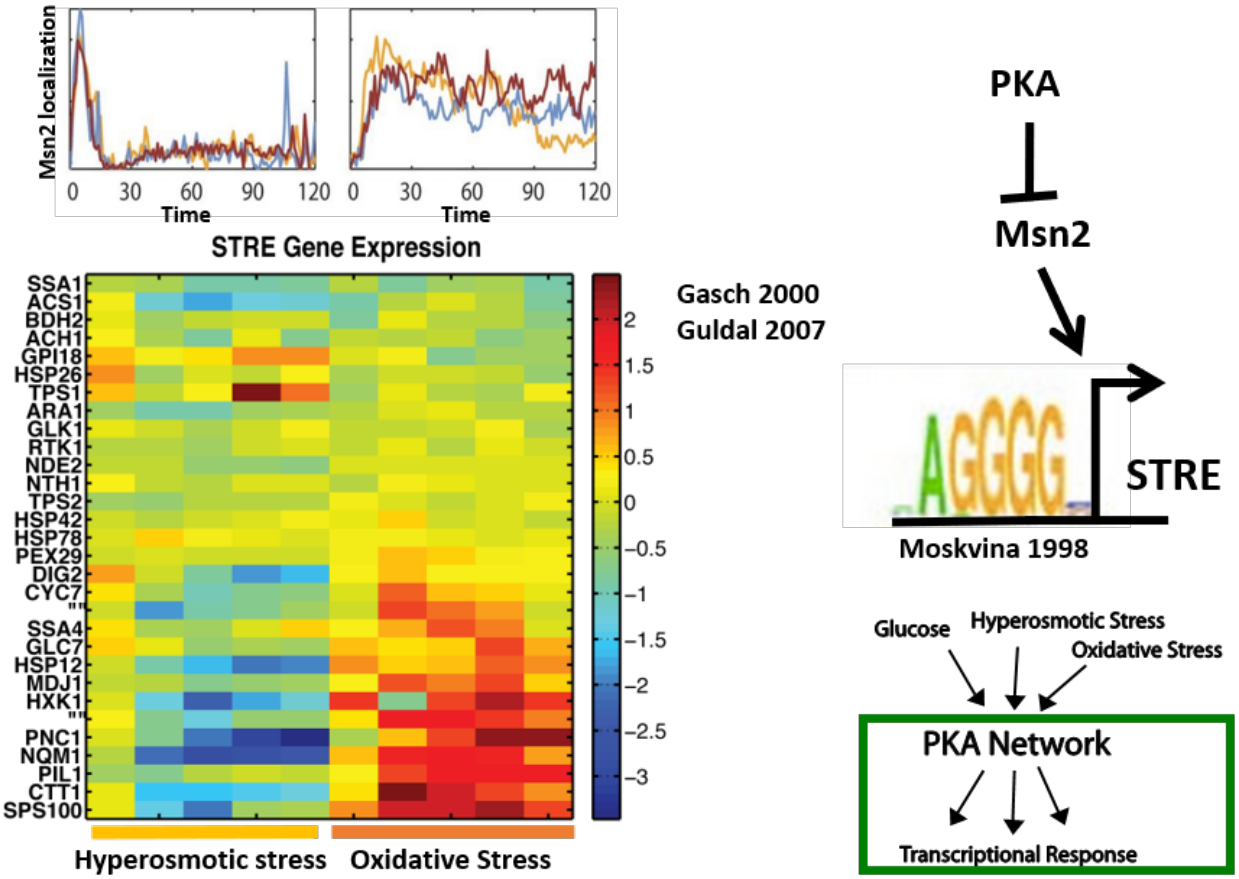


Figure 4.14: **PKA-mediated Msn2 nuclear localization dynamics correlate with distinct gene expression patterns** (left top) Msn2 nuclear localization in response to hyperosmotic stress and oxidative stress as published in Hao 2011. (left bottom) Clustered gene expression of known Msn2 target genes in response to hyperosmotic and oxidative stress. The gene expression datasets were taken from Gausch 2000 and Guldal 2007. (right) PKA is known to de-phoshporylate Msn2 and cause Msn2 nuclear localization to the nucleus. In the nucleus, Msn2 is known to bind STRE promoter elements, as published by Moskvina 1998. Diagram of PKA/Msn2 potentially integrating different stress inputs and producing different transcriptional responses.

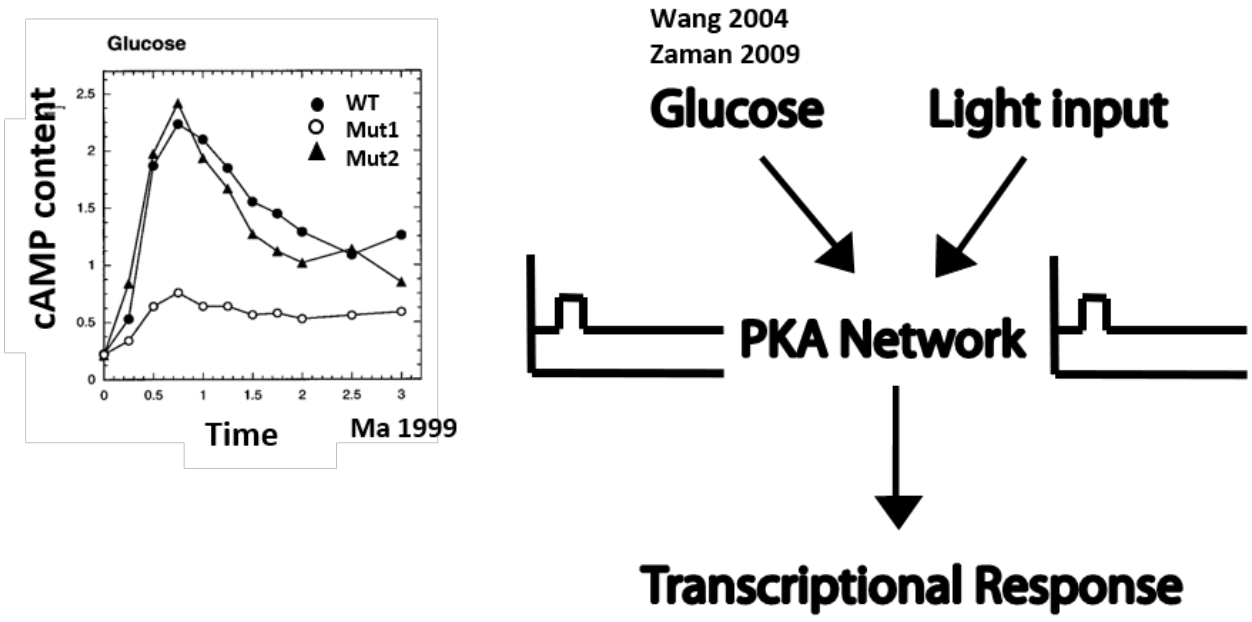


Figure 4.15: **cAMP/PKA time course in response to glucose addition** Plot of cAMP over time in response to glucose addition from Ma 1999. Schematic of glucose and light addition impinging on PKA to result in a transcriptional response.

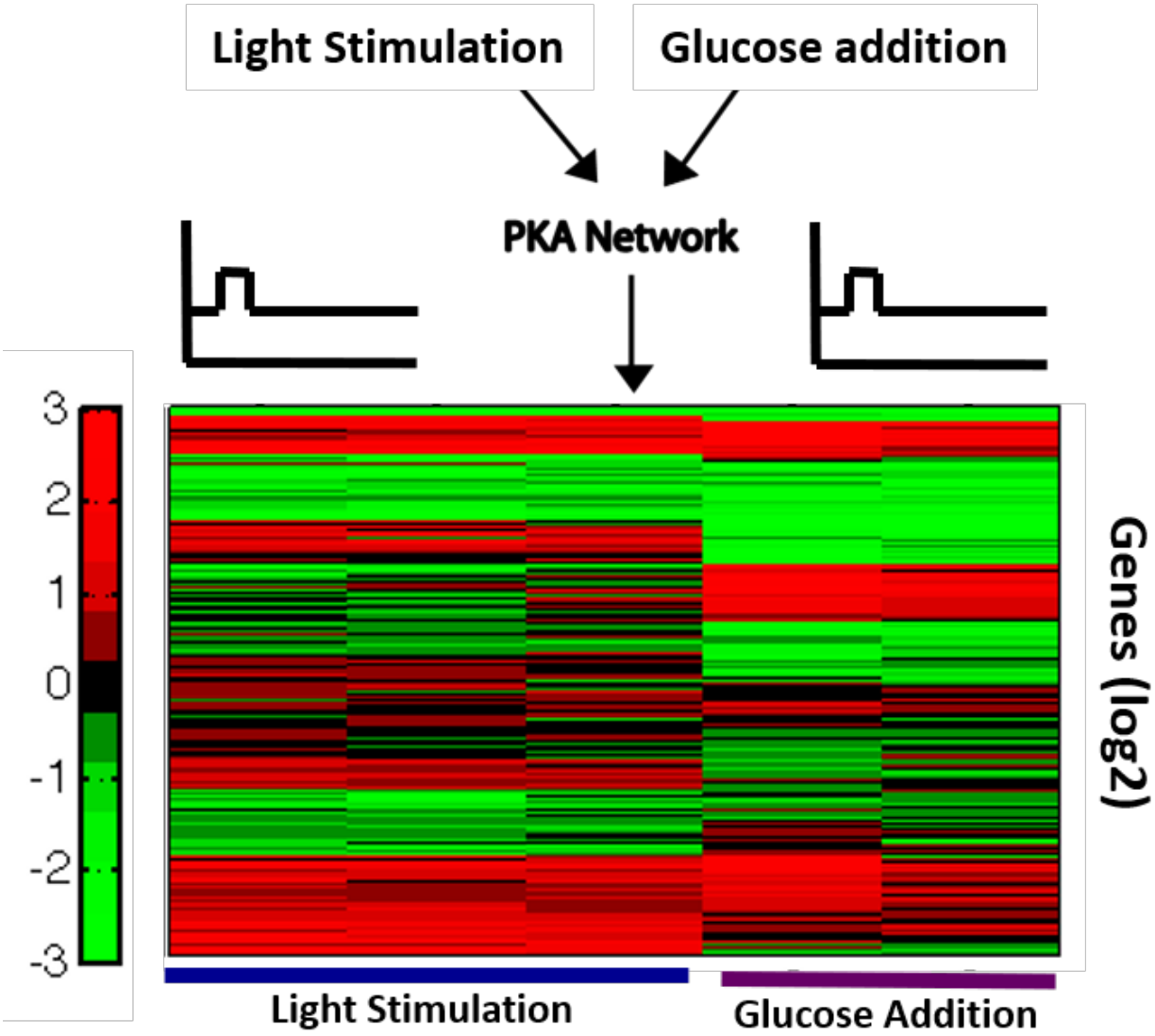


Figure 4.16: **RNAseq in response to light stimulation and glucose addition** Clustered heatmap of light and glucose addition. The different columns are biological replicates.

4.6 Controlling transcription factor nuclear localization with 1-NMPP1 was unreliable

It was shown that PKA is an upstream regulator of many transcription factors. Hence, it is plausible that a small molecule inhibitor of PKA, 1-NMPP1 would be able to control the nuclear localization of PKA regulated transcription factors[21] (Figure 4.17, left). To test this, 500uM of 1-NMPP1 was added to cells with RFP tagged transcription factors. Time courses of nuclear localization for these four different PKA regulated transcription factors, Msn2, Tod6, Gln3, and Stb3, were obtained. The data shows that indeed, in response to a step input of 1-NMPP1, all four transcription factors localized to the nucleus (Figure 4.17, right).

Next, to control the dynamics of nuclear localization, the 1-NMPP1 was delivered in a microfluidic chamber, making the washout and re-delivery of 1-NMPP1 possible. To initially test 1-NMPP1 control with this method, a single transient pulse of 1-NMPP1 is deliver through microfluidics and the nuclear localization of Tod6 was observed over time (Figure 4.18). Interestingly neither the lab stock or the Shokat lab stock of 1-NMPP1 was able to maintain nuclear localization for the duration of the 1-NMPP1 pulse. Glucose depletion, however, followed the dynamics of the media input. This suggests that it was not due to a technical error of the microfluidic plate. The localization of Tod6 in response to 1uM of rapamycin, moreover, followed a similar trend to 500uM 1-NMPP1. This suggested that the properties of the small molecules 1-NMMP1 and rapamycin were interfering with the delivery of input through the microfluidic plate. A search of the literature found that another lab that used 1-NMPP1 had built their own custom microfluidic setup because the PDMS material of the microfluidic plate absorbed the small molecule[89]. This revelation was consistent with the dampened nuclear localization experimentally observed in the presence of a 30 minute pulse.

It became clear that a commercial microfluidic device was not suitable for dynamic delivery of 1-NMPP1 to cells. In collaboration with another labmate, Patrick Harrigan, a mesofluidic device was built that might be able to dynamically deliver 1-NMPP1. The device consisted of two fluidic pumps (that was re-purposed from an HTS on a flow cytometer), one of which was used to pump media into the sample and the other to pump media out of the sample (Figure 4.19). Media was pumped out in one line and split into four lines. Similarly, media was removed from four different wells and combined to one line that fed into a waste bottle. Hence, there were two sets of tubing in every well.

An initial test of the media exchange capabilities of the device was carried out. YPD was exchanged with SDC, and fluorescence in the RFP channel was imaged (Figure 4.20). The dynamics of media exchange was very rapid, with an upper bound of 10 seconds, as the images were taken every 10 seconds. Next, the mesofluidic device was use to exchange glucose rich and glucose depleted media. SDC and SDC with 0.05% glucose were alternatively delivered to four separate samples with RFP tagged Msn2 with a period of 10 minutes (Figure 4.21, left). The nuclear localization of Msn2 rose in the presence of glucose depletion media and fell in the presence of SDC. Finally, a 1-NMPP1 washout experiment was performed to determined the speed at which the small molecule could be removed from the media (Figure 4.21, right). The washout of the 1-NMPP1 in SDC media occurred within 90 seconds as denoted by the blue line. The red and green lines are the controls for no 1-NMMP1 and constant 1-NMPP1. This experiment demonstrated the ability of the device to remove 1-NMPP1.

There were two main factors that rendered the device non-functional for dynamically delivering 1-NMPP1. First, the flow rates from the four tubes that delivered media and the four tubes that removed media were not uniform. And it was not clear based on the existing capabilities in the lab how to devise uniform flow rates. Second, it was apparent after the device was built that the tubes used to deliver and remove media could also be a

sticky surface for the 1-NMPP1 molecule, similar to the PDMS. Hence, the device ineffective for the dynamic study of transcription factors even though the media exchange properties of the device were functional.

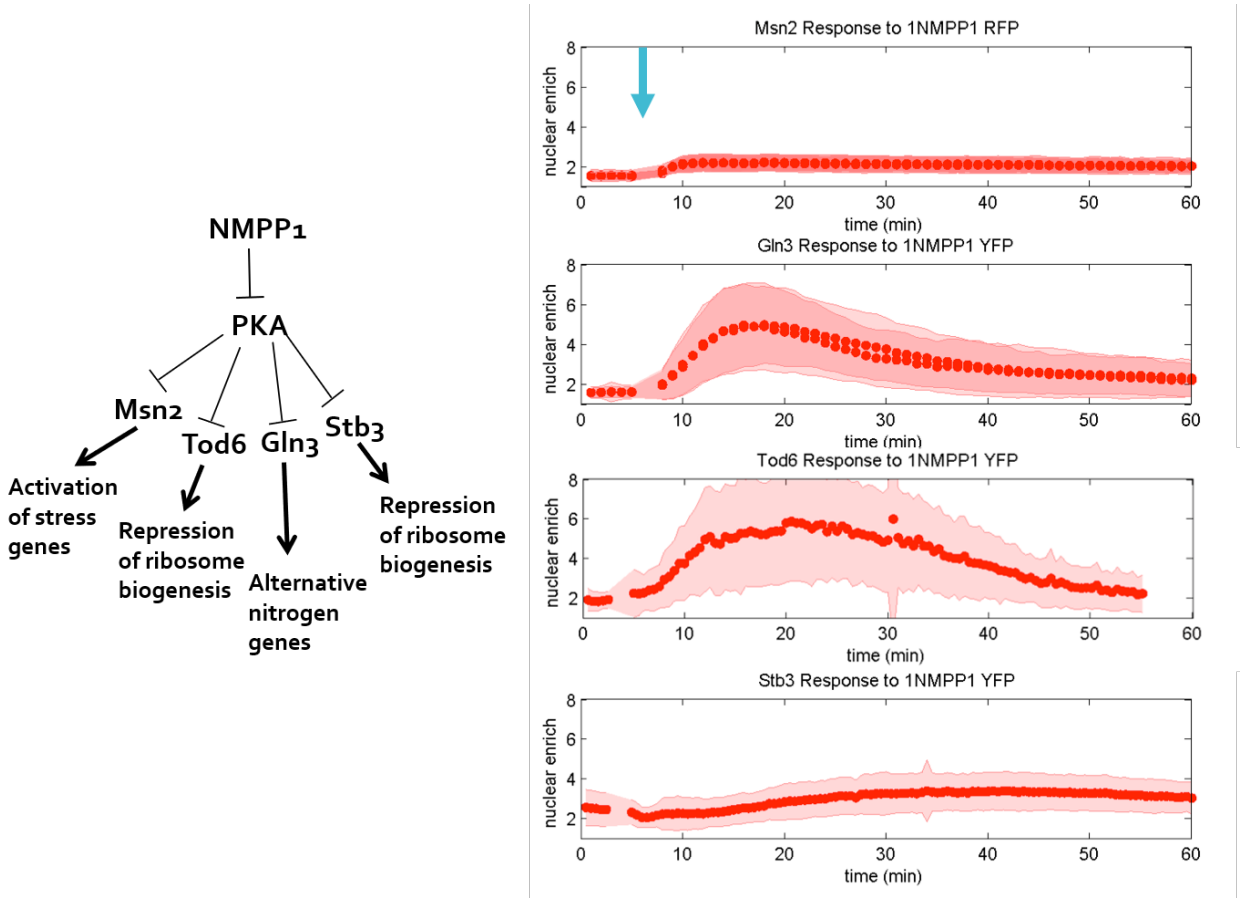


Figure 4.17: **1-NMPP1 can control transcription factor nuclear localization** (left) It is known that PKA regulates different downstream transcription factors, and that 1-NMPP1 is a synthetic inhibitor of PKA. (right) Time course microscopy of nuclear localization in response to a step input of 1-NMPP1 (blue arrow) is plotted. The red dots represent the mean and the envelop represents the standard deviation.

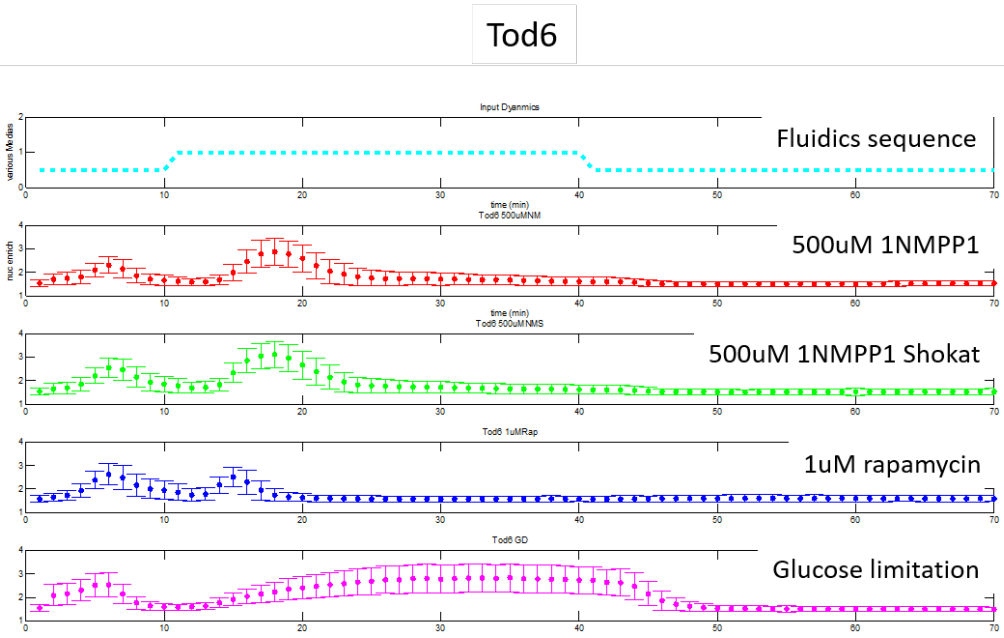


Figure 4.18: **1-NMPP1 delivery was not well controlled by CellAsics microfluidics** For a transcription factor Tod6, a transient pulse of input was delivered for 30 minutes. These inputs included 500uM 1-NMPP1, 500uM 1-NMPP1 from a Shokat stock, 1uM rapamycin, and 0.05% glucose SDC. The time course of nuclear localization was assayed every 1 minute. The error bars represent standard deviation.

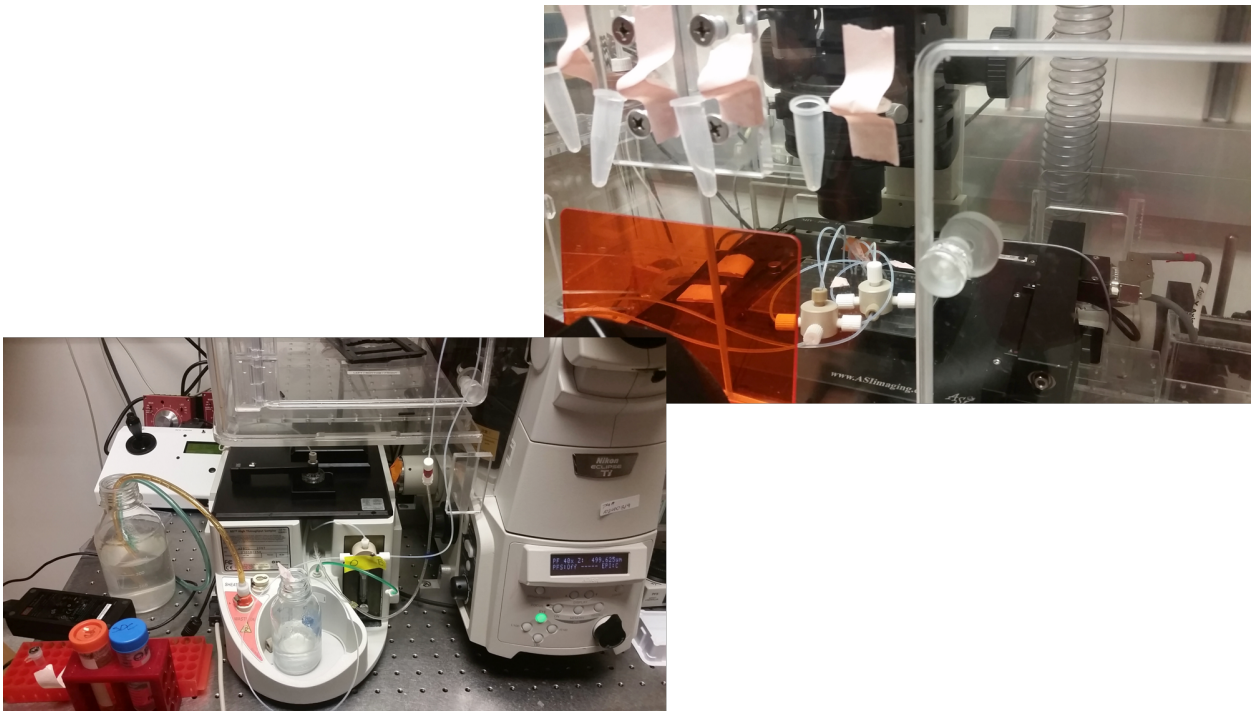


Figure 4.19: **Meso-fluidic contraption physical set-up** A meso-fluidic device controlled by LabView. The High-Throughput Sampler was responsible for delivering and removing media or input from a 96-well imaging plate. Each well contained an input and output tube. The sequence of input delivery was setup in LabView and not synched to the imaging of the microscope.

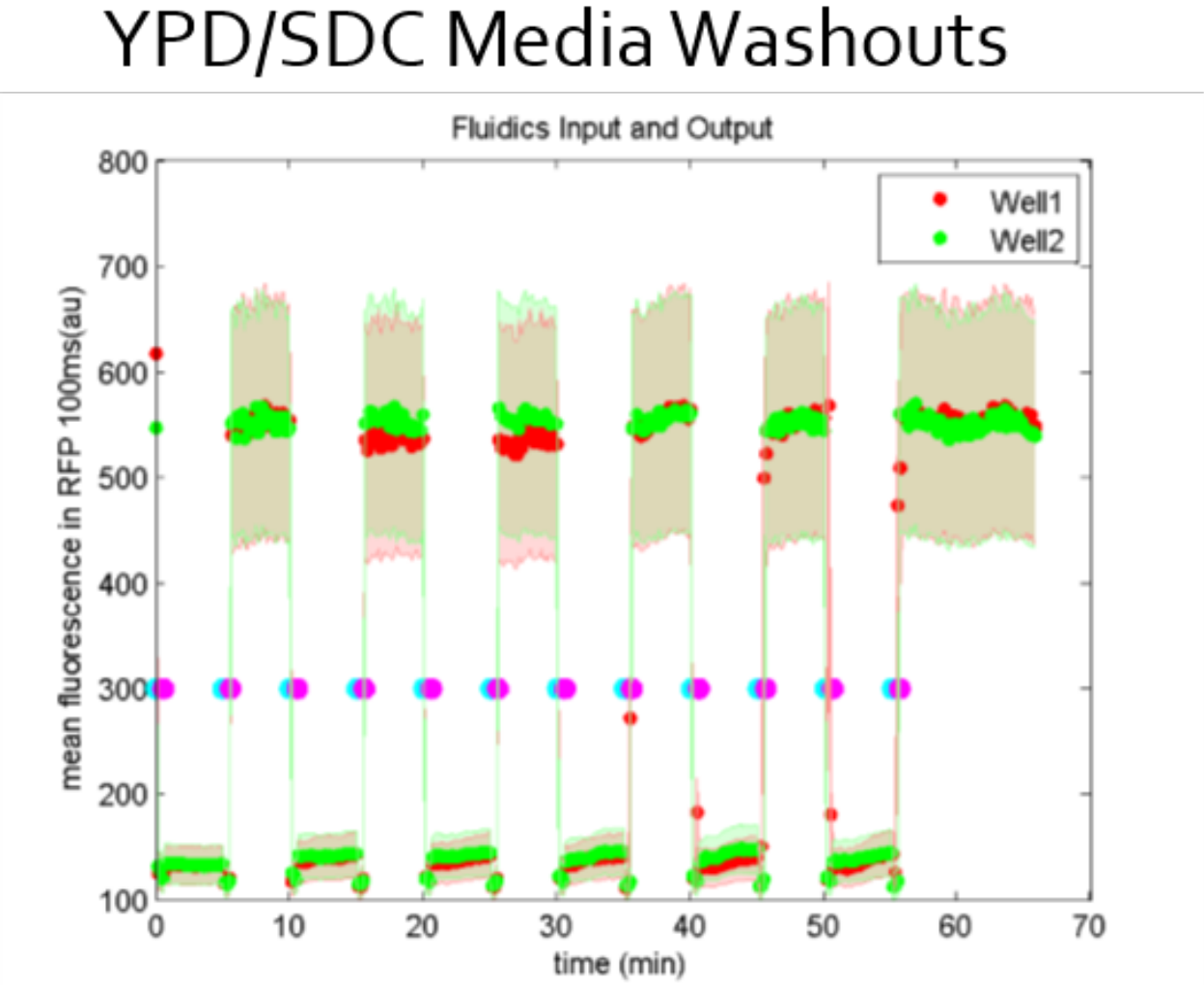


Figure 4.20: **Mesofluidic device able to rapidly exchange media** Time course imaging of RFP every 10 seconds. A sample well with YPD was removed and replaced with SDC periodically. The blue and magenta dots represent when the YPD is removed and SDC is added, respectively. The blue and red traces represent two technical replicates. The shading represents standard deviation.

Glucose Dropout/Drop in

1-NMPP1 Washout test

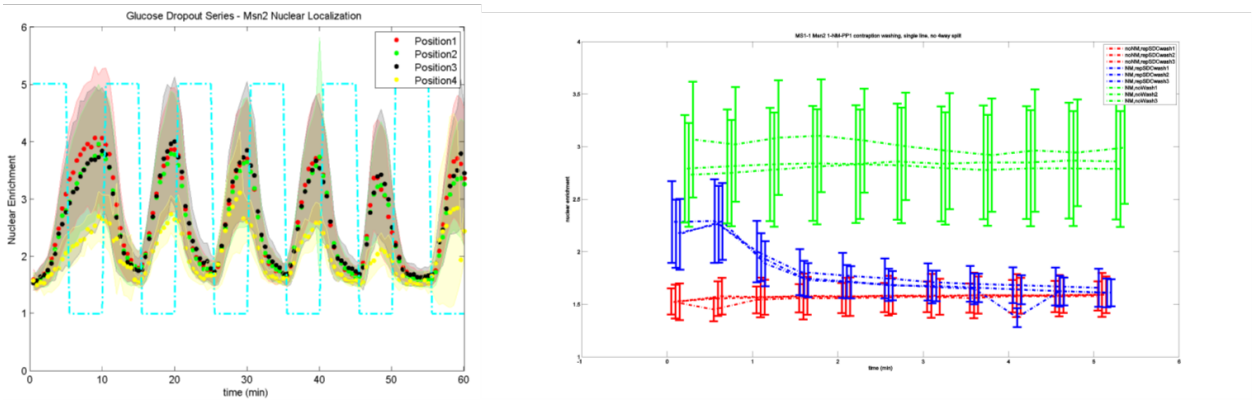


Figure 4.21: **Mesofluidic device able to deliver glucose dropout and 1-NMPP1 washout** (left) Time course of nuclear enrichment of Msn2-RFP is plotted. The different colored dots represent technical replicates and the shaded envelops represent standard deviation. The light blue line represents the delivery of SDC to glucose starved cells. (right) The time course of a 1-NMPP1 washout is plotted. The red lines represent a sample with no 1-NMPP1. The green lines represent sampled induced with 1-NMPP1. The blue lines represent a sample with 1-NMPP1 that is then washed out with SDC at time, t=0. The clustered lines are technical replicates. The error bars represent standard deviation.

4.7 yeLANS can precisely control nuclear localization of TFs but show weak activation of gene expression

Prior to CLASP, there was a lot of work on a previous iteration of an optogenetic tool to control transcription factor. This tool, yeast optimized LANS (yeLANS), utilized the same LOV2 uncaging of an NLS to localize a tagged cargo to the nucleus[13] (Figure 4.22A). A small library of NLSs of different classes[56] were screened for optimized dynamic range of nuclear localization, with a class III NLS being used in the final yeLANS construct (Figure 4.22B). The nuclear localization of an mCherry-yeLANS was assayed. Specifically, different

periods of light were delivered to test the speed and time resolution of nuclear localization (Figure 4.22C). The mCherry-yeLANS has a time resolution of about 2min.

yeLANS was then tagged to several stress transcription factors such as Rtg1, Rtg3, and Dot6. These yeLANS tagged transcription factors were able to localize to the nucleus for the duration of the light input, which was 60 minutes (Figure 4.23).

Similarly, Crz1-yeLANS was also able localize to the nucleus for the duration of the light input, which was 40 minutes (Figure 4.24). Crz1-yeLANS was of particular interest because Crz1 localization differs based on the environmental input. In natural Crz1 signaling, Crz1 de-phorylates and enters the nucleus in response to both $CaCl_2$ and PH8 (Figure 4.25, left top). It is possible that these two input produce different gene expression programs, as suggested by the published data sets of gene expression in response to $CaCl_2$[62] and PH8[90] (Figure 4.25, left bottom). The nuclear localization of Crz1 in response to $CaCl_2$ and PH8 were measured and quantified. It was found that $CaCl_2$ translocated more transiently to the nucleus compared to PH8, which exhibited more sustained nuclear localization (Figure 4.25, right). One hypothesis that arose from this observation was that the differences in gene expression could be due to the dynamic differences of Crz1 nuclear localization in response to the two inputs.

To test this hypothesis, qPCR of seven known Crz1 target genes was performed for Crz1-yeLANS strains that were exposed to light, to 0.2M $CaCl_2$, PH8, or no input (Figure 4.26). The most striking observation was that both $CaCl_2$ and PH8 exhibited large fold changes to target gene expression. Gene expression in response to 40 minutes of light was weak at best, with the highest fold change being 3.34 and many other under 2 fold. This suggested that Crz1-yeLANS was not able to activate downstream genes to nearly the same extent as natural inputs. To further confirm this hypothesis, time course RNA sequencing was performed in the conditions of 0.2M $CaCl_2$ and light for 100 minutes. The gene expression profiles partially confirmed the hypothesis that Crz1-yeLANS could not activate

downstream genes to the same extent as $CaCl_2$. At 60 minutes, expression of Crz1 target genes was low compared to 60 minutes with 0.2M $CaCl_2$. However, at 100 minutes, the gene expression profiles of light and 0.2M $CaCl_2$ inputs were more comparable, suggesting that Crz1-yeLANS activates Crz1 target genes more slowly than 0.2M $CaCl_2$ (Figure 4.27, top panel). This phenotype is plotted more explicitly with gene expression as a function of time for the mean of and individual Crz1 target genes for the two inputs of $CaCl_2$ and light (Figure 4.27, bottom panel).

One reason for this muted and delayed expression in response to Crz1-yeLANS is that the Crz1 in this construct is not de-phosphorylated and that de-phosphorylation could be play a role in activation of gene expression in the natural system. A Crz1 S19A mutant was constructed that confirmed this hypothesis (see Supplementary Figure 3.9 in Chapter 3). This finding was the main rationale for using a Crz1 S19A mutant. The necessity of adding a second sequestration domain to Crz1*-yeLANS to form Crz1*-CLASP was due to the constitutive nuclearly localized phenotype of the Crz1 S19A mutant.

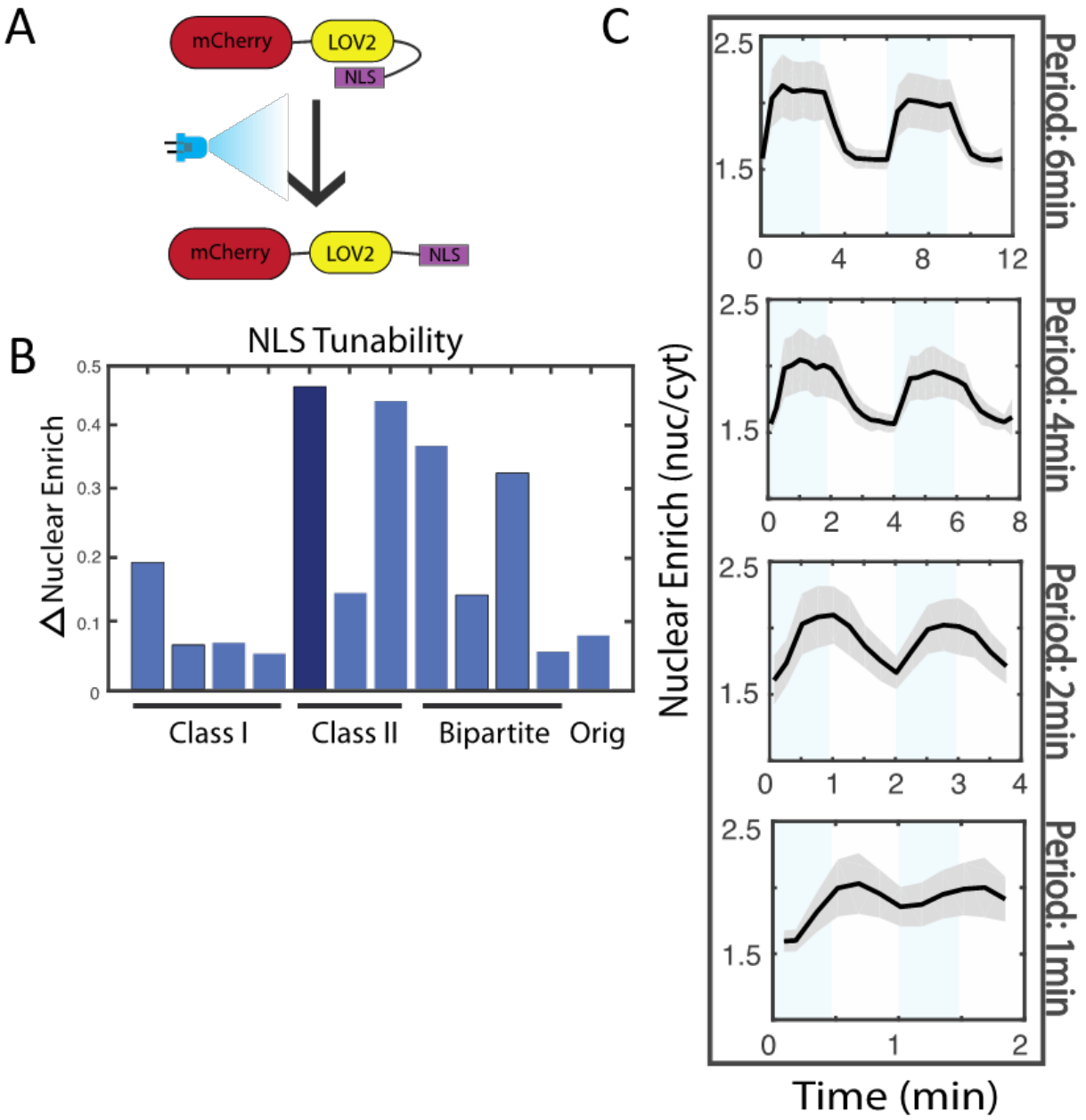


Figure 4.22: **Characterize optogenetically induced nuclear localization with a fluorescent protein (FP)** A) Schematic of light-gated Nuclear Localization Sequence (NLS) sequestration mechanism B) Using an mCherry reporter, a panel of NLSs were tested in response to light. C) Decreasing periods of blue light (ranging from 6 to 1 min) were applied. The mean nuclear enrichment (solid line) and standard deviations (shaded area) was measured via microscopy.

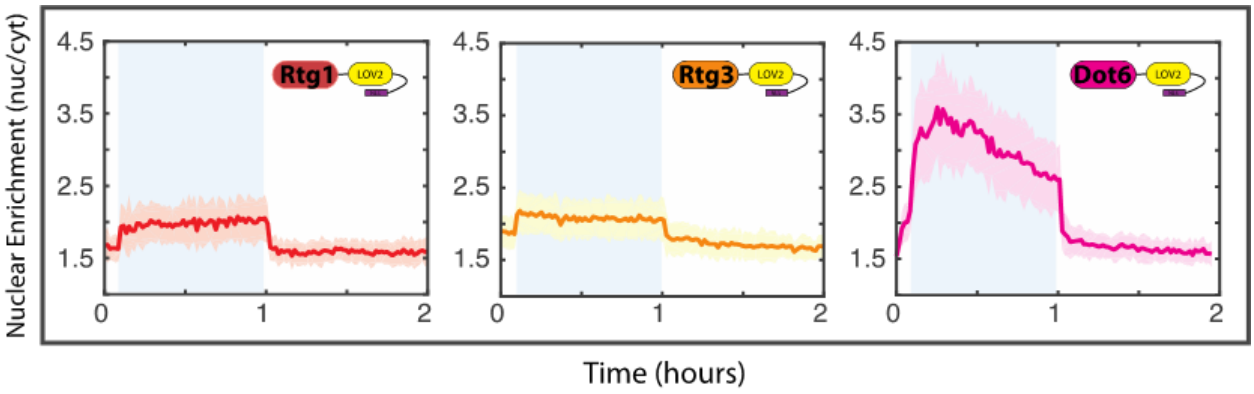


Figure 4.23: **Optogenetically induced nuclear localization of naturally pulsing TFs** A 60 minute pulse of blue light at 450nm at approx. 40uW/mm2 was applied to cells with TF tagged with mCherry and the optogenetic tag. The light input holds TFs in the nucleus over a long period of light durations.

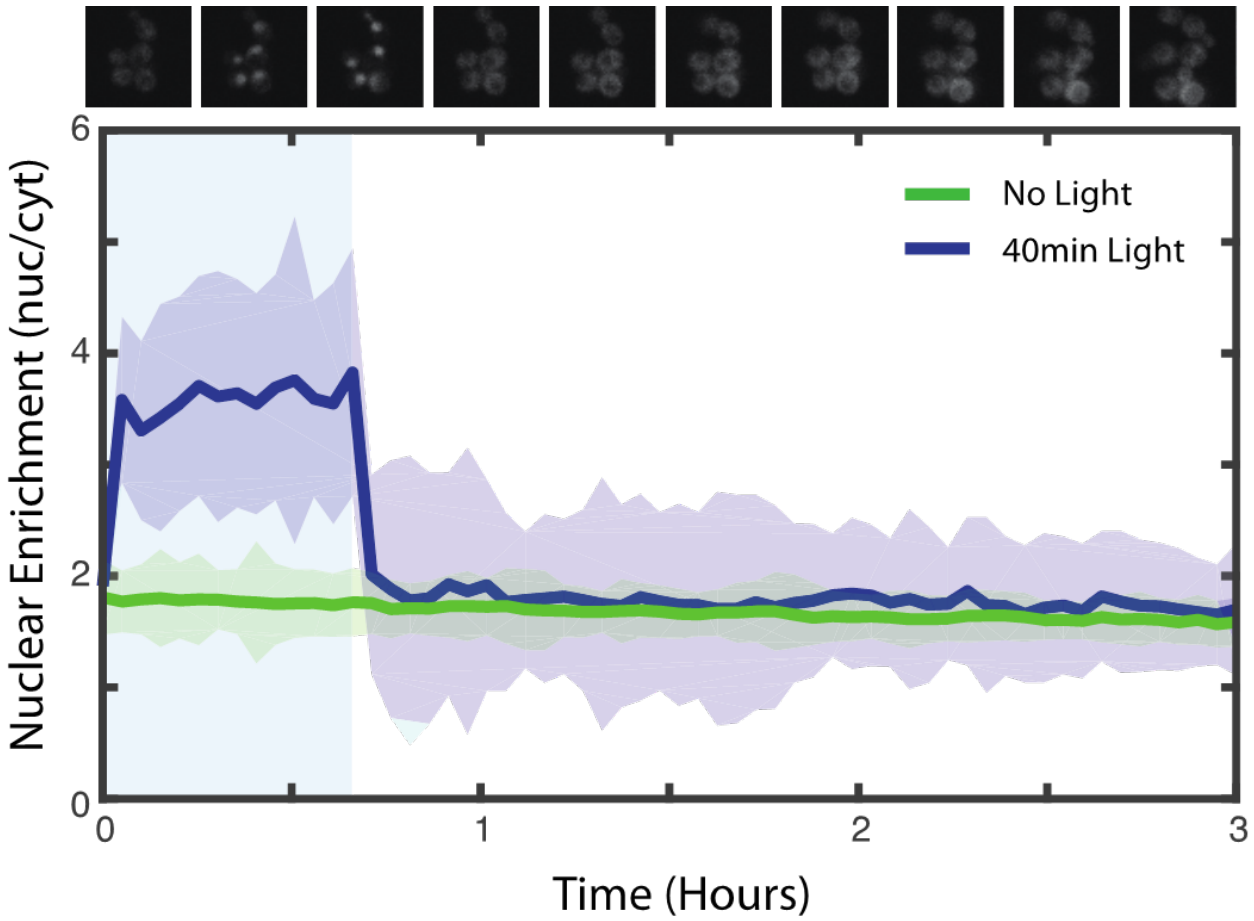


Figure 4.24: **yeLANs can precisely control Crz1 nuclear localization** Crz1 localization in response to 40min light is measured.

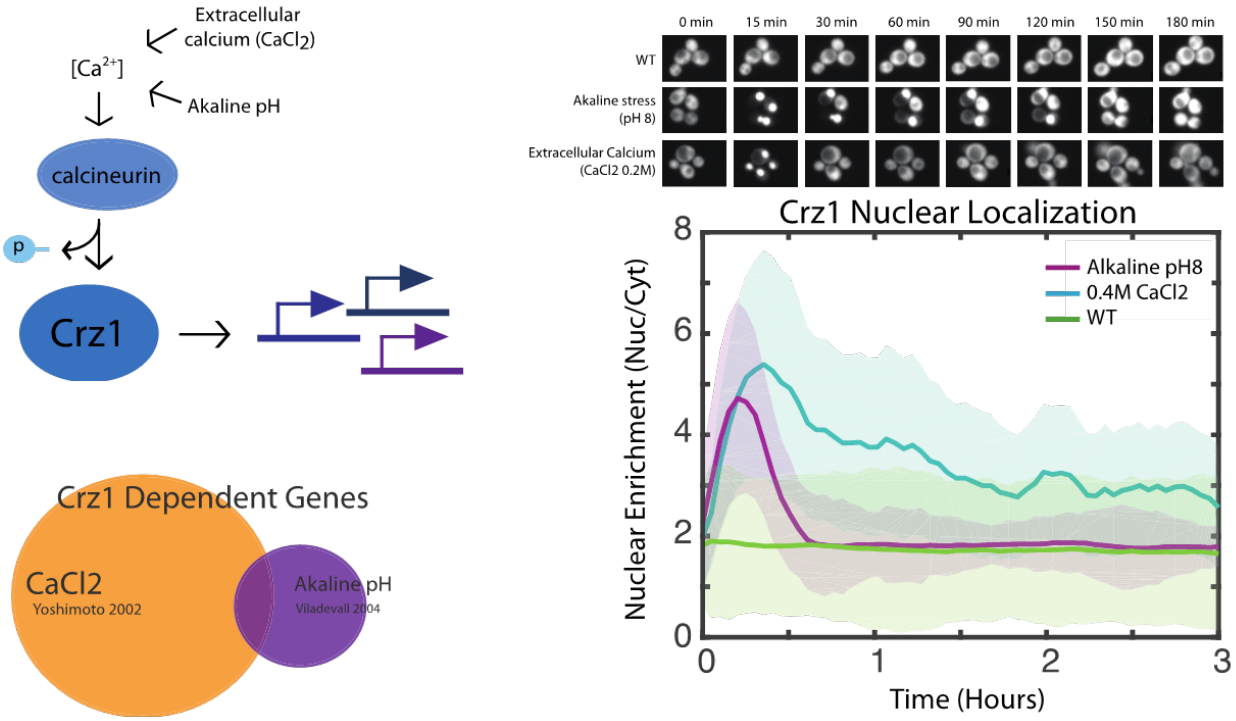
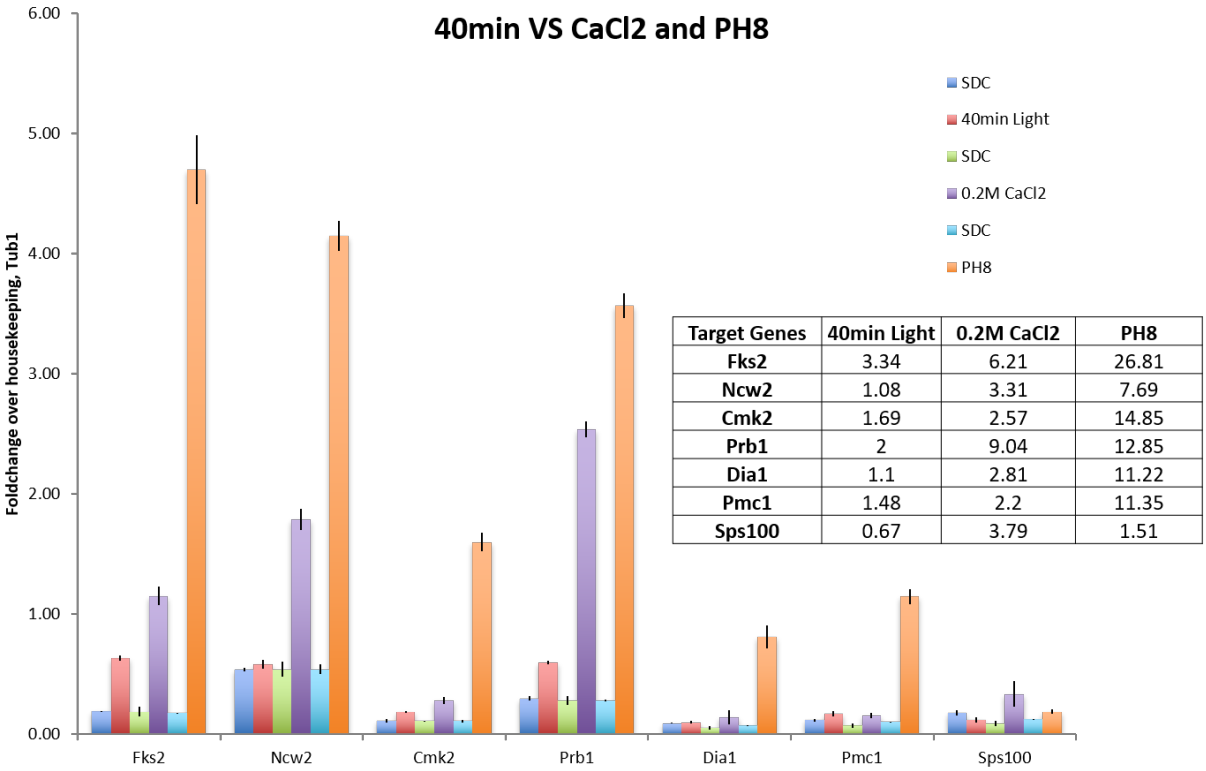


Figure 4.25: **Crz1 nuclear localization in response to natural inputs** Crz1 is a key signaling mediator for transcriptional activity in response to Ca2+ changes (CaCl2 and alkaline pH). Upon these natural inputs, Crz1 translocates to the nucleus with distinct localization profiles. These Crz1 profiles may contain regulatory information for activating distinct gene sets.



Target Genes	40min Light	0.2M CaCl2	PH8
Fks2	3.34	6.21	26.81
Ncw2	1.08	3.31	7.69
Cmk2	1.69	2.57	14.85
Prb1	2	9.04	12.85
Dia1	1.1	2.81	11.22
Pmc1	1.48	2.2	11.35
Sps100	0.67	3.79	1.51

Figure 4.26: **Crz1-yeLANs either weakly or does not activate Crz1 target genes** qPCR mRNA expression is assayed from 7 Crz1 target genes after 40minutes of light induction or 40minutes of $CaCl_2$ or PH8 natural input. The inset shows a table of expression fold changes in response to different inputs.

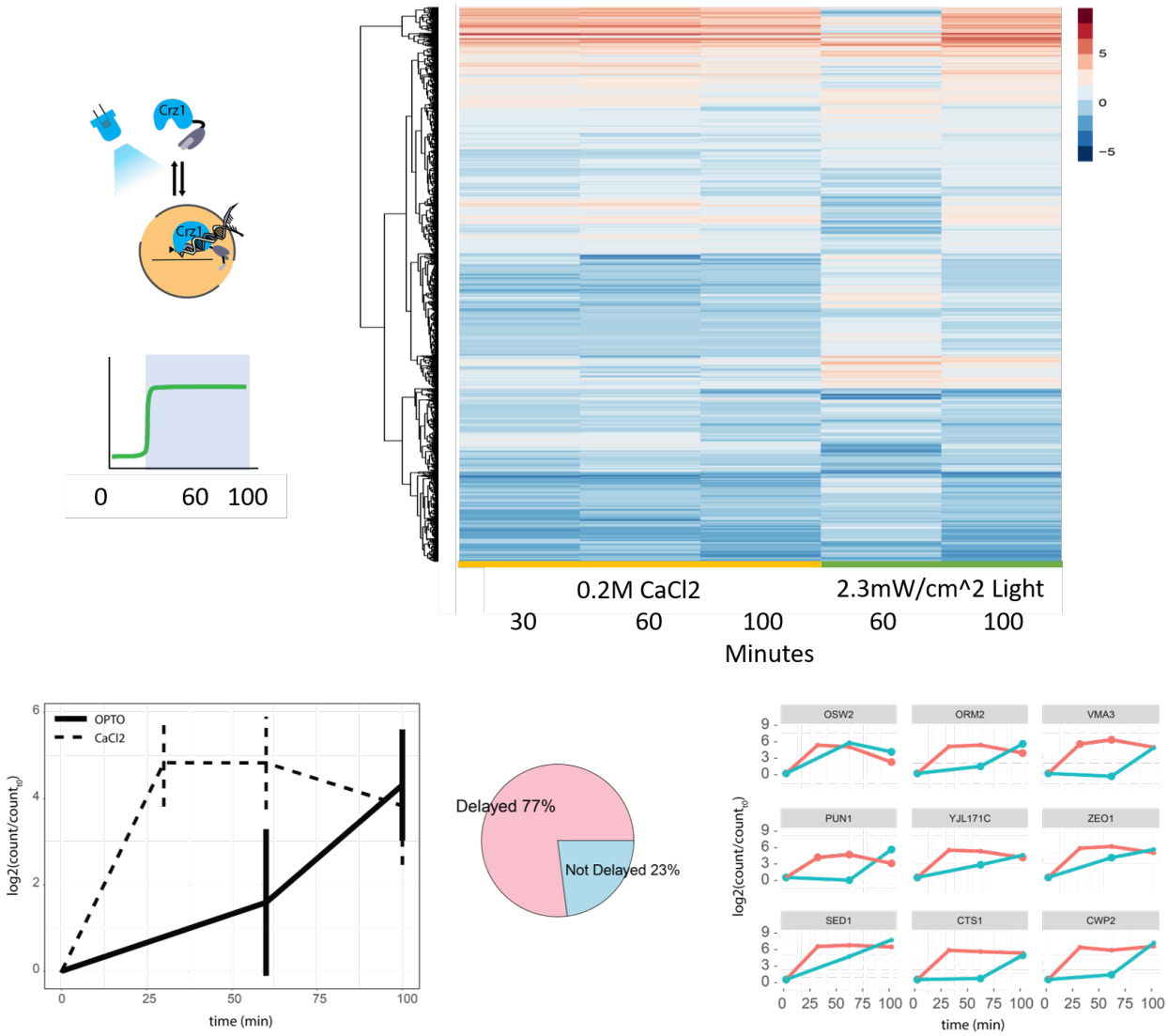


Figure 4.27: **Expression profiling shows delayed gene expression of Crz1-yeLANS compared to 0.2M** $CaCl_2$ (top left) Schematic of light experiment in which 100min of blue light is delivered to cells. (top right) Gene expression profiling of time courses for 0.2M $CaCl_2$ at 30, 60, and 100 minutes, and for 100 minutes of light exposure (2.3mW/cm^2) at 60 and 100 minutes.The intensities are fold-change over no input. The profiles are hierarchically clustered. (bottom left) Log2 fold-change of the average of all Crz1 target genes are plotted over time. The solid line denotes light induction while dotted line denotes $CaCl_2$ induction. The errorbars represent standard deviation. (bottom middle) Pie chart showing the percentage of Crz1 target genes that are delayed and not delayed in the strains exposed to light. (bottom right) Example Crz1 target gene log2 fold-change time traces with turquoise line denoting light induced samples while salmon line denotes $CaCl_2$ induced samples.

4.8 Gene expression profiling of Crz1-yeLANS identifies "fast" and "slow" genes

From the previous qPCR and transcriptome profiling studies of gene expression in response to Crz1-yeLANS, it is clear that the Crz1-yeLANS is a significantly less efficient and delayed version of the natural Crz1 transcription factor. While the study of Crz1 dynamics proceeded with the Crz1 S19A mutant, Crz1-yeLANS could still serve to identify target genes that are easily or not easily activated by dynamic Crz1-yeLANS inputs. The caveat to this approach is that a "blunted" Crz1 could alter more than just the timing of gene expression response to Crz1. In fact, this study cannot be compared to natural Crz1 gene expression. With this caveat in mind, use of the Crz1-yeLANS could still be used as a case study for how target genes might respond to transient transcription factor nuclear localization inputs. Crz1-yeLANS was exposed to either a transient 40 minute pulse of light or continuous light for 120 minutes. The nuclear localization of Crz1-yeLANS showed that nuclear localization remained for the duration of the light input (Figure 4.28A). Cells were collected at time 0, 30, 60, and 100 minutes, and gene expression profiling using RNA sequencing was performed. The resulting gene expression profiles are then analyzed for differential gene expression. The fold change of gene expression were then clustered in a heatmap (Figure 4.28B). A clear result was that most genes tended to follow the dynamics of the light input, meaning that they increased in expression when the light was on and returned to basal when the light was off. The inset shows examples of dynamic traces. There was also a small subset of genes that did not turn on in the transient input but turned on after 60 minutes or greater of light illumination. The gene expression profiling showed that most genes were "fast" in that they responded to the transient input.

The two groups of genes, "fast" and "slow", could be analyzed for Gene Ontology Enrichment. The "fast" or "linear" genes have two significantly enriched GO Term which

are Cellular Differentiation and Reproductive Process. The "slow" genes were enriched in Trehalose Metabolism in response to stress (Figure 4.29A). Hence, there may be some functional differentiation between the "fast" and "slow" genes.

The growth of Crz1-yeLANS cells were also assayed for transient and continuous (or sustained) light inputs. A growth defect was observed for the continuous light input, suggesting that continued expression of Crz1 target genes is deleterious for the cell (Figure 4.29B). In summary, this study used an optogenetic tool called yeLANS to mimic the nuclear shuttling of Crz1 in response to either $CaCl_2$ or PH8 stress inputs, which are stereotyped as transient and continuous, respectively. Gene expression profiling was then done for light-induced transient and continuous localization of Crz1. The experiment identified two groups of genes that either followed the pulse, termed "fast" (or "linear"), or "slow" genes that did not turn on in the transient pulse. It will be interesting explore further the potential mechanism for the "fast" and "slow" behavior of the genes. One hypothesis that was already put forth by the Crz1*-CLASP study is the presence of nucleosomes in the "slow" genes (Figure 4.30). By promoter swapping and further promoter bashing, it will be exciting not only to identify the mechanistic components that underlie these two behaviors but also to forward engineer promoters sensitive to dynamic nuclear localization.

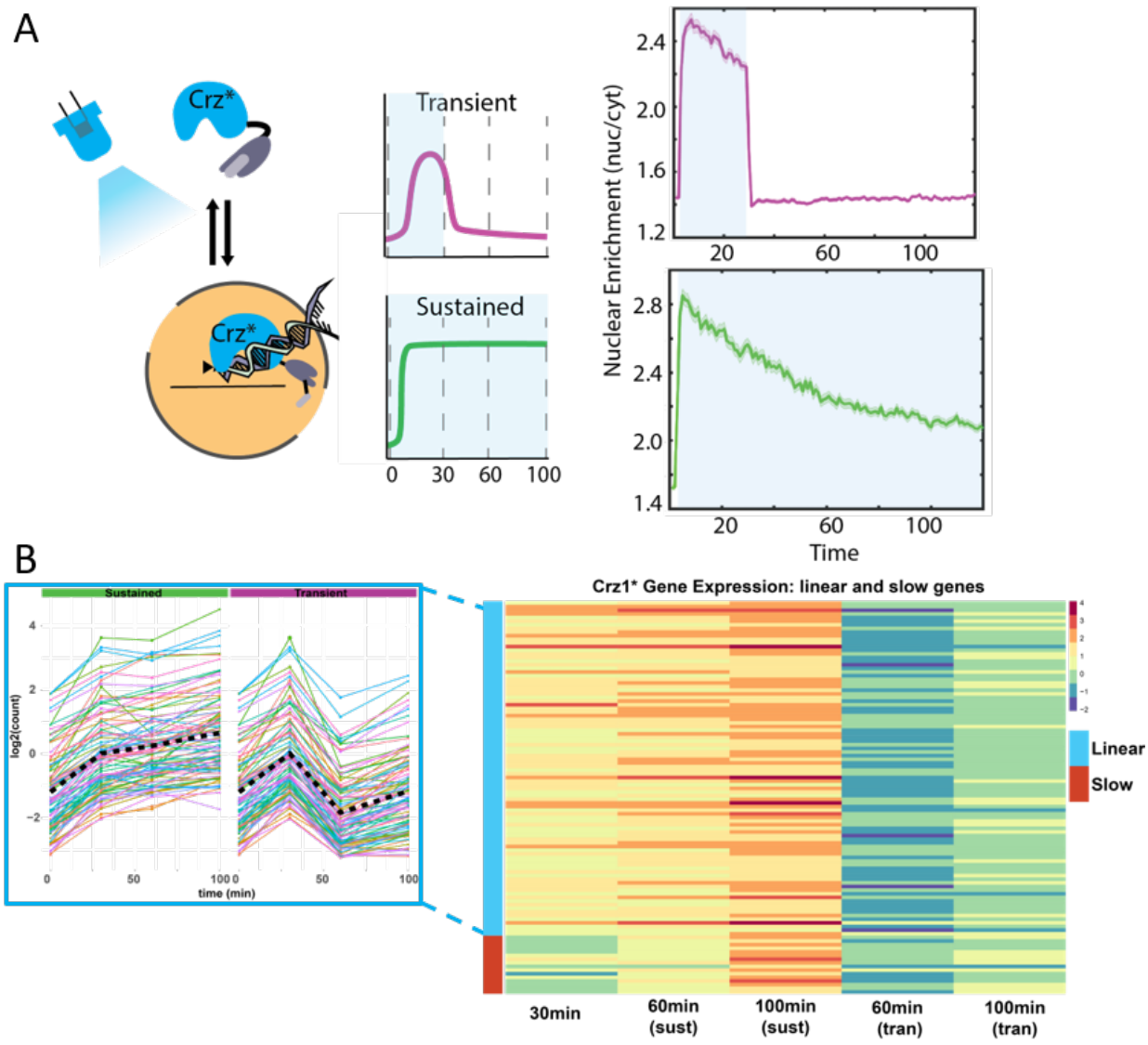


Figure 4.28: **Crz1 Nuclear shuttling and RNAseq of gene expression** A) Experimental setup with optogenetically shuttling Crz1. RNAseq samples are collected at 0, 30, 60, and 100 minutes. B) Heatmap of gene expression grouped by time trajectories. Genes are grouped into "linear" and "slow" sets. The individual time traces for the "linear" genes are shown (inset).

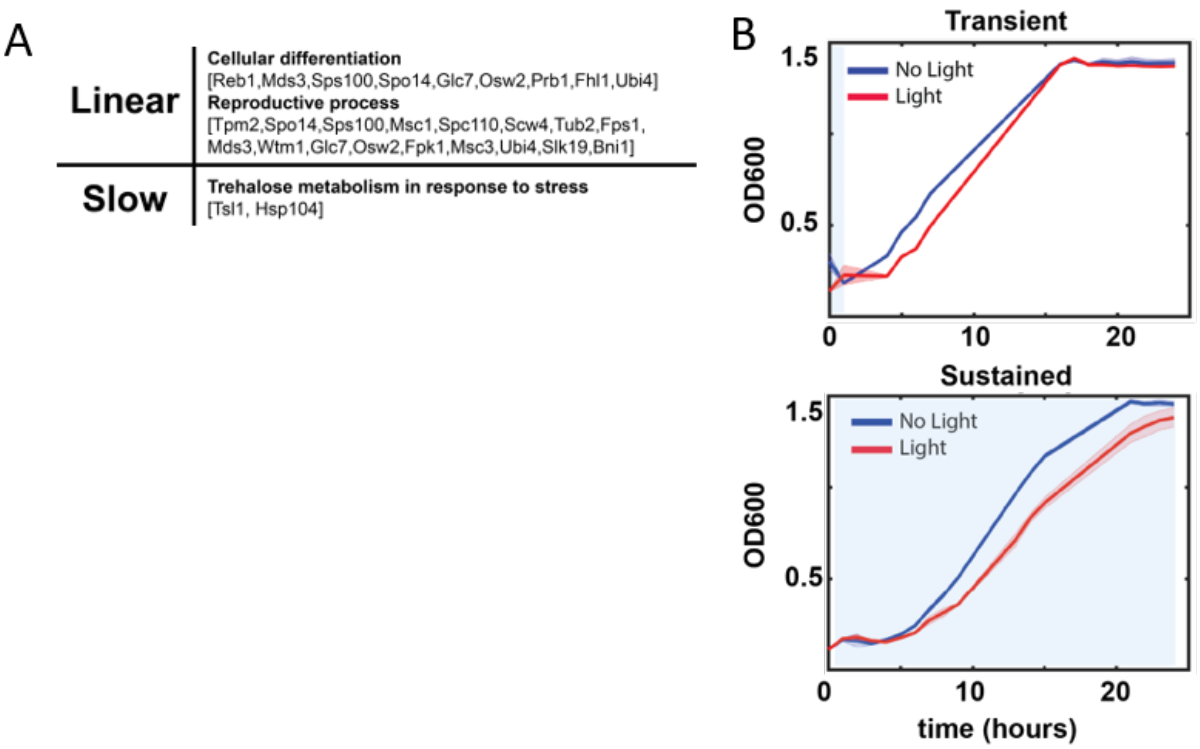


Figure 4.29: [**Functional interpretation of linear and slow genes** A) Table of GO terms and genes that are enriched in the two cluster of genes termed "linear" and "slow". B) Growth as measured by OD600 for cells grown with a transient 40 minute blue light exposure compared to cells with continuous 24 hour light exposure.

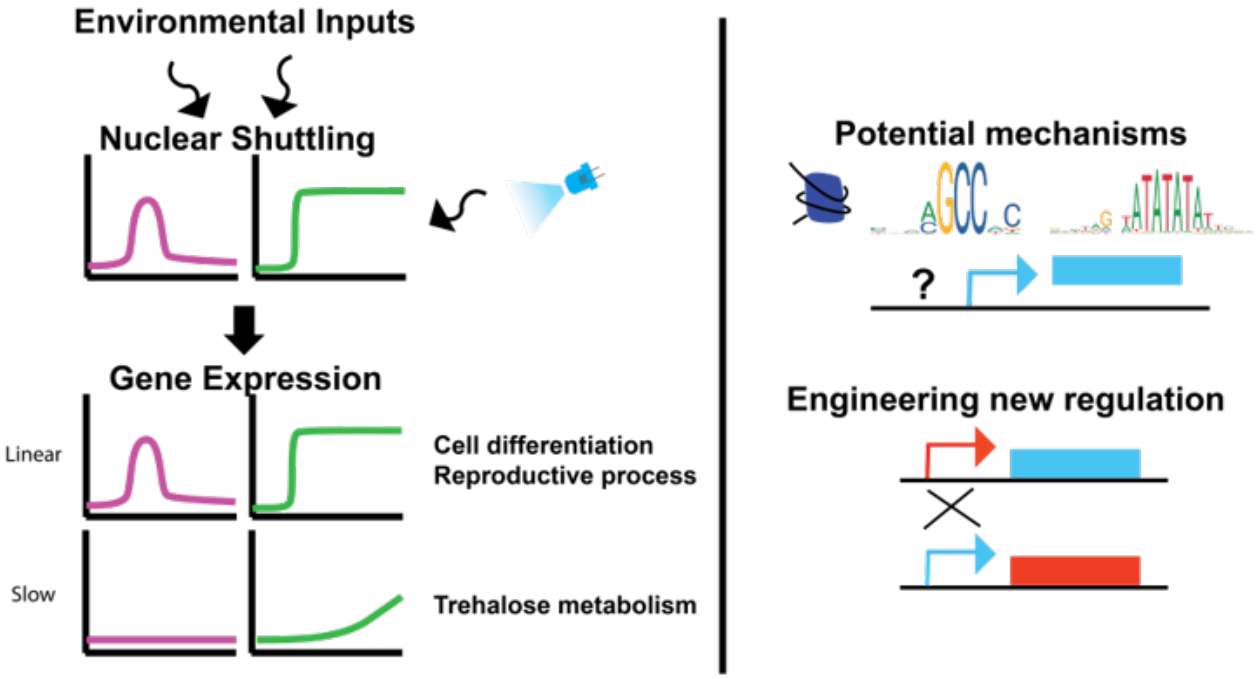


Figure 4.30: **Summary of findings** A schematic for the findings in this study, potential next steps to further identify the mechanism of fast or slow genes, and possibility of engineering new regulation.

4.9 Other Notes on Crz1*-CLASP

4.9.1 Crz1*-CLASP is inactive in the dark even when in the nucleus

When characterizing CLASP, the measured dose response of Crz1 target genes pCMK2-venus and pGYP7-venus at different promoter expressions of Crz1*-CLASP did not match the gene expression measured in the presence of light (Figure 4.32, right panels). Intriguingly, for a given nuclear amount of Crz1*-CLASP, gene expression was higher when that amount was achieved through light induction than through an increase in the

expression of Crz1*-CLASP from a stronger promoter in the absence of light.

One hypothesis was that adaption to constitutive amounts of nuclear Crz1 at different promoter levels causes the lowered dose response in the absence of light induction. To test this hypothesis, target gene expression was measured in the presence of constant light that maintained Crz1*-CLASP or synTF-CLASP in the nucleus for 10 hours (Figure 4.31). If there was adaptation, the expectation was that gene expression would reach a maximum then decrease to a lower gene expression. This is not observed, and hence adaptation of target genes was ruled out as a possible cause for the lowered target gene dose responses in the absence of light.

Another hypothesis was that the CLASP construct prevented the transcription factor from being active in the dark. This would be the case if the synTF target gene expression also showed the same phenotype as the Crz1 target genes. Indeed, the synTF target gene dose response was saturated for increasing synTF-CLASP expression levels from different promoters in the dark, but was much more dose responsive when the nuclear amount was delivered by light activation (Figure 4.32, left panel). Hence, increased gene expression due to increased TF nuclear localization levels only occurs with light-activated of TF-CLASP.

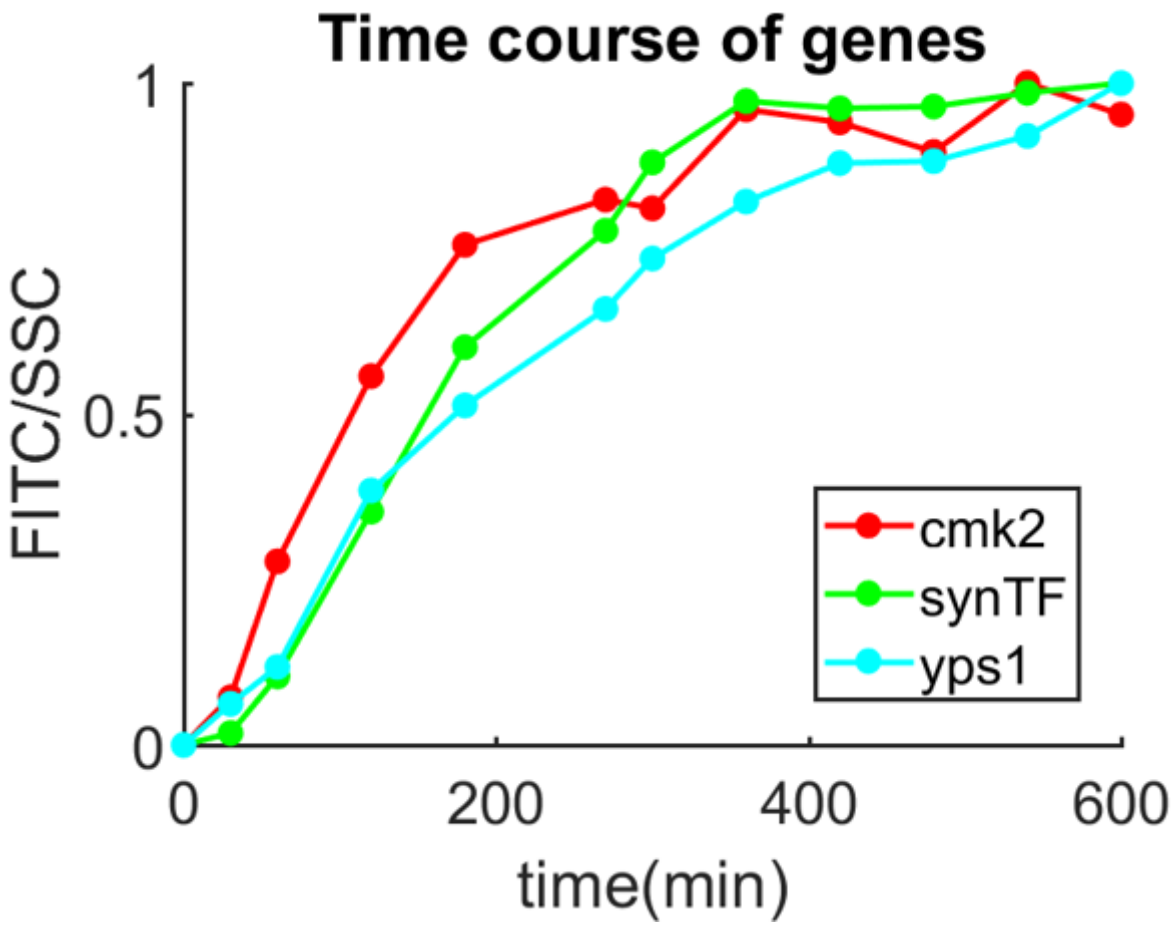


Figure 4.31: **Time course gene expression in constant blue light** Target gene expressions of Crz1*-CLASP and synTF-CLASP are plotted as a function of time in constant blue light.

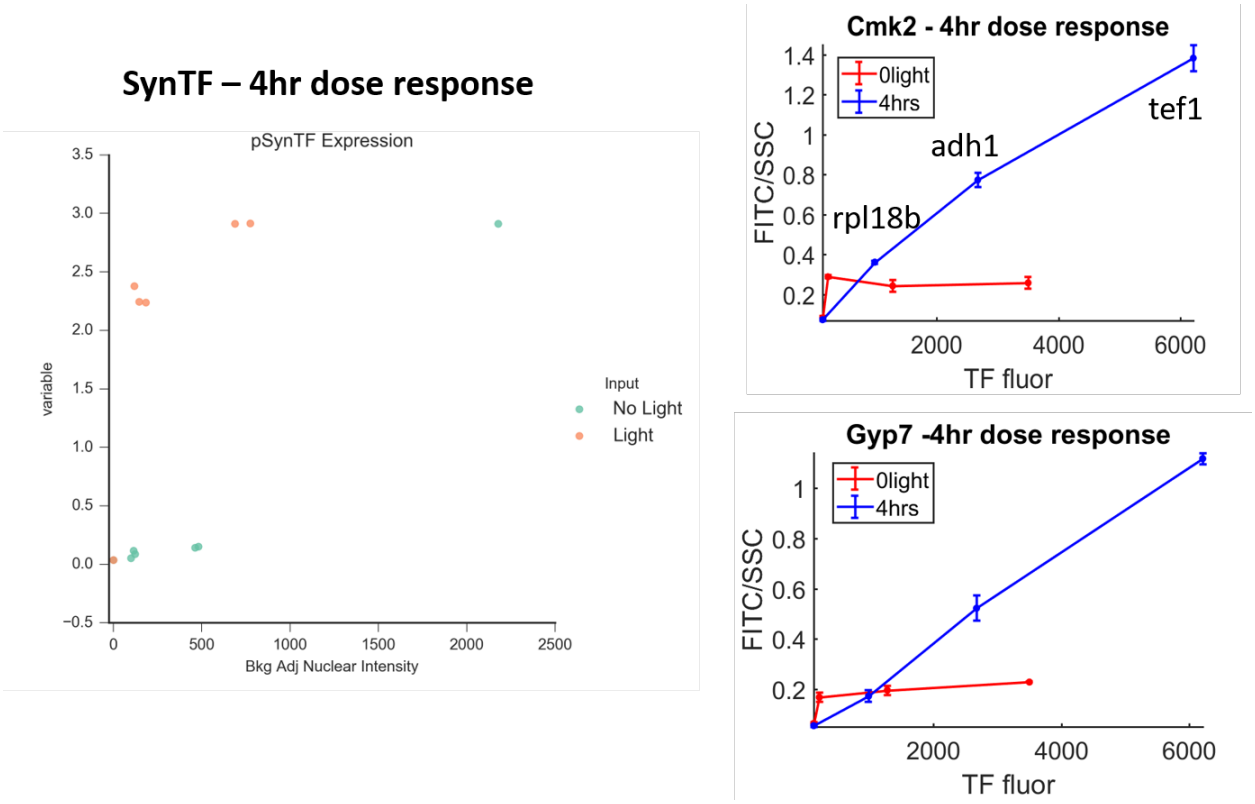


Figure 4.32: **Dose response to target genes with and without light activation** (left) SynTF dose response at different promoter expression levels (pRpl18B, pADH1, and pTEF1) of synTF-CLASP induced with 4 hours of light (orange dots) or no light (green dots) measured at 5 hours after light induction. Each cluster of dots represents biological replicates. (right, top panel) Cmk2 dose response obtained in the same manner as synTF. (right, bottom panel) Gyp7 dose response obtained in the same manner as synTF.

4.9.2 Crz1*-CLASP can override natural $CaCl_2$ signal

In addition to using CLASP to mimic natural TF localization events, CLASP can also be used to study nuclear localization and gene expression in the context of natural inputs. For instance, Crz1*-CLASP can be prevented from entering the nucleus even in the presence of different concentrations of $CaCl_2$ (Figure 4.33). To test whether the sequestration of Crz1*-CLASP can prevent gene expression of Crz1 target genes in the presence of $CaCl_2$, pRPL18B-Crz1*-CLASP was exposed to increasing concentrations of $CaCl_2$ in the absence

or presence of light, and the resulting gene expression was measured using RNA sequencing (Figure 4.34). The log2 fold-change expression over no input, showed that sequestration is not complete even in the 0.1M $CaCl_2$, and sequestration is less effective as the concentration of $CaCl_2$ increases. Despite this, in the presence of light, which localizes pRPL18B-Crz1*-CLASP to the nucleus, increased expression of $CaCl_2$ genes were observed (Figure 4.34). This indicates that sequestration is able to partially control gene expression in the presence of a natural stress.

The potential of CLASP, with optimization of its sequestering capabilities, could enable its use in dynamic TF knockouts and rescues in the context of stress inputs. It could precisely test questions such as the contribution of Crz1 in $CaCl_2$ signaling. It could also probe the effects of changing the dynamic signaling and timing of Crz1 (and other TFs) in the combinatorial control of genes that occur in the background of natural stress inputs. This property of CLASP further demonstrates its versatility as a molecular tool.

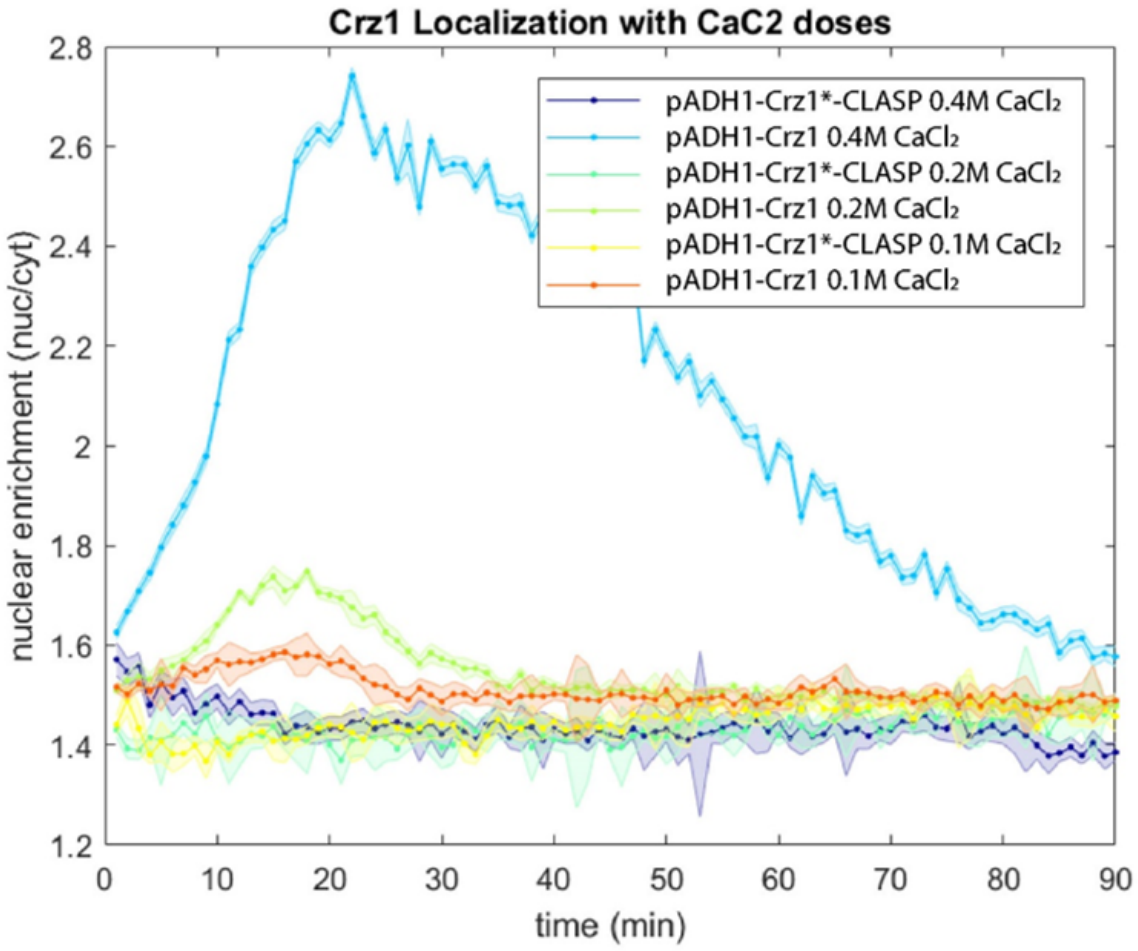


Figure 4.33: **Crz1*-CLASP sequestration in the presence of** $CaCl_2$ Time course of Crz1 or Crz1*-CLASP nuclear localization in response to increasing concentrations of $CaCl_2$. Mean localization is represented by the solid line and the shaded envelop represents the standard error.

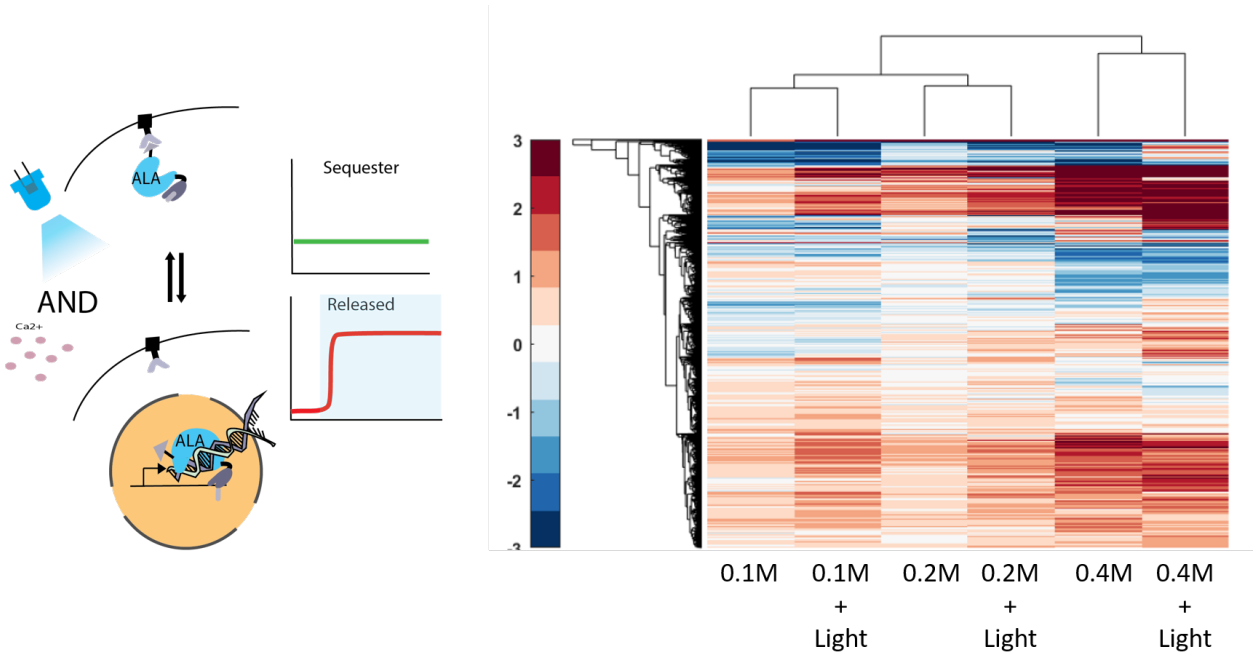


Figure 4.34: **pRPL18B-Crz1*-CLASP sequestration from $CaCl_2$ and release with light** (left) Schematic of experimental setup in which pRPL18B - Crz1 S19A - CLASP is exposed first varying concentrations of $CaCl_2$ and then induced with light. (right) Log2 fold-change over no input of gene expression for cells exposed to increasing 0.1M, 0.2, and 0.4M $CaCl_2$ in the absence and presence of light.

4.10 Target gene interpretation of natural Crz1 pulse in response to 0.2M $CaCl_2$

The motivation for using light to nuclear localize transcription factors was inspired by the observation that many transcription factor localize dynamically to the nucleus in response to natural inputs. A light based approached was used and documented extensively in this work to study the transcriptional consequences of dynamic nuclear localization. It would be interesting to study the transcriptional responses of known TF target genes to natural inputs. Using Crz1 again as a test bed, 0.2M $CaCl_2$ was added and the transcriptional

outputs of Crz1 target genes were measured over time. It was found that a large portion of target genes (65%) follow the dynamics of the initial concerted Crz1 nuclear localization pulse, while 26% of the genes remained on even when the Crz1 localization had receded and 9% of the genes initially increased but decreased below the initial levels before returning to the pre-perturbation levels (Figure 4.35). These results suggest that target genes are able to differentially interpret the initial concerted and transient pulse of Crz1 nuclear localization induced by natural $CaCl_2$ input.

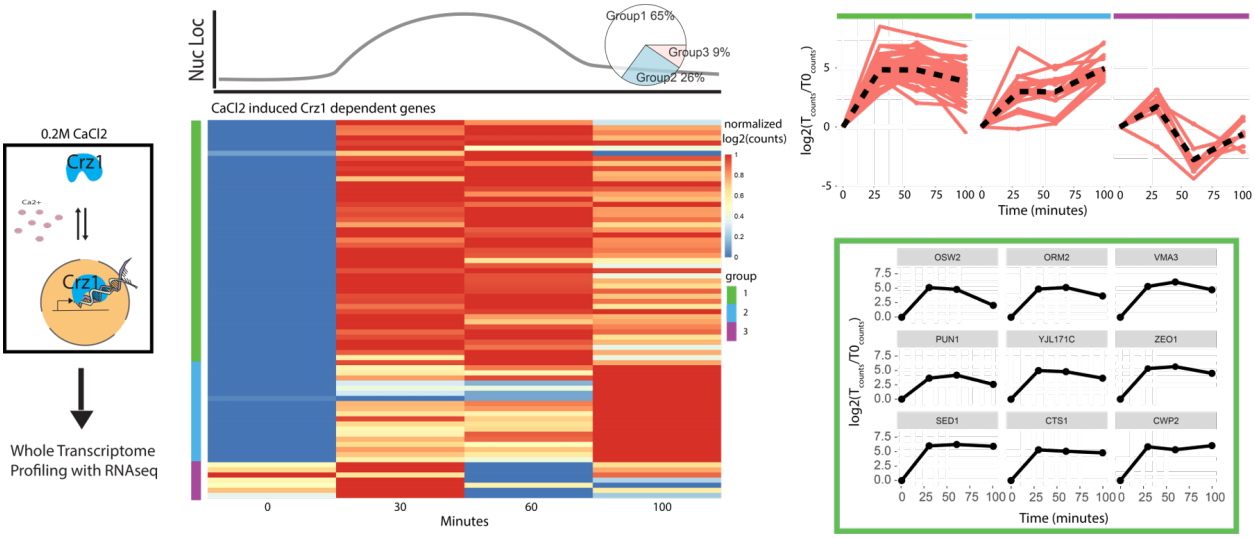


Figure 4.35: **Time-course profiling of gene expression in response to 0.2M** $CaCl_2$ (left) Schematic of experimental setup (middle, top) Cartoon trace of Crz1 nuclear localization pulse in response to $CaCl_2$. A pie chart of Crz1 targets that belongs to 1 of 3 groups of responses. (middle, bottom) Normalized log2(read counts) of $CaCl_2$ induced Crz1 dependent genes at 0, 30, 60, and 100 minutes. Dynamic behaviors are categorized into 3 groups denoted by 3 colors. (right, top) Gene expression over time for individual genes separated by groups. The black dotted line is the mean of the time courses while the salmon lines are individual genes. (right, bottom) Gene expression over time of sample Crz1 target genes in group 1 (green).

Bibliography

[1] Stefan Hohmann. Osmotic stress signaling and osmoadaptation in yeasts. 66(2): 300–372. ISSN 1092-2172. doi: 10.1128/MMBR.66.2.300-372.2002. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC120784/.

[2] Kevin A. Morano, Chris M. Grant, and W. Scott Moye-Rowley. The response to heat shock and oxidative stress in saccharomyces cerevisiae. 190(4):1157–1195. ISSN 0016-6731. doi: 10.1534/genetics.111.128033. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3316637/.

[3] Lee Bardwell. A walk-through of the yeast mating pheromone response pathway. 26 (2):339–350. ISSN 0196-9781. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3017506/.

[4] Michaela Conrad, Joep Schothorst, Harish Nag Kankipati, Griet Van Zeebroeck, Marta Rubio-Texeira, and Johan M. Thevelein. Nutrient sensing and signaling in the yeast saccharomyces cerevisiae. 38(2):254–299. ISSN 1574-6976. doi: 10.1111/1574-6976.12065.

[5] A. M. Stathopoulos and M. S. Cyert. Calcineurin acts through the CRZ1/TCN1-encoded transcription factor to regulate gene expression in yeast. 11(24): 3432–3444. ISSN 0890-9369.

[6] E. Moskvina, C. Schüller, C. T. Maurer, W. H. Mager, and H. Ruis. A search in the

genome of saccharomyces cerevisiae for genes regulated via stress response elements. 14(11):1041–1050. ISSN 0749-503X. doi: 10.1002/(SICI)1097-0061(199808)14:11¡1041::AID-YEA296¿3.0.CO;2-4.

[7] A. Stathopoulos-Gerontides, J. J. Guo, and M. S. Cyert. Yeast calcineurin regulates nuclear localization of the crz1p transcription factor through dephosphorylation. 13 (7):798–803. ISSN 0890-9369.

[8] Chiraj K. Dalal, Long Cai, Yihan Lin, Kasra Rahbar, and Michael B. Elowitz. Pulsatile dynamics in the yeast proteome. 24(18):2189–2194. ISSN 0960-9822. doi: 10.1016/j.cub.2014.07.076. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4203654/.

[9] Nan Hao and Erin K. O'Shea. Signal-dependent dynamics of transcription factor translocation controls gene expression. 19(1):31–39. ISSN 1545-9993. doi: 10.1038/nsmb.2192. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3936303/.

[10] Shadia Zaman, Soyeon I. Lippman, Lisa Schneper, Noam Slonim, and James R. Broach. Glucose regulates transcription in yeast through a network of signaling pathways. 5:245. ISSN 1744-4292. doi: 10.1038/msb.2009.2.

[11] Nan Hao, Bogdan A. Budnik, Jeremy Gunawardena, and Erin K. O'Shea. Tunable signal processing through modular control of transcription factor translocation. 339 (6118):460–464. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1227299. URL http://science.sciencemag.org/content/339/6118/460.

[12] Manuela Stierl, Patrick Stumpf, Daniel Udwari, Ronnie Gueta, Rolf Hagedorn, Aba Losi, Wolfgang Gärtner, Linda Petereit, Marina Efetova, Martin Schwarzel, Thomas G. Oertner, Georg Nagel, and Peter Hegemann. Light modulation of cellular cAMP by a

small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium beggiatoa. 286(2):1181–1188. ISSN 1083-351X. doi: 10.1074/jbc.M110.185496.

[13] Hayretin Yumerefendi, Daniel J. Dickinson, Hui Wang, Seth P. Zimmerman, James E. Bear, Bob Goldstein, Klaus Hahn, and Brian Kuhlman. Control of protein activity and cell fate specification via light-mediated nuclear translocation. 10(6):e0128443, . ISSN 1932-6203. doi: 10.1371/journal.pone.0128443.

[14] Hui Wang, Marco Vilela, Andreas Winkler, Miroslaw Tarnawski, Ilme Schlichting, Hayretin Yumerefendi, Brian Kuhlman, Rihe Liu, Gaudenz Danuser, and Klaus M Hahn. LOVTRAP, an optogenetic system for photo-induced protein dissociation. 13 (9):755–758, . ISSN 1548-7091. doi: 10.1038/nmeth.3926. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5137947/`.

[15] J. J. Tyson and J. D. Murray. Cyclic AMP waves during aggregation of dictyostelium amoebae. 106(3):421–426. ISSN 0950-1991.

[16] George G. Holz, Emma Heart, and Colin A. Leech. Synchronizing ca2+ and cAMP oscillations in pancreatic beta-cells: a role for glucose metabolism and GLP-1 receptors? focus on "regulation of cAMP dynamics by ca2+ and g protein-coupled receptors in the pancreatic beta-cell: a computational approach". 294(1):C4–6. ISSN 0363-6143. doi: 10.1152/ajpcell.00522.2007.

[17] Jason Ptacek, Geeta Devgan, Gregory Michaud, Heng Zhu, Xiaowei Zhu, Joseph Fasolo, Hong Guo, Ghil Jona, Ashton Breitkreutz, Richelle Sopko, Rhonda R. McCartney, Martin C. Schmidt, Najma Rachidi, Soo-Jung Lee, Angie S. Mah, Lihao Meng, Michael J. R. Stark, David F. Stern, Claudio De Virgilio, Mike Tyers, Brenda Andrews, Mark Gerstein, Barry Schweitzer, Paul F. Predki, and Michael Snyder. Global analysis of protein phosphorylation in yeast. 438(7068):679–684. ISSN 1476-4687. doi: 10.1038/nature04187.

[18] J. Nikawa, S. Cameron, T. Toda, K. M. Ferguson, and M. Wigler. Rigorous feedback control of cAMP levels in saccharomyces cerevisiae. 1(9):931–937. ISSN 0890-9369.

[19] A. Smith, M. P. Ward, and S. Garrett. Yeast PKA represses msn2p/msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation. 17 (13):3556–3564. ISSN 0261-4189. doi: 10.1093/emboj/17.13.3556.

[20] Long Cai, Chiraj K. Dalal, and Michael B. Elowitz. Frequency-modulated nuclear localization bursts coordinate gene regulation. 455(7212):485–490. ISSN 0028-0836. doi: 10.1038/nature07292. URL http://www.nature.com/nature/journal/v455/n7212/full/nature07292.html.

[21] Anders S Hansen and Erin K O'Shea. Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. 9:704, . ISSN 1744-4292. doi: 10.1038/msb.2013.56. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4039373/.

[22] Anders S. Hansen and Erin K. O'Shea. cis determinants of promoter threshold and activation timescale. 12(8):1226–1233. ISSN 2211-1247. doi: 10.1016/j.celrep.2015.07.035. URL https://www.cell.com/cell-reports/abstract/S2211-1247(15)00795-0.

[23] Mineo Iseki, Shigeru Matsunaga, Akio Murakami, Kaoru Ohno, Kiyoshi Shiga, Kazuichi Yoshida, Michizo Sugai, Tetsuo Takahashi, Terumitsu Hori, and Masakatsu Watanabe. A blue-light-activated adenylyl cyclase mediates photoavoidance in euglena gracilis. 415(6875):1047–1051. ISSN 0028-0836. doi: 10.1038/4151047a.

[24] W. Görner, E. Durchschlag, M. T. Martinez-Pastor, F. Estruch, G. Ammerer, B. Hamilton, H. Ruis, and C. Schüller. Nuclear localization of the c2h2 zinc finger protein msn2p is regulated by stress and protein kinase a activity. 12(4):586–597. ISSN 0890-9369.

[25] Cecilia Garmendia-Torres, Albert Goldbeter, and Michel Jacquet. Nucleocytoplasmic oscillations of the yeast transcription factor msn2: Evidence for periodic PKA activation. 17(12):1044–1049. ISSN 0960-9822. doi: 10.1016/j.cub.2007.05.032. URL https://www.cell.com/current-biology/abstract/S0960-9822(07)01411-X.

[26] Min-Hyung Ryu, Oleg V. Moskvin, Jessica Siltberg-Liberles, and Mark Gomelsky. Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. 285(53):41501–41508. ISSN 1083-351X. doi: 10.1074/jbc.M110.177600.

[27] Jacob Stewart-Ornstein, Christopher Nelson, Joe DeRisi, Jonathan S. Weissman, and Hana El-Samad. Msn2 coordinates a stoichiometric gene expression program. 23(23): 2336–2345. ISSN 1879-0445. doi: 10.1016/j.cub.2013.09.043.

[28] Sonia Colombo, Daniela Ronchetti, Johan M. Thevelein, Joris Winderickx, and Enzo Martegani. Activation state of the ras2 protein and glucose-induced signaling in saccharomyces cerevisiae. 279(45):46715–46722. ISSN 0021-9258. doi: 10.1074/jbc.M405136200.

[29] Dong Jian, Zhang Aili, Bai Xiaojia, Zhao Huansheng, and Hu Yun. Feedback regulation of ras2 guanine nucleotide exchange factor (ras2-GEF) activity of cdc25p by cdc25p phosphorylation in the yeast saccharomyces cerevisiae. 584(23):4745–4750. ISSN 1873-3468. doi: 10.1016/j.febslet.2010.11.006.

[30] Paolo Cazzaniga, Dario Pescini, Daniela Besozzi, Giancarlo Mauri, Sonia Colombo, and Enzo Martegani. Modeling and stochastic simulation of the ras/cAMP/PKA pathway in the yeast saccharomyces cerevisiae evidences a key regulatory function for intracellular guanine nucleotides pools. 133(3):377–385. ISSN 0168-1656. doi: 10.1016/j.jbiotec.2007.09.019.

[31] J. A. Nelder and R. Mead. A simplex method for function minimization. 7(4):

308–313. ISSN 0010-4620. doi: 10.1093/comjnl/7.4.308. URL https://academic.oup.com/comjnl/article/7/4/308/354237.

[32] P. Sass, J. Field, J. Nikawa, T. Toda, and M. Wigler. Cloning and characterization of the high-affinity cAMP phosphodiesterase of saccharomyces cerevisiae. 83(24): 9303–9307. ISSN 0027-8424.

[33] M Czyz, M M Nagiec, and R C Dickson. Autoregulation of GAL4 transcription is essential for rapid growth of kluyveromyces lactis on lactose and galactose. 21(18): 4378–4382. ISSN 0305-1048. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC310076/.

[34] Amit Sadeh, Natalia Movshovich, Misha Volokh, Larisa Gheber, and Amir Aharoni. Fine-tuning of the msn2/4–mediated yeast stress responses as revealed by systematic deletion of msn2/4 partners. 22(17):3127–3138. ISSN 1059-1524. doi: 10.1091/mbc.E10-12-1007. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3164460/.

[35] Michael Springer, Dennis D. Wykoff, Nicole Miller, and Erin K. O'Shea. Partially phosphorylated pho4 activates transcription of a subset of phosphate-responsive genes. 1(2):e28. ISSN 1545-7885. doi: 10.1371/journal.pbio.0000028. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0000028.

[36] Eric Batchelor, Alexander Loewer, Caroline Mock, and Galit Lahav. Stimulus-dependent dynamics of p53 in single cells. 7:488. ISSN 1744-4292. doi: 10.1038/msb.2011.20. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3130553/.

[37] Markus W. Covert, Thomas H. Leung, Jahlionais E. Gaston, and David Baltimore. Achieving stability of lipopolysaccharide-induced NF-kappaB activation. 309(5742): 1854–1857. ISSN 1095-9203. doi: 10.1126/science.1112304.

[38] Y. Gotoh, E. Nishida, T. Yamashita, M. Hoshi, M. Kawakami, and H. Sakai. Microtubule-associated-protein (MAP) kinase activated by nerve growth factor and epidermal growth factor in PC12 cells. identity with the mitogen-activated MAP kinase of fibroblastic cells. 193(3):661–669. ISSN 0014-2956.

[39] Alexander Hoffmann, Andre Levchenko, Martin L. Scott, and David Baltimore. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. 298(5596):1241–1245, . ISSN 1095-9203. doi: 10.1126/science.1071914.

[40] D. E. Nelson, A. E. C. Ihekwaba, M. Elliott, J. R. Johnson, C. A. Gibney, B. E. Foreman, G. Nelson, V. See, C. A. Horton, D. G. Spiller, S. W. Edwards, H. P. McDowell, J. F. Unitt, E. Sullivan, R. Grimley, N. Benson, D. Broomhead, D. B. Kell, and M. R. H. White. Oscillations in NF-kappaB signaling control the dynamics of gene expression. 306(5696):704–708. ISSN 1095-9203. doi: 10.1126/science.1099962.

[41] T. T. Nguyen, J. C. Scimeca, C. Filloux, P. Peraldi, J. L. Carpentier, and E. Van Obberghen. Co-regulation of the mitogen-activated protein kinase, extracellular signal-regulated kinase 1, and the 90-kDa ribosomal s6 kinase in PC12 cells. distinct effects of the neurotrophic factor, nerve growth factor, and the mitogenic factor, epidermal growth factor. 268(13):9803–9810. ISSN 0021-9258.

[42] Jeremy E. Purvis and Galit Lahav. Encoding and decoding cellular information through signaling dynamics. 152(5):945–956. ISSN 0092-8674. doi: 10.1016/j.cell.2013.02.005. URL http://www.sciencedirect.com/science/article/pii/S0092867413001530.

[43] Jeremy E. Purvis, Kyle W. Karhohs, Caroline Mock, Eric Batchelor, Alexander Loewer, and Galit Lahav. p53 dynamics control cell fate. 336(6087):1440–1444. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1218351. URL http://science.sciencemag.org/content/336/6087/1440.

[44] Savaş Tay, Jacob J. Hughey, Timothy K. Lee, Tomasz Lipniacki, Stephen R. Quake, and Markus W. Covert. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. 466(7303):267–271. ISSN 1476-4687. doi: 10.1038/nature09145.

[45] S. Traverse, N. Gomez, H. Paterson, C. Marshall, and P. Cohen. Sustained activation of the mitogen-activated protein (MAP) kinase cascade may be required for differentiation of PC12 cells. comparison of the effects of nerve growth factor and epidermal growth factor. 288 (Pt 2):351–355. ISSN 0264-6021.

[46] Shannon L. Werner, Jeffrey D. Kearns, Victoria Zadorozhnaya, Candace Lynch, Ellen O'Dea, Mark P. Boldin, Averil Ma, David Baltimore, and Alexander Hoffmann. Encoding NF-kappaB temporal control in response to TNF: distinct roles for the negative regulators IkappaBalpha and a20. 22(15):2093–2101. ISSN 0890-9369. doi: 10.1101/gad.1680708.

[47] Anders S Hansen and Erin K O'Shea. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. 4:e06559, . ISSN 2050-084X. doi: 10.7554/eLife.06559. URL https://doi.org/10.7554/eLife.06559.

[48] Yolanda T. Chong, Judice L. Y. Koh, Helena Friesen, Supipi Kaluarachchi Duffy, Kaluarachchi Duffy, Michael J. Cox, Alan Moses, Jason Moffat, Charles Boone, and Brenda J. Andrews. Yeast proteome dynamics from single cell imaging and automated analysis. 161(6):1413–1424. ISSN 1097-4172. doi: 10.1016/j.cell.2015.04.051.

[49] Alejandro A. Granados, Julian M. J. Pietsch, Sarah A. Cepeda-Humerez, Iseabail L. Farquhar, Gašper Tkačik, and Peter S. Swain. Distributed and dynamic intracellular organization of extracellular information. 115(23):6088–6093. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1716659115. URL https://www.pnas.org/content/115/23/6088.

[50] Zohreh AkhavanAghdam, Joydeb Sinha, Omar P. Tabbaa, and Nan Hao. Dynamic control of gene regulatory logic by seemingly redundant transcription factors. URL https://elifesciences.org/articles/18458.

[51] Yen Ting Lin and Charles R. Doering. Gene expression dynamics with stochastic bursts: Construction and exact results for a coarse-grained model. 93(2):022409. doi: 10.1103/PhysRevE.93.022409. URL https://link.aps.org/doi/10.1103/PhysRevE.93.022409.

[52] David Pincus, Andrés Aranda-Díaz, Ignacio A. Zuleta, Peter Walter, and Hana El-Samad. Delayed ras/PKA signaling augments the unfolded protein response. 111 (41):14800–14805. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1409588111. URL https://www.pnas.org/content/111/41/14800.

[53] Dominik Niopek, Dirk Benzinger, Julia Roensch, Thomas Draebing, Pierre Wehler, Roland Eils, and Barbara Di Ventura. Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. 5:4404. ISSN 2041-1723. doi: 10.1038/ncomms5404.

[54] Taras A. Redchuk, Evgeniya S. Omelina, Konstantin G. Chernov, and Vladislav V. Verkhusha. Near-infrared optogenetic pair for protein regulation and spectral multiplexing. 13(6):633–639. ISSN 1552-4469. doi: 10.1038/nchembio.2343.

[55] Hayretin Yumerefendi, Hui Wang, Daniel J. Dickinson, Andrew M. Lerner, Per Malkus, Bob Goldstein, Klaus Hahn, and Brian Kuhlman. Light-dependent cytoplasmic recruitment enhances the dynamic range of a nuclear import photoswitch. 19(12):1319–1325, . ISSN 1439-7633. doi: 10.1002/cbic.201700681.

[56] Shunichi Kosugi, Masako Hasebe, Nobutaka Matsumura, Hideaki Takashima, Etsuko Miyamoto-Sato, Masaru Tomita, and Hiroshi Yanagawa. Six classes of nuclear

localization signals specific to different binding grooves of importin alpha. 284(1): 478–485. ISSN 0021-9258. doi: 10.1074/jbc.M807017200.

[57] S. P. Heximer, H. Lim, J. L. Bernard, and K. J. Blumer. Mechanisms governing subcellular localization and function of human RGS2. 276(17):14195–14203. ISSN 0021-9258. doi: 10.1074/jbc.M009942200.

[58] Ahmad S. Khalil, Timothy K. Lu, Caleb J. Bashor, Cherie L. Ramirez, Nora C. Pyenson, J. Keith Joung, and James J. Collins. A synthetic biology framework for programming eukaryotic transcription functions. 150(3):647–658. ISSN 0092-8674. doi: 10.1016/j.cell.2012.05.045. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653585/.

[59] A. P. Gasch, P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein, and P. O. Brown. Genomic expression programs in the response of yeast cells to environmental changes. 11(12):4241–4257. ISSN 1059-1524. doi: 10.1091/mbc.11.12.4241.

[60] Noam Vardi, Sagi Levy, Yonat Gurvich, Tamar Polacheck, Miri Carmi, Diego Jaitin, Ido Amit, and Naama Barkai. Sequential feedback induction stabilizes the phosphate starvation response in budding yeast. 9(3):1122–1134. ISSN 2211-1247. doi: 10.1016/j.celrep.2014.10.002.

[61] Alexander Hoffmann, Andre Levchenko, Martin L. Scott, and David Baltimore. The IB-NF-b signaling module: Temporal control and selective gene activation. 298(5596): 1241–1245, . ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1071914. URL http://science.sciencemag.org/content/298/5596/1241.

[62] Hiroyuki Yoshimoto, Kirstie Saltsman, Audrey P. Gasch, Hong Xia Li, Nobuo Ogawa, David Botstein, Patrick O. Brown, and Martha S. Cyert. Genome-wide analysis of

gene expression regulated by the calcineurin/crz1p signaling pathway in saccharomyces cerevisiae. 277(34):31079–31088. ISSN 0021-9258. doi: 10.1074/jbc.M202718200.

[63] Romain Christiano, Nagarjuna Nagaraj, Florian Fröhlich, and Tobias C. Walther. Global proteome turnover analyses of the yeasts s. cerevisiae and s. pombe. 9(5): 1959–1965. ISSN 2211-1247. doi: 10.1016/j.celrep.2014.10.065. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4526151/.

[64] Yulei Wang, Chih Long Liu, John D. Storey, Robert J. Tibshirani, Daniel Herschlag, and Patrick O. Brown. Precision and functional specificity in mRNA decay. 99(9): 5860–5865, . ISSN 0027-8424. doi: 10.1073/pnas.092538799. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC122867/.

[65] Michael E. Lee, William C. DeLoache, Bernardo Cervantes, and John E. Dueber. A highly characterized yeast toolkit for modular, multipart assembly. 4(9):975–986. ISSN 2161-5063. doi: 10.1021/sb500366v.

[66] Arnaud Gautier, Duy P. Nguyen, Hrvoje Lusic, Wenlin An, Alexander Deiters, and Jason W. Chin. Genetically encoded photocontrol of protein localization in mammalian cells. 132(12):4086–4088. ISSN 0002-7863. doi: 10.1021/ja910688s. URL https://doi.org/10.1021/ja910688s.

[67] Sae Shimizu-Sato, Enamul Huq, James M. Tepperman, and Peter H. Quail. A light-switchable gene promoter system. 20(10):1041–1044. ISSN 1546-1696. doi: 10.1038/nbt734. URL https://www.nature.com/articles/nbt734.

[68] Devin Strickland, Yuan Lin, Elizabeth Wagner, C. Matthew Hope, Josiah Zayner, Chloe Antoniou, Tobin R. Sosnick, Eric L. Weiss, and Michael Glotzer. TULIPs: Tunable, light-controlled interacting protein tags for cell biology. 9(4):379–384. ISSN 1548-7091. doi: 10.1038/nmeth.1904. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3444151/.

[69] Jared E. Toettcher, Christopher A. Voigt, Orion D. Weiner, and Wendell A. Lim. The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. 8:35–38, . ISSN 1548-7105. doi: 10.1038/nmeth.f.326. URL https://www.nature.com/articles/nmeth.f.326.

[70] Jared E. Toettcher, Orion D. Weiner, and Wendell A. Lim. Using optogenetics to interrogate the dynamic control of signal transmission by the ras/erk module. 155(6): 1422–1434, . ISSN 0092-8674. doi: 10.1016/j.cell.2013.11.004. URL http://www.sciencedirect.com/science/article/pii/S0092867413014177.

[71] Niall Dillon and Richard Festenstein. Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. 18(5):252–258. ISSN 0168-9525.

[72] François Spitz and Eileen E. M. Furlong. Transcription factors: from enhancer binding to developmental control. 13(9):613–626. ISSN 1471-0064. doi: 10.1038/nrg3207.

[73] Colin R. Lickwar, Florian Mueller, Sean E. Hanlon, James G. McNally, and Jason D. Lieb. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. 484(7393):251–255. ISSN 1476-4687. doi: 10.1038/nature10985.

[74] Jesse M. Platt, Paul Ryvkin, Jennifer J. Wanat, Greg Donahue, M. Dan Ricketts, Steven P. Barrett, Hannah J. Waters, Shufei Song, Alejandro Chavez, Khaled Omar Abdallah, Stephen R. Master, Li-San Wang, and F. Brad Johnson. Rap1 relocalization contributes to the chromatin-mediated gene expression profile and pace of cell senescence. 27(12):1406–1420. ISSN 1549-5477. doi: 10.1101/gad.218776.113.

[75] Chao Cheng, Chong Shou, Kevin Y Yip, and Mark B Gerstein. Genome-wide analysis of chromatin features identifies histone modification sensitive and insensitive yeast

transcription factors. 12(11):R111. ISSN 1465-6906. doi: 10.1186/gb-2011-12-11-r111. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3334597/.

[76] Ann-Christin Gaupel, Thomas Begley, and Martin Tenniswood. High throughput screening identifies modulators of histone deacetylase inhibitors. 15:528. ISSN 1471-2164. doi: 10.1186/1471-2164-15-528.

[77] Israel Steinfeld, Ron Shamir, and Martin Kupiec. A genome-wide analysis in saccharomyces cerevisiae demonstrates the influence of chromatin modifiers on transcription. 39(3):303–309. ISSN 1061-4036. doi: 10.1038/ng1965.

[78] Pulak Kar, Gary R. Mirams, Helen C. Christian, and Anant B. Parekh. Control of NFAT isoform activation and NFAT-dependent gene expression through two coincident and spatially segregated intracellular ca2+ signals. 64(4):746–759. ISSN 1097-2765. doi: 10.1016/j.molcel.2016.11.011. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5128683/.

[79] Jennifer M. Gardner and Sue L. Jaspersen. Manipulating the yeast genome: Deletion, mutation, and tagging by PCR. In Jeffrey S. Smith and Daniel J. Burke, editors, *Yeast Genetics: Methods and Protocols*, Methods in Molecular Biology, pages 45–78. Springer New York. ISBN 978-1-4939-1363-3. doi: 10.1007/978-1-4939-1363-3$_5$.URL

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

3/18/19

Date