

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Nodal Gradient Interpretation During Zebrafish Germ Layer Patterning

Permalink

<https://escholarship.org/uc/item/9jk1f2hk>

Author

Reade, Anna Lorraine

Publication Date

2015

Supplemental Material

<https://escholarship.org/uc/item/9jk1f2hk#supplemental>

Peer reviewed|Thesis/dissertation

Nodal Gradient Interpretation During Zebrafish
Germ Layer Patterning

by

Anna Lorraine Reade

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Developmental Biology

in the

Copyright 2015

by

Anna Reade

for my husband,
Ian Thomas Foe

ACKNOWLEDGEMENTS

First and foremost I would like to express my gratitude to my advisors, Orion Weiner, Didier Stainier, and Stephanie Woo. Orion, I truly appreciate your constant and unwavering excitement for my endeavors and confidence in my abilities to achieve them. These qualities were especially valuable in times of doubt and fatigue. Didier, thank you for the support, resources and perspective you provided whenever and wherever it was needed. And finally, Stephanie, I would have been lost were you not so generous with your time, patience and advice. With no obligations towards me whatsoever, you befriended me during my intimidating first days in the Stainier lab and soon after became my closest mentor, both scientifically and career-wise. In addition, I want to thank my undergraduate research advisor, Kathy Foltz, who inspired me to embark on this journey in the first place.

To my qualification and thesis committee members: Wallace Marshall, Jeremy Reiter, Cynthia Kenyon, Rich Schneider, thank you for the great conversations and moral support. To my former and current Weiner and Stainier labmates, especially Delquin Gong, Jared Toettcher, and Oliver Hoeller, thank you for your guidance and company throughout so many years. A special thank you to our lab technicians, aka “lab mommies”, Grace Peng, Anna Goldberg, and Sana Chintamen, disaster would have befallen us all if it wasn’t for all your hard work and patience. I have had the privilege of working with a wonderful set of collaborators. Clare Buckley and Jon Clarke, it was a pleasure working with you to finally get the phytochrome system to work in zebrafish. Laura Motta-Mena and Kevin Gardner, thank you for the opportunity to work with you on EL222 and for the advice in developing TAEL. To Jan Huisken and lab, especially Gopi Shah, Benjamin Schmid, Nico Scherf, without you my thesis work would not be possible. Working with you, I discovered how rewarding a truly collaborative endeavor can be.

My fellow graduate students: Theresa Loveless, Brian Caliando, Justin Farlow, Silvi Rouskin, Kelly Nissen, Adrian Contreras, Hayley Pemble, Joshua Dunn, Mark Slabodnick, Vanessa Van Voorhis, and Karmela Ramos, from Bioreg mania, Granli skits, insightful scientific and personal conversations, nights out on the town, Tahoe trips, venting about the stresses of graduate school, weddings and babies, I couldn't imagine better company to be in. And to the strong, independent and brilliant women I have the honor of calling my friends and peers: Amina Massey, Jessica Grochowsky, Fanny Ho, Dannan Hodge, Kaitlin Foe, Andrea Medina-Ramirez, Maura Madou, and Robin Elgart, the support and inspiration that you provide me is invaluable, thank you.

Tean-gau Wang and David Reade, my parents and "number one cheerleaders", I am so grateful for your unconditional love and support. The confidence, discipline, and passion for learning you instilled in me were essential in this achievement. To my grandmother, Mary Reade, I am continually in awe of your ever curious mind. From music and art, to travel and culture, and even science, thank you for a lifetime of wonderful conversations. Mary and Chris, I couldn't ask for a better second set of parents. Thank you for your acceptance and love and for always providing a much needed retreat from work and the city. I am also immensely grateful for all the support and love I've received from my brother and all my aunties and cousins. There is nothing quite like being part of a big, gregarious, slightly crazy, but above all, loving family.

And finally, my dear husband, you have been through the worst of it with me. Yet you remain supportive, encouraging and patient through it all. This is for the both of us.

STATEMENT REGARDING AUTHOR CONTRIBUTIONS

Chapter two is from the unpublished article by Clare E Buckley, myself, Anna Goldberg, Orion D Weiner, Jonathan DW Clarke. Clare Buckley performed the majority of the experiments with Anna Goldberg, Orion Weiner and I providing reagents and the groundwork knowledge and experience to help with trouble-shooting various problems that arose from transferring the phytochrome technology into zebrafish. For the paper itself, I contributed Figure 1e. Clare Buckley, Orion Weiner, and Jon Clarke analyzed the data. Clare Buckley and Jon Clarke wrote the paper.

Chapter three is a reprint of a *Nature Chemical Biology* article published in 2014 (volume 10, issue 3, pp 196–202) by Laura B Motta-Mena, myself, Michael J Mallory, Spencer Glantz, Orion D Weiner, Kristen W Lynch & Kevin H Gardner. Laura B Motta-Mena, myself, Orion D Weiner, Kristen W Lynch & Kevin H Gardner conceived and designed the experiments. Laura B Motta-Mena performed the experiments for Figures 1 and 2. Spencer Glantz and Kevin H Gardner generating the kinetic model (Fig. 3). Michael J Mallory performed the experiments for Figure 4. I performed the experiments for Figure 5. Laura B Motta-Mena, myself, Michael J Mallory, Orion D Weiner, Kristen W Lynch and Kevin H Gardner analyzed the data. Laura B Motta-Mena and Kevin H Gardner wrote the paper.

Chapter four is an adaptation of the technology from chapter three for use in zebrafish by Stephanie Woo and myself. Stephanie Woo, Orion Weiner, Didier Y Stainier and I conceived and designed the experiments. Stephanie Woo and I performed all the experiments and analyzed the data equally. I wrote the chapter, which will soon be converted to a journal article format for publishing.

Chapter five is the introduction and partial methods of my main thesis project, to investigate the interpretation of Nodal morphogen signaling in zebrafish germ layer determination. Data analysis is ongoing and the full methods, results and discussion will be added to complete a paper to be published in the near future. Orion Weiner, Didier Stainier and I conceived and designed the biological experiments. Stephanie Woo, Gopi Shah and I performed the biological experiments. Jan Huisken, Gopi Shah, Benjamin Schmid, and Orion Weiner helped me with image acquisition development and optimization. Benjamin Schmid wrote the code to optimize the microscope for my experiments. Benjamin Schmid and Nico Scherf built the data viewer and are helping me perform the data analysis.

Nodal Gradient Interpretation During Zebrafish Germ Layer Patterning

Anna Lorraine Reade

This thesis presents work toward understanding how cells within an embryo receive fate and positional information. It concentrates on the Nodal morphogen gradient, which specifies germ layer identity through differing levels of Nodal signaling. Although much progress has been made in understanding the components and outcomes of the Nodal signaling pathway, there is very poor quantitative understanding of how cells translate the duration and concentration of Nodal signaling into fate/positional information. To address this question, several new technologies were employed or developed.

Light sheet fluorescence microscopy was used to image Nodal signaling dynamics *in vivo* (Chapter 5). This new imaging technology allowed us to: determine which cells are exposed to the Nodal signal, quantitate the duration and level of this signaling, and directly correlate this with cellular response, i.e. which Nodal target genes are turned on and what germ layer fate is adopted. We will use this information to propose an initial model for how Nodal signaling is interpreted at a cellular level. We will test our model with additional data taken in the presence of an inhibitor, to assess whether the system behaves as expected when the input is modified.

In order to develop finer control of the spatial and temporal aspects of perturbing a biological system, we simultaneously developed optogenetic tools for use in zebrafish embryos. Most optogenetic systems, especially those that control protein localization and expression, have been developed at the tissue-culture level and do not transfer directly to the multicellular organism level. We were successfully able to optimize two optogenetic systems for use in zebrafish, the phytochrome system for protein localization control (Chapter 2) and the LOV system for transcriptional control (Chapters 3 and 4).

TABLE OF CONTENTS

Title Page	i
Acknowledgements	iv
Statement regarding author contributions	vi
Abstract	viii
Table of Contents	ix
List of Figures	x
Chapter One	1
Introduction	
Chapter Two	6
Reversible optogenetic control of subcellular protein localisation in a vertebrate embryo <i>in vivo</i> using the Phytochrome system	
Chapter Three	53
An optogenetic gene expression system with rapid activation and deactivation kinetics.	
Chapter Four	110
TAEL: A Non-toxic, zebrafish-optimized optogenetic gene expression system with large dynamic range and rapid kinetics	
Chapter Five	140
Nodal Gradient Interpretation During Zebrafish Germ Layer Patterning	

LIST OF FIGURES

Chapter Two

Figure 1: Adapting the phytochrome system for use in a vertebrate embryo	31
Figure 2: Light-controlled shuttling of protein between cytoplasm and membrane in vivo	33
Figure 3: Rapid kinetics of PhyB-PIF binding and unbinding	35
Figure 4: Subcellular control of protein localisation	37
Figure 5: Mislocalising Pard3 using the phyochrome system	39
Figure S1: Statistical analysis of EGFP intensity	42
Figure S2: Gradual midline localisation of Pard3-EGFP_PIF6	44
Figure S3: Mislocalisation of Pard3 during C-division	46

Chapter Three

Figure 1: Model for the EL222-based light-inducible gene expression system	77
Figure 2: Dose-dependent activation and photoreversibility of gene expression by VP- EL222	79
Figure 3: Kinetic modeling of VP-EL222 activation	81
Figure 4: Light-regulated gene expression of the splicing factor CELF2 using VP-EL222 in the T-cell derived JSL1 cell line	83
Figure 5: VP-EL222 robustly activates reporter gene expression in the developing zebrafish embryo in a light-dependent manner	85
Figure S1: VP-EL222 is expressed in 293T cells and can function to activate luciferase expression in response to light	87
Figure S2: VP-EL222 is not activated by red light and that illumination with blue light has no negative effect on cell viability	89
Figure S3: Light-triggered expression of luciferase protein by VP-EL222	91

Figure S4: Full Western blots and denaturing polyacrylamide gel of RT- PCR samples	93
Figure S5: VP-EL222 is only moderately toxic in zebrafish	95
Table S1: Occurrence of C120 sequence in human, mouse and zebrafish genomes	97

Chapter Four

Figure 1: Optimization of a optogenetic gene expression system for zebrafish, TAEL	127
Figure 2: Characterization of TAEL system, dynamics and range	129
Figure 3: Ectopic endoderm induction via over-expression of sox32	131
Figure 4: Spatial control of TAEL induction	133
Figure S1: Cell ablation via diphtheria toxin induction	135

CHAPTER ONE

Introduction

"Its only crazy when you first hear it, but the more you think about it the less crazy it sounds"

-Wallace Marshall

"I'd forgotten how much work you've done on this, fantastic job, really exciting!"

- Orion Weiner

In order to form the complex structure of multicellular organisms it is essential for cells within an embryo to receive fate and positional information. An extensively studied source of fate/positional information is the Nodal morphogen gradient (1). Nodal is essential for germ layer patterning in vertebrates. High levels of Nodal specify endoderm, low levels mesoderm, and lack of Nodal results in ectoderm. Although much progress has been made in understanding the components and outcomes of the Nodal signaling pathway, there is very poor quantitative understanding of how cells translate the duration and concentration of Nodal signaling into fate/positional information.

To address this question, we have employed light sheet fluorescence microscopy to image Nodal signaling dynamics *in vivo* (2,3). This new imaging technology has allowed us to image a real-time activity reporter of Nodal signaling (4) at a whole embryo scale, with sub-cellular resolution. Combined with a nuclear fluorescent marker to track cells throughout our imaging and live Nodal target gene transcriptional readouts (5), we are able to: determine which cells are exposed to the Nodal signal, quantitate the duration and level of this signaling, and correlate this with cellular response, i.e. which Nodal target genes are turned on and what germ layer fate is adopted. We will use this information to propose an initial model for how Nodal signaling is interpreted at a cellular level. Possible models of input interpretation include, integration of signaling over time, cellular memory of maximal signaling levels achieved, or change in signaling over time or space (6). We imaged Nodal signaling dynamics under wild-type conditions as well as with low levels of a small molecule inhibitor (7) that globally decreased Nodal signaling levels. Perturbing the system with a small molecule inhibitor will allow us to test whether our initial model, made with data acquired under wild-type conditions, behaves as expected when the input to the system changes.

Our work, as described above, emphasized to us the need for biotechnology that would provide finer control of the spatial and temporal aspects of perturbing a biological system. Although we will learn a great deal from comparing Nodal signaling interpretation under wild-type and globally decreased input levels, our project, as well as many others, would benefit greatly from the ability to specify exactly where, when, and how much a certain perturbation should be made. With this in mind, we simultaneously developed optogenetic tools (genetically encoded systems with light-gated control of protein function) for use in zebrafish embryos. Most optogenetic systems, especially those that control protein localization and expression, have been developed at the tissue-culture level and do not transfer directly to the multicellular organism level. We were successfully able to optimize two optogenetic systems for use in zebrafish, the phytochrome system (8) for protein localization control (Chapter 2) and the LOV system (9) for transcriptional control (Chapters 3 and 4). Although we did not develop this technology in time to use in our dissection of Nodal gradient interpretation, we believe they will prove to be a powerful set of tools for the zebrafish community.

REFERENCES

1. Schier, Alexander F. "Nodal Morphogens." *Cold Spring Harbor Perspectives in Biology* 1, no. 5 (November 2009). doi:10.1101/cshperspect.a003459.
2. Huisken, Jan, and Didier Y. R. Stainier. "Selective Plane Illumination Microscopy Techniques in Developmental Biology." *Development (Cambridge, England)* 136, no. 12 (June 15, 2009): 1963–75. doi:10.1242/dev.022426.
3. Schmid, Benjamin, Gopi Shah, Nico Scherf, Michael Weber, Konstantin Thierbach, Citlali Pérez Campos, Ingo Roeder, Pia Aanstad, and Jan Huisken. "High-Speed Panoramic Light-Sheet Microscopy Reveals Global Endodermal Cell Dynamics." *Nature Communications* 4 (July 25, 2013). doi:10.1038/ncomms3207.
4. Harvey, Steven A, and James C Smith. "Visualisation and Quantification of Morphogen Gradient Formation in the Zebrafish." *PLoS Biology* 7, no. 5 (May 2009). doi:10.1371/journal.pbio.1000101.
5. Bennett, James T., Katherine Joubin, Simon Cheng, Pia Aanstad, Ralf Herwig, Matthew Clark, Hans Lehrach, and Alexander F. Schier. "Nodal Signaling Activates Differentiation Genes During Zebrafish Gastrulation." *Developmental Biology* 304, no. 2 (April 15, 2007): 525–40. doi:10.1016/j.ydbio.2007.01.012.
6. Nahmad, Marcos, and Arthur D. Lander. "Spatiotemporal Mechanisms of Morphogen Gradient Interpretation." *Current Opinion in Genetics & Development* 21, no. 6 (December 2011): 726–31. doi:10.1016/j.gde.2011.10.002.
7. Hagos, Engda G, and Scott T Dougan. "Time-Dependent Patterning of the Mesoderm and Endoderm by Nodal Signals in Zebrafish." *BMC Developmental Biology* 7 (March 28, 2007): 22. doi:10.1186/1471-213X-7-22.

8. Levskaya, Anselm, Orion D. Weiner, Wendell A. Lim, and Christopher A. Voigt.
“Spatiotemporal Control of Cell Signalling Using A Light-Switchable Protein Interaction.”
Nature 461, no. 7266 (October 15, 2009): 997–1001. doi:10.1038/nature08446.
9. Motta-Mena, Laura B., Anna Reade, Michael J. Mallory, Spencer Glantz, Orion D. Weiner,
Kristen W. Lynch, and Kevin H. Gardner. “An Optogenetic Gene Expression System with
Rapid Activation and Deactivation Kinetics.” Nature Chemical Biology 10, no. 3 (March
2014): 196–202. doi:10.1038/nchembio.1430.

CHAPTER TWO

Reversible optogenetic control of subcellular protein localisation in a vertebrate embryo *in vivo* using the Phytochrome system

“Easy things don’t like you”

- Gopi Shah

Adapted from Clare E Buckley, Anna Reade, Anna Goldberg, Orion D Weiner, Jonathan DW Clarke (2015), Reversible optogenetic control of subcellular protein localisation in a vertebrate embryo *in vivo* using the Phytochrome system. In preparation.

SUMMARY

We demonstrate, for the first time, light-induced heterodimerization using the Phytochrome system in living multicellular organisms. This system is used to rapidly and reversibly recruit proteins to specific subcellular regions within the zebrafish embryo hindbrain. As proof of principle we also alter subcellular localization of the apical polarity protein Pard3. The phytochrome system therefore offers an unprecedented level of experimental control within whole organisms. Our optimizations of optogenetic component expression and chromophore purification and delivery should significantly lower the barrier for establishing this powerful optogenetic system in other multicellular organisms.

INTRODUCTION

While gene knockout, overexpression, and mutation have been used to reveal the involvement of proteins in biological systems, more subtle manipulation of proteins are needed to interrogate their precise roles. For example, manipulation of protein localisation within cells will be critical for understanding how intracellular asymmetries and polarity are established. The use of optogenetics to precisely control cell processes such as gene transcription, signalling activation and protein dimerization has begun to transform biomedical research. It is allowing researchers to probe specific components of signalling pathways and cellular function often with subcellular resolution and with rapid temporal control. While this approach has had widespread success in cell culture and in single celled organisms, there is a need to adapt current technologies to multicellular organisms such as the vertebrate embryo. In this work we develop a reversible optogenetic control of protein localisation in the zebrafish embryo.

There are several strategies for light-dependent control of signalling in cells, many of which are built from natural photoreceptors. The main optogenetic systems have recently been comprehensively reviewed^{1,2} and further adaptations and improvements are continually making these systems even more versatile. Each system has a different combination of properties such as reversibility, dynamic range and chromophore requirement², which must be considered carefully when choosing which system to use in living whole organisms.

Light-oxygen voltage (LOV) domains are found in many organisms such as plants, algae, fungi and bacteria and are a type of Per-ARNT-SIM (PAS) photosensor, which binds flavins³. Following blue-light illumination, a covalent bond is made between flavins and a cysteine in LOV, therefore altering the structure of LOV with downstream consequences such as kinase activation⁴. LOV-based strategies have so far had the most success of the optogenetic systems in transferring from cell culture to whole organisms. Allosteric LOV-based photoactivation of

Rac1⁵ was successfully replicated in zebrafish embryos, where it was used to control the protrusion and migration of zebrafish neutrophils⁶ and in drosophila ovary border cells to control collective migration⁷. An engineered LOV-based system has also recently been used to drive tissue-specific transcription of MCherry in zebrafish⁸. LOV-based systems are therefore highly adaptable. However, they often require high levels of optimisation. The cryptochrome system comprises the Arabidopsis cryptochrome (CRY2) protein, which is induced to bind to CRY-interacting bHLH1 protein (CIB1) under blue light. This binding is lost naturally approximately 10 minutes after illumination is ceased⁹. There is a precedent for using the cryptochrome-based systems *in vivo* in living vertebrate systems for protein transcription regulation both in whole homogenates of zebrafish embryos¹⁰ and in the mouse cortex¹¹.

Arguably the most important property when considering the level of temporal and spatial resolution that a system offers is the speed of reversibility. One disadvantage of both the LOV and cryptochrome-based systems is that it is not possible to actively reverse their light-induced interactions. Instead this happens naturally in the dark, and is relatively slow. This makes levels of basal activation higher and sets significant constraints on the spatial and temporal precision that can be achieved with these systems. Only two optogenetics systems can be actively reversed. The first is the photoactivatable protein Dronpa. This protein switches between its monomeric and dimeric states, using 490nm and 390nm light, respectively. Although successfully used in cell culture both to shuttle a protein on and off the membrane and to switch on and off Cdc42 by releasing and inducing steric occlusion of Cdc42¹², Dronpa has so far not been used in multicellular organisms. A much faster reversal can be achieved with the Arabidopsis red light-inducible Phytochrome (PHYB-PIF) system, which comprises the phytochrome B (PHYB) protein and the basic helix-loop-helix (bHLH) transcription factor phytochrome interaction factor (PIF). These 2 domains are induced to bind under far-red light and the binding is reversed within seconds of exposure to infrared light but is otherwise stable

for hours in the dark¹³. The Phytochrome system has a 10-100x larger dynamic range (respectively) than the Cry and LOV-based systems (PMID: 25350266) and the affinity of its light gated interaction is 100x tighter than Dronpa^{12,14}. The phytochrome system therefore offers the highest level of control of the currently available optogenetics systems. An added advantage of the phytochrome system is that the wavelengths that it requires for photomodulation (red and infrared) are easier to use in combination with common fluorescent proteins such as BFP, CFP, YFP, and MCherry and is less toxic than the blue light activated systems due to the low intensity of light required.

The spatial precision of the phytochrome system has been used to test the spatial sufficiency of Rac for directed cell migration *in vitro*¹⁴ and it has also been successfully used to specifically and reversibly recruit proteins to several different organelles within yeast, therefore dynamically controlling the activation and inactivation of signalling pathways¹⁵. The phytochrome system's temporal precision has been used to understand the cell signalling pathways that are differentially activated by transient versus sustained Ras activation¹⁶. The tunability of this system has been used to perform full dose-response curves for isolated signalling modules in single cells¹⁶. Since its equilibrium can be rapidly set by adjusting the ratio of red to infrared light, the phytochrome system is compatible with active feedback control to maintain precise signal input despite variations in expression of the optogenetic components or to maintain a fixed input despite cellular feedback^{17,18}.

Due to this high level of control, the phytochrome system would be an optimal system for dissecting the complex interconnected signalling networks present in whole multicellular organisms. However, so far the phytochrome system has not been successfully ported from plants to other live multicellular organisms. This is partly because, unlike the cryptochrome and LOV-based systems, which use ubiquitously occurring flavin as a chromophore, the

phytochrome system requires an external chromophore, Phycocyanobilin (PCB). External exposure to PCB is sufficient to mediate binding of PCB in mammalian cell culture¹⁴ and it is also possible to induce cells to make their own PCB by genetically introducing genes responsible for PCB synthesis (heme oxygenase 1 (HO1) and PCB:ferredoxin oxidoreductase (PcyA))^{19,20}. However, this has not yet been demonstrated in whole multicellular organisms. Previous unpublished attempts to use the phytochrome system in living organisms have also found that unpurified PCB has toxic effects and that genetically inducing cells to express the *Arabidopsis* phytochrome protein is problematic.

In the current study we have optimised PHYB for robust expression within the optically tractable zebrafish embryo. We have also purified PCB and successfully delivered it to cells deep within the developing zebrafish embryo without toxic effects. We demonstrate high binding and reversal kinetics between PHYB and PIF in cells within live zebrafish as well as highly specific subcellular recruitment of PIF tagged protein. As a proof of principle within the zebrafish embryo, we also mislocalised the apical polarity protein Partitioning defective 3 (Pard3) subcellularly. As well as developing the phytochrome system for use in zebrafish embryos, we envision that our optimisations of PCB and PHYB, as well as the methodology described here (such as the use of the Y276H mutant of PHYB as a sensor for PCB delivery) will also facilitate the use of the phytochrome system in other organisms.

RESULTS

1. Removal of the plant C-terminal PAS repeated domain is necessary for the successful expression of PHYB in zebrafish

We first determined how to efficiently express the plant PHYB protein in zebrafish cells. Arabidopsis PHYB protein is an 1172 amino acid protein made up of PAS₂, GAF, PHY, PAS and HKRD domains: PAS (Period/ARNT/Single-minded) domains function within many signalling proteins, where they sense cues from the environment²¹. GAF (cGMP phosphodiesterase/adenylate cyclase/FhlA) domains regulate catalytic activity via ligand binding, and form the PCB attachment site^{22,23}. The PHY (phytochrome) domain is a member of the GAF family, and contains a hairpin loop thought to be important in stabilising the photoactivated state (Pfr) of Phytochrome²². Therefore, together, the N-terminal PAS-GAF-PHY acts as the sensory part of Phytochrome protein²². The C-terminal comprises a PAS repeated domain and a HKRD (Histidine kinase related domain), which are important for signal transduction, although the HKRD appears to be dispensable²⁴.

Cell culture studies found the HKRD domain to be dispensable for reversible PHY-PIF binding but found that the PAS repeated domain is necessary for allowing reversal of PHY-PIF binding under 750nm light¹⁴. However, we find that the inclusion of the PAS repeated domain prevents the expression of PHYB in zebrafish embryos (Figure 1Ai). In contrast, a more truncated version of PHYB, without the C-terminal HKRD or PAS repeated domain, is robustly and reproducibly expressed in zebrafish tissue (Figure 1Aii).

2. Microinjection of 1.5pMoles of HPLC purified PCB is sufficient to mediate PHYB-PIF6 binding

In order to generate red light-dependent binding of PHYB and PIF domains, Phycocyanobilin (PCB) chromophore must be bound to PHYB protein. PCB is not present in eukaryotic cells and so must be added before the system will function. In order to test how efficiently PCB can be delivered to cells throughout the zebrafish embryo, we took advantage of a PCB biosensor, the Y276H constitutively active mutant of PHYB, which fluoresces under far-red light when bound to PCB^{14,25} (Figure 1Bii). The Y276H mutant is a useful tool to help optimise the binding requirements initially without light modulation. We generated a version of this mutant without the C-terminal HKRD or PAS repeated domain. RNA encoding Y276H PHYB and nuclear-labelling H2A-GFP was microinjected into 1-cell stage embryos, which were then dechorionated at 4 hours post fertilisation (h.p.f.) and exposed to 300 μ M PCB in the media overnight. When imaged under 633nm light, the enveloping layer (EVL) cells of the embryo exhibited strong fluorescence, indicating that PCB had bound to Y276H PHYB successfully. However, the underlying neuroepithelial (NE) cells did not show evidence of PCB binding (Figure 1Ci), demonstrating that PCB does not penetrate beyond the first layer of zebrafish tissue. We therefore injected 1.5pMoles of PCB directly into 1 cell of 16 cell stage embryos, along with RNA encoding a CAAX-linked Y276H PHYB and PIF6-EGFP. 18h.p.f. embryos were then imaged under 633nm light. The cell membranes of neuroepithelial cells fluoresced brightly throughout the depth of the neural rod, indicating that PCB had bound successfully to Y276H PHYB (Figure 1Cii), significantly more efficiently than following external PCB exposure. Cell morphology appeared normal and there was no apparent toxicity associated with PCB injection at this concentration. In this experiment, Y276H PHYB-CAAX also successfully recruited PIF6-EGFP to cell membranes (Figure 1Cii), demonstrating that PCB binding was sufficient to drive PIF6 recruitment. As expected, this recruitment was not reversible when embryos were exposed to 750nm wavelength light for 5 minutes because the Y276H PHYB acts in a constitutive-active manner upon PCB binding, and therefore binds PIF6 in both red and infrared light (Figure 1D).

For these experiments, we found that the standard crude PCB obtained from *Spirulina*¹⁴ proved phototoxic to the embryo even when delivered at concentrations as low as 10uM (Figure 1E). However, further HPLC purification of PCB (see methods) removed this toxicity. When embryos were bathed in purified PCB for 20 hours embryo survival was not significantly different between treated and controls, even at PCB concentrations up to 300uM (data not shown).

3. EGFP can be reversibly shuttled between the cytoplasm and cell membrane

To test whether the truncated version of PHYB (missing the C-terminal HKRD or PAS repeated domain) was capable of reversibly recruiting PIF6-tagged proteins to the cell membrane of neuroepithelial cells within the developing zebrafish embryo brain, RNA encoding PHYB-MCherry-CAAX (the same construct used in Figure 1Aii) and PIF6-EGFP was co-injected into 1 cell of 4-16 cell stage embryos along with PCB. Embryos were then raised in the dark. At the 15-somite stage, embryos were globally exposed to alternating 5-minute intervals of 650nm (ON) and 750nm (OFF) filtered bright-field light, and the expression of PHYB-MCherry-CAAX and PIF6-EGFP was assessed by confocal microscopy at the end of each interval. This resulted in the robust and repeatable shuttling of EGFP between the cell membrane and the cytoplasm (Figure 2). The relative differences in EGFP intensity between cytoplasm and membrane under the different light conditions were quantified (Supplementary figure 1). These analyses demonstrated a large and significant reduction of 39% in EGFP intensity within the volume of the cytoplasm when PHYB-PIF binding was activated under 650nm light (Supplementary figure 1A). A concurrent increase of 24% in EGFP intensity within the membrane was also seen (Supplementary figure 1B). This resulted in a 43% higher level of EGFP intensity at the membrane when compared to the cytoplasm under 650nm light (supplementary figure 1C).

To test the speed of PHYB-PIF binding and unbinding, we quantified PIF6-EGFP localisation over time following bright field (ON) and 740nm (OFF) filtered bright-field light (Figure 3). This

data demonstrates that PHYB-PIF binding is rapid: A time constant (τ) of 6.5s was found for decreasing cytoplasmic EGFP levels, which plateaued 33% lower than starting levels (Figure 3B). Unbinding of PHYB-PIF was slower than binding but still relatively rapid: A τ of 46.9s was found for increasing cytoplasmic EGFP levels (Figure 3C). Cytoplasmic readings may be more robust than membrane readings due to the difficulty of restricting sampling of fluorescence intensity from membrane alone (discussed later).

This is the first demonstration of the experimental movement of proteins between different cell compartments within a multicellular organism using the phytochrome system and demonstrates a high degree of temporal resolution.

4. Subcellular control of protein localisation

In order to probe specific signalling and polarity pathways within cells, it is necessary to control protein localisation at a subcellular level. We therefore tested the capability of the phytochrome system to recruit a cytoplasmically localised EGFP protein (PIF6-EGFP) to a specific region of the plasma membrane. As before, RNA encoding PHYB-MCherry-CAAX and PIF6-EGFP was co-injected into 1 cell of 8-16 cell stage embryos along with PCB. Embryos were then raised in the dark. At approximately the 14-15 somite-stage, selected embryos were globally exposed to bright field light, causing PIF6-EGFP to be uniformly recruited to the cell membranes (Figure 4A). The embryos were then globally exposed to 740nm filtered bright-field light causing PIF6-EGFP to leave the membrane (Figure 4B). A region of interest was then specified over a restricted part of the plasma membrane (Figure 4C) and low levels of 633nm laser light were specifically localised to this region in the presence of global 740nm light. This resulted in the robust and specific recruitment of PIF6-EGFP to the ROI (Figure 4D). The embryos were then globally exposed to low levels of 633nm laser or bright field light, resulting in the recruitment of PIF6-EGFP to all cell membranes (Figure 4E). Subcellular recruitment was rapid, occurring

within 1 minute of exposure to 633nm light (Figure 4Di). Longer exposures to 633nm light still resulted in robust subcellular recruitment, without obvious bleaching(15 minutes, figure 4Dii). Subcellular recruitment was successful both in superficial large EVL cells and in the deeper and smaller NE cells (Figure 4).

To quantify subcellular recruitment dynamics, we then analysed the mean EGFP intensity from different areas within the single cell shown in Figure 4Di. The restricted subcellular exposure to 633nm illumination resulted in a relatively larger increase in membrane recruitment when compared to cells receiving a global exposure to 633nm (Figure 4F and Supplementary Figure 1).

These results demonstrate that the phytochrome system is an efficient method to accurately, rapidly and reversibly recruit proteins to specific subcellular regions within a vertebrate embryo. It therefore has the potential for highly specific control of subcellular signalling and polarity pathways, both temporally and spatially.

5. Mislocalising Pard3 using the phytochrome system

In order to determine whether the phytochrome system can be used to alter polarity protein distribution within the living zebrafish brain, we sought to mislocalize the apical polarity protein Pard3. Pard3 is a key protein during epithelial development in many systems. Our previous work has suggested that the spatial localisation of zebrafish Pard3 may play important roles in establishing neuroepithelial polarity²⁶⁻²⁸ and in regulating asymmetric divisions in the neural tube²⁹. Pard3 localises gradually but specifically to the tissue midline prior to lumen formation during zebrafish neural tube development²⁶⁻²⁸ (summarized in Figure 5A). To manipulate Pard3

localisation, we generated a fusion construct of Pard3 with EGFP and PIF6. This fusion protein localised to apically positioned rings of epithelial junctions within the zebrafish neural rod (Figure 5B), thus faithfully reproducing endogenous Pard3 expression. Initially we sought to use the phytochrome system to distribute Pard3 to the whole cell membrane of neuroepithelial cells, rather than just to the apical domain (Figure 5Ci). To maximize the likelihood of success we first did this using the constitutively active Y276H version of PHYB. RNA encoding Y276H PHYB-CAAX and Pard3-EGFP-PIF6 was co-injected into 1 cell of 4 cell stage embryos along with PCB. Embryos were then raised in the dark and imaged at approximately the 17-somite stage. Since Y276H PHYB-CAAX is constitutively active (figure 1D), PHYB-PIF6 binding was induced from very early in development, without the need for specific light exposure. Interestingly, at the tissue level, both Y276H PHYB-CAAX and Pard3-EGFP-PIF6 were partially mislocalized: Y276H PHYB-CAAX was not uniformly localised throughout the cell membranes but was preferentially localised towards the more apical ends of cells, whilst Pard3-EGFP-PIF6 was found both at its normal midline location but also in more basolateral domains. At a cellular level, this some cells had a higher accumulation of both Pard3-EGFP-PIF6 and Y276H PHYB-CAAX at or near the apical midline of the neural rod (as is normal for Pard3 localisation) (Figure 5Di). In other cells both proteins co-localized along the lateral edges of neuroepithelial cells (demonstrating a mislocalization of Pard3) (Figure 5Dii). This suggests that a ‘tug of war’ occurred between the intended anchor (PHYB) and bait (PIF6) and resulted in a ‘streaky’ distribution of Pard3 throughout the neural rod but still with an apical bias (Figure 5Diii). When embryo development was followed over time, Pard3-EGFP-PIF6 was found to gradually localise to the midline, dragging Y276H PHYB-CAAX with it (Supplementary Figure 2A). Thus both the CAAX anchor and Pard3 bait ended up at the midline. Eventually Pard3-EGFP-PIF6-rich apical end feet formed and the lumen was able to open (Supplementary Figure 2B). These results demonstrate that it is possible to link Pard3 to the membrane within neuroepithelial cells using

the phytochrome system. However, the 'pull' of endogenous Pard3 to the midline is eventually stronger than the 'pull' of CAAX-linked PHYB. We suspect the initial mislocalization of the Pard3-EGFP-PIF6 and its subsequent correct localisation to the midline may reflect the timing of the endogenous mechanism that traffics Pard3 to the midline.

We next attempted to use the phytochrome system to mislocalize Pard3 to targeted domains within individual cells (Figure 5Cii). RNA encoding PHYB-MCherry-CAAX and Pard3-EGFP-PIF6 was co-injected into 1 cell of 16-32 cell stage embryos along with PCB. Embryos were then raised in the dark. EVL and NE cells were illuminated with 740nm light, to switch off PHYB-PIF binding, and subsequently illuminated with 633nm light in specific ROIs within cells (5E and F). Within the EVL this resulted in the specific upregulation of Pard3-EGFP-PIF6 to the ROI and its concurrent depletion from its normal location in the surrounding cell membrane (figure 5Eiii). Following further 740nm illumination, Pard3-EGFP-PIF6 redistributed back into the cell membrane (figure 5Eiv). Within mature NE cells, 633nm illumination within specific non-apical ROIs resulted in the specific upregulation of Pard3-EGFP-PIF6 to the ROI and its concurrent depletion from the cytoplasm (Figure 5Fiii). The Pard3-EGFP-PIF6 that had already accumulated at the apical midline was maintained, demonstrating either that the endogenous mechanism of apical localisation is stronger than the competing PHYB-PIF6 binding or that there is a low turnover between apical Pard3 and the cytoplasm. Continued 633nm illumination within the ROI resulted in the NE cell entering division with two Pard3 positive poles (Figure 5Fiv); one at the normal apical midline (arrow) and one as a result of mislocalized Pard3 at the basal pole (star). We then sought to mislocalize Pard3 to one pole of a cell undergoing mitosis during midline crossing division (C-division). This resulted in the majority of Pard3-EGFP-PIF6 being recruited into one of the 2 daughter cells thus generating an asymmetric inheritance of Pard3-EGFP-PIF6 (Supplementary figure 3B). This raises the exciting possibility of being able to directly manipulate and test the importance of asymmetric inheritance of potential fate

determinants in vertebrate neural progenitors.

These results show, for the first time, the subcellular mislocalization of a polarity complex protein inside a living multicellular organism.

DISCUSSION

Adapting the phytochrome system for use in a vertebrate embryo

Here we achieve functionality of the phytochrome system in vertebrate embryos by solving the challenges of expressing the optogenetic components and delivering the PCB chromophore. With regards to expressing the protein components, the removal of the C-terminal PAS repeat domain in PHYB made a marked difference to the success of expression of PHYB protein in living zebrafish embryos (Figure 1A), and the resulting truncated PHYB was still able to highly efficiently and reversibly recruit PIF proteins (Figures 2 and 3). The C-terminal PAS repeat domain of PHYB protein has been shown to be necessary and sufficient for regulating the light dependent shuttling of the Pfr (activated) form of PHYB into the nucleus³⁰. The localisation of PHYB in the nucleus allows the regulation of gene expression within its native Arabidopsis. Therefore removal of the PAS repeat domain may abrogate the regulation of native signalling pathways. However, our current study suggests that when using the phytochrome system as a reversible heterodimerizer tool, the C-terminal signalling domain is not necessary. This is in line with earlier observations that the C-terminal is not actually directly involved in signal transduction and that, if made to localise in the nucleus, the isolated N-terminal actually mediates more sensitive signalling responses than full length PHYB³¹.

Injection of HPLC purified PCB chromophore directly into 1 cell of 16 cell stage embryos, along with RNA encoding PHYB and PIF proteins, allowed successful penetration of PCB deep into the embryo, without any obvious toxicity (Figure 1Cii). A more uniform level of PCB could be achieved if zebrafish cells were engineered to synthesise their own PCB by genetically introducing genes responsible for PCB synthesis (heme oxygenase 1 (HO1) and PCB:ferredoxin oxidoreductase (PcyA)). This has previously been carried out in bacteria¹⁹ and has now recently been successfully achieved in mammalian cells²⁰ but has so far not been

achieved in whole multicellular organisms. However we have found that, at least up to 24h.p.f., injection of PCB protein is sufficient for optogenetic experiments, and the transgenic generation of PCB is not required.

Protein expression levels and efficiency of binding/unbinding

As noted previously in cell culture³³, we also found that achieving a proper ratio of PHYB and PIF protein inside an individual cell are critical in assaying PHYB-PIF interactions. A relatively higher level of PHYB is necessary to robustly recruit PIF protein to the membrane and to deplete it from the cytoplasm. Following RNA injection, there is always a variance in the levels of fusion proteins that are expressed between embryos and even between individual cells in one embryo, so this resulted in some embryos and some cells within embryos that did not significantly recruit PIF protein. If a more standardised level of protein expression (and therefore PHYB-PIF binding) is required, stable transgenic lines of zebrafish should be made. Once PHY-PIF binding is achieved, a solution to modulate recruitment levels despite cell-to-cell variability in protein levels would be the use of a computational feedback controller, as has been used in cell culture to vary the amount of light being delivered to each cell, therefore controlling the level of PHYB-PIF binding over time and space¹⁸.

The efficiency of PHY-PIF binding and unbinding achieved *in vivo* within the zebrafish embryo was high. The speed of binding was slightly slower than in cell culture, with a $\tau=6.5s$ for cytoplasmic depletion during PHYB-MCh-CAAX:PIF6-EGFP binding in zebrafish (figure 3B), compared with 1.3s during PHYB-MCh-CAAX:PIF6-YFP binding in NIH3T3 cells¹⁴. However, this still allowed us to reproduce the rapid and reversible shuttling of a fluorescent protein between the cytoplasm and membrane within a few seconds as seen in cell culture (figures 2

and 3). The relative difference in EGFP levels between recruited membrane and the rest of the cell is much more marked when binding illumination is restricted to subcellular ROIs (62% higher EGFP levels in the membrane compared with the cytoplasm under binding light in figure 4 compared with 43% in supplementary figure 1). Unbinding speed was slower than binding, with $\tau=46.9\text{s}$ for cytoplasmic increase in EGFP levels (Figure 3C). This may partly be due to the differences in method used for measuring τ_{on} and τ_{off} rates (see methods). However, we suggest that the main reason is ambient light causing low levels of PHYB-PIF binding, therefore slowing down unbinding rates under 740nm light.

Since the cells being analysed in this work are within a living whole organism they are not static, instead moving in all dimensions between time frames. This makes a standardised approach to sampling from a single z-level challenging. Due to the different scales of membrane and cytoplasm pools seen in single z-levels, sampling error from hand-drawn ROIs is much greater when specifying membrane ROIs than when specifying cytoplasmic ROIs. Therefore, due to contamination of the ROI by cytoplasmic voxels, the time constants (τ) calculated for membrane binding and unbinding of EGFP are artificially slowed and the variance is higher than in cytoplasmic readings (Figure 3B and C). For the same reason, the readings for relative changes in EGFP intensity in the membrane following different light exposures are artificially lower than cytoplasmic readings (e.g. compare supplementary figure 1A and B). Therefore the cytoplasmic readings should be viewed as truer kinetics measurements in these experiments.

It should be noted that we found PHYB-PIF binding to be very sensitive to ambient light. Even background light from computer monitors was enough to initiate some PHYB-PIF binding, and bright field light very effectively caused binding. As previously discussed, the more blue-shifted the light wavelength, the less binding it will cause but even wavelengths as low as 488nm cause some PHY-PIF binding¹⁸. We also found this to be the case in zebrafish embryos. Conversely,

only wavelengths of light very close to 740-750nm were sufficient to initiate unbinding. This sensitivity necessitates careful experimental design and control. However, it also makes it highly unlikely that PHY-PIF binding will be reversed by exposure to background light, making the use of the phytochrome system for subcellular recruitment robust.

Altering apico-basal polarity in live zebrafish embryos

To demonstrate that the phytochrome system can be used to alter polarity pathways within the living zebrafish brain, we chose to manipulate Pard3 localisation during neuroepithelial development. Pard3 is a key apical polarity protein in epithelia and has been shown to be instrumental in normal zebrafish neural tube development. For example, it has been shown to be important in mediating mirror-symmetric cell division during zebrafish neural rod development²⁶, allowing proper lumen formation²⁷, positioning centrosomes³⁴, and promoting neurogenic divisions in the brain²⁹. Global knock out of Pard3 is informative, but a more targeted manipulation of Pard3 localization would elucidate the role of these proteins in individual cells and at individual times. We therefore used the phytochrome system to recruit Pard3-EGFP-PIF6 to specific ectopic areas of the cell membrane. While mislocalization within the EVL resulted in the concomitant depletion of Pard3-EGFP-PIF6 from its normal location around the cell periphery (Figure 5E), the endogenous drive for Pard3 to localize to the apical midline in NE cells was stronger than the ‘pull’ to ectopic localizations generated by subcellular illumination with 633nm laser light and Pard3-EGFP-PIF6 was still able to localise at the apical midline (Figure 5 and supplementary figures 2 and 3). This result raises several experimental considerations that need to be met before the phytochrome system can be used for robust subcellular manipulation of polarity proteins within multicellular organisms. First, these experiments were carried out using the overexpression of fusion proteins in a wild-type

background. This means that endogenous Pard3 should not have been affected by PHYB binding and may have made it more permissive for Pard3-EGFP-PIF6 to reach the tissue midline via endogenous mechanisms. A genetic approach to replace endogenous Pard3 with PIF-tagged Pard3 may make Pard3-EGFP-PIF6 mislocalization more robust and would standardise protein expression levels. Secondly, while specifying an ROI using a point scanner approach, as in figures 4 and 5, is a simple and effective way in which to pattern light illumination, it is only possible to do this at a single z-level. Due to the cone effect of conventional confocal microscopes, the PHYB-PIF binding induced via this method may not have generated a sufficiently accurate binding area to generate a robust 'pulling force' and to therefore mislocalize all of the Pard3-EGFP-PIF6 present within cells. Third, while only the ROI was specifically illuminated with binding 633nm light, the whole embryo was bathed in unbinding 740nm light. This is likely to have reduced the binding efficiency of the ROI. Fourth, as mentioned phytochrome binding is very sensitive to ambient light and even 488nm causes some binding. Therefore, every time an image was taken, global binding occurred briefly, again disrupting specific recruitment to the ROI. Ultimately, a more sophisticated method to create inverted light patterns with a much higher z-specificity would be to employ a spatial light modulation system³⁵. This type of structured illumination will be necessary to allow 3D ROIs to be made as well as to allow simultaneous imaging and control of binding/unbinding of PHYB and PIF.

These experiments have demonstrated that the phytochrome system can be used in live multicellular organisms to investigate the role of particular components of cell polarity pathways at different times at a subcellular level in a whole organism. The phytochrome system therefore has the potential for an unprecedented level of experimental specificity within whole organisms, both temporally and spatially.

CONCLUSION

We have successfully developed the use of light-induced heterodimerization to control subcellular protein movement within the *in vivo* zebrafish brain using the phytochrome system. This system is rapidly reversible and allows a very high level of temporal and spatial control, with binding rates close to those seen in cell culture. As a proof of principle, we also show that the phytochrome system can be used to alter polarity pathways within the living zebrafish brain by manipulating subcellular Pard3 location within both EVL and NE cells. The phytochrome system can therefore be used in whole multicellular organisms in order to specifically manipulate protein location at a specific tissue or subcellular level in a temporally controlled and reversible manner, in combination with high-resolution time-lapse microscopy. This removes the need to globally remove proteins or to abrogate their function, with the potential to make experimental approaches much more targeted. The phytochrome heterodimerization technique will be particularly useful to studies in which protein localisation and timing is critical, such as studies of cell polarity and migration. The system is also optimal when exact titrations of optogenetic input are important, since ratiometric activation by red/infrared is possible^{16,18}, and allows the use of multiple activity probes (such as FRET sensors) in conjunction with optogenetic control.

This is the first time that the phytochrome system has been successfully used in live multicellular organisms, and allows an unprecedented level of experimental control. We envision that the specific optimizations and general methodology detailed here will enable similar successes in other organisms.

METHODS

Phycocyanobilin (PCB) extraction and HPLC purification

PCB was extracted as described previously³³. The PCB was then further purified: The PCB-DMSO solution was spun at 5000xG 10min at room temp. The supernatant was passed through glass wool and a 0.45um filter. The filtered PCB-DMSO solution was run on a C18 column (Waters Atlantis Prep T3 OBD 5 um, 19x100mm column). Solution A: 0.5% formic acid in water. Solution B: 0.5% formic acid in acetonitrile. The sample was run with a gradient to 5% solution B in 5 min, followed by a gradient to 100% solution B in 11 minutes. Liquid chromatography-mass spectrometry confirmed only the major peak contained PCB (587.6 g/mol). PCB from the major peak was pooled, lyophilized and resuspended in a minimal amount of DMSO. To obtain the final PCB concentration, a small amount of the final preparation was diluted 1:100 into 1 mL 95:5% MeOH:HCl (37.5%) solution and the absorbance at 680 nm measured. In 95:5% MeOH:HCl, the extinction coefficient of PCB is 37,900 M⁻¹ cm⁻¹. Typical final concentrations from this procedure were between 3 and 15 mM. The PCB was diluted to 12.5mM in DMSO and stored at -20 °C until use.

Confocal imaging

Embryos were mounted in low melting point agarose and imaged on upright Leica LSM microscopes (SP5 and SP8) using high Na water dipping x25 and x65 objectives. Images were processed using Volocity and Image J software.

Far red and infra-red light sources

Whole embryo illumination experiments were carried out under bright light filtered through a 650nm band-pass filter (Edmund Optics) and a 750nm long-pass filter (Newport, RG9). Subcellular recruitment experiments were carried out using a low intensity of the 633nm laser on a laser scanning confocal microscope (Leica SP5) within a region of interest (ROI). The

background was illuminated with filtered 740nm light from a Schott KL 1500 LCD light source, filtered with bandpass glass (Envin Scientific Ltd.). When taking confocal images of EGFP localization during subcellular experiments it was necessary to briefly switch off both the 633nm laser and the 740nm unbinding light.

Binding and unbinding assays

Binding and unbinding assays were carried out on a laser scanning confocal microscope (Leica SP5) at 28.5°C. For the binding assay (Figure 3B) we replaced 740nm light with BF light at 0 seconds and took an image under 488nm light every 4 seconds. Since even 488nm light caused binding in our experiments we took a different approach for the unbinding assay (Figure 3C). At 0 seconds we replaced BF light with 740nm light and took an image under 488nm light after 15 seconds. We then recruited GFP to the membrane under BF light for approximately 1 minute before again replacing BF light with 740nm light and taking an image under 488nm light after 30 seconds. This process was repeated for all timepoints assessed.

Image analysis

To calculate relative EGFP intensity, mean EGFP intensity readings from 3-4 ROIs were taken for each condition from both the cell membrane and the cytoplasm using Volocity image analysis software, as well as from the background of the image. The mean background reading was then subtracted from each of the ROI readings. ROI readings were normalised to overall EGFP levels, as described in figure legends. Statistical analyses are described in the figure legends and were carried out using Graphpad Prism software.

Fusion constructs

Fusion constructs were made using the PCS2+ plasmid backbone using an enzymatic assembly method³⁶. 10 amino acid polyglycine-serine linkers were used in most constructs. RNA was then synthesised with the Ambion mMessage mMachin System from the sp6

promoter (AM1340) and injected into early embryos as described in the text using standard injection protocols³⁷. Details of constructs are below:

Construct name	Composition	Approximate amount injected
1. N-PAS2-GAF-PHY-PAS	1-917 amino acids of Arabidopsis PHYB (Gene ID: 816394), tagged with MCherry fluorescence protein and the CAAX membrane moiety. Mammalian codon optimised	0.2ng
2. N-PAS2-GAF-PHY/PHYB-MCherry-CAAX	1-621 amino acids of Arabidopsis PHYB (Gene ID: 816394), tagged with MCherry fluorescence protein and the CAAX membrane moiety. Mammalian codon optimised	0.2ng-0.25ng
3. Y276H PHYB-CAAX and Y276H PHYB	1-621 amino acids of Arabidopsis PHYB (as in construct 2) with and without the CAAX membrane moiety tag. Tyrosine residue 276 of PHYB was mutated to Histidine.	0.2ng-0.25ng (CAAX tagged) 0.15ng (without CAAX tag)
4. H2A-GFP	Histone 2A tagged with GFP	0.04ng
5. PIF6-EGFP	1-100 amino acids of Arabidopsis PIF6 (Gene ID: 825382), tagged with EGFP fluorescent protein. Mammalian codon optimised.	0.05ng
6. Pard3-EGFP-PIF6	Full length zebrafish Pard3 (Gene ID: 403050), tagged with EGFP fluorescent protein and Arabidopsis PIF6 (as in construct 5).	0.1-0.15ng

Embryo Care

Embryos were collected, staged and cultured according to standard protocols^{37,38}. All procedures were carried out with Home Office approval and were subject to local Ethical Committee review.

ACKNOWLEDGEMENTS

Thank you to the rest of the Clarke and Weiner labs for help and advice, especially to Andy Symonds, Doug Tischer, Jared Toettcher and Stephanie Woo. Thanks to Jared Toettcher, Delquin Gong, Doug Tischer, and the Weiner and Degrado labs for their help with PCB purification. Thank you to Andrew Lowe for advice on analysis. Thank you to the European Molecular Biology Organisation for awarding a short-term fellowship to allow this collaborative work to take place. Thank you to our other funding sources; BBSRC, MRC, Wellcome Trust (JDWC) GM084040, GM096164, and AHA Established Investigator Award (ODW)

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

Figure 1: Adapting the phytochrome system for use in a vertebrate embryo

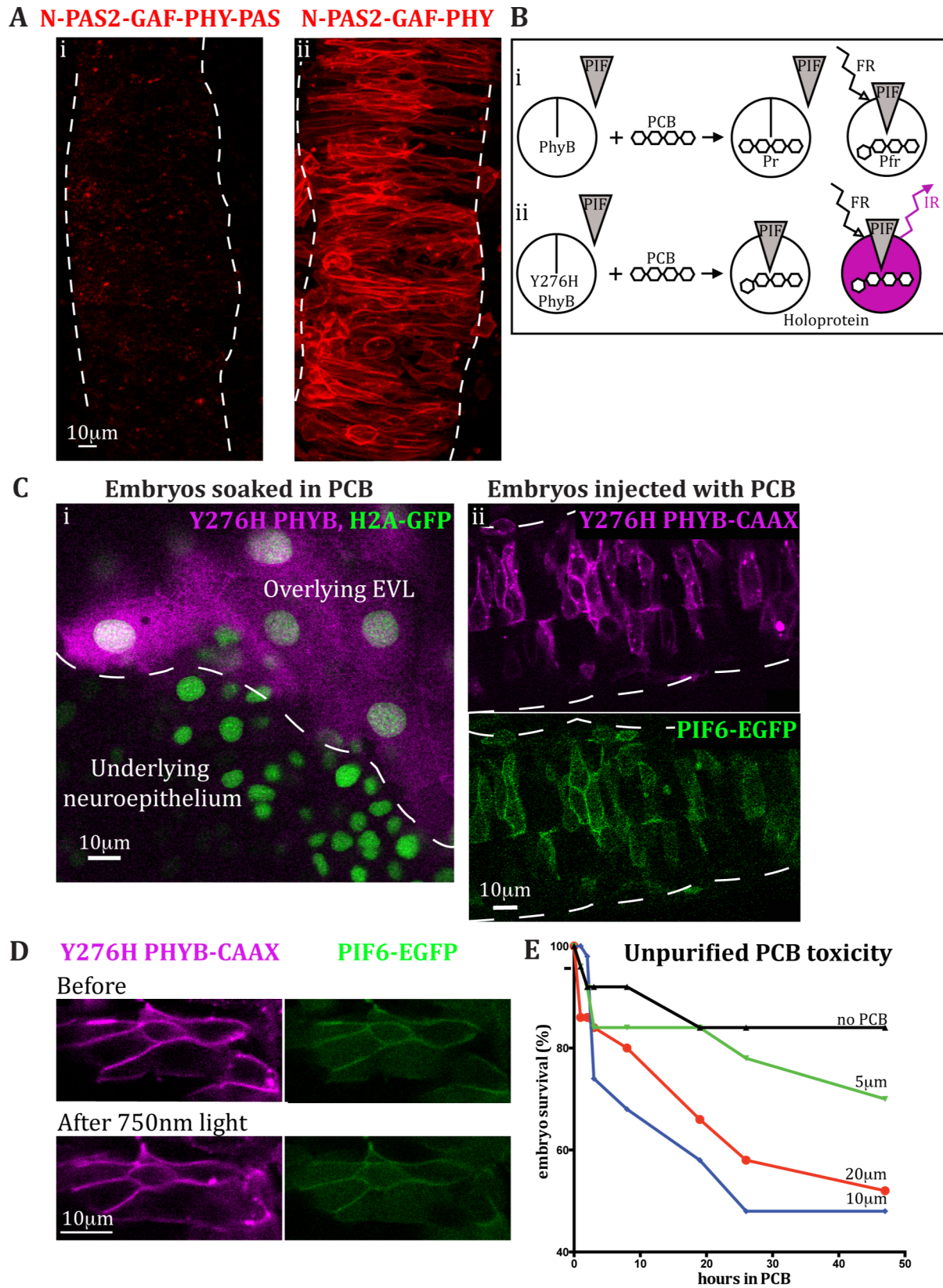


Figure 1

A. 50µm Z-projections through the hindbrains of approximately 14-somite stage zebrafish embryos mosaically labelled with PHYB-MCherry-CAAX fusion proteins with different PHYB truncations. Dotted lines denote basal edges. **(i)** N-PAS2-GAF-PHY-PAS contains a PAS domain at the C-terminal and was not successfully expressed in 6/6 embryos. **(ii)** N-PAS2-GAF-PHY does not contain a PAS domain at the C-terminal and was robustly expressed in cell membranes in 9/9 embryos. **Bi.** Normal PHYB covalently binds PCB chromophore. Energy from far-red light causes photoisomerisation of PCB and the allosteric transition of PHYB from its inactive (Pr) to its active (Pfr) state. This is reversible by infrared light exposure (not shown). The Pfr state can bind PIF. **Bii.** Conjugation of PCB with Y276H mutant PHYB creates an activated holoprotein that can directly bind PIF, without the need for far-red light illumination²⁵. Energy from far-red illumination causes infra-red fluorescence. **Ci.** An oblique single z-slice through the overlying enveloping layer (EVL) and the underlying developing neuroepithelium of a 14h.p.f. embryo, labelled with Y276H PHYB and H2A-GFP. Only the EVL fluoresces under far red light (magenta), demonstrating binding of PCB to Y276H PHYB. **Cii.** A single horizontal z-slice through the neuroepithelium of an 18h.p.f. embryo, mosaically labelled with Y276H PHYB-CAAX and PIF6-EGFP. Basal edges are labelled with dotted lines. The membranes of all labelled cells fluoresced under far red light, demonstrating binding of PCB to Y276H PHYB-CAAX. PIF6-EGFP was also recruited to the membrane in these cells. **D.** A rotated and magnified view of cells from the same embryo as in 'Cii'. Anterior is up. PIF6-EGFP recruitment to the membrane was not reversed after 5 minutes exposure to 750nm light. **E.** Time-course of percentage embryo survival after 3-5h.p.f. embryos were dechorionated and bathed in increasing concentrations of unpurified PCB. 50 embryos were used for each concentration.

Figure 2: Light-controlled shuttling of protein between cytoplasm and membrane in vivo

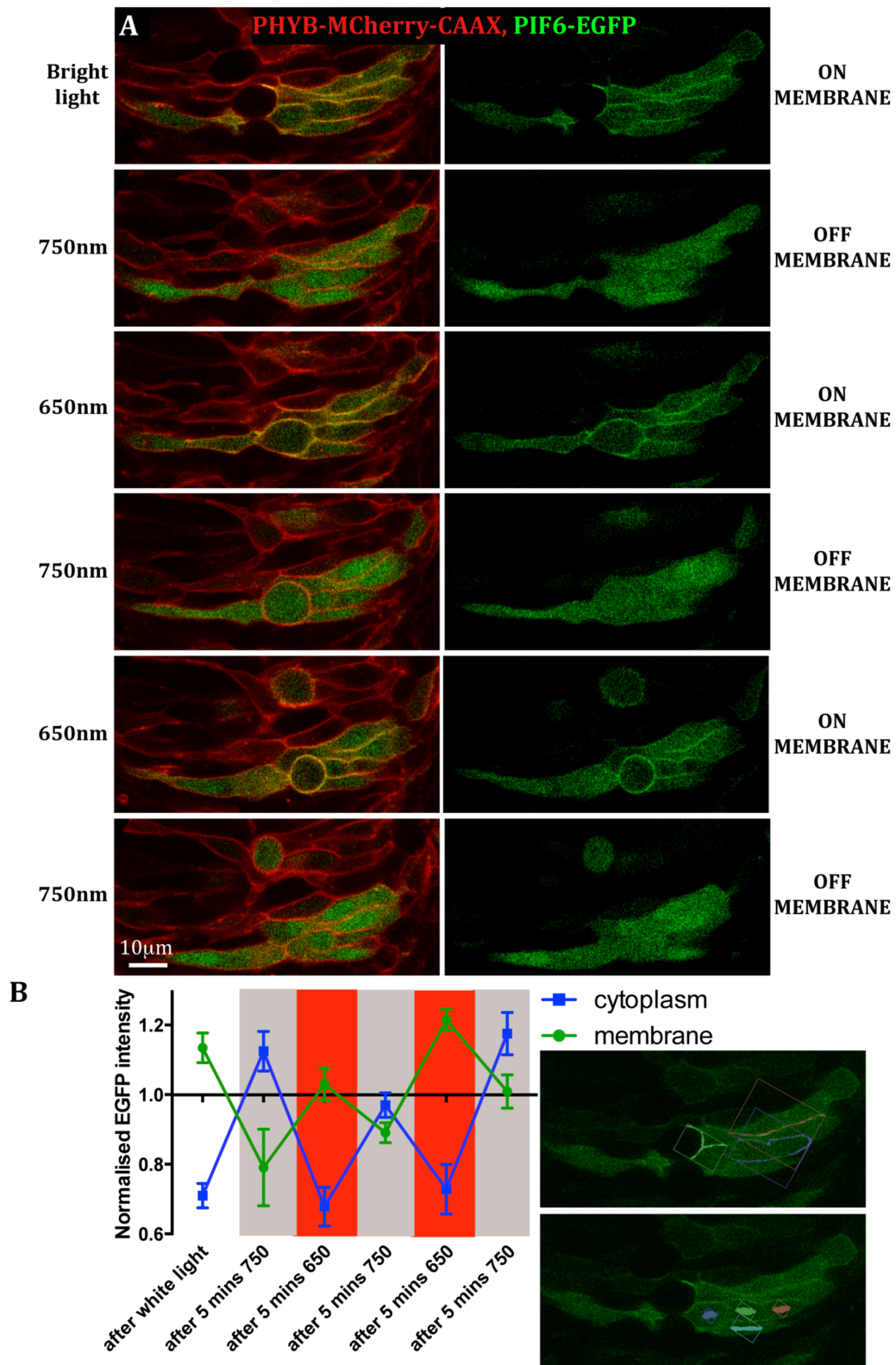


Figure 2

A. Sequential images of a horizontal single z-slice through the developing neuroepithelium of a 15 somite embryo mosaically labelled with PHYB-MCherry-CAAX and PIF6-EGFP. The embryo was illuminated with alternating 5-minute exposures to 650nm and 750nm light. PIF6-EGFP was recruited to the membrane after 650nm illumination and shuttled back to the cytoplasm after 750nm illumination. **B.** Relative EGFP intensity after alternating 650nm and 750nm illumination, normalised to mean 750nm levels from combined membrane and cytoplasm readings. EGFP intensity readings for each area were taken from 3-4 ROIs, totalling between 26 and 45 μm^2 (see images on right of graph as examples). The mean background EGFP intensity was calculated in the same manner for each image and subtracted from each membrane or cytoplasmic value. Error bars denote standard error of the mean.

Figure 3: Rapid kinetics of PhyB-PIF binding and unbinding

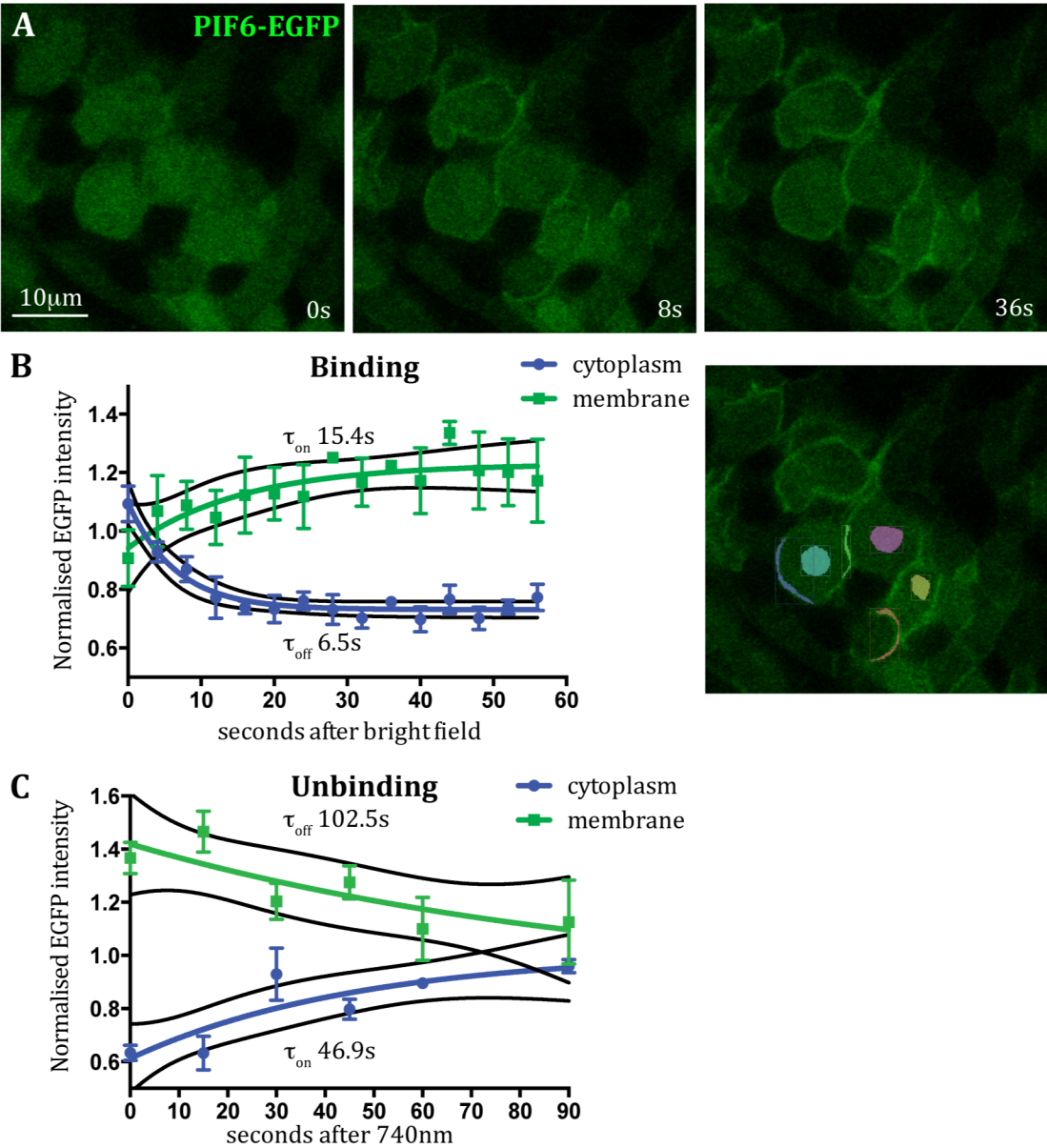


Figure 3

A. A sagittal z-slice through neuroepithelial cells of an approximately 15-somite stage embryo mosaically labelled with PHYB-MCherry-CAAX and PIF6-EGFP. Only the green channel is shown. At 0 seconds, 740nm illumination was replaced with bright field illumination and PIF6-EGFP was rapidly recruited to the membrane. **B and C.** Quantification of EGFP intensity over time following bright field (B) and 740nm (C) illumination. EGFP intensity readings for membrane and cytoplasm area were taken from 3 ROIs per timepoint, similar to those described in figure 3 (shown on the image on the right). The mean background EGFP intensity was subtracted from each timepoint and EGFP intensity was normalised to mean levels at t=0 seconds. One-phase exponentials were fitted to EGFP levels, generating time constants (τ) for binding and unbinding of PIF6 to PHYB. τ = time taken for EGFP levels to either decrease by a factor of 1/e (approximately 36.8% of the original amount) or to increase by a factor of 1-1/e (approximately 63.2% of the asymptotic value). Error bars denote standard error of the mean. Black lines denote 95% confidence bands. **B.** Quantification was from the cells shown in 'A'. Time constants (τ) of 6.5 seconds were generated for cytoplasmic EGFP decrease and 15.4 seconds for membrane EGFP increase during the binding and PIF6 to PHYB. **C.** Quantification was from NE cells from an approximately 22 somite stage embryo. At 0 seconds, bright field illumination was replaced with 740nm illumination and PIF6-EGFP was lost from the membrane. Time constants (τ) of 46.9 seconds were generated for cytoplasmic EGFP increase and 102.5 seconds for membrane EGFP decrease during the unbinding of PIF6 from PHYB.

4. Subcellular control of protein localisation

PHYB-MCherry-CAAX, PIF6-EGFP

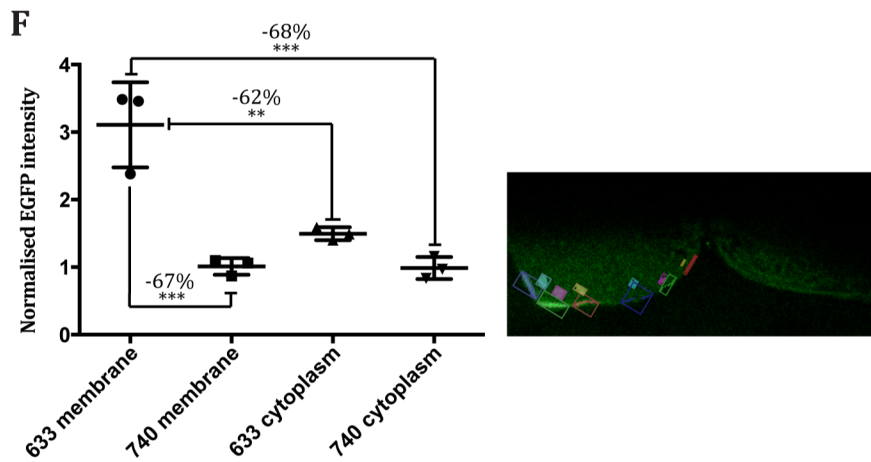
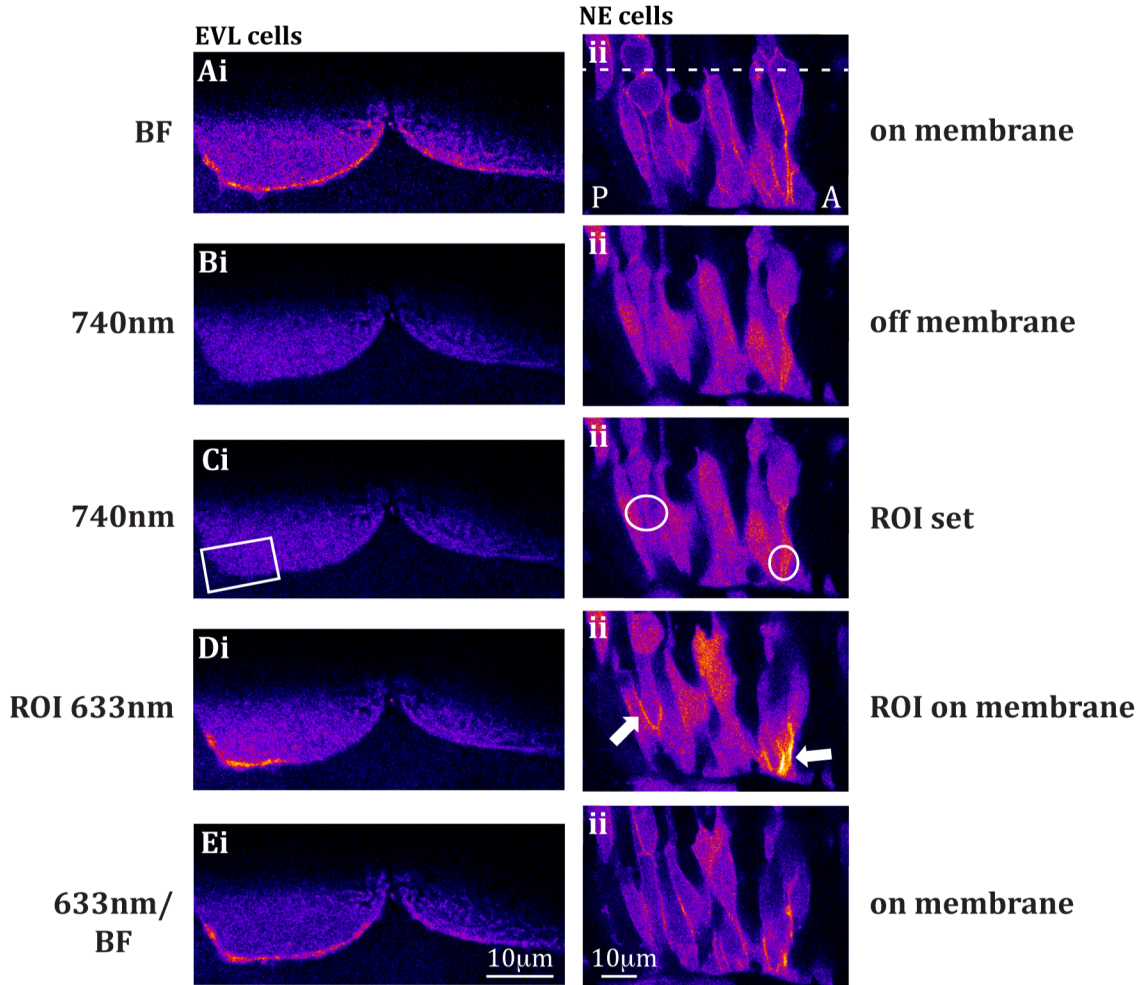


Figure 4

Sequential single z-slices through (i) two enveloping layer (EVL) cells of an approximately 14 somite-stage embryo and (ii) a collection of horizontally viewed neuroepithelial (NE) cells, mosaically labelled with PIF6-EGFP and PHYB-MCherry-CAAX. Dotted line denotes developing midline of the neural rod. Anterior and posterior orientation is indicated with 'A' and 'P'. Only the green channel is shown and is pseudo-coloured with the 'Fire' look-up table. The embryos were exposed to different patterns of low-level 633nm laser light and 740nm filtered bright-field light. **A.** Uniform bright-field (BF) light. **B.** Uniform 740nm light for 2 minutes (EVL) and 4 minutes (NE). **C.** ROI region was specified whilst illuminated with 740nm light. **D.** 633nm light was specifically delivered within the ROI with a background uniform 740nm light. EVL cells were illuminated for 1 minute. NE cells were illuminated for 15 minutes. PIF6-EGFP was specifically recruited to the areas of cell membrane that were illuminated with 633nm light. Successful subcellular PIF6-EGFP membrane recruitment was seen in 3 EVL cells and 8 examples of NE cells. **E.** Uniform 633nm light for 1 minute (EVL) or BF light for a few seconds (NE). **F.** Quantification of EGFP intensity from the image depicted in Di. EGFP intensity readings for membrane and cytoplasm area were taken from 3 ROIs for each region, similar to those described in figure 3 (shown on the image on the right). ROIs were placed at the membrane and in the cytoplasm both from regions illuminated by 633nm light and from those illuminated by 740nm light only. A one-way ANOVA with Tukey's multiple comparison test was carried out. This demonstrated that the EGFP intensity in the membrane illuminated with 633nm light was significantly higher when compared to all other regions of the cell. However, EGFP intensity was not significantly different between the different areas of cytoplasm or the membrane illuminated by 740nm only.

Figure 5: Mislocalising Pard3 using the phytochrome system

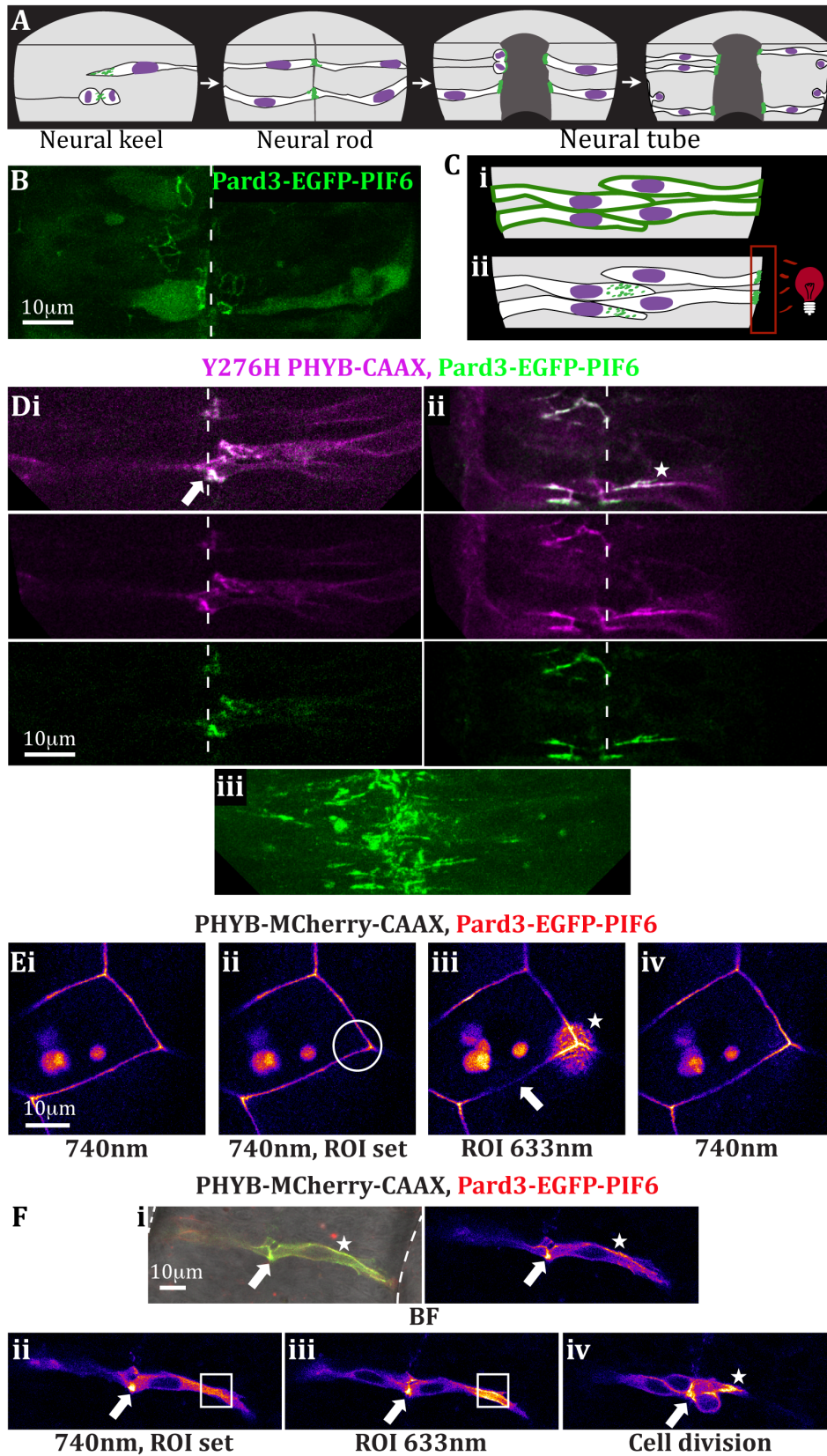
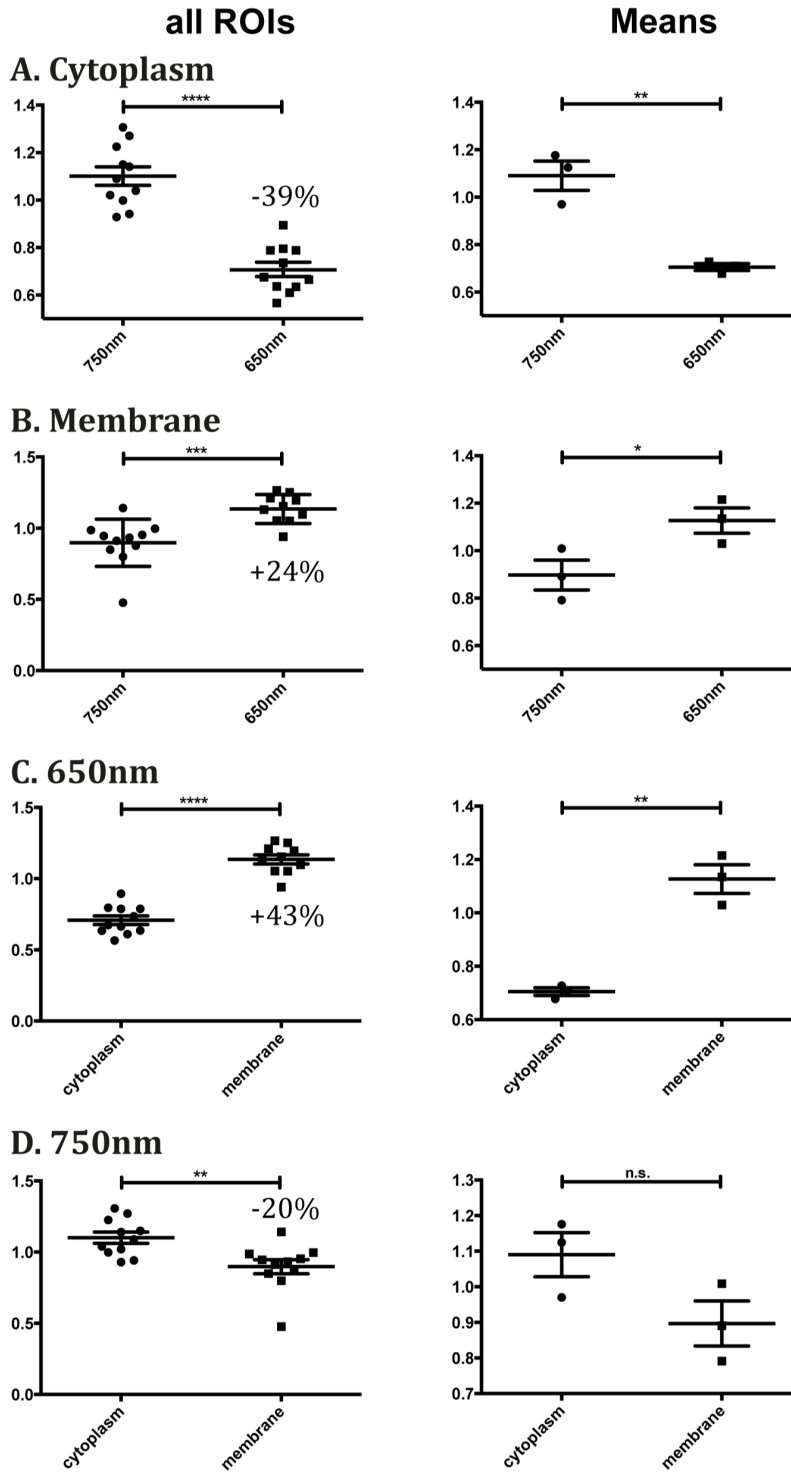


Figure 5

A. Cartoon illustrating the localisation of Pard3 (green) during the maturation of neural epithelial cells within the developing neural tube. At keel stages Pard3 broadly localises to the tissue centre before division. Midline crossing divisions (C-divisions) occur at the centre of the tissue and localise Pard3 more specifically to their cleavage planes, helping to produce a precise localisation of Pard3 at the tissue midline at rod stages. The neural tube lumen then cavitates from this midline, resulting in the opening of the neural tube. NE cells undergo further rounds of division at the Pard3-labelled apical surface during interkinetic nuclear migration, resulting in the production of both NE cells and neurons. **B.** An 84um horizontal z-projection through the neuroepithelium of a mosaically labelled 20h.p.f. embryo (approximately 22 somite stage). Pard3-EGFP-PIF6 localises normally to the apical end feet of neuroepithelial progenitor cells, which are visible as rings close to the tissue midline (dotted line). **C.** Cartoon depicting the experimental aims. Example cells are from the neural keel stage. **i.** By linking Pard3-EGFP-PIF6 to Y276H PHYB-CAAX, we aimed to mislocalize Pard3 to the whole cell periphery. **ii.** By illuminating a subcellular ROI within NE cells with 633nm light, we attempted to link Pard3-EGFP-PIF6 to PHYB-MCh-CAAX specifically within this region, therefore mislocalizing it to the basal ends of NE cells. **Di and ii.** Two examples of single horizontal z-slices through an approximately 17-somite stage embryo, showing colocalization between Y276H PHYB-CAAX and Pard3-EGFP-PIF6 (white). The two different channels are shown separately below the merged image. **i.** Pard3-EGFP-PIF6 has mainly reached the midline, dragging Y276H PHYB-CAAX with it (e.g. arrow). **ii.** Pard3 is mislocalized along the lateral edges of neuroepithelial cells (e.g. star). **Diii.** A 71um horizontal z-projection of the same embryo as in Bi and ii. Pard3-EGFP-PIF6 is partially mislocalized along the lateral edges of neuroepithelial cells but is still broadly localized towards the midline. **E and F.** Sequential single z-slices through **(E)** an enveloping layer (EVL) cell of an approximately 18 somite-stage embryo and **(F)** a pair of

neuroepithelial (NE) cells of an approximately 21 somite-stage embryo, labelled with Pard3-EGFP-PIF6 and PHYB-MCherry-CAAX. The green channel is shown alone and is pseudo-coloured with the 'Fire' look-up table. The embryos were exposed to different patterns of low-level 633nm laser light and 740nm filtered bright-field light. **Ei.** In the unbound state under 740nm light, Pard3-EGFP-PIF6 is distributed around the whole cell membrane. **ii.** ROI region was specified whilst illuminated with 740nm light. **iii.** 633nm light specifically within the ROI with a background uniform 740nm light for 23 minutes. Pard3-EGFP-PIF6 specifically accumulated to the ROI (star) and was depleted from surrounding cell membrane (e.g. arrow). **iv.** Uniform 740nm light for 9 minutes. Pard3-EGFP-PIF6 was redistributed along the cell membrane. Internal EGFP signal is an artefact. **Fi.** Mature NE cells without and with pseudo-colouring. Basal edges of the neural rod are indicated by dotted lines. Following BF illumination some Pard3-EGFP-PIF6 has accumulated normally at the apical surface (arrow) and some has been mislocalize to the lateral cell membrane (star). **ii.** Uniform illumination with 740nm light. Ectopic Pard3-EGFP-PIF6 shuttled from the cell membrane to the cytoplasm. Apically accumulated Pard3-EGFP-PIF6 remained (arrow). ROI region was specified while illuminated with 740nm light. **iii.** 633nm light specifically within the ROI with a background uniform 740nm light for 5 minutes. Pard3-EGFP-PIF6 specifically accumulated to the ROI and was depleted from surrounding cytoplasm. Apically accumulated Pard3-EGFP-PIF6 remained (arrow). **iv.** Specific 633nm illumination within the ROI was maintained for another 14 minutes, after which the cell started to divide with ectopic Pard3 at its basal pole. Successful subcellular recruitment of Pard3-EGFP-PIF6 was seen in 21 NE cells of varying levels of maturity.

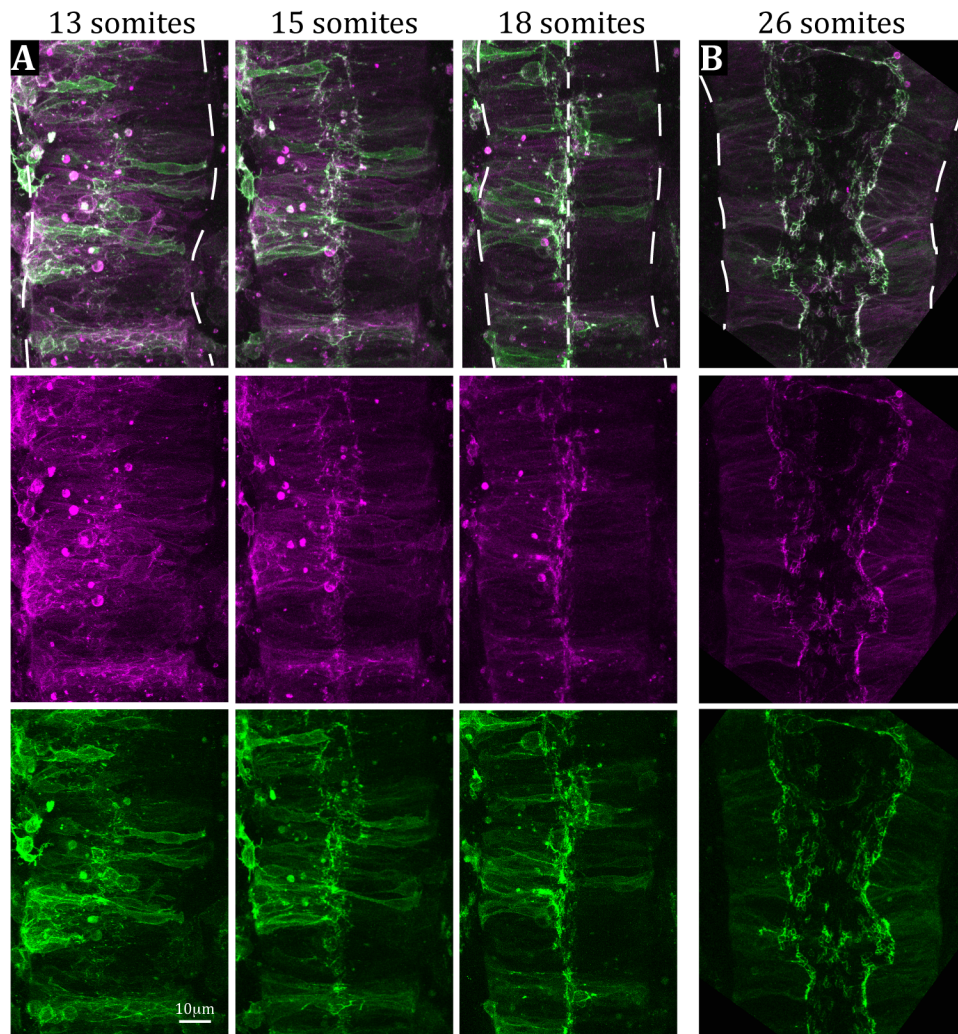
**Supplementary Figure 1:
Statistical analysis of EGFP intensity**



Supplementary Figure 1

The ROIs from figure 2B were combined and used to compare the mean EGFP intensity under different light conditions and at different regions of the cell using unpaired 2-way t-tests. Since the total area sampled varied between each image, the mean EGFP intensity from the ROIs in each image were also compared in the same way. Error bars denote standard error of the mean. **A.** There was a 39% higher EGFP intensity in the cytoplasm under 750nm light than under 650nm light ($P=0.0009$ for all ROIs and 0.5 for mean ROIs). **B.** There was a 24% lower EGFP intensity in the membrane under 750nm light than under 650nm light ($P<0.0001$ for all ROIs and 0.0038 for mean ROIs). **C.** There was a 43% lower EGFP intensity in the cytoplasm than in the membrane under 650nm light ($P<0.0001$ for all ROIs and 0.0016 for man ROIs). **D.** There was a 20% higher EGFP intensity in the cytoplasm than in the membrane under 750nm light ($P=0.0043$ for all ROIs and 0.939 for mean ROIs).

Supplementary Figure 2:
Gradual midline localisation of Pard3-EGFP-PIF6
Y276H PHYB-CAAX, Pard3-EGFP-PIF6

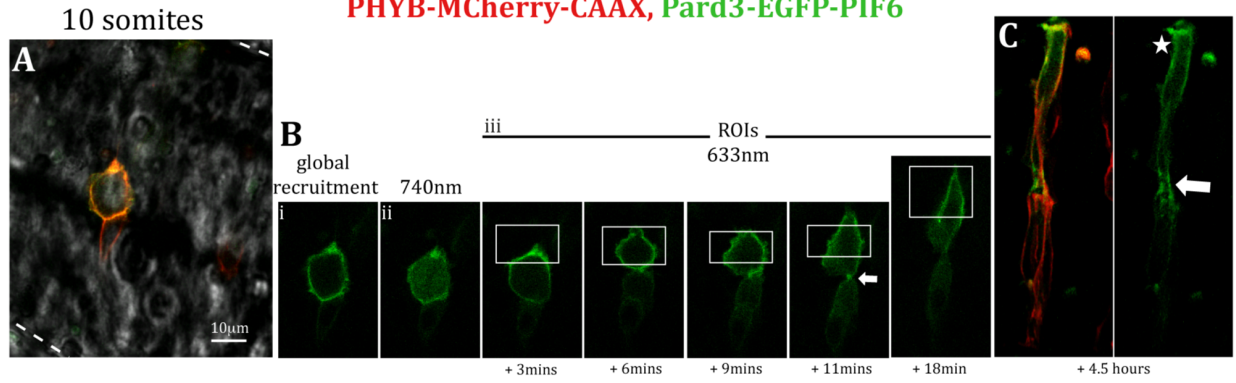


Supplementary Figure 2

Basal edges of the developing neural tube are shown by dotted lines. Developing midline is shown by small dotted line. **A.** Time series of a 73um horizontal z-stack through the hindbrain of a developing zebrafish embryo. Although Pard3-EGFP-PIF6 was initially mislocalized around the membrane of NE cells, it gradually localised to the midline, dragging Y276H PHYB-CAAX with it. **B.** Horizontal 64um z-stack through the hindbrain of a 26-somite stage embryo showing colocalization of Pard3-EGFP-PIF6 and Y276H PHYB-CAAX at the apical end-face of NE cells lining the open lumen. The two different channels are shown separately below the merged images.

Supplementary Figure 3: Mislocalisation of Pard3 during C-division

PHYB-MCherry-CAAX, Pard3-EGFP-PIF6



Supplementary Figure 3

A. A single horizontal z-slice through a NE cell undergoing C-division within a 13-somite zebrafish embryo. The cell is labelled with Pard3-EGFP-PIF6 and PHYB-MCherry-CAAX. **B.** The same cell as in 'A' after: (i) global recruitment to the membrane. Pard3-EGFP-PIF3 has been recruited to the cell membrane around the whole periphery of the cell. (ii) 740nm light illumination. The majority of Pard3-EGFP-PIF3 has shuttled into the cytoplasm. (iii) 633nm illumination specifically within the ROI (indicated by white box) with a background uniform 740nm light over 18 minutes. The majority of Pard3 localised within the ROI and therefore in the topmost cell. Some Pard3-EGFP-PIF6 was able to localise to the cleavage plane between the two cells after 11 minutes (white arrow). Both top and bottom cells integrated into their respective sides of the neuroepithelium. **C.** 3D reconstruction of the pair of sister cells after 4.5 hours of recruitment to the basal end of the top-most cell. Some Pard3-EGFP-PIF6 was localised basally (star). However, some was still able to localise apically (arrow). Image is shown with both red and green channels followed by green channel only.

REFERENCES

- 1 Pathak, G. P., Vrana, J. D. & Tucker, C. L. Optogenetic control of cell function using engineered photoreceptors. *Biology of the cell / under the auspices of the European Cell Biology Organization* **105**, 59-72, doi:10.1111/boc.201200056 (2013).
- 2 Tischer, D. & Weiner, O. D. Illuminating cell signalling with optogenetic tools. *Nature reviews. Molecular cell biology* **15**, 551-558, doi:10.1038/nrm3837 (2014).
- 3 Christie, J. M., Salomon, M., Nozue, K., Wada, M. & Briggs, W. R. LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 8779-8783 (1999).
- 4 Harper, S. M., Neil, L. C. & Gardner, K. H. Structural basis of a phototropin light switch. *Science* **301**, 1541-1544, doi:10.1126/science.1086810 (2003).
- 5 Wu, Y. I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461**, 104-108, doi:10.1038/nature08241 (2009).
- 6 Yoo, S. K. *et al.* Differential regulation of protrusion and polarity by PI3K during neutrophil motility in live zebrafish. *Developmental cell* **18**, 226-236, doi:10.1016/j.devcel.2009.11.015 (2010).
- 7 Wang, X., He, L., Wu, Y. I., Hahn, K. M. & Montell, D. J. Light-mediated activation reveals a key role for Rac in collective guidance of cell movement in vivo. *Nature cell biology* **12**, 591-597, doi:10.1038/ncb2061 (2010).
- 8 Motta-Mena, L. B. *et al.* An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature chemical biology* **10**, 196-202, doi:10.1038/nchembio.1430 (2014).

- 9 Kennedy, M. J. *et al.* Rapid blue-light-mediated induction of protein interactions in living cells. *Nature methods* **7**, 973-975, doi:10.1038/nmeth.1524 (2010).
- 10 Liu, H., Gomez, G., Lin, S., Lin, S. & Lin, C. Optogenetic control of transcription in zebrafish. *PloS one* **7**, e50738, doi:10.1371/journal.pone.0050738 (2012).
- 11 Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, doi:10.1038/nature12466 (2013).
- 12 Zhou, X. X., Chung, H. K., Lam, A. J. & Lin, M. Z. Optical control of protein activity by fluorescent protein domains. *Science* **338**, 810-814, doi:10.1126/science.1226854 (2012).
- 13 Ni, M., Tepperman, J. M. & Quail, P. H. Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature* **400**, 781-784, doi:10.1038/23500 (1999).
- 14 Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997-1001, doi:10.1038/nature08446 (2009).
- 15 Yang, X., Jost, A. P., Weiner, O. D. & Tang, C. A light-inducible organelle-targeting system for dynamically activating and inactivating signaling in budding yeast. *Molecular biology of the cell* **24**, 2419-2430, doi:10.1091/mbc.E13-03-0126 (2013).
- 16 Toettcher, J. E., Weiner, O. D. & Lim, W. A. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* **155**, 1422-1434, doi:10.1016/j.cell.2013.11.004 (2013).
- 17 Hoeller, O., Gong, D. & Weiner, O. D. How to understand and outwit adaptation. *Developmental cell* **28**, 607-616, doi:10.1016/j.devcel.2014.03.009 (2014).

- 18 Toettcher, J. E., Gong, D., Lim, W. A. & Weiner, O. D. Light-based feedback for controlling intracellular signaling dynamics. *Nature methods* **8**, 837-839, doi:10.1038/nmeth.1700 (2011).
- 19 Levskaya, A. *et al.* Synthetic biology: engineering *Escherichia coli* to see light. *Nature* **438**, 441-442, doi:10.1038/nature04405 (2005).
- 20 Muller, K. *et al.* Synthesis of phycocyanobilin in mammalian cells. *Chem Commun (Camb)* **49**, 8970-8972, doi:10.1039/c3cc45065a (2013).
- 21 Taylor, B. L. & Zhulin, I. B. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiology and molecular biology reviews : MMBR* **63**, 479-506 (1999).
- 22 Hughes, J. Phytochrome three-dimensional structures and functions. *Biochemical Society transactions* **38**, 710-716, doi:10.1042/BST0380710 (2010).
- 23 Ponting, C. P. & Aravind, L. PAS: a multifunctional domain family comes to light. *Current biology : CB* **7**, R674-677 (1997).
- 24 Krall, L. & Reed, J. W. The histidine kinase-related domain participates in phytochrome B function but is dispensable. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8169-8174, doi:10.1073/pnas.140520097 (2000).
- 25 Su, Y. S. & Lagarias, J. C. Light-independent phytochrome signaling mediated by dominant GAF domain tyrosine mutants of *Arabidopsis* phytochromes in transgenic plants. *The Plant cell* **19**, 2124-2139, doi:10.1105/tpc.107.051516 (2007).
- 26 Tawk, M. *et al.* A mirror-symmetric cell division that orchestrates neuroepithelial morphogenesis. *Nature* **446**, 797-800, doi:10.1038/nature05722 (2007).
- 27 Buckley, C. E. *et al.* Mirror-symmetric microtubule assembly and cell interactions drive lumen formation in the zebrafish neural rod. *The EMBO journal* **32**, 30-44, doi:10.1038/emboj.2012.305 (2013).

- 28 Girdler, G. C., Araya, C., Ren, X. & Clarke, J. D. Developmental time rather than local environment regulates the schedule of epithelial polarization in the zebrafish neural rod. *Neural development* **8**, 5, doi:10.1186/1749-8104-8-5 (2013).
- 29 Alexandre, P., Reugels, A. M., Barker, D., Blanc, E. & Clarke, J. D. Neurons derive from the more apical daughter in asymmetric divisions in the zebrafish neural tube. *Nature neuroscience* **13**, 673-679, doi:10.1038/nn.2547 (2010).
- 30 Chen, M., Tao, Y., Lim, J., Shaw, A. & Chory, J. Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals. *Current biology : CB* **15**, 637-642, doi:10.1016/j.cub.2005.02.028 (2005).
- 31 Matsushita, T., Mochizuki, N. & Nagatani, A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* **424**, 571-574, doi:10.1038/nature01837 (2003).
- 32 Khanna, R. *et al.* A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors. *The Plant cell* **16**, 3033-3044, doi:10.1105/tpc.104.025643 (2004).
- 33 Toettcher, J. E., Gong, D., Lim, W. A. & Weiner, O. D. Light control of plasma membrane recruitment using the Phy-PIF system. *Methods in enzymology* **497**, 409-423, doi:10.1016/B978-0-12-385075-1.00017-2 (2011).
- 34 Hong, E., Jayachandran, P. & Brewster, R. The polarity protein Pard3 is required for centrosome positioning during neurulation. *Developmental biology* **341**, 335-345, doi:10.1016/j.ydbio.2010.01.034 (2010).
- 35 Lutz, C. *et al.* Holographic photolysis of caged neurotransmitters. *Nature methods* **5**, 821-827, doi:10.1038/nmeth.1241 (2008).
- 36 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).

- 37 Westerfield, M. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio)*. 4th ed. edn, (Univ. of Oregon Press, 2000).
- 38 Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310, doi:10.1002/aja.1002030302 (1995).

CHAPTER THREE

An optogenetic gene expression system with rapid activation and deactivation kinetics.

“Well, you should probably stay as long as you can stand it.”

-Didier Stainier

Adapted from Laura B Motta-Mena, Anna Reade, Michael J Mallory, Spencer Glantz, Orion D Weiner, Kristen W Lynch & Kevin H Gardner (2014), An optogenetic gene expression system with rapid activation and deactivation kinetics. In *Nature Chemical Biology* volume 10, issue 3, pp 196–202. Nature Publishing Group. Terry Sheppard and Catherine Goodman, editors.

SUMMARY

Optogenetic gene expression systems can control transcription with spatial and temporal detail unequaled with traditional inducible promoter systems. However, current eukaryotic light-gated transcription systems are limited by toxicity, dynamic range, or slow activation/deactivation. Here we present an optogenetic gene expression system that addresses these shortcomings and demonstrate its broad utility. Our approach utilizes an engineered version of EL222, a bacterial Light-Oxygen-Voltage (LOV) protein that binds DNA when illuminated with blue light. The system has a large (>100-fold) dynamic range of protein expression, rapid activation (< 10 s) and deactivation kinetics (< 50 s), and a highly linear response to light. With this system, we achieve light-gated transcription in several mammalian cell lines and intact zebrafish embryos with minimal basal gene activation and toxicity. Our approach provides a powerful new tool for optogenetic control of gene expression in space and time.

INTRODUCTION

Essential for many research applications in biomedical research, inducible promoter systems have enabled the artificial control of gene transcription in eukaryotic cells (1,2). While these approaches have generated widely useful tools, their reliance on small molecule inducer compounds (e.g. doxycycline) limits their utility when precisely timed or localized induction is needed. Once inside a cell, chemical inducers are also limited by their rate of diffusion (slowing activation), difficult removal (slowing deactivation), and potential off-target effects on normal cellular function. In contrast, light is a rapid and nontoxic stimulus that naturally regulates many different cellular processes in diverse settings (3). To take advantage of these favorable properties, a variety of natural photosensitive proteins have recently been engineered into light-controlled transcriptional activators (4-11). Such newly developed light-dependent regulatory systems have the potential to enable the activation of gene expression with previously unattainable spatial and temporal control.

Nevertheless, these existing systems have significant drawbacks that limit their use in a wide range of experiments. Major disadvantages include toxicity (12), low levels of transcriptional activation (< 20-fold) (4-8,11), long deactivation times on the order of hours (10), use of exotic chromophores not found in vertebrate systems (4,5), potential interference of the active photoreceptor with endogenous signaling pathways (8), and the need for multiple protein components (4,6,7,9,11).

To address these limitations, we developed a new inducible promoter system that is based on the bacterial transcription factor EL222 (13), which consists only of the minimal elements needed for light-dependent transcriptional activation: a photosensor and a DNA-binding domain.

The photosensor is a flavin-binding LOV (14) domain, which controls a standard Helix-Turn-

Helix (HTH) DNA-binding domain in this protein. We previously demonstrated that the EL222 LOV domain binds to the HTH domain in the dark, blocking access to the HTH 4 α helix that is essential for dimerization and DNA binding (13,15). Blue light illumination (450 nm) triggers the photochemical formation of a protein/flavin adduct within the LOV domain, leading to conformational changes that disrupt the inhibitory LOV/HTH interactions and allow the protein to dimerize and bind DNA. These structural changes spontaneously reverse in the dark, rapidly reverting EL222 to its inactive state ($\tau \sim 11$ s at 37°C) (16). Within the native *Erythrobacter litoralis* HTCC2594 host, we observed the light-dependent activation of genes adjacent to EL222-binding sites identified by ChIP-Seq, strongly implicating this protein as a photosensitive transcription factor (15).

Our mechanistic understanding of EL222 paves the way for its use in a novel, broadly applicable, single component system for light-dependent gene activation useful in a broad array of settings. Here we report that a minimally-engineered variant of EL222 activates transcription in several different kinds of eukaryotic cells upon stimulation with moderate levels of blue light. With this system, we demonstrate over 200-fold upregulation of gene expression from an EL222-responsive luciferase reporter in 293T cells illuminated with levels of blue light compatible with robust cellular growth. In contrast, dark state and red light controls show < 2-fold changes, establishing minimal leakiness under non-inducing conditions. Our system has rapid activation (< 10 s) and deactivation kinetics (< 50 s), which compare favorably to the > 2 hr turn-off kinetics of a recently developed LOV-based transcriptional system (10). Furthermore, our system can achieve functional regulation of cellular processes, as we demonstrate for light-gated regulation of splicing in T-cells. Finally, we demonstrate that EL222 can be used for applications with either global or tissue-specific gene expression in zebrafish in a light-dependent manner and with minimal toxicity, further expanding the repertoire of this expression

system. In sum, these data highlight the broad utility of the EL222 system and its unique strengths as an optogenetic tool.

RESULTS

Development of an inducible gene expression system based on EL222

EL222, a small (222 residue) transcription factor from *Erythobacter litoralis* HTCC2594, serves as the foundation of our engineered expression system (Fig. 1a). While this protein functions as a photosensitive transcriptional activator in its native prokaryotic host (15), two N-terminal additions adapt it for eukaryotic applications: a VP16 transcriptional activation domain (AD) (17) and a canonical nuclear localization signal (NLS) sequence (Fig. 1a). Western blot analysis confirmed that the resulting VP-EL222 fusion protein was expressed in 293T cells and distributed between the nucleus and cytoplasm (Supplementary Results, Supplementary Fig. 1a).

To test the potency of the VP-EL222 transactivator system, we constructed a reporter vector containing the firefly luciferase (Fluc) gene under the control of five copies of the EL222-binding Clone 1-20 bp (C120) (15) sequence and a canonical TATA box (pC120-Fluc) (Fig. 1a). In transient transfection experiments, 293T cells expressing pVP-EL222 showed higher levels of luciferase when illuminated with pulsed blue light for 24 hr (20 s on, 60 s off; 8 W/m² at 465 nm) compared to transfected cells kept in the dark (Fig. 1b). The light-driven upregulation of luciferase levels required EL222, since transfection of a vector containing only the VP16 AD (empty vector) showed effectively no activation of the pC120-Fluc reporter in any condition. Importantly, the luciferase levels observed with VP-EL222 in the dark are quite similar to those measured from cells expressing VP16 AD alone, indicating that the VP-EL222 has minimal dark state activity as seen with *in vitro* DNA binding assays (13,15). To quantify the fold change (FC) in transcriptional activation of pC120-Fluc, we normalized the firefly luciferase values to an internal vector control (co-transfected with pVP-EL222 and pC120-Fluc) that used the constitutive CMV promoter to drive expression of Renilla luciferase (Rluc). After normalization,

we calculated that cells transfected with VP-EL222 showed 216-fold upregulation in luciferase relative to cells transfected with empty vector when illuminated with blue light. In contrast, cells kept in the dark showed only a two-fold change (Fig. 1c), as expected from the lack of an intact C120 EL222 binding site in the human genome (Supplementary Table 1) and the low affinity of EL222 for DNA in the dark (13,15). These factors lead to a net 108-fold increase in luciferase expression that is directly attributable to illumination.

For maximum utility, specificity in inducible promoter systems is essential – both with target DNA sequences and with input stimuli – to avoid off-target effects. For the former, we have previously investigated the in vitro specificity of EL222 with the C120 DNA sequence, showing that changes to single base pairs within the sequence could produce over ten-fold reductions in EL222/DNA affinity(15). Here we show that the VP-EL222 chimera retains its ability to bind to C120 DNA with high specificity in 293T cells, as measured by its ability to selectively drive luciferase expression from the pC120-Fluc reporter compared to two other control reporter constructs (Fig. 1d,e). The first control contains three copies of a lower affinity EL222 substrate called AN45(13) (p(AN45)3-Fluc); gel shift assays suggest that AN45 binds EL222 with an EC50 of approximately 30-fold more weakly than C120(13,15). A second control contained five copies of the GAL4-binding Upstream Activation Sequence (UAS) (p(UAS)5- Fluc). In both cases, luciferase expression levels were relatively low and invariant between dark and light conditions when co-transfected with pVP-EL222. Additional experiments confirmed that expression of pVP-EL222 in 293T cells did not significantly affect cell viability as compared to the pVP-empty control (Supplementary Fig. 1b), which suggests that neither VP- EL222 itself, VP-EL222 driven transcription nor our illumination protocol alter normal cellular function with any gross off-target activities. Lastly, we found that VP-EL222 is specifically activated by blue light, consistent with the flavin photochemistry used within LOV domains. This is supported by the inability of

continuous red light to significantly upregulate luciferase from the pC120-Fluc reporter above the level observed in the dark (Supplementary Fig. 1c,d).

Together these data demonstrate that the functional and photochemical properties inherent to EL222 are suitable for use in heterologous expression in cultured mammalian cells.

Characterization of VP-EL222 in stably transfected 293T cells

To reduce variability in our experimental results caused by differential transfection efficiencies, we stably expressed VP-EL222 in 293T cells and transfected only the pC120-FLuc plasmid (Fig. 1b,c). Consistent with our results obtained in transient transfections of pVP-EL222, luciferase levels were greatly higher in cells illuminated with blue light relative to cells left in the dark (162-fold enhanced dark- to light in cells transfected with pC120-Fluc with 20 s on, 60 s off; 12 hr illumination). From these results, we conclude that the VP-EL222 system can produce a high level of reporter gene expression in both transient and stable transfection experiments with minimal leakiness.

One of the major advantages of a light-switchable promoter system is the ability to change the level of gene expression by tuning the amount of light inducer. Using the VP-EL222 stable cell line and pC120-Fluc, we investigated how luciferase expression levels were affected by varying the illumination duty cycle within a constant 80 s period (Fig. 2a). As expected, we observed a dose-dependent increase in luciferase with increasing duration of illumination during each cycle. For periods of 5 s or greater, luciferase levels correlated linearly with illumination times, indicating that the VP-EL222 system can be used to tune luciferase expression by varying the illumination protocol. In addition to producing luciferase at levels that can be easily quantitated by enzymatic output, VP-EL222 is capable of expressing proteins at levels sufficient to be

detected by Western blot (luciferase, Fig. 2b) or fluorescence microscopy (mCherry, Fig. 2c). In both cases, signals were readily detectable only following blue light illumination.

Turning from steady-state measurements to the kinetics of gene expression, we examined the time course of protein and mRNA levels. To do so, we measured the luciferase activities of VP-EL222-expressing cells transfected with pC120-Fluc and incubated in the dark or illuminated with a 20 s on, 60 s off protocol for the indicated times (Fig. 2d). We observed that luciferase activity was rapidly induced, with 80-fold increases observed over dark state control after only 3 hr. Induction levels increased slightly to around 90-fold by 6 hr, and plateaued after 9 hr at 100-fold. Interestingly, luciferase activity began to drop slightly between 12-24 hr; nevertheless, after 24 hr luciferase expression was > 20-fold above background. While we detected an increase in luciferase activity over time in dark-incubated cells, indicative of leaky expression due to spontaneous low-level VP-EL222 activation or other factors, these levels were nevertheless significantly lower than from comparable illuminated cells (Fig. 2d, $p < 0.0001$). In parallel, we also measured the mRNA levels for luciferase over time using real-time qPCR (Fig. 2e). These analyses showed that luciferase mRNA reached peak levels after 3 hr of illumination (~ 22-fold increase over dark control). After 6 hr of illumination, luciferase mRNA levels leveled off at 20-fold over dark control, subsequently decreasing back down to approximately 2-fold up at 24 hr. We note that the plateau and subsequent decrease in luciferase transcript levels after prolonged illumination precedes the corresponding trend in protein levels (Fig. 2d). This implied a loss of VP-EL222 transcriptional activity with extended illumination, potentially attributable to a moderate decrease in cellular VP-EL222 protein levels (Fig. 2b).

This effect led us to investigate the ability of VP-EL222 to be reactivated after an initial round of illumination followed by a dark state recovery period to allow for synthesis of fresh transcription factor. For this experiment, we monitored luciferase activity in stable VP-EL222 cells

(transfected with the pC120-Fluc reporter) that were initially illuminated with blue light pulses (20 s on, 60 s off) for 3 hr, incubated in the dark for 21 hr, and then illuminated again with pulsed blue light for 3 hr (Fig. 2f). After the initial 3 hr activation, luciferase levels were approximately 75-fold above dark state levels. Once the light was removed, we observed a sustained increase in the absolute amount of luciferase activity in the illuminated cells; however, because their luciferase background levels also increased in the dark control cells, the dark-to-light fold induction dropped to near background levels over 21 hr (< 1.5-fold difference dark-to-light). After this dark state recovery period, cells were again illuminated for 3 hr with pulsing blue light with a concomitant increase in luciferase back up to 11-fold over dark state controls. Following another 21 hr dark recovery period, luciferase induction decreased to around 3.7-fold. Therefore, we conclude that light-triggered activation of VP-EL222 can generate multiple rounds of gene expression, enabling novel experiments utilizing transient increases of protein level.

Modeling the activation/deactivation kinetics of VP-EL222 in cells

Key aspects of these applications – such as the linear dose-response (Fig. 2a) and favorable kinetic properties (Fig. 2f) – are enabled by the intrinsically fast activation and deactivation kinetics of EL222 itself. In vitro measurements reveal LOV domains activate via microsecond-timescale adduct formation and millisecond timescale conformational changes(18,19), while deactivation by adduct cleavage takes seconds to hours depending on the LOV domain(16,20). Measuring comparable activation/deactivation parameters for VP-EL222 driven transcription within the cell is complicated by slow reporter mRNA and protein turnover (e.g. luciferase mRNA half-life: 3-5 hr(21,22), protein: 3-4 hr(23)), limiting the temporal resolution of experiments that simply yield this information. As an alternative, we developed a kinetic model that correlates gene expression to the times required for VP-EL222 to initiate transcription upon illumination (τ_{On}) or cease after returning to the dark (τ_{Off}) (Fig. 3a; Supplementary Notes 1, 2), and used

this to examine the dependence of luciferase expression on the length of dark/light cycles (Fig. 2a). We emphasize that these data – final end-point measurements of luciferase levels – depend on the cumulative effects of repeatedly activating VP-EL222 (> 500 80 s cycles in a 12 hr experiment), each time gaining a burst of mRNA at levels determined by second- timescale activation/deactivation events. Hence, this model (and accompanying least-squared error analysis; Fig. 3b), provides us a framework to determine how well various combinations of τ_{on} and τ_{off} values explain our experimental measurements.

With this approach, we were able to define ranges for the VP-EL222 τ_{on} and τ_{off} parameters that best modeled our data. For activation, the observed lag in luciferase expression at short illumination times implies an initial τ_{on} delay of approximately several seconds for VP- EL222 to activate, bind DNA and initiate transcription. Our model exhibits an expected inverse relationship between gene expression and τ_{on} , with values near 5 s recapitulating our data most accurately (Fig. 3c). On deactivation, our model recapitulates a direct link between τ_{off} and luciferase levels, with optimal τ_{off} values of approximately 30 s (Fig. 3d). A more complete grid search of all values of τ_{on} and τ_{off} (each varied independently between 1-100 s) reveals a range of values compatible with our data (Fig. 3e). This range includes activation times of approximately 3-5 s, compatible with single molecule measurements of transcriptional initiation rates (24) and the high level of VP-EL222 within 293T cells. The same analysis indicates τ_{off} values between 10-40 s; notably, the shortest of these delays is consistent with our in vitro measurements of EL222 adduct cleavage at this temperature ($\tau_{adduct} \sim 11$ s at 37°C(16)), suggesting limited effects of cellular properties on this critical step. Taken together, these data suggest that VP-EL222 functions with rapid on/off kinetics in cells, a key advantage over other comparable systems (10).

Light-inducible expression of the splicing factor CELF2 in T-cells

To test the utility of the VP-EL222 system in other cultured cell lines, we investigated its ability to drive the expression of a functionally-active protein within the T-cell derived Jurkat splicing line 1 (JSL1) cell line (25). The JSL1 cell line has been extensively used to study changes in alternative pre-mRNA splicing that occur in response to T-cell activation (25-27). One protein implicated in such control is CUGBP and ETR-2 like factor 2 (CELF2) (28), an RNA-binding protein with a known role in splicing regulation in JSL1 cells and thymocytes following cellular stimulation (29). Phorbol myristate acetate (PMA)-induced activation of JSL1 cells results in an increase in CELF2 expression, promoting CELF2 binding to regulatory sequences in target pre-mRNAs and affecting their processing (29). Indeed, at high levels CELF2 protein represses the inclusion of exon 6 in its own pre-mRNA (26,29,30), providing an assay to ascertain if the VP-EL222 system could drive CELF2 overexpression sufficiently to confer light-dependent control of pre-mRNA splicing.

To examine this, we created a JSL1 cell line stably integrated with both the pVP-EL222 activator and a FLAG-tagged CELF2 under the control of the EL222-specific C120 promoter (called VP-EL222/CELF2 cells). VP-EL222/CELF2 cells incubated in the dark showed nearly no expression of FLAG-CELF2 protein by immunoblot analysis (Fig. 4a). In contrast, VP-EL222/CELF2 cells exposed to pulsed blue light for 24 hr (20 s on, 60 s off) showed moderate FLAG-CELF2 expression, indicating functional light-triggered activation of VP-EL222 in JSL1 cells. We note that the levels of VP-EL222 protein itself decreased markedly (> 50%) with light exposure (Fig. 4a), more substantially than we observed with 293T cells (Fig. 2b). Nevertheless, the small amount of VP-EL222 is sufficient to produce amounts of CELF2 protein easily detectable by Western blot.

To address the functional significance of light-induced FLAG-CELF2 protein upregulation, we analyzed the splicing pattern of exon 6 of the endogenous CELF2 transcript by reverse transcription-PCR (RT-PCR) analysis, using primers that specifically recognize the endogenous transcript and not the transfected cDNA. We found that blue light induced a moderate, but statistically significant, decrease in the inclusion of the CELF2 exon 6 in VP-EL222/CELF2 cells ($p < 0.05$, Fig. 4b), demonstrating light-regulated alternative splicing. Importantly, blue light treatment had no discernible effects on exon 6 inclusion in wildtype JSL1 cells, underscoring the necessity of photosensitive VP-EL222 in this process. Previous studies determined that PMA stimulation of JSL1 cells leads to a two-fold increase in CELF2 protein levels (29), changing percent exon inclusion values by 20-30% (26). In our experiment, light increases the amount of FLAG-CELF2 protein five-fold over dark control; however, this does not detectably increase total CELF2 levels as seen by Western blot (data not shown). This observation is consistent with known autoregulatory mechanisms which maintain stable CELF2 expression (29). In addition, the fact that induction of FLAG-CELF2 only marginally increases overall CELF2 protein is consistent with the relative effect we observed in CELF2 exon 6 inclusion (80% to 75%).

VP-EL222 functions as a transcriptional activator in vivo

To transition from cell culture into intact multicellular organisms, we examined the capability of VP-EL222 to drive light-triggered gene expression in the zebrafish (*Danio rerio*). To do so, we microinjected the pC120-mCherry reporter plasmid into zebrafish embryos at the one-cell stage with or without 50 pg VP-EL222 mRNA (Fig. 5a). When embryos microinjected with VP-EL222 mRNA and pC120-mCherry were illuminated with constant blue light (14 mW/m²), mCherry fluorescence was readily detected after illuminating for only 5 hr (70% epiboly stage). After 22 hr illumination (24 hr post-fertilization [h.p.f.] stage), 100% of the 50 embryos analyzed had marked mCherry fluorescence. A z-stack series of an embryo at 70% epiboly showing expression of

mCherry better illustrates the robust level of protein expression induced by VP-EL222 (Supplementary Video 1). In contrast, when VP-EL222/pC120-mCherry embryos were left in the dark or did not receive VP-EL222 mRNA, no fluorescence could be detected (0/50 embryos in each set). These results show that VP-EL222 can rapidly and robustly activate transcription in developing zebrafish in a light-dependent manner.

We next examined whether the VP-EL222 system could provide light-inducible gene expression in a tissue-specific manner. For this, we constructed a dual-promoter plasmid encoding the VP-EL222 ORF controlled by the zebrafish cardiac *myl7* (myosin light polypeptide 7) promoter (31) and the mCherry ORF controlled by the EL222-specific C120 promoter. Zebrafish embryos that were microinjected with *pmyl7*-VP-EL222/C120-mCherry plasmid and illuminated with constant blue light, but not those kept in the dark, showed noticeable mCherry fluorescence that was specifically localized in the developing heart (Fig. 5b; Supplementary Video 2).

Finally, we titrated different amounts of VP-EL222 mRNA into zebrafish embryos to examine the toxicity of the protein (Fig. 5c). As a control, varying amounts of GFP mRNA were also injected. Our results show that a small amount of VP-EL222 (50 pg mRNA), was sufficient to elicit high levels of mCherry expression (Fig. 5a) with only minimal morphological effects (81% unaffected embryos for VP-EL222 vs. 93% for GFP) or toxicity (<10% severely affected or dead) when compared with comparable GFP controls. Increasing amounts of VP-EL222 mRNA affected more embryos up to 150 pg microinjected mRNA, but these effects remained constant (~60% embryos unaffected; [severely affected + dead] < 25%) above this level. These results suggest that even at high levels VP-EL222 protein is only moderately toxic to zebrafish, much less than that observed for a cryptochrome-based light-driven transcription system (12).

DISCUSSION

In this study, we describe a new inducible gene expression system that successfully confers high-level, blue light-sensitive transcriptional initiation to various human cell lines and zebrafish embryos. Based on the naturally-occurring EL222 transcription factor (13,15,16), this system utilizes the photochemical generation of a protein/flavin adduct to trigger conformational changes that control its function, as seen in other LOV domains and the broader family of PAS (Per-ARNT-Sim) environmental sensors (32,33). Here, these conformational changes produce a light-dependent breakage of intramolecular interactions that inhibit the ability of EL222 to dimerize and bind DNA. This evolved linkage between sensor and effector domain enables more substantial change in DNA binding affinity upon activation than observed with engineered photoactive DNA binding proteins (34,35).

This system has several benefits compared with alternative methods of photocontrolled gene expression. First, the LOV photosensory domain utilizes flavin chromophores that are widely distributed in many eukaryotic cells, eliminating the need to supply exogenous chromophores or precursors (4,5). This enables sensitive photochemistry which allows VP-EL222 to be triggered with modest intensity blue light ($8 \text{ W/m}^2 = 0.008 \text{ mW/mm}^2$), less than that required to activate light-regulated channelrhodopsins used in other optogenetic applications (e.g. 5 mW/mm^2) (36). Second, VP-EL222 relies on only a single 33 kDa protein with a directly regulated DNA binding step, simplifying genetic manipulation and tuning compared to light-dependent two-hybrid systems (4,6,7,9,11) or those which tie into existing cellular signaling pathways upstream of transcriptional initiation (8). Third, VP-EL222 shows low levels of overt toxicity and basal transcriptional activity in cell lines and zebrafish, which may be related to the absence of intact C120 targets within these organisms as indicated by genomic BLAST searches (Supporting Table 1). Finally, the VP-EL222 activates and resets within short periods of time (Fig. 3e), both

facilitating a variety of kinetic experiments and providing excellent and predictable dose-response behavior.

These hallmarks are best exemplified by comparing VP-EL222 to two other single-component light-inducible systems that have been reported for regulating gene expression in mammalian cells (8,10). One of these couples the light-triggered activation of an exogenously expressed melanopsin receptor to gene expression via a GPCR signaling cascade that activates the NFAT sequence-specific transcription factor via changes in calcium levels (8). Although this ingenious system utilizes intracellular components common to most cells, its level of transcriptional activation is low (~ 20-fold) and may have potential off-target effects via its reliance on altering levels of the Ca²⁺ second messenger. A more directly-regulated system is based on an extensively-modified version of the Vivid LOV protein, which can strongly activate gene transcription (> 200-fold) (10). However, because this engineered form of Vivid has an especially long-lived photoactive state (half-life = 2 hr (10)), the system has non-ideal dose-response behavior and deactivation kinetics.

This last feature – kinetics of induction and deactivation for photoactivated transcriptional regulation – is germane for VP-EL222 given the rapid on/off kinetics provided by its quick activation/deactivation (Fig. 3) and the straightforward regulation of DNA binding.

While speed has been viewed as essential for controlling intracellular localization and cytoskeletal remodeling (37,38) known to act on the timescale of seconds, we suggest the increasing appreciation for the speed of native transcriptional control (24) indicates that comparably fast artificial control may be useful as well. Rapid on/off kinetics ensure a maximally linear response between input light levels and output gene expression, and further enable novel classes of experiments which utilized kinetically-pulsed waves of transcription that are difficult to achieve with other artificially-inducible systems. Our data indicate that VP-EL222 driven

transcription is quickly activated and deactivated, in good agreement with comparable transcription factors in cells (24) or in vitro for EL222 itself (16).

We close by noting that this linkage between in vitro and cell-based data is critical for future engineering of this and other photosensitive systems. It is clear that various applications will have different requirements for artificial regulators, such as maximum kinetic resolution or sensitivity. Mechanism-based approaches – such as EL222 variants we have engineered with shorter- and longer-lived photoactive states (16) – are essential to developing the reagents that are optimal for each of these applications. Coupled with the potential to adapt other types of protein domains to EL222 to recruit epigenetic or other cellular machinery to specific DNA sites in a light-dependent manner, we anticipate that this protein will enable an even wider range of applications in the future.

ACKNOWLEDGEMENTS

This work was funded by grants from the US National Institutes of Health (R01 GM081875 to K.H.G.; R01 GM103383 to K.W.L.; R01 GM096164 to O.D.W.), DARPA (Living Foundries HR0011-12-C-0068 to Brian Chow [Univ. Pennsylvania], supporting S.G.) and the Robert A. Welch Foundation (I-1424 to K.H.G.). K.H.G. is the Virginia Lazenby O'Hara Chair in Biochemistry and W.W. Caruth Scholar in Biomedical Research. A.R. was supported by a National Science Foundation Graduate Research Fellowship. We thank S.L. McKnight and P.R. Potts (UT Southwestern) for generously providing constructs.

AUTHOR CONTRIBUTIONS

L.B.M.M., A.R., O.D.W., K.W.L. and K.H.G. conceived and designed the experiments. L.B.M.M., A.R., M.J.M. performed the experiments. L.B.M.M., A.R., M.J.M., O.D.W., K.W.L. and K.H.G. analyzed the data, with S.G. and K.H.G. generating the kinetic model (Fig. 3). L.B.M.M. and K.H.G. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests; details are provided in the online version of this paper.

METHODS

Vector construction.

DNA containing residues 14-222 of EL222 (13) was cloned into the pVP16 (Clontech) vector to obtain pVP-EL222. The pVP-EL222-puro plasmid was created by PCR amplification of VP16-EL222 ORF and cloning into pIRESpuro (Clontech). Five tandem copies of the 20-bp Clone-1 (C120) sequence (15) were chemically synthesized (GeneArt) and inserted into the pGL4.23[luc2/minP] (Promega) to make pGL4-C120-Fluc. pcDNA-C120-Fluc was created by PCR amplification of five copies of C120 sequence together with the firefly luciferase ORF and cloning into pcDNA3.1+ (Invitrogen). pGL4-C120-mCherry was created by PCR amplification of mCherry ORF and sub-cloning into pGL4-C120-Fluc to replace the luciferase ORF.

Cell culture, transfections, light induction, and cell viability assay.

293T (ATCC) and JSL1 cells were cultured at 37°C in 5% CO₂ in DMEM (Thermo Scientific) and RPMI (Gibco) respectively. Both did not contain phenol red and were supplemented with 5-10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin solution.

To make the 293T VP-EL222 stable cell line, cells were transfected with pVP-EL222- puro plasmid and allowed to recover for 3 days. Afterwards, cells were serially diluted into medium containing 2 µg/ml of puromycin (Gibco) and grown for 1-2 weeks. Puromycin- resistant clones were expanded and analyzed for VP-EL222 expression by Western blotting. To make a JSL1 VP-EL222 stable cell line cells were diluted in medium containing 0.2 mg/ml zeocin (Gibco). The JSL1 VP-EL222/C120-FLAG-CELF2 double stable cell line was made by transfecting the JSL1 VP-EL222 stable cell line with pC120-FLAG-CELF2 plasmid. After transfection, cells were diluted in medium containing 0.6 mg/ml G418 sulfate (Gibco). Drug- resistant clones were screened by RT-PCR for genomic integration of the pC120-FLAG-CELF2.

For transient transfections, 293Ts were plated at 2×10^5 cells/well in 24-well plates and transfected with 0.5 μg pGL4-C120-Fluc DNA the same day using lipofectamine (Invitrogen). pVP-EL222 or pVP-empty, pGL4-C120-Fluc, and pGL4.75[hRluc/CMV] (Promega) constructs were transfected using 5 : 1 : 0.04 ratio, respectively. 24 hr post-transfection, a blue LED panel (465 nm, 2501BU, LED Wholesalers) was placed above the plate. The intensity of the light received by cells was measured to be $39.7 \text{ mol s}^{-1} \text{ m}^{-2}$ (equivalent to 8 W/m^2) using the LI-190 Quantum Sensor and LI-250A light meter (LI-COR Biosciences). The LED panel was connected to an electronic intervalometer (Model 451, GraLab) and set to a cycle of 20 s on and 60 s off. The control plate was kept in the dark throughout the experiment. 48 hr post-transfection, firefly and Renilla luciferase activities were measured using the Dual-Glo luciferase assay kit (Promega) according to the manufacturer's instructions. The following equation was used to determine the normalized fold change in transcription in the dark and with light between cells expressing pVP-EL222 and pVP-empty: $\text{FC} = (\text{Fluc/Rluc})_{\text{VP-EL222}}/(\text{Fluc/Rluc})_{\text{empty}}$.

For transfection of 293T VP-EL222 stable line, cells were plated one day before transfection at 1×10^5 cells/ml in 24-well plates. The next day, 0.8 μg of pcDNA-C120-Fluc DNA were transfected using lipofectamine. Immediately afterwards, the cells were illuminated using LED panel (20 s on, 60 s off) for 12 hr, unless otherwise indicated in figure legend. For experiments done with JSL1 cells, wild type or VP-EL222/C120-FLAG-CELF2 stable cells were plated at 6×10^5 cells/well in 6-well plates. The next day, cells were illuminated (20 s on, 60 s off) for 24 hr and subsequently harvested for Western and RT-PCR analysis. Viability of 293T cells was evaluated using the Cell Titer Blue assay (Promega) according to the manufacturer's instructions.

Nuclear/cytoplasmic extract isolation and Western blotting.

293T cells (5.4×10^6 cells total) were grown in 10 cm dishes and were transfected with pVP-EL222 or left untreated. 48 hr post- transfection, nuclear and cytoplasmic extracts were purified using the following protocol. Cells were harvested and pelleted by centrifugation for 5 min at $3,220 \times g$. The pellet was resuspended in 1 ml of ice-cold PBS, centrifuged for 5 min at $100 \times g$, and resuspended in 1 ml ice-cold Buffer A (10 mM Tris-Cl pH 7.5, 1.5 mM $MgCl_2$, 10 mM KCl). The cell suspension was incubated on ice for 5 min, in a dry ice/ethanol bath for 5 min, and in a $37^\circ C$ water bath for 5 min. This incubation series was repeated 2 more times. Afterwards, the cells were centrifuged at $4^\circ C$ for 15 min at $15,900 \times g$. The supernatant (nuclear extract) was moved to a new tube and the pellet (cytoplasmic extract) was resuspended in 1 ml Buffer A.

For Western blotting, equal protein amounts of total cell lysates were separated on a 10% Mini-PROTEAN TGX precast gel (BioRad) and then transferred to a polyvinylidene fluoride (PVDF) membrane (Amersham). The protein signal was detected using the Pierce ECL Western Blotting Substrate (Thermo Scientific) according to the manufacturer's instructions. The antibodies used were as follows: anti-VP16 AD (ab4808, Abcam), anti-luciferase (L0159, Sigma), anti- β -actin (A5441, Sigma), anti-ARNT (sc-17811, Santa Cruz Biotechnology), anti- hnRNP L (ab6106, Abcam), anti-FLAG (2368, Cell Signaling).

Pre-mRNA splicing analysis.

RNA isolation and analysis of pre-mRNA splicing by RT-PCR were done as described previously for JSL1 cells (25,39). Primers for the analysis of the endogenous CELF2 gene are as follows: forward primer in the 5'-UTR region, 5'-TCTGCTCGACAGCAGCACGCAGTG-3'; reverse primer downstream of variable exon 6, 5'- CAGGTGGCAGTGTTGAGCTGC-3'.

Quantitative PCR (qPCR) analysis.

Total RNA was isolated from transfected 293T VP- EL222 cells using an RNeasy kit (Qiagen) following the manufacturer's instructions. 4 µg of total RNA were treated with DNase I (NEB) to remove genomic DNA. 1 µg of each treated RNA sample was reverse transcribed using iScript cDNA synthesis kit (BioRad). qPCR was performed on a Applied Biosystems 7300 real-time PCR system using TaqMan Fast Advanced Master Mix and TaqMan Gene Expression Assays for Luciferase and GAPDH (Applied Biosystems) with 100 ng of cDNA as template. Samples were run in triplicate and the average C_T was calculated. The average luciferase C_T value for each sample was normalized to the corresponding average GAPDH C_T value and the comparative C_T ($\Delta\Delta C_T$) method (40) was used to calculate the fold change in luciferase between dark and light-treated samples.

Live cell fluorescence microscopy.

Wild type or VP-EL222 stable 293T cells were transfected with pC120-mCherry plasmid, immediately after the cells were illuminated for 24 hr (20 s on, 60 s off) or left in the dark. Cells were examined on a Nikon Eclipse TS100 epifluorescence microscope running NIS Elements and equipped with Photometrics Coolsnap HQ camera. Images were taken with a 10x/0.25 NA Achromat Ph1 objective and mCherry fluorescence was imaged with a G2A filter. Image processing and analyses were performed using ImageJ software (41).

Zebrafish strains.

Adult zebrafish, both TL and AB wild-type strains, were maintained under standard laboratory conditions (42).

Transient expression and light induction in zebrafish.

Expression plasmid pCS2-VP-EL222 was created by PCR amplification of VP-EL222 ORF and then cloned into pCS2+ (gift from Stephanie Woo). Capped messenger RNA was synthesized using the mMMESSAGE mMACHINE SP6 kit (Ambion). 50 pg of VP-EL222 mRNA and/or 20 pg of pGL4-C120- mCherry plasmid DNA were injected at the one-cell stage. For heart-specific expression of VP- EL222 and light-induced induction of mCherry reporter, VP-EL222 ORF and C120-mCherry promoter and ORF were PCR amplified and cloned into pminiTol2-myl7 (43,44) to create the dual promoter construct, pminiTol2- myl7-VP-EL222-C120-mCherry. 20 pg of pminiTol2-myl7-VP- EL222-C120-mCherry plasmid DNA along with 50 pg of Tol2 transposase mRNA were injected at the one-cell stage.

Constant blue light was applied at approximately 2 h.p.f. with a blue LED panel (465 nm, 2501BU, LED Wholesalers). Actual power of light received by embryos was measured to be ~1 mW using a PM100D Laser Power and Energy Meter Console (Thorlabs). Dark controls were placed in a lightproof box in the same 29°C incubator as light-treated samples. The light was turned off at 24 h.p.f. for imaging and analysis of embryos. For heart specific induction of mCherry, constant light was applied from 10 h.p.f. to 24 h.p.f.

Microscopy and image processing of zebrafish embryos.

Fluorescent and brightfield images at 70% epiboly were taken on a Digital Scanned Laser Light Sheet Microscope (45). Embryos were mounted in a 1.5% low-melt agarose cylinder using 3 mm O.D./2 mm I.D. FEP tubing (Bola). Z stacks of 2.58 μ m intervals were taken with a 10x/0.5 NA objective. mCherry fluorescence was imaged with 561 nm laser line and a 561LP filter. Brightfield images were acquired using room light. Fluorescent and brightfield images at 24 h.p.f. were taken on a Nikon Eclipse Ti microscope running NIS Elements and equipped with a

Lambda XL Broad Spectrum Light Source (Sutter) and an iXon DU-897 EMCCD camera (Andor).

Dechorionated embryos were embedded in 1.5% low-melt agarose within glass-bottom Petri dishes (MatTek Corporation). Whole embryo images were taken with a 4x/0.13 NA Plan-Fluor objective and heart-specific images were taken with a 20x/0.75 NA Plan Apo objective. Standard filter settings were applied. Image processing and analysis was performed using ImageJ software (41). For the 70% epiboly images, maximum intensity projections of the fluorescent Z-stack were performed and merged with a corresponding brightfield image. For whole embryo and heart-specific 24 h.p.f. images, mCherry and brightfield channels were merged.

Toxicity curves in zebrafish.

50, 100, 150, 200 or 300 pg of VP-EL222 or GFP (control) mRNA per embryo were injected at the one- to two-cell stage. Unfertilized embryos were removed on day 0, and phenotypes of each group were scored alongside uninjected control embryos from the same clutch on day 1 after manual dechoriation. Each group had at least $n = 100$ embryos. Embryos were scored as follows: normal - unaffected, wild-type phenotype; mildly deformed - presence of a slightly curved tail and/or mild edema; severely deformed - presence of smaller heads, major curve or kink in tail, and/or severe edema.

FIGURE 1

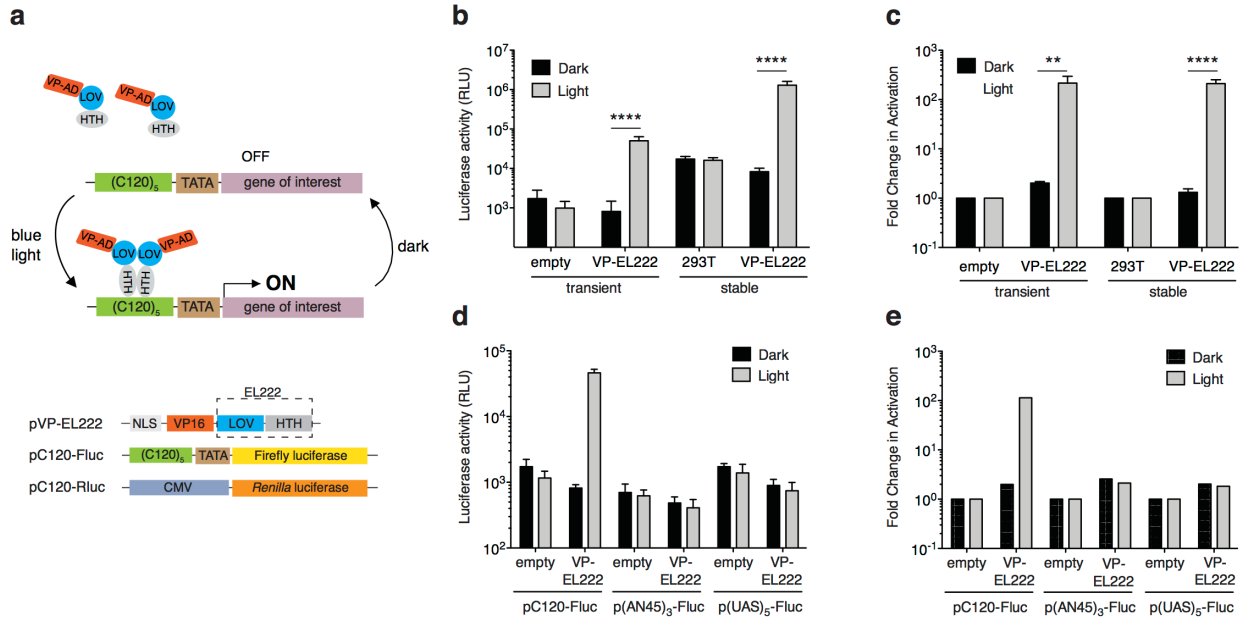


Figure 1. Model for the EL222-based light-inducible gene expression system.

(a) Top, the VP-EL222 protein consists of the transcriptional activation domain from the VP16 protein fused to EL222. In the dark VP-EL222 is unable to bind DNA; however, exposure to blue light triggers a photochemical reaction between the LOV domain and its flavin chromophore, which activates the attached HTH domain to bind DNA and turn on gene transcription. Bottom, schematic representations of the DNA constructs used in this work. **(b, c)** For transient transfections, cells were transfected with either empty vector (containing only VP16 AD) or pVP-EL222 and pC120-Fluc plus an internal control plasmid pCMV-*Renilla*. For stable transfections, cells stably expressing VP-EL222 protein and wildtype 293T cells were transiently transfected with pC120-Fluc and pCMV-*Renilla*. Cells were kept in the dark or illuminated with blue light pulses (465 nm; 20 s on, 60 s off; 8 W/m²) for 24 hr (transient) or 12 hr (stable). Luciferase levels (b) were normalized to *Renilla* luciferase levels to calculate the fold change (FC) (c) in transcription in control versus VP-EL222 cells under dark and light conditions. $FC_{\text{dark or light}} = ((\text{Firefly}/\text{Renilla})_{\text{VP-EL222}} / (\text{Firefly}/\text{Renilla})_{\text{empty or 293T}})$. (n = 3, triplicates for each condition). **(d, e)** Cells were transiently transfected with either empty vector or pVP-EL222 and one of three reporter constructs, pC120-Fluc, p(AN45)₃-Fluc, or p(UAS)₅-Fluc and kept in the dark or illuminated with blue light pulses (465 nm; 20 s on, 60 s off; 8 W/m²) for 24 hr. Levels of luciferase activity (d) and the fold change in expression (e) are shown (n = 1, triplicates for each condition). ** P < 0.01; **** P < 0.0001 using two-tailed Student's t-test. All data are represented as mean ± s.d.

FIGURE 2

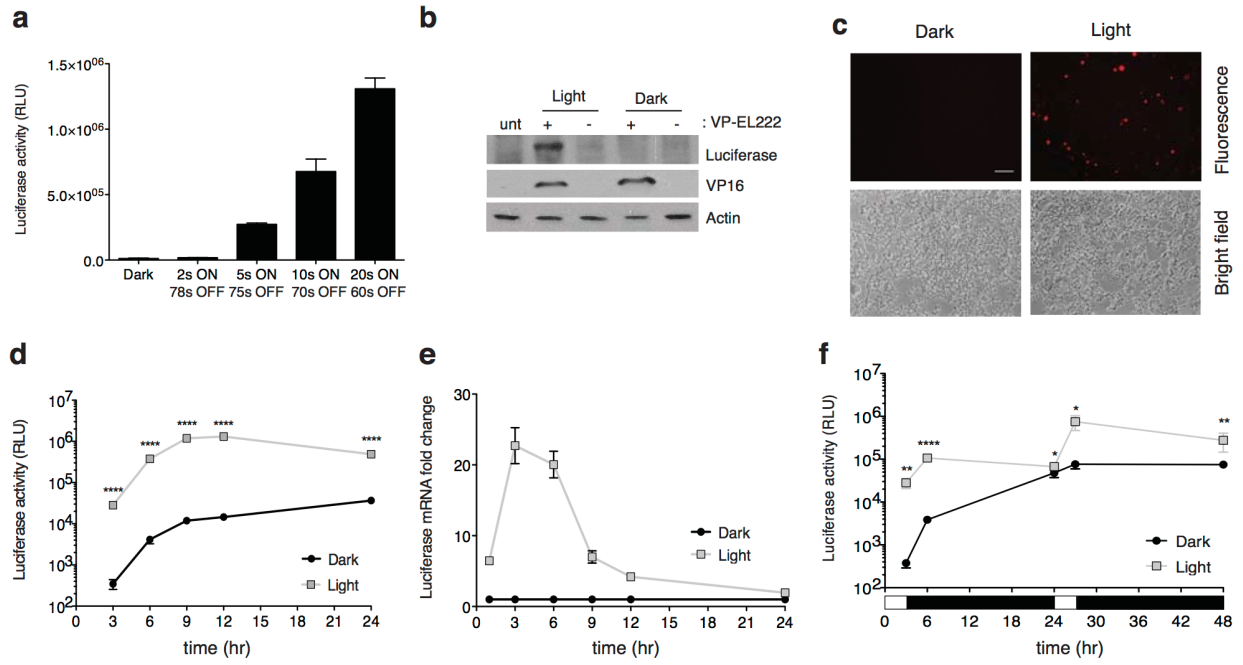


Figure 2. Dose-dependent activation and photoreversibility of gene expression by VP-EL222.

(a-f) 293T cells that stably express VP-EL222 protein (VP-EL222 cells) were transiently transfected with pC120-Fluc or pC120-mcherry and illuminated with blue light pulses (20 s on, 60 s off) for 12 hr, unless otherwise indicated. **(a)** Luciferase activity levels measured in VP-EL222 cells treated with blue light pulses of varying duration for 12 hr or kept in the dark. **(b)** Western blot analysis of luciferase expression in VP-EL222 cells (+) or wild type 293T cells (-) after illuminating with blue light or kept in the dark. Untransfected control (unt). **(c)** Representative images of VP-EL222 cells transiently transfected with pC120-mCherry. Cells were left in the dark or illuminated with blue light pulses for 24 hr. Scale bar, 100 μ m. **(d)** Time course analysis of luciferase expression in VP-EL222 cells illuminated with blue light or kept in the dark for the indicated times. **(e)** Luciferase mRNA transcript levels were quantified by qPCR from VP-EL222 cells transfected with pC120-Fluc and treated with blue light for indicated times or kept in the dark. **(f)** VP-EL222 cells were illuminated with pulsing blue light for two separate 3 hr periods (white box) each separated by a 21 hr dark period (black box). Controls were kept in the dark for the entire experiment. Luciferase activity was measured at the indicated timepoints. (a, d, e, f, n = 1, triplicates for each condition). *P < 0.05; ** P < 0.01; **** P < 0.0001 using two-tailed Student's t-test. All data are represented as mean \pm s.d.

FIGURE 3

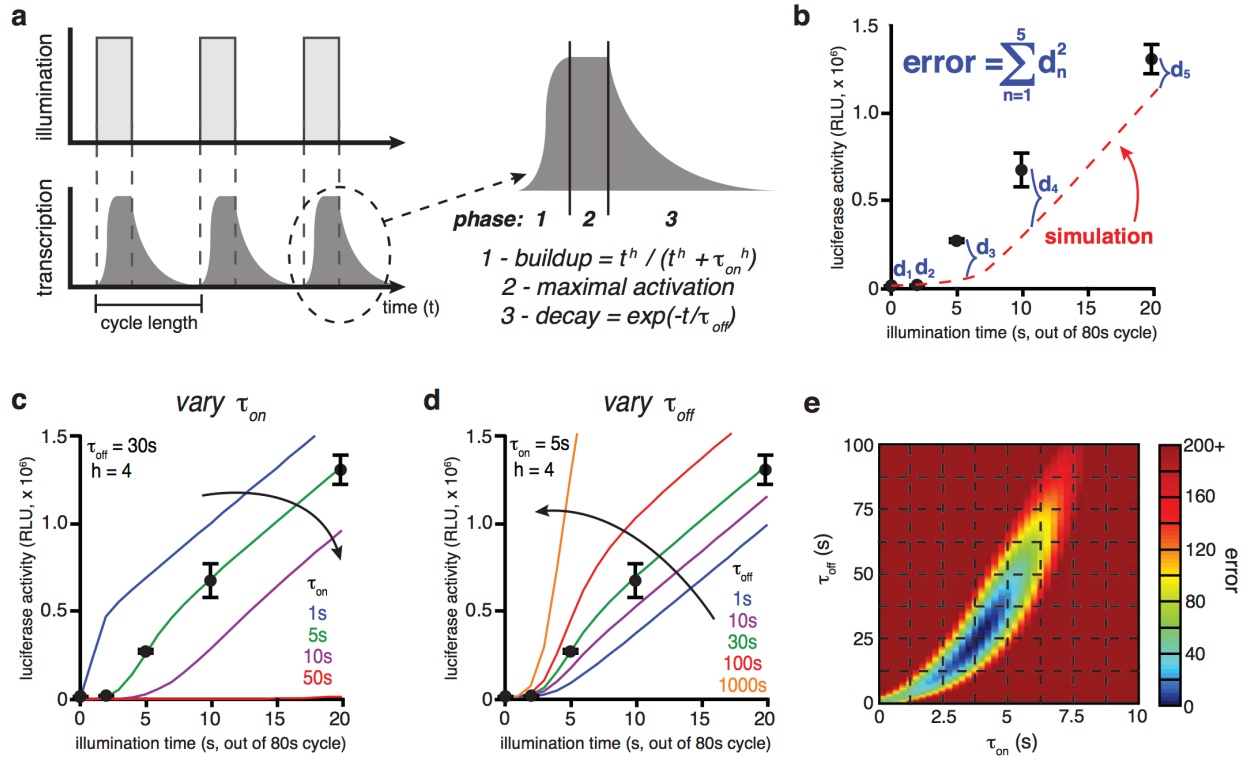


Figure 3. Kinetic modeling of VP-EL222 activation.

(a) Summary of the model used to describe VP-EL222 transcriptional activation, including three phases of transcriptional activity. Additional details are provided in Supplementary Methods. **(b)** Data used for kinetic modeling (replotted showing mean \pm s.d. from Fig. 2a) along with definition of least squared error function. **(c)** Effect of varying τ_{on} on transcriptional activity. For a given τ_{off} of 30 s (estimated from in vitro measurements of EL222 deactivation (16)) and Hill coefficient of 4, average steady- state transcriptional activities at τ_{on} values between 1-50 s were calculated using our model. Best agreement with experimental data were obtained with $\tau_{on} \sim 5$ s. **(d)** Effect of varying τ_{off} on transcriptional activity. For a given τ_{on} of 5 s (based on panel c), average transcriptional activities at τ_{off} values between 1-1000 s were calculated. Best agreement with experimental error were obtained with $\tau_{off} \sim 30$ s. **(e)** Grid search of τ_{on} , τ_{off} values (independently iterated for $\tau_{on} < 100$ s, $\tau_{off} < 100$ s), using the model, data and error function described above. Heatmap indicates value of error function; only the region with error function < 200 (τ_{on} 1-10s, τ_{off} =1- 100s) is shown here.

FIGURE 4

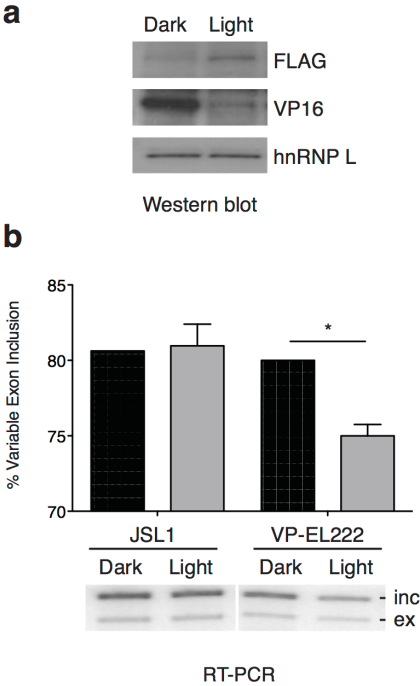


Figure 4. Light-regulated gene expression of the splicing factor CELF2 using VP-EL222 in the T-cell derived JSL1 cell line.

(a) Western blot analysis of protein lysates derived from JSL1 cells stably expressing both VP-EL222 and pC120-FLAG-CELF2 (JSL1 VP-EL222/CELF2 cells) were illuminated for 24 hr with either pulsing blue light (20 s on, 60 s off) or kept in the dark. An hnRNP L antibody was used as a loading control. **(b)** Percent inclusion of exon 6 in endogenous CELF2 gene for wild type and VP-EL222/CELF2 JSL1 cells (VP-EL222) that were illuminated with pulsing blue light (20 s on, 60 s off) for 24 hr or kept in the dark. The calculated exon inclusion in illuminated samples was normalized relative to that measured in dark samples for each cell line. Below, representative RT-PCR gel showing the drop in exon inclusion. (n = 2 for each condition). *P < 0.05 using two-tailed Student's t-test. Data are represented as mean \pm s.e.m.

FIGURE 5

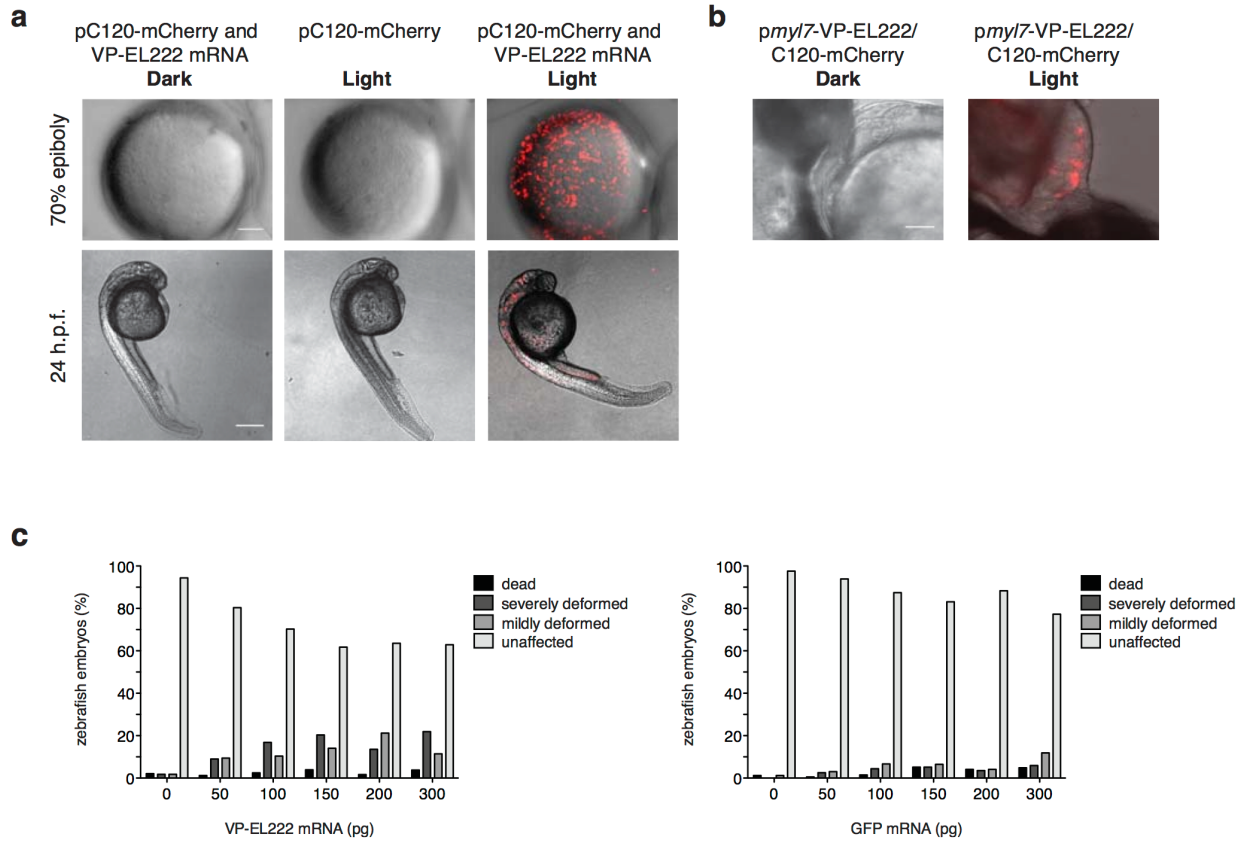
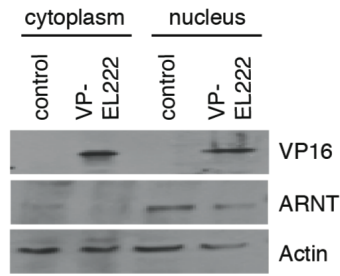


Figure 5. VP-EL222 robustly activates reporter gene expression in the developing zebrafish embryo in a light-dependent manner.

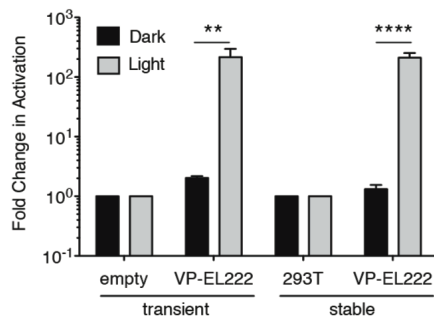
(a) Representative images of zebrafish embryos after injection with both VP-EL222 mRNA and pC120-mCherry DNA or pC120- mCherry DNA alone. The embryos were kept in the dark or illuminated with constant blue light for 5 hr (70% epiboly stage) or 22 hr (24 h.p.f. stage) beginning at 2 h.p.f. (n = 50 per condition). Each image is a maximum intensity projection of a fluorescent Z-stack merged with its corresponding brightfield image. Top scale bar, 100 μm ; bottom scale bar, 300 μm . **(b)** Visualization of mCherry fluorescence in the zebrafish heart in illuminated (as described above) versus control embryos after injection with a dual promoter vector that contains VP-EL222 and mCherry under the control of the cardiomyocyte-specific myl7 promoter and the EL222-specific C120 promoter; respectively. mCherry and brightfield channels were merged. Scale bar, 50 μm . **(c)** Dose-response curve showing the effect of VP-EL222 mRNA (left) expression on zebrafish development as compared to expression of a control GFP mRNA (right). Embryos were injected at the one to two-cell stage and scored at 24 h.p.f. (at least n = 100 per condition).

SUPPLEMENTARY FIGURE 1

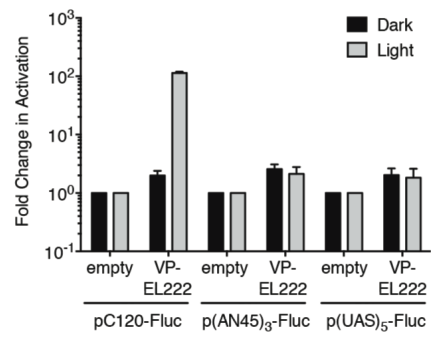
a



b



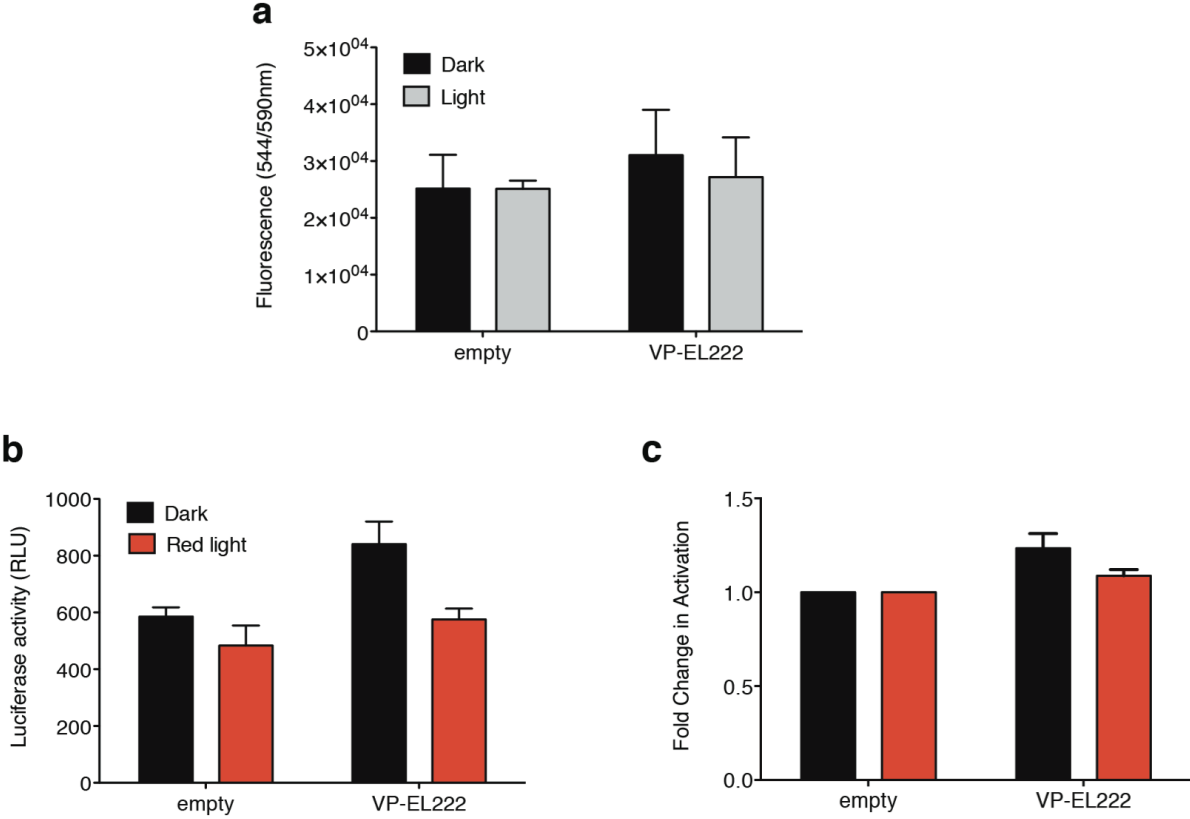
c



Supplementary Figure 1. VP-EL222 is expressed in 293T cells and can function to activate luciferase expression in response to light.

(a) Western blot analysis of VP-EL222 expression in cytoplasmic and nuclear fractions made from 293T cells transiently transfected with pVP- EL222. An ARNT antibody was used as a marker for nuclear localization. **(b)** The fold change (FC) in transcription in control versus VP-EL222 cells under dark and light conditions was calculated by normalizing the luciferase levels (Fig. 1) to Renilla luciferase levels [FC_{dark or light} = (Firefly/Renilla)_{VP-EL222} / (Firefly/Renilla)_{empty or 293T}]. (n = 3 independent experiments, each performed in triplicate per condition).

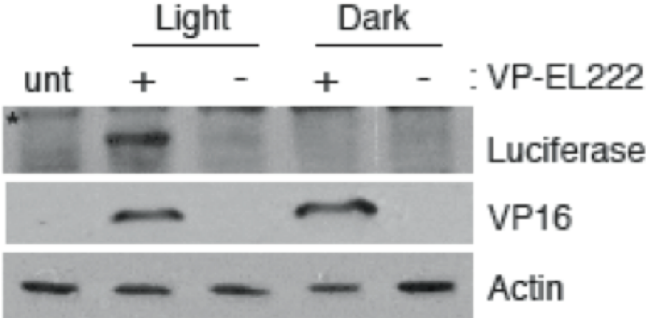
SUPPLEMENTARY FIGURE 2



Supplementary Figure 2. VP-EL222 is not activated by red light and that illumination with blue light has no negative effect on cell viability.

(a) Results of cell viability assays using CellTiter-Blue reagent. 293T cells were transiently transfected with either empty vector or pVP- EL222 and pC120-Fluc and kept in the dark or illuminated with pulsing blue light (20 s on, 60 s off) for 24 hr. Subsequently, the cells were incubated with CellTiter-Blue for 2 hr and the resultant fluorescence ($544_{Ex}/590_{Em}$) was recorded (n = 2 independent experiments, each performed with six replicates per condition). **(b, c)** 293T cells were transiently transfected with either empty vector or pVP-EL222 and pC120-Fluc, at 24 hr post-transfection cells were kept in the dark or illuminated with continuous red light for 24 hr. Afterwards, luciferase levels were measured **(c)** and then used to calculate the fold change in gene expression **(d)** in cells expressing VP-EL222 versus empty vector control (n = 1 independent experiment performed with three replicates per condition). All data are represented as mean \pm s.d.

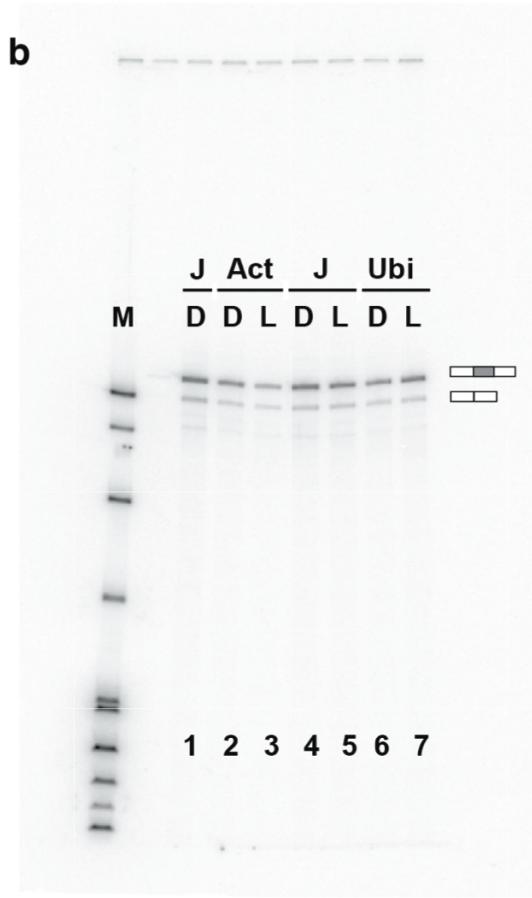
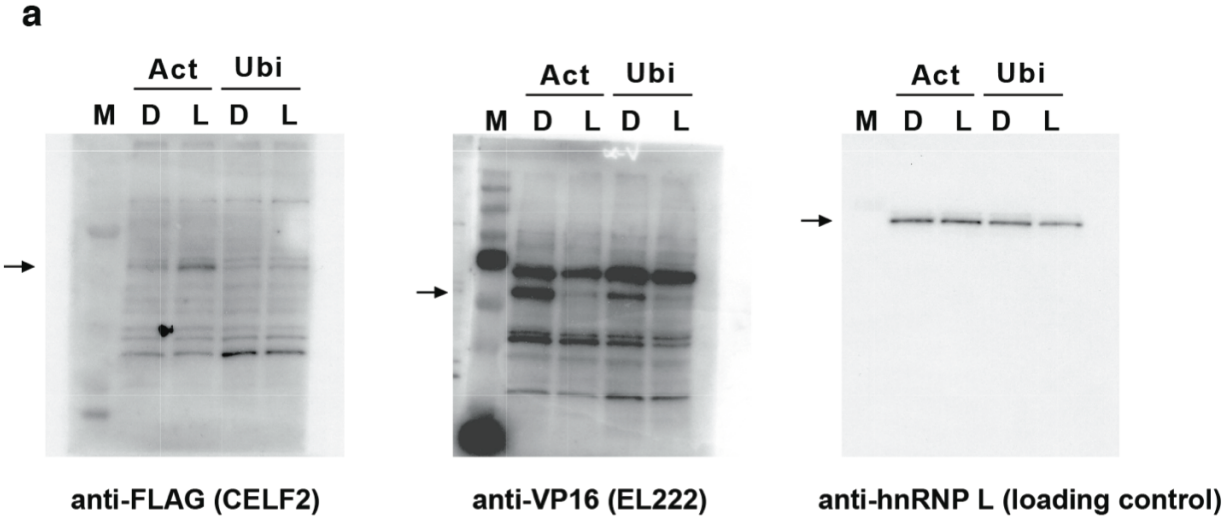
SUPPLEMENTARY FIGURE 3



Supplementary Figure 3. Light-triggered expression of luciferase protein by VP-EL222.

Western blot analysis of luciferase expression in VP-EL222 cells (+) or wild-type 293T cells (-) after illuminating with blue light or kept in the dark. Untransfected control (unt).

SUPPLEMENTARY FIGURE 4

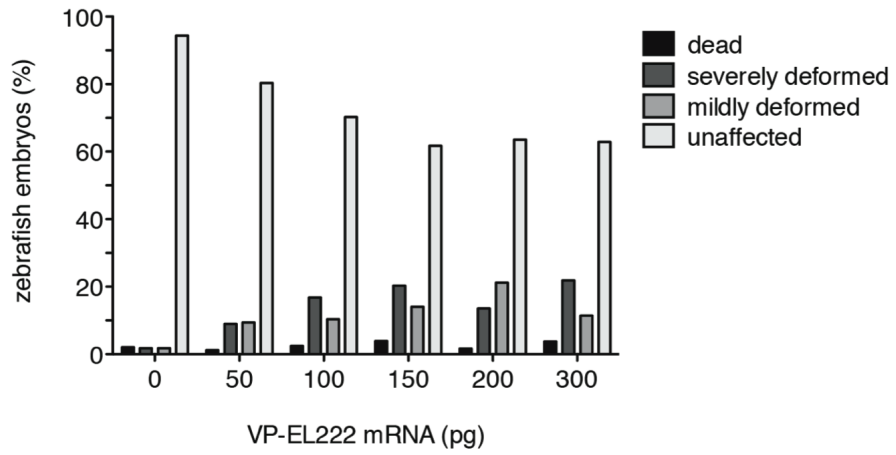


Supplementary Figure 4. Full Western blots and denaturing polyacrylamide gel of RT-PCR samples.

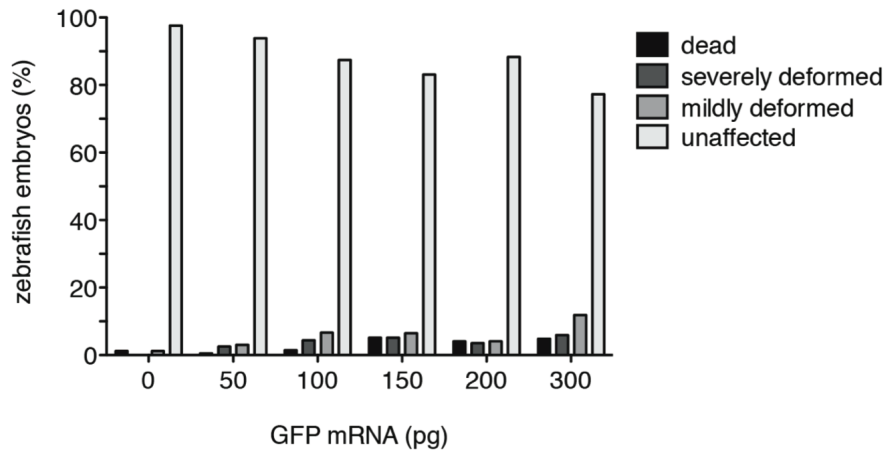
(a) Blots were probed with antibodies against FLAG, VP16, and hnRNP L. EL222-VP16 was driven from an Actin promoter (Act; used in Fig. 4a) or Ubiquitin promoter (Ubi, not used in subsequent studies). Arrow indicates band of interest. **(b)** RT-PCR to quantify splicing of endogenous CELF2. Lane 1, wild-type JSL1 cells. Lanes 2-7, JSL1 cells containing VP16-EL222 driven from Actin promoter (Act). Lanes 4-5, wild-type JSL1 cells. Lanes 6-7, JSL1 cells containing VP16-EL222 driven from Ubiquitin promoter (Ubi). Lanes 2-5 are shown in Fig. 4b (lanes 2-3, right side of panel; lanes 4-5, left side of panel). Dark conditions (D); Light conditions (L); Molecular weight marker (M).

SUPPLEMENTARY FIGURE 5

a



b



Supplementary Figure 5. VP-EL222 is only moderately toxic in zebrafish.

Dose-response curve showing **(a)** the effect of VP-EL222 mRNA expression on zebrafish development as compared to **(b)** expression of a control GFP mRNA (right). Embryos were injected at the one to two-cell stage and illuminated with constant blue light until they were scored at 24 h.p.f. (at least n = 100 embryos per condition).

Supplementary Table 1. Occurrence of C120 sequence in human, mouse and zebrafish genomes.

The BLASTN 2.2.28+[3] program was used to search the indicated nucleotide databases for the EL222 binding site Clone-1 20 bp (C120) sequence. For comparison, the same databases were searched for the GAL4-specific Upstream Activation Sequence (UAS). To identify the top hits in each search, we chose a cutoff Expect (E) value < 10 and required that the sequences have no gaps or mismatches to original query sequence.

Search sequences:

C120 (20 bp) TAGGTAGCCTTTAGTCCATG

UAS (20 bp) GGAGGACAGTACTCCGCTCg* (one extra base added to make 20 bp query)

Organism	Database searched	# of sequences in database	C120 # of hits (E value <1000 / E value <10)	UAS # of hits (E value <1000 / E value <10)
<i>Homo sapiens</i>	Ref Seq Genomic	22,540	200 / 14	426 / 7
<i>Mus musculus</i>	Ref Seq Genomic	20,034	200 / 22	486 / 0
<i>Danio rerio</i>	Genome (ref only)	4,560	217 / 6	131 / 0

Top hits with E value <10, no mismatches, no gaps

Organism	C120	UAS
<i>Homo sapiens</i>	8 hits match 17 bp	7 hits match 17 bp
<i>Mus musculus</i>	18 hits match 16 bp	0 hits match 15 bp
<i>Danio rerio</i>	6 hits match 15 bp	0 hits match 14 bp

SUPPLEMENTARY VIDEOS

Supplementary Video 1.

Z-stack of 70% epiboly embryo showing mosaic expression of mCherry after illumination with blue light. Representative zebrafish embryo after injection with both VP-EL222 mRNA and pC120-mCherry DNA and illumination with constant blue light for 5 hr beginning at 2 h.p.f. Fluorescent and brightfield images were acquired every 2.58 μm on a Digital Scanned Laser Light Sheet Microscope (45). The two channels were merged and the z- stack converted into a video on ImageJ (41). Mosaic expression is due to random incorporation of the pC120-mCherry DNA into cells as the embryo develops.

Supplementary Video 2.

Z-stack of 70% epiboly embryo showing no expression of mCherry under dark conditions. Representative zebrafish embryo after injection with both VP-EL222 mRNA and pC120-mCherry DNA and kept in the dark for 7 hr. Fluorescent and brightfield images were acquired every 2.58 μm on a Digital Scanned Laser Light Sheet Microscope (45). The two channels were merged and the z-stack converted into a video on ImageJ (41).

Supplementary Video 3.

Localization of fluorescent mCherry in the heart of a developing zebrafish embryo at 24 h.p.f. Representative zebrafish embryo's heart after injection of both pminiTol2-myl7-VP-EL222-C120-mCherry DNA and transposase mRNA and illumination with constant blue light for 14 hr beginning at 10 h.p.f. Fluorescent images of a single plane within the zebrafish heart were acquired by time-lapse epifluorescent microscopy (Nikon Ti-E). Frames were acquired every 144 ms for 10 s. Playback is 7 frames/s.

SUPPLEMENTARY NOTES

Supplementary Note 1. Algorithm for kinetic modeling of VP-EL222 activation parameters.

To examine the dependence of EL222-based transcriptional activation on properties of the engineered VP-EL222 protein, we developed a conceptual model for the effect of pulsed blue light activation on EL222-driven transcription. We assumed that each pulse of light triggers 3 phases of gene expression (Fig. 3a):

- 1). A sigmoidal buildup phase characterized by EL222 activation and dimerization, DNA binding and transcriptional activation through the point of RNA Pol II promoter clearance, described by the collective rate constant τ_{on} . We checked for the need for cooperativity in this process, as implemented by Hill coefficients (h) between 1-5; while it was essential to have some degree of cooperativity ($h > 1$), we found minimal variation in the quality of fitting our experimental data for values between 2-5. Of these, $h=4$ provided the most optimal fit and was chosen for these simulations.
- 2). Once $t > \tau_{on}$, transcriptional activity saturates and enters a steady state phase, where transcription occurs at a maximal rate for as long as the cell is illuminated.
- 3). Once illumination ceases, active VP-EL222 decays as a first order exponential at a rate of τ_{off} , a process we assume to be likely dominated by cleavage of the cysteine/flavin adduct within the EL222 LOV domain. Transcriptional activity falls subsequently as EL222 reverts to the monomeric dark state, free from DNA.

For the purposes of the simulation, luciferase activity was taken to reflect the sum of all transcriptional activity over time, where each EL222 binding event generates one or more luciferase mRNAs provided that it remains on the promoter long enough for RNA polymerase

commitment to generating a full-length transcript. In order for transcriptional activity to move away from initialized values and reach steady state, the model was allowed to run for 10 light-dark cycles, each 80 s long, with an illumination time between 0 and 20 s. The EL222 concentration was assumed to be at steady-state and unaffected during these ten cycles. Predicted luciferase activities for a given τ_{on} , τ_{off} pair were normalized and compared to experimental values. For illumination times of 0, 2, 5, 10 and 20 s, the difference between the predicted and actual value for luciferase activity was computed and squared (Fig. 3b). The sum of these five errors was assumed to reflect the overall quality of the prediction. To determine τ_{on} , τ_{off} pairs that gave the least squared error, a simple grid search was implemented in MATLAB version R2012a (code provided as Supplementary Note 2), where the error function was evaluated at all combinations of τ_{on} and τ_{off} for values of $\tau_{on} = 1-100$ s and $\tau_{off} = 1-100$ s (Fig. 3e shows expansion of values between $\tau_{on} = 1-10$ s and $\tau_{off} = 1-100$ s).

Supplementary Note 2. MATLAB code for kinetic model

```
clc
clear

%% Motta-Mena et al., Supplementary Note 2
%% MATLAB R2012a code implementing kinetic model for VP-EL222
transcriptional activation
%%
%% developed by Spencer Glantz (Univ. Pennsylvania) in collaboration with
Kevin Gardner (UT Southwestern Medical Center)
%% contact: Kevin.Gardner@utsouthwestern.edu

%%Initialize Parameter Values

t = 1; %t is time in seconds for
tau_off = 0; %tau_off is the rate constant for exponential decay from the
excited state in the dark
tau_on = 0; %tau_on is the rate constant for the sigmoidal activation of
EL222 mediated transcription
endtime = 800; %endtime is the total time of the simulation required to reach
steady state
light_on = 0; %light_on is the duration of time in seconds during which the
system is illuminated at the beginning of each cycle
cycle_time = 80; %cycle_time is the length of each illumination cycle in
seconds
A_on = 0; %transcriptional activity (percent) (starts at 0 at simulation
beginning)
A_total = 100; %The maximum transcriptional activity possible (percent)
Amax = 50000; %A normalization factor used for comparison of simulated data
to experimental data

%%Counting variables that keep track of progress in the coode
s = 1; %the number of illumination times tested
n = 0; %time within a cycle
tau_off_count = 1; %the number of tau values tested
tau_on_count = 1; %the number of tau_on values tested

while tau_on <= 15 % Run simulation for 0 < tau_on <= 15 seconds
light_on = 0;
tau_off = 0;
tau_off_count = 1;
s = 1;

while tau_off <= 100 %Run simulation for 0 < tau_off <= 100 seconds
light_on = 0;
s = 1;

while light_on <= 20 %Generate prediction of transcriptional activity for 0 <
```

```

light_on < 20 seconds

t = 1;
n = 0;
A_on = 0;

while t<endtime %Simulate transcription until steady state is reached

    if n<=light_on %While the cycle time is less than the illumination time,
the light is on

        status = 'lit';
        A_on = A_total*(n^4/(n^4+tau_on^4)); %Transcriptional activation by
light is cooperative and is a function of tau_on

    else %for cycle times longer than the illumination time, light is off
        status = 'dark';
        A_on = A_on*exp(-(n-light_on)/tau_off); %Transcriptional activity
exponentially decays as a function of tau_off
    end

    if n == cycle_time-1 %restart the cycle clock once a cycle has ended
        n = 0;
    end

    A_t(t) = A_on; %Keep track of transcriptional activity at every time t
    t = t+1; %Advance overall simulation time
    n = n+1; %Advance time within an individual cycle

end

%After the end of the simulation is reached, record data and change the
illumination time

A_integral(s) = trapz(A_t); %Approximate the area under the transcriptional
activity curve for a given illumination time

if light_on < 5
light_on = light_on + 1;

else
light_on = light_on + 2.5;

end
s = s+1;

end

% Once all illumination times have been tested for a given tau_on, tau_off
% pair, compare the simulated and experimental average transcription

```

```

% activity v. illumination time curves

A_avg = A_integral/(endtime); %find average transcriptional activity over the
simulation

A_avg = A_avg*Amax; %Scale the avg transcriptional activity by a
normalization parameter

%Experimental Data
mean_y = [13675.333 18499.6667 272269.333 676027.667 1308505.67];
error_y = [4253 685 10878 95300 82860];

%Compute the sum of the squares of the difference between predicted and
measured values
%for all illumination times
sq1 = (A_avg(1)-13675.333)^2;
sq2 = (A_avg(3)-18499.6667)^2;
sq3 = (A_avg(6)-272269.333)^2;
sq4 = (A_avg(8)-676027.667)^2;
sq5 = (A_avg(12)-1308505.67)^2;

sum = sq1+sq2+sq3+sq4+sq5;

quality_total(tau_off_count,tau_on_count) = sum; %Record that error and the
tau_on/tau_off pair associated with it

tau_off_count = tau_off_count + 1;
tau_off = tau_off + .5; %Test a new tau_off

end

tau_on = tau_on + .25
tau_on_count = tau_on_count + 1; %Test a new tau_on

end

```

REFERENCES

1. Weber, W. & Fussenegger, M. Inducible product gene expression technology tailored to bioprocess engineering. *Curr. Opin. Biotechnol.* 18, 399-410 (2007).
2. Weber, W. & Fussenegger, M. Emerging biomedical applications of synthetic biology. *Nat. Rev. Genet.* 13, 21-35 (2012).
3. Briggs, W. R. & Spudich, J. L. *Handbook of Photosensory Receptors.* (Wiley-VCH, 2005).
4. Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. *Nat. Biotechnol.* 20, 1041-1044 (2002).
5. Levskaya, A. et al. Synthetic biology: engineering *Escherichia coli* to see light. *Nature* 438, 441-442 (2005).
6. Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol.* 27, 941-945 (2009).
7. Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* 7, 973-975 (2010).
8. Ye, H., Daoud-El Baba, M., Peng, R. W. & Fussenegger, M. A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science* 332, 1565-1568 (2011).

9. Ohlendorf, R., Vidavski, R. R., Eldar, A., Moffat, K. & Moglich, A. From dusk till dawn: one-plasmid systems for light-regulated gene expression. *J. Mol. Biol.* 416, 534-542 (2012).
10. Wang, X., Chen, X. & Yang, Y. Spatiotemporal control of gene expression by a light-switchable transgene system. *Nat. Methods* 9, 266-269 (2012).
11. Polstein, L. R. & Gersbach, C. A. Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J. Am. Chem. Soc.* 134, 16480-16483 (2012).
12. Liu, H., Gomez, G., Lin, S., Lin, S. & Lin, C. Optogenetic control of transcription in zebrafish. *PLoS One* 7, e50738 (2012).
13. Nash, A. I. et al. Structural basis of photosensitivity in a bacterial light-oxygen- voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. *Proc Natl Acad Sci U S A* 108, 9449-9454 (2011).
14. Huala, E. et al. Arabidopsis NPH1 - A protein kinase with a putative redox-sensing domain. *Science* 278, 2120-2123 (1997).
15. Rivera-Cancel, G., Motta-Mena, L. B. & Gardner, K. H. Identification of natural and artificial DNA substrates for light-activated LOV-HTH transcription factor EL222. *Biochemistry* 51, 10024-10034 (2012).
16. Zoltowski, B. D., Nash, A. I. & Gardner, K. H. Variations in protein-flavin hydrogen bonding in a light, oxygen, voltage domain produce non-Arrhenius kinetics of adduct decay. *Biochemistry* 50, 8771-8779 (2011).

17. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* 335, 563-564 (1988).
18. Chen, E., Swartz, T. E., Bogomolni, R. A. & Kliger, D. S. A LOV story: the signaling state of the phot1 LOV2 photocycle involves chromophore-triggered protein structure relaxation, as probed by far-UV time-resolved optical rotatory dispersion spectroscopy. *Biochemistry* 46, 4619-4624 (2007).
19. Kennis, J. T. et al. Primary reactions of the LOV2 domain of phototropin, a plant blue-light photoreceptor. *Biochemistry* 42, 3385-3392 (2003).
20. Harper, S. M., Neil, L. C., Day, I. J., Hore, P. J. & Gardner, K. H. Conformational changes in a photosensory LOV domain monitored by time-resolved NMR spectroscopy. *J. Am. Chem. Soc.* 126, 3390-3391 (2004).
21. Pan, Y. X., Chen, H. & Kilberg, M. S. Interaction of RNA-binding proteins HuR and AUF1 with the human ATF3 mRNA 3'-untranslated region regulates its amino acid limitation-induced stabilization. *J. Biol. Chem.* 280, 34609-34616 (2005).
22. Qian, X. et al. Posttranscriptional regulation of IL-23 expression by IFN-gamma through tristetraprolin. *J. Immunol.* 186, 6454-6464 (2011).
23. Thompson, J. F., Hayes, L. S. & Lloyd, D. B. Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene* 103, 171-177 (1991).
24. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. & Singer, R. H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332, 475- 478 (2011).

25. Lynch, K. W. & Weiss, A. A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. *Mol. Cell. Biol.* 20, 70-80 (2000).
26. Ip, J. Y. et al. Global analysis of alternative splicing during T-cell activation. *RNA* 13, 563-572 (2007).
27. Martinez, N. M. et al. Alternative splicing networks regulated by signaling in human T cells. *RNA* 18, 1029-1040 (2012).
28. Faustino, N. A. & Cooper, T. A. Identification of putative new splicing targets for ETR-3 using sequences identified by systematic evolution of ligands by exponential enrichment. *Mol. Cell. Biol.* 25, 879-887 (2005).
29. Mallory, M. J. et al. Signal- and development-dependent alternative splicing of LEF1 in T cells is controlled by CELF2. *Mol. Cell. Biol.* 31, 2184-2195 (2011).
30. Dembowski, J. A. & Grabowski, P. J. The CUGBP2 splicing factor regulates an ensemble of branchpoints from perimeter binding sites with implications for autoregulation. *PLoS Genet.* 5, e1000595 (2009).
31. Yelon, D., Horne, S. A. & Stainier, D. Y. Restricted expression of cardiac myosin genes reveals regulated aspects of heart tube assembly in zebrafish. *Dev. Biol.* 214, 23-37 (1999).
32. Harper, S. M., Neil, L. C. & Gardner, K. H. Structural basis of a phototropin light switch. *Science* 301, 1541-1544 (2003).

33. Scheuermann, T. H. et al. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat. Chem. Biol.* 9, 271-276 (2013).
34. Fan, H. Y. et al. Improving a designed photocontrolled DNA-binding protein. *Biochemistry* 50, 1226-1237 (2011).
35. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *Proc Natl Acad Sci U S A* 105, 10709-10714 (2008).
36. Mattis, J. et al. Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nat. Methods* 9, 159-172 (2012).
37. Strickland, D. et al. TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* 9, 379-384 (2012).
38. Wu, Y. I. et al. A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461, 104-108 (2009).
39. Lynch, K. W. & Weiss, A. A CD45 polymorphism associated with multiple sclerosis disrupts an exonic splicing silencer. *J. Biol. Chem.* 276, 24341-24347 (2001).
40. Bookout, A. L. & Mangelsdorf, D. J. Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nucl Recept Signal* 1, e012 (2003).
41. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671-675 (2012).

42. Westerfield, M. The zebrafish book. A guide for the laboratory use of zebrafish (*Danio rerio*). 5th edn, (The University of Oregon Press, 2007).
43. Balciunas, D. et al. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet.* 2, e169 (2006).
44. Huang, C. J., Tu, C. T., Hsiao, C. D., Hsieh, F. J. & Tsai, H. J. Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Dev. Dyn.* 228, 30-40 (2003).
45. Maizel, A., von Wangenheim, D., Federici, F., Haseloff, J. & Stelzer, E. H. High-resolution live imaging of plant growth in near physiological bright conditions using light sheet fluorescence microscopy. *Plant J.* 68, 377-385 (2011).

CHAPTER FOUR

TAEL: Non-toxic, zebrafish-optimized optogenetic gene expression system with large dynamic range and rapid kinetics

"Maybe there is a lever on the gene that can turn it on...Don't laugh!"

-Wallace Marshall

Manuscript in preparation.

SUMMARY

Current inducible expression systems in zebrafish are limited in fine-grained spatial and temporal control. Here we address this by developing an optogenetic gene expression system that is optimized for use in zebrafish. We tested two potential optogenetic gene expression systems, LightOn and EL222, and show that they are both able to strongly induce expression of fluorescent reporter constructs. However we observed significant toxicity at mid-to-high concentrations of LightOn and EL222's transcription factors in zebrafish embryos. This was remedied by replacing the VP16 or p65 domain of EL222 and LightOn, respectively, with KalTA4, a transactivating domain optimized to be non-toxic in zebrafish. This switch removed all toxicity up to 200pg of injected mRNA/embryo. We then tested each system and found only the reengineered EL222 system, TAEL, to still be functional. Characterization of TAEL dynamics shows the system capable of a large range of induction with relatively quick on/off kinetics. As a functional example, we demonstrate that TAEL is able to induce ectopic endodermal cells via *sox32* induction. Finally, we provide several methods of how to spatially restrict gene induction with the TAEL system.

INTRODUCTION

Extensive insight into pathways involved in biological phenomena and the components that make up these pathways has been achieved through various traditional genetic methods. However biological systems are notorious for their ability to compensate for mutagenesis that occurs on the time scale that is required to create stable transgenic or mutant lines. Additionally, when working with multicellular organisms it is often difficult to limit a perturbation to a specific set of cells or to study any gene or pathway that is used more than once or that is essential for survival of the organism as a whole. The ability to precisely control the location, amount and timing of gene expression has been a long sought after powerful tool for experimental biologists.

In zebrafish, as well as other multicellular model organisms, current inducible gene expression systems are quite limited in spatial and temporal control. For temporal control they either rely on the administration of small molecule compounds, such as in the tetracycline-inducible expression system (1), or activation of heat shock promoters (2). In the case of small molecule dependent activation, the timing and penetration can be very variable and hard to control. The use of the heat shock promoter improves upon these issues, but introduces the complication of inducing a heat shock response in your organism as well. Tissue-specific promoters in conjunction with either of these inducible systems is the finest level of spatial control that can currently be achieved, which is only sufficient if there exist a reliable tissue-specific promoter for the region of organism that you are interested in or if there is not a need for sub-tissue level control.

By far the most sophisticated form of gene expression control that exists today is optogenetics-based. Because optogenetic gene expression systems rely on a light for activation, they allow for much finer spatial and temporal control. Several useful optogenetic systems have been developed, but so far, these systems are not ideally suited for use in zebrafish. In order to be

truly useful for studying zebrafish or any other multicellular organism, a system would need to be: genetically encoded, not require complicated optics or exogenous small molecules, have a large range of induction, be reversible with fairly quick kinetics and, most importantly, have little to no toxicity.

Recently, two systems, LightOn (3) and EL222 (4), showed the promise to fulfill all of these qualifications if optimized for zebrafish use. Both systems make use of an engineered bacterial light-oxygen-voltage (LOV) protein domain that dimerizes when illuminated with blue light (Figure 1a). The LightOn system developed in Yi Yang's lab is based off a synthetically constructed transcriptional activator, GAVPO; a LOV domain fused to the Gal4 DNA binding domain, which can only bind its corresponding UAS promoter sequence when dimerized, and a transactivating domain to induce transcription upon binding. When a UAS promoter is placed in front of your gene of interest, induction of its transcription can be activated by GAVPO dimerization and binding upon exposure to blue light. The EL222 system developed in Kevin Gardner's lab is similarly designed, though based off a naturally occurring light-responsive transcription factor, EL222, in *Erythrobacter litoralis*. EL222 also contains a LOV domain fused to a helix-turn-helix (HTH) DNA binding domain. To complete the system, the Gardner lab added a VP16 transactivating domain and nuclear localization sequence to EL222 and then located EL222's binding sequence, C120, to create a promoter with five copies of C120 in front of a TATA box.

We have tested both systems in zebrafish and found them both easily expressed and able to strongly induce expression of fluorescent reporter constructs. However we observed significant toxicity at mid-to-high concentrations of LightOn and EL222's transcription factors in zebrafish embryos. Studies have shown that high levels of strong transactivating domains are toxic for zebrafish development (5). We therefore replaced the VP16 or p65 domain of EL222 and

LightOn, respectively, with KalTA4, a zebrafish optimized transactivating domain. This switch removed all toxicity up to 200pg of injected mRNA/embryo. We then tested each system and found only the reengineered EL222 system, TAEL, to still be functional.

Characterization of TAEL dynamics shows the system capable of a large range of induction with relatively quick on/off kinetics. As a functional example, we show that TAEL is able to generate ectopic endodermal cells via *sox32* induction. We then demonstrate several different methods to restrict blue light application to activate the TAEL system in a spatially restricted manner.

RESULTS

Optimization of a optogenetic gene expression system for zebrafish: TAEI.

To test for LightOn function in early zebrafish development, GAVPO mRNA was injected into single cell Tg(UAS:kaede) embryos. Embryos were then exposed to constant global blue light from approximately 4 h.p.f. to 24 h.p.f. GAVPO injected embryos showed robust induction of kaede fluorescent protein at 24 h.p.f., with no visible induction of kaede expression in dark or uninjected controls. To test for EL222 function in early zebrafish development, EL222 mRNA along with a plasmid containing the C120 promoter driving an mCherry reporter was injected into single cell wild-type embryos. Embryos were then exposed to constant global blue light from approximately 4 h.p.f. to 24 h.p.f. EL222/C120:mCherry injected embryos showed robust induction of mCherry fluorescent protein at 24 h.p.f., with no visible induction of mCherry expression in dark or C120:mCherry only controls. To test for potential toxicity caused by expressing these light-responsive transcriptional activators, a toxicity curve for each was performed. Embryos were injected with increasing amounts of EL222 or GAVPO mRNA and then scored as unaffected or affected (deformed or dead) at d1. Both transcriptional activators were significantly more toxic to zebrafish development than a GFP control, with GAVPO having the strongest negative effect (Figure 1c).

Studies have shown that high levels of strong transactivating domains are toxic for zebrafish development. Therefore we replaced the VP16 or p65 domain of EL222 and GAVPO, respectively, with the zebrafish-optimized transactivating domain of KAT4 (Figure 1b). Toxicity curves were repeated for both modified transcriptional activators, TA4-EL222 and GAVPO-TA4 (shortened to TAEI and GAVTA, respectively). Both TAEI and GAVTA showed little-to-no toxicity over the control even up to 200pg of injected mRNA/embryo (Figure 1d). Both transcriptional activators were then tested for any alterations in function due to the change in

transactivating domain. TAEI mRNA injection into Tg(C120:mCherry) embryos still induces strong mCherry expression under blue light conditions with no visible expression in controls (Figure 1f). However, injection of GAVTA mRNA into Tg(UAS:kaede) embryos and subsequent exposure to blue light was no longer able to induce kaede expression over dark and uninfected controls (Figure 1e).

Characterization of TAEI system, dynamics and range.

In order to examine the kinetics and range of TAEI-induced transcription, a qPCR timecourse of mCherry transcript levels was performed on Tg(C120:mcherry) embryos injected with 100pg TAEI mRNA and exposed to global 465nm light at 3.5 hpf (Figure 2). Constant blue light illumination resulted in a relatively quick and strong induction of mCherry transcription; with an approximately 200 fold increase in transcript levels from 30 min to 3 hours post activation (h.p.a.). Transcription peaks at ~ 3 h.p.a., decreasing steadily under continued constant illumination. This is most likely from photobleaching of the LOV domain of the TAEI transcriptional activator. Once photobleached the LOV domain would be unable to dimerize in response to blue light illumination. Pulsing illumination off and on at hour intervals lessens this photobleaching, allowing for longer and steadier induction kinetics. To examine the kinetics of transcriptional “turn off” of the TAEI system, we performed a qPCR timecourse of mCherry transcript levels on embryos removed from blue light illumination after 3 hours of activation. No new mCherry transcription was observed after 30 minutes following removal from blue light illumination, showing TAEI “turn off” kinetics to be relatively rapid.

Ectopic endoderm induction via over-expression of sox32.

The TAEI system’s ability to induce expression of a gene of interest when and where desired makes it a powerful tool for studying zebrafish development. An example is the ability to

ectopically express a master transcription factor in order to change the fate of a cell or set of cells during development. Sox32 is a master transcription factor that is responsible for specifying endoderm fate in early zebrafish development (6). We therefore decided to test if activating the expression of sox32 via the TAEL system could induce ectopic endodermal cells (Figure 3a). We first created a C120:sox32 stable transgenic line and crossed it to sox17:GFP transgenic fish, as sox17 is a marker for endodermal fate. We then injected TAEL mRNA, along with H2B-mCherry mRNA as a nuclear marker, into C120:sox32/sox17:GFP embryos and activated under constant blue light immediately. Activated embryos showed a significant number of GFP positive cells outside of sox17's normal expression profile, which is restricted to a band of cells around the margin in the inner most layer of the embryo from 60-80% epiboly. Activated embryos showed GFP positive cells located both above this band all the way up to the animal pole (Figure 3b, bracket) and in the outermost layer of the embryo (Figure 3b, arrow). Dark and no-TAEL controls show normal sox17:GFP expression.

Cell ablation via diphtheria toxin induction.

In an effort to develop an advanced cell ablation method we tested the ability to ablate cells via diphtheria toxin (dtx) induction with the TAEL system (7). We first injected wild-type embryos with a DNA plasmid containing dtx under the C120 promoter, along with TAEL mRNA. The injected embryos were kept in the dark until approximately 4 h.p.f. and at this point assessed for any toxicity due to leakiness of the C120 promoter; none was detected. The embryos were then exposed to 1 hour of global blue light activation. As a result of the mosaic expression of dtx from plasmid injection, various levels of cell death occurred, as was ascertained by cell extrusion from the embryo (Sup. Figure 1). These positive results encouraged us to proceed with making a stable transgenic line containing C120:dtx. However, all founders encountered produced either unhealthy F1's or were unaffected by TAEL injection and activation, most likely through

silencing of the C120 promoter. This suggests that the C120 promoter has a tight enough off state to not activate and kill embryos kept in the dark for the transient experiments, but is slightly leaky in that a healthy and functional stable line was not able to be made.

Spatial control of TAEEL induction.

A major advantage of the TAEEL system over previous gene expression systems is the ability to control its activation spatially by restricting the application of activating blue light. Here we show several examples of how this can be done. Since the range of blue light that can activate TAEEL is from 450-490 nm, the GFP channel's excitation light can be used on any microscope as an activating light source. The simplest way to restrict blue light exposure is to close down the field diaphragm on an epifluorescent microscope (Figure 4a). This restricts the light coming through the objective and hitting your sample to a small hexagon column. Light inputs can also be specifically applied to defined regions of the embryo with a digital micromirror device (DMD). A DMD has thousands of microscopic mirrors arranged in a rectangular array that correspond to pixels of an image to be projected. "On" and "Off" pixels can be set by adjusting the angle of each mirror to either reflect light onto or off the sample, respectively. We use a custom DMD illuminated with a 470nm LED to project various sized squares of blue light onto our samples (Figure 4b), though almost any shape could be projected in this manner. Finally we restricted the range in which the 488nm laser was scanned on a Digital Scanned Laser Light Sheet Microscope (DSLMS) to shine a thin beam of blue light through our sample (Figure 4c). In addition to controlling the height of the beam, the sample can be moved in the z direction to increase the area of activation from a thin 2-dimensional beam to a 3-dimensional rectangular column. We exposed the eye region of 10-somite stage C120:mCherry (+ TAEEL mRNA) embryos to a single 2 minute pulse of blue light by each of these methods and were able to activate expression of mCherry in the same region. Off target activation was seen to varying

levels due to scattering of blue light through the thick tissue of the head. As one would expect, the more restricted the application of blue light, the less off target activation is seen. We suspect that the use of a 2-photon microscope would give us 3-dimensional control of TAEL induction and are currently testing the system's ability to be activated by 2-photon technology.

DISCUSSION

Here we show how we have developed a powerful new tool for controlling gene expression in zebrafish. The TAEEL system offers several advantages over current inducible gene expression systems in zebrafish. First, this system does not require an exogenous component in order to function. This is because the flavin chromophore that the LOV domain requires to respond to light (8) is endogenous to zebrafish. This is not the case in phytochrome-based optogenetics, which requires the application of exogenous phycocyanobilin in order to function (9). Additionally it is genetically encoded, nontoxic to zebrafish development, and composed of a single homodimerizing transcriptional activator. This minimal component aspect has the benefit of eliminating the need for expression optimization of multiple proteins.

We also demonstrate that the TAEEL system has a large range of induction and, on the time scale of gene expression, relatively quick on and off kinetics. With constant blue light activation, transcript levels peaked at approximately 3 hours post activation at 200 fold over levels at 30 minutes post activation. Higher and faster induction could also be possible with a higher concentration of TAEEL or stronger blue light application. There appears to be an inactivation of the system when activated continually for long periods of time, which is possibly due to photobleaching of the LOV domain. In an effort to compensate for this we tried pulsing the blue light on and off at one hour intervals. This lessened the photobleaching and allowed for a more even and longer induction event. We also showed that once returned to a dark state, the TAEEL system stops all new transcription within 30 minutes. All of this suggests varying the amount of TAEEL present and the strength and timing of blue light activation should allow one to achieve almost any level of induction desired.

The greatest advantage to this system is its high level of spatial control. First we showed that complicated optics are not required to activate the TAEL system in a spatially restricted manner. Spatial restriction of induction can be as simple as closing down the field diaphragm on an epifluorescent microscope. Higher levels of control can also be achieved through more advanced methods of blue light application. For example, a DMD fitted with a LED in the 450-490nm range can restrict activation to a specific areas of interest. If one has access to a light sheet microscope, restricting the sheet height or the area in which the laser is scanned is also a way to spatially control activation. We believe that 2-photon microscopy would give the highest level of spatial control and are currently testing this method of activation to assess whether we can achieve 3-dimensional control of gene induction.

With the versatility that we have demonstrated, the TAEL system could serve a multitude of different functions. It can be used to express genes of interest at any time and location, which is extremely useful for studying a gene or pathway that is expressed at multiple locations or times during development. It could also be used for targeted cell ablation, though we would recommend a less potent toxicity factor than diphtheria toxin, such as a caspase. And finally, we are currently testing our ability to use the TAEL system for spatiotemporal control of gene knockdown via Cas9 induction (10). If functional, this TAEL/Crisper combination could be used to improve our ability to study traditionally hard to perturb genes, such as those essential for survival.

METHODS

Vector construction and mRNA synthesis.

Expression plasmid pCS2-(VP16)EL222 is previously described in (4). All sub-cloning was done by Gibson assembly (11). Expression plasmid pCS2-GAVPO was created by PCR amplification of GAVPO ORF (3) and then cloned into pCS2+. Expression plasmid pCS2-TAEL was created by PCR amplification of TA4 ORF of *KalTA4*(5) and then cloned into pCS2-(VP16)EL222 cut with *EcoRI* and *NcoI* to remove the VP16 domain. Expression plasmid pCS2-GAVTA was created by PCR amplification of TA4 ORF and then cloned into pCS2-GAVPO cut with *EcoRI* and *StuI* to remove the p65 domain. Transgene plasmid mTol2-cryaa:Venus;5xC120:mCherry was created by separate PCR amplification of (cryaa promoter and Venus ORF) and (5xC120 promoter and mCherry ORF) and then cloned into pminiTol2 (12). Transgene plasmid mTol2-crya:Venus;5xC120:dtx was created by PCR amplification of dtx ORF (gift from Dan Hesselson) and then cloned into mTol2-cryaa:Venus;5xC120:mCherry cut with *NcoI* and *EcoRV* to remove the mCherry ORF. Transgene plasmid mTol2-cry:Venus;5xC120:sox32-P2A-tBFP was created by separate PCR amplification of sox32 ORF and tagBFP ORF then cloned into mTol2-cryaa:Venus;5xC120:mCherry cut with *NcoI* and *EcoRV* to remove the mCherry ORF. Transgene plasmid mTol2-cry:Venus;5xC120:Cas9 was created by PCR amplification of Cas9 ORF (Addgene) and then cloned into mTol2-cryaa:Venus;5xC120:mCherry cut with *NcoI* and *EcoRV* to remove the mCherry ORF.

Capped messenger RNA was synthesized using the mMESSAGE mMACHINE SP6 kit (Ambion) with pCS2 constructs cut with *NotI* as linear template.

Zebrafish strains.

Adult zebrafish were maintained under standard laboratory conditions. Tg(UAS:kaede) transgenic line courtesy of Herwig Baier's lab (rk8Tg in ZFIN: <http://zfin.org/>). Tg(sox17:GFP) created as previously described (13). Tg(cryaa:Venus;C120:mCherry), Tg(cry:Venus;C120:sox32-P2A-tBFP) and Tg(cry:Venus;C120:Cas9) transgenic lines were created by injecting 20 pg of corresponding transgene plasmid DNA along with 100 pg of Tol2 transposase mRNA at the one-cell stage. Injected embryos were then sorted by eye marker on d2, raised to adulthood and then screened for founders by outcrossing to wild-type.

Toxicity curves in zebrafish.

To access the toxicity of the transcriptional activators, 50, 100, 150, 200 or 300 pg of VP-EL222, GAVPO, TA4-EL222, GAVPO(TA4), or GFP (control) mRNA per embryo were injected at the one- to two-cell stage. Unfertilized embryos were removed on day 0, and phenotypes (unaffected versus developmentally deformed or dead) of each group were scored alongside uninjected control embryos from the same clutch on day 1 after manual dechoriation. Each group had at $n > 100$ embryos.

Microscopy and image processing of zebrafish embryos.

Figure 1's fluorescent and brightfield images were taken at 24 h.p.f. on a Zeiss Axio Observer.Z1 running Zen Blue, equipped with a X-Cite 120Q fluorescence lamp (Lumen Dynamics), Coolsnap ES2 CCD camera and a 5x 0.25NA Fluar Zeiss objective. Dechorionated embryos were embedded in 1.5% low-melt agarose within glass-bottom Petri dishes (MatTek Corporation). Standard filter settings were applied and then brightfield and fluorescent images merged. Figure 3's fluorescent images were taken at 70% epiboly on a Digital Scanned Laser Light Sheet Microscope, DSLM (14). Embryos were mounted in a 1.5% low-melt agarose cylinder using 3 mm O.D./2 mm I.D. FEP tubing (Bola). Z stacks of 2.58 μm intervals were taken

with a 10x 0.5NA Zeiss objective. mCherry fluorescence was imaged with 561 nm laser line and a 561LP filter. GFP fluorescence was imaged with a 488 nm laser line and a 488LP filter. Maximum intensity projections of the Z-stacks for the GFP channel were performed for the top panel of Fig. 3b and a representative z slice of each condition was chosen to show a merge of the mCherry and GFP channel images for the bottom panel. Figure 4's fluorescent and brightfield images of spatial activation modes were taken on their corresponding microscopes:

1. Epifluorescence: Nikon Eclipse Ti microscope running NIS Elements, equipped with a Sutter Lambda XL Broad Spectrum Light Source, an electron multiplying charge-coupled device (EM-CCD) camera (Andor iXon DU-897) and a 10x Plan Fluor 0.3NA or a 20x Plan Apo 0.75NA Nikon objective. A 760LP filter was placed in the brightfield path to prevent unwanted activation from brightfield illumination. Standard filter settings were applied and then brightfield and fluorescent images merged.
2. DSLM: Embryos were mounted in a 1.5% low-melt agarose cylinder using 3 mm O.D./2 mm I.D. FEP tubing (Bola). A 10x 0.5NA Zeiss objective was used to image the 488nm activating beam (488nm laser line and a 488LP filter) and brightfield image (?).
3. DMD: Nikon Eclipse Ti microscope running Nikon Elements, equipped with a custom digital micro mirror device (Andor Technologies), an EM-CCD camera (Evolve, Photometrics), and a Nikon 10x Plan Fluor 0.3NA DIC objective. 'On' pixels (regions to be stimulated with activating light) were illuminated with 470nm light (Lumileds), whereas 'off' pixels were unexposed.

Fluorescent and brightfield images of mCherry reporter expression after spatial activation were taken at 24 h.p.f. on the Zeiss Axio Observer (see above). All image processing and analysis was performed using ImageJ software (15).

Light induction.

For global light induction, a TaoTronics TT-AL09 120W Dimmable LED Aquarium Hood was used apply constant or pulsed light (GraLab Model 451 High-Accuracy Digital Electronic Timer). Actual power of light received by embryos (lids of plates removed) was measured to be ~1.5 mW using a PM100D Laser Power and Energy Meter Console (Thorlabs). Dark controls were placed in a lightproof box in the same 30°C incubator as light-treated samples. For spatially restricted light induction, several different methods were used: closed down field diaphragm of an epifluorescent microscope, 488nm channel (Nikon Eclipse Ti described above), restricted scanning of the 488nm laser on a DSLM (described above), small DMD created square, 470nm LED (see above). After activation, PTU was added to final concentration of 0.003% and embryos were returned to the incubator (in the dark) for another 6 hours before imaging for mCherry reporter expression.

Real-time Quantitative PCR

To examine the kinetics of TAEL-induced transcription, Tg(5xC120:mCherry) embryos were injected with 100pg TAEL mRNA at the one-cell stage. At 3-3.5 hpf (just after the mid-blastula transition), embryos were illuminated with 465nm light; for negative controls, injected embryos were kept in the dark by covering dishes with aluminum foil. At the indicated time points, total RNA was extracted by Trizol extraction (16) or using the RNEasy Kit (Qiagen). 500ng RNA was used for reverse transcription using the Quantitect cDNA synthesis kit (Qiagen). The qPCR reaction mixture contained 2X SYBR green PCR master mix (Applied Biosystems), 10-fold diluted cDNA, and 714 nM each primer. Reactions were carried out in an Applied Biosystems 7900HT Fast Real-Time PCR System (Applied Biosystems) as follows: Initial activation at 95°C

for 10 min, followed by 40 cycles of 30s at 95°C, 1 min at 60°C, and 1 min at 72°C. Once the PCR was completed, a melt curve analysis was performed to determine reaction specificity. Samples were run in triplicate and data presented in Fig. 2 represent averages from two biological replicates each with two technical replicates. The housekeeping gene ef1a was used as a reference.

Primers used in this study:

mCherry forward: 5'-GACCACCTACAAGGCCAAGA-3'

mCherry reverse: 5'-CTCGTTGTGGGAGGTGATGT-3'

ef1a forward: 5'-CAAGAAGAGTAGTACCGCTAGCAT-3'

ef1a reverse: 5'-CACGGTGACAACATGCTGGAG-3'

FIGURE 1

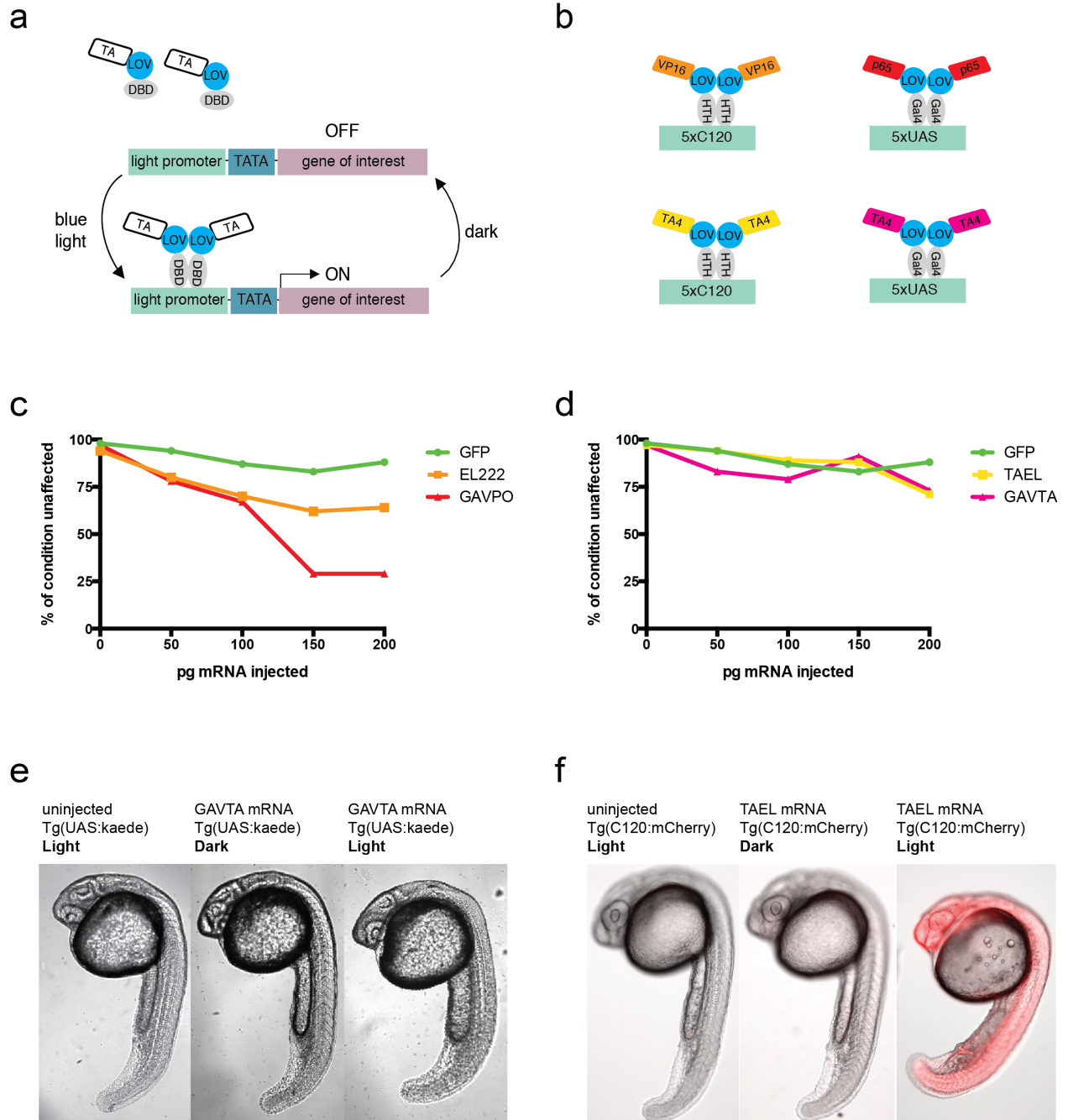
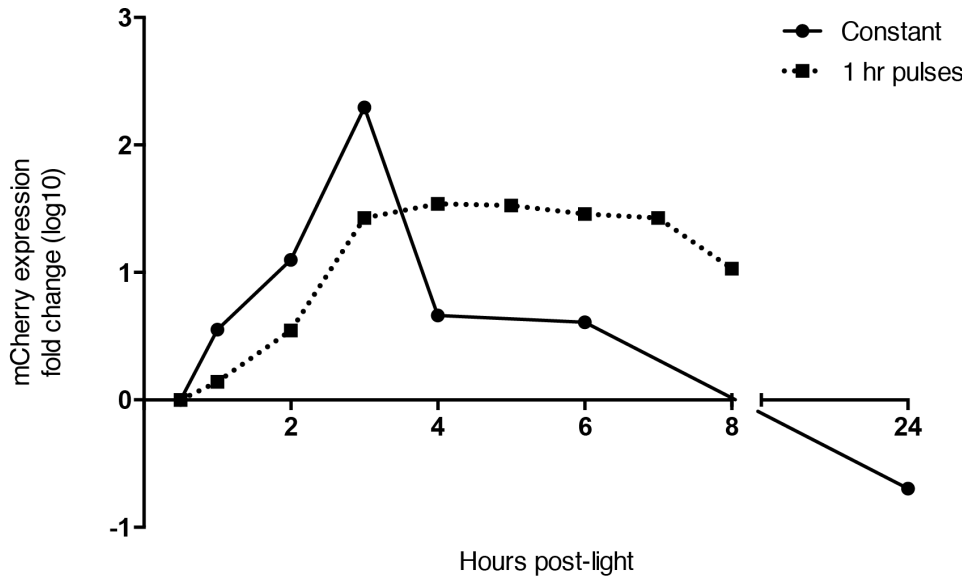


Figure 1: Optimization of a optogenetic gene expression system for zebrafish, TAEI.

(a) Cartoon depicting general setup for LOV-based, light-responsive gene expression systems. **(b)** Cartoons depicting transcriptional activators: EL222 (top left), GAVPO (top right), TAEI (bottom left), GAVTA (bottom right). Toxicity curves showing percentage of unaffected/healthy embryos per injection condition. **(c)** Both original GAVPO and EL222 transcriptional activators show toxicity over the GFP control. **(d)** Optimization of these transcriptional activators, GAVPO to GAVTA and EL222 to TAEI, causes a significant decrease in their toxicity **(e)** GAVTA is unable to induce transcription of kaede from the UAS promoter **(f)** TAEI is still able to induce transcription of mCherry from the C120 promoter under activating conditions.

FIGURE 2

a



b

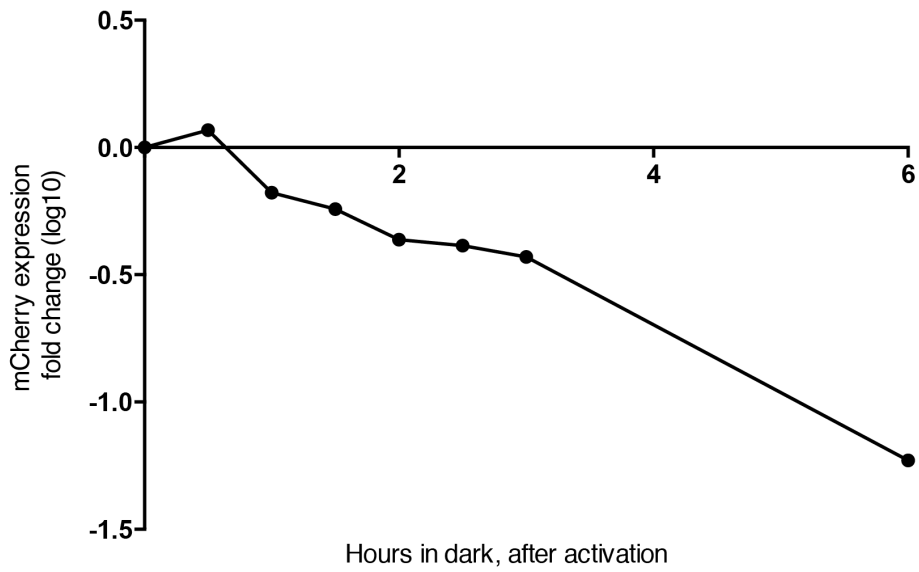
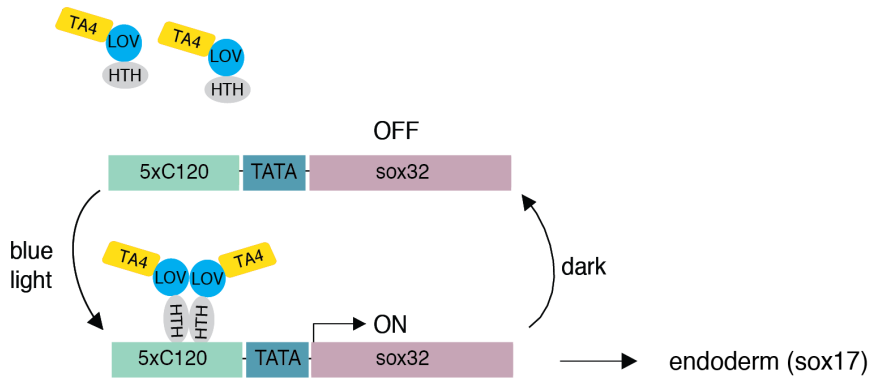


Figure 2: Characterization of TAEEL system, dynamics and range.

A qPCR timecourse of mCherry transcript levels was performed on Tg(C120:mcherry) embryos injected with TAEEL mRNA and illuminated with 465nm light at 3.5 hpf to examine the kinetics of TAEEL-induced transcription. (a) mCherry expression fold change (\log_{10}), normalized to 30 min post illumination, for both constant illumination and 1 hour ON/OFF pulsing conditions (continuous and dotted line, respectively). Timecourse shows fast and strong induction of mCherry transcription under constant illumination conditions, with a ~ 200 fold increase in transcript levels from 30 min to 3 hours post activation (h.p.a.). Transcription decreases after 3 h.p.a., most likely from photobleaching/inactivation of the LOV domain. Pulsing illumination off and on at hour intervals lessens this photobleaching allowing for longer and steadier induction kinetics. (b) No new mCherry transcription is observed by 30 minutes after activated embryos are returned to dark conditions, showing TAEEL “turn off” kinetics to be relatively rapid.

FIGURE 3

a



b

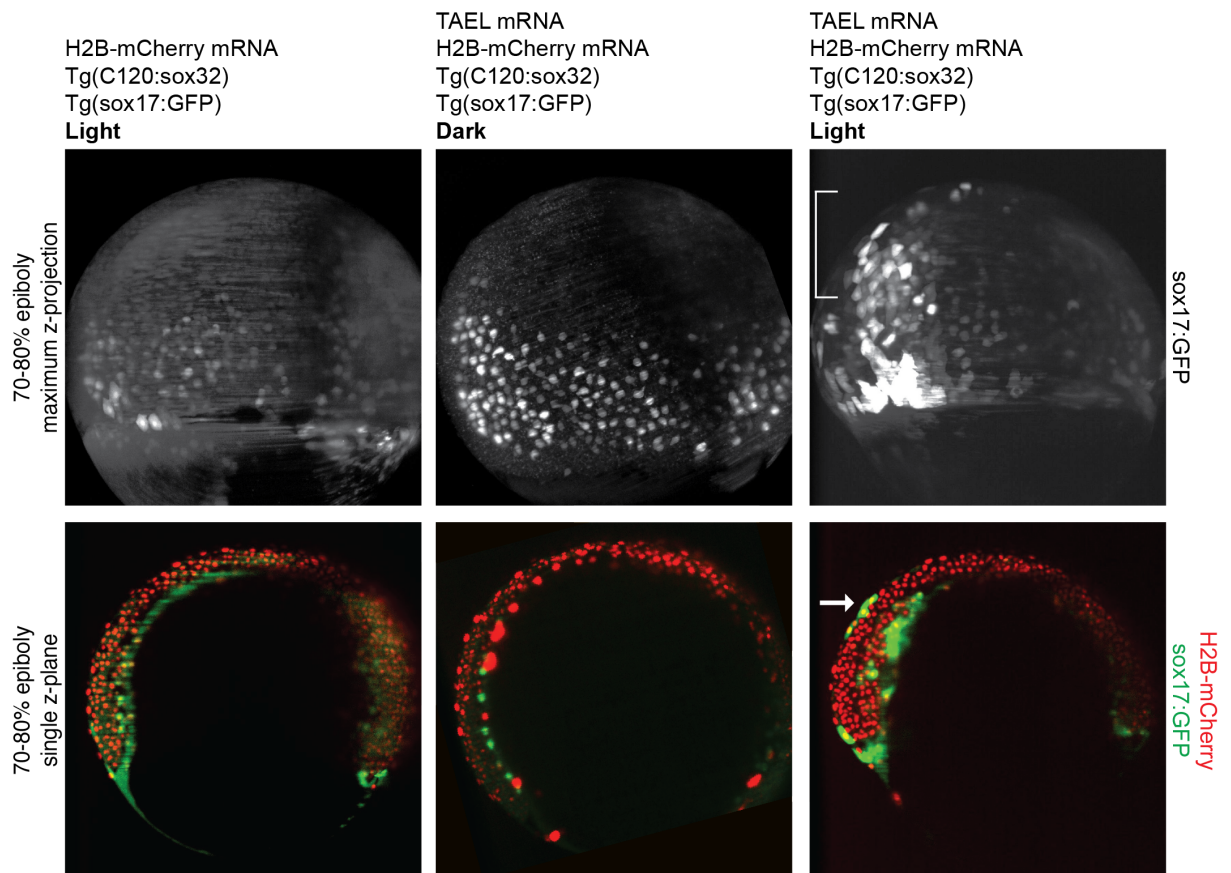


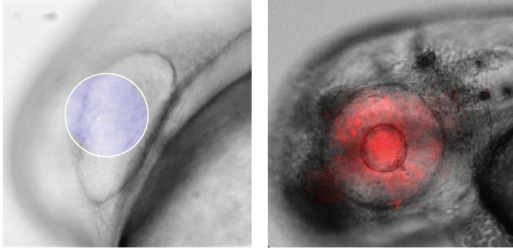
Figure 3: Ectopic endoderm induction via over-expression of sox32.

(a) Cartoon depicting experimental set up. Once exposed to activating blue light, TAE1 mRNA injected Tg(C120:sox32/sox17:GFP) embryos will mis-express sox32 which will result in ectopic induction of endoderm cells, labeled by sox17:GFP **(b)** Maximum z-projections of sox17:GFP on top panel, bracket marks sox17:GFP positive cells outside the normal expression range in TAE1 activated embryos. Bottom panel shows representative z planes, with mCherry as a nuclear marker to show layer location of sox17:GFP positive cells, arrow marks sox17:GFP positive cells located outside the inner-most layer of the embryo in TAE1 activated embryos.

FIGURE 4

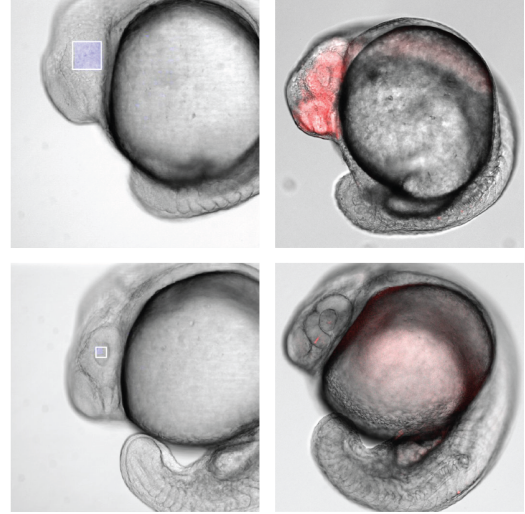
a

closed down field diaphragm
epifluorescence, 20x



b

mosaic square
DMD, 470nm LED, 5x



c

minimal scanning range
DSL, 488nm laser

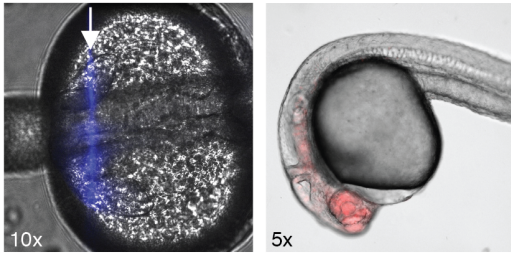


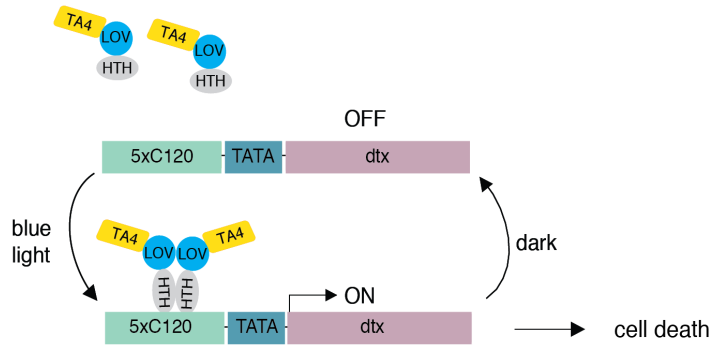
Figure 4: Spatial control of TAEL induction.

We demonstrate three different ways to spatially restrict TAEL-based transcriptional induction.

(a) Closing down the field diaphragm on an epifluorescent microscope (488nm, GFP excitation setting) restricts the light coming through the objective and hitting a sample to a small hexagon column. **(b)** DMD illuminated with a 470nm LED to project various sized squares of blue light onto a samples **(c)** Restricted scanning range of the 488nm laser on a DSLM to shine a thin beam of blue light through a sample. We exposed the eye region of 10-somite stage C120:mCherry (+ TAEL mRNA) embryos to a single 2 minute pulse of blue light by each of these methods and were able to activate expression of mCherry in the same region.

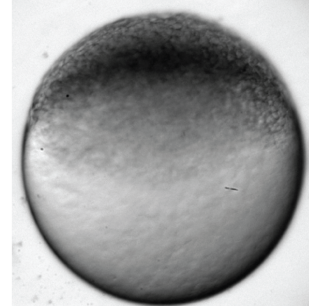
SUPPLEMENTARY FIGURE 1

a



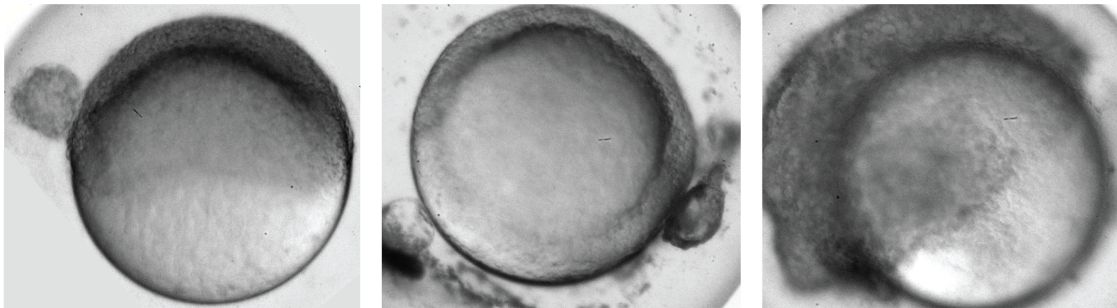
b

TAEL mRNA
pmTol2-C120:dtx DNA
Before Activation



c

TAEL mRNA
pmTol2-C120:dtx DNA
After Activation



Supplementary Figure 1: Cell ablation via diphtheria toxin induction.

(a) Cartoon depicting experimental set up. Once exposed to activating blue light, embryos injected with TAEL mRNA and mTol2-C120:dtx plasmid DNA will express dtx which will result in cell death **(b)** Injected embryos immediately before light activation, kept in the dark until ~4 h.p.f. **(c)** Injected embryos after 1 hour of TAEL activation. Various levels of cell death occurred, marked by cell extrusion from the embryo.

REFERENCES

1. Knopf, Franziska, Kristin Schnabel, Christa Haase, Katja Pfeifer, Konstantinos Anastassiadis, and Gilbert Weidinger. "Dually Inducible TetON Systems for Tissue-Specific Conditional Gene Expression in Zebrafish." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 46 (November 16, 2010): 19933–38. doi: 10.1073/pnas.1007799107.
2. Shoji, Wataru, and Mika Sato-Maeda. "Application of Heat Shock Promoter in Transgenic Zebrafish." *Development, Growth & Differentiation* 50, no. 6 (August 2008): 401–6. doi: 10.1111/j.1440-169X.2008.01038.x.
3. Chen, Xianjun, Xue Wang, Zengmin Du, Zhengcai Ma, and Yi Yang. "Spatiotemporal Control of Gene Expression in Mammalian Cells and in Mice Using the LightOn System." *Current Protocols in Chemical Biology* 5, no. 2 (2013): 111–29. doi:10.1002/9780470559277.ch120267.
4. Motta-Mena, Laura B., Anna Reade, Michael J. Mallory, Spencer Glantz, Orion D. Weiner, Kristen W. Lynch, and Kevin H. Gardner. "An Optogenetic Gene Expression System with Rapid Activation and Deactivation Kinetics." *Nature Chemical Biology* 10, no. 3 (March 2014): 196–202. doi:10.1038/nchembio.1430.
5. Distel, Martin, Mario F. Wullimann, and Reinhard W. Köster. "Optimized Gal4 Genetics for Permanent Gene Expression Mapping in Zebrafish." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 32 (August 11, 2009): 13365–70. doi: 10.1073/pnas.0903060106.

6. Kikuchi, Y., A. Agathon, J. Alexander, C. Thisse, S. Waldron, D. Yelon, B. Thisse, and D. Y. Stainier. "Casanova Encodes a Novel Sox-Related Protein Necessary and Sufficient for Early Endoderm Formation in Zebrafish." *Genes & Development* 15, no. 12 (June 15, 2001): 1493–1505. doi:10.1101/gad.892301.
7. Kurita, Ryo, Hiroshi Sagara, Yutaka Aoki, Brian A. Link, Ken-ichi Arai, and Sumiko Watanabe. "Suppression of Lens Growth by alphaA-Crystallin Promoter-Driven Expression of Diphtheria Toxin Results in Disruption of Retinal Cell Organization in Zebrafish." *Developmental Biology* 255, no. 1 (March 1, 2003): 113–27.
8. Crosson, Sean, Sudarshan Rajagopal, and Keith Moffat. "The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains." *Biochemistry* 42, no. 1 (January 14, 2003): 2–10. doi:10.1021/bi026978l.
9. Shimizu-Sato, Sae, Enamul Huq, James M. Tepperman, and Peter H. Quail. "A Light-Switchable Gene Promoter System." *Nature Biotechnology* 20, no. 10 (October 2002): 1041–44. doi:10.1038/nbt734.
10. Ablain, Julien, Ellen M. Durand, Song Yang, Yi Zhou, and Leonard I. Zon. "A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish." *Developmental Cell* 32, no. 6 (March 23, 2015): 756–64. doi:10.1016/j.devcel.2015.01.032.
11. Gibson, Daniel G., Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison, and Hamilton O. Smith. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." *Nature Methods* 6, no. 5 (May 2009): 343–45. doi:10.1038/nmeth.1318.

12. Clark, Karl J., Mark D. Urban, Kimberly J. Skuster, and Stephen C. Ekker. "Transgenic Zebrafish Using Transposable Elements." *Methods in Cell Biology* 104 (2011): 137–49. doi: 10.1016/B978-0-12-374814-0.00008-2.
13. Mizoguchi, Takamasa, Heather Verkade, Joan K. Heath, Atsushi Kuroiwa, and Yutaka Kikuchi. "Sdf1/Cxcr4 Signaling Controls the Dorsal Migration of Endodermal Cells during Zebrafish Gastrulation." *Development* 135, no. 15 (August 1, 2008): 2521–29. doi:10.1242/dev.020107.
14. Maizel, Alexis, Daniel von Wangenheim, Fernán Federici, Jim Haseloff, and Ernst H.K. Stelzer. "High-Resolution Live Imaging of Plant Growth in near Physiological Bright Conditions Using Light Sheet Fluorescence Microscopy." *The Plant Journal* 68, no. 2 (October 1, 2011): 377–85. doi:10.1111/j.1365-313X.2011.04692.x.
15. Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. "NIH Image to ImageJ: 25 Years of Image Analysis." *Nature Methods* 9, no. 7 (July 2012): 671–75.
16. Peterson, Samuel M., and Jennifer L. Freeman. "RNA Isolation from Embryonic Zebrafish and cDNA Synthesis for Gene Expression Analysis." *Journal of Visualized Experiments : JoVE*, no. 30 (August 7, 2009). doi:10.3791/1470.

CHAPTER FIVE

Nodal Gradient Interpretation During Zebrafish Germ Layer Patterning

“Really, you should check with me before you start anything, we’re most likely already working on it.” - anonymous

“If there is a bug in the software, Anna will find it”
- Benjamin Schmid

SUMMARY

In order to form the complex structure of multicellular organisms it is essential for cells within an embryo to receive fate and positional information. An extensively studied source of positional information is the Nodal morphogen gradient. Although much progress has been made in understanding the components and outcomes of the Nodal signaling pathway, there is very poor quantitative understanding of how cells translate the duration and concentration of Nodal signaling into positional information. To address these questions, we employed Selective Plane Illumination Microscopy (SPIM) to image whole embryos *in vivo* with sub-cellular resolution. With this technology we were able to image and quantify Nodal morphogen gradient dynamics over time. Single cell tracking in conjunction with live Nodal target gene reporters allowed us to directly correlate Nodal signaling inputs with subsequent target gene induction and cell fate decisions. We combined these measurements with perturbations of Nodal signaling to elucidate how developing organisms interpret this key morphogen.

INTRODUCTION

A single egg, once fertilized, begins rapid and synchronous cell divisions that divide the enormous volume of zygote cytoplasm into numerous smaller cells. Further development requires a transition from a fairly homogenous population of cells to cells with different structures, movements, and fates. This period is characterized by the formation of the three germ layers, specification of the anterior-posterior and dorsal-ventral axis, morphogenetic movements, and finally specification of tissue and organ progenitors (1). Positional information within the embryo is essential for these complex patterning events (2), and a principal mechanism for specifying positional information is provided by morphogen gradients (3). Morphogens act directly on cells to produce specific cellular responses in a concentration-dependent manner. Several mechanisms have been proposed for how morphogen gradients are interpreted by cells, but lack of sophisticated tools for perturbing and measuring morphogen gradients limit our understanding of these key regulators of positional information.

Nodal is an extensively studied and developmentally important morphogen (4). Part of the TGF- β superfamily, its signaling is essential for germ layer and left-right patterning in vertebrates. Excellent genetic and developmental studies have uncovered the core components of the Nodal signaling pathway. In zebrafish, Nodal ligands are expressed at the margin and yolk syncytial layer (YSL) during the blastula stage and the resulting signaling gradient specifies germ layer fate along the animal-vegetal axis by inducing endoderm at high levels of Nodal signaling and mesoderm at low levels. The two zebrafish Nodal ligands, Cyclops (Cyc) and Squint (Sqt), have short- and long-range effects, respectively. The signaling pathway [5] is activated by a ligand binding to a type II TGF- β receptor, inducing interaction with a EGF-CFC co-receptor, One-eyed-pinhead (Oep), and the type I TGF- β receptor, TARAM-A. TARAM-A becomes phosphorylated and activated by the constitutively active type II receptor and induces the subsequent phosphorylation of the transcription factors Smad2 and Smad3. This

phosphorylation facilitates complex formation with the co-transcription factor Smad4. The Smad complexes accumulate in the nucleus, where they bind DNA and other transcription factors to regulate the expression of target genes. A few examples of the many Smad target genes include: *Cyc* and *Sqt* themselves, their antagonists, *Lefty1* and *Lefty2*, the zebrafish *Brachyury* homolog *no-tail* (*Ntl*), and the homeobox gene *gooseoid* (*Gsc*).

Despite this knowledge, we still lack a systems-level understanding of how this complex signaling cascade converts gradients of Nodal into proper positional information. This is mostly due to a lack in a quantitative understanding of: the *in vivo* dynamics of the gradient and the mechanisms by which signal concentration and duration are translated into positional information by responding cells. This knowledge is crucial to understanding how the Nodal ligand's linear input is converted into discrete stepped outputs, such as germ layer fate, and how the robustness and precision required of this event is achieved.

Response to morphogen gradients during development is a spatially and temporally dynamic process[6]. Despite this, many of the tools used to study Nodal signaling have relied on fixed time points. A comprehensive understanding of Nodal-dependent patterning requires quantitative measurements of the spatiotemporal dynamics of: the morphogen gradient, its intracellular signaling cascades, and the regulation of its downstream target genes. Accordingly, we combined: recent technologies in imaging, *in vivo* Nodal activity reporters and a quantitative approach to data collection and analysis to address: How is the Nodal morphogen gradient converted to the positional information responsible for discrete changes in gene expression that ultimately determine germ layer fate/patterning?

METHODS

Vector construction

All sub-cloning was done by Gibson assembly (7). Transgene plasmid mTol2-ubiq:GFP-Smad2 was created by separate PCR amplification of the ubiquitin promoter and GFP ORF and then cloned into pmTol2-ef1a:Venus-Smad2 (gift from Steve Harvey) cut with EcoRV and AgeI to remove the ef1a promoter and Venus ORF. Transgene plasmid mTol2-gsc:tRFP was created by separate PCR amplification of the gsc promoter (gift from Dirk Meyer) and the tagRFP ORF and then cloned into pmTol2. Transgene plasmid mTol2-sox17:tBFP was created by separate PCR amplification of the sox17 promoter and tagBFP ORF and then cloned into pmTol2.

Zebrafish strains

Tg(ubiq:GFP-Smad2) and Tg(gsc:tRFP) transgenic lines were created by injecting 20 pg of corresponding transgene plasmid DNA along with 100 pg of Tol2 transposase mRNA at the one-cell stage. Injected embryos were then sorted by fluorescence on d0, raised to adulthood and then screened for founders by outcrossing to wild-type.

Histone labeling

CF405-histone dimers were made as described in Shahian and Narlikar, 2012 *Methods Mol Bio* (8), with the following modifications: Cy5-maleimide was replaced with CF405-maleimide, once H2A was labeled with CF405 dye it was refolded with equal concentration of H2B to form H2A/H2B histone dimers instead of full histone octamers, and labeled histone dimers were HPLC purified to remove any free dye still present after initial filtering through a Microcon-10 filter. Bright, uniform labeling of histones was achieved by injection of 5ng of CF405-histone dimer directly into single celled embryos.

Microscopy and Imaging Setup

SPIM microscope was set up as described in Schmid, et. al. 2013 Nat. Commun (9). Embryos were mounted in a cleaned FEP tube (inner diameter: 2.0 mm, wall thickness: 0.5 mm) with 1.5% low-melt agarose in E3 medium and a 1:20,000 dilution of fluorescent microspheres (F-Y 050, Millipore), for bead-based-registration. The embryos with their chorions intact were positioned on top of each other with minimum gap between them. Once the FEP tube was mounted inside the SPIM chamber, bead-based-registration was used to register the samples in their 0 and 45 degree angle positions. A data acquisition mask was then set to acquire a relevant 320um radius shell of information around the embryo. Z-stacks composed of 2um steps were taken every 2 minutes at both 0 and 45 degree rotations. For gsc final reporter imaging, GFP-Smad2/gsc:tRFP embryos injected with 5ng CF405-histone dimer were imaged with the 405nm channel for the duration of the time-lapse, 200 time points (laser power ramped up to a final of 150% of original power), the 488nm channel was imaged for the first 116 timepoints, and the 561nm channel for the final time point. For sox17 final reporter imaging, GFP-Smad2 embryos injected with 100pg H2B-mCherry mRNA and 15pg of sox17:tBFP DNA were imaged with the 405nm channel for the final time point, the 488nm channel for the first 116 time points and the 561nm channel for the duration of the time-lapse, 200 timepoints (laser power ramped down to a final of 60% of original power).

REFERENCES

1. Schier, Alexander F., and William S. Talbot. "Molecular Genetics of Axis Formation in Zebrafish." *Annual Review of Genetics* 39 (2005): 561–613. doi:10.1146/annurev.genet.37.110801.143752.
2. Wolpert, L. "Positional Information and the Spatial Pattern of Cellular Differentiation." *Journal of Theoretical Biology* 25, no. 1 (October 1969): 1–47.
3. Tabata, Tetsuya, and Yuki Takei. "Morphogens, Their Identification and Regulation." *Development (Cambridge, England)* 131, no. 4 (February 2004): 703–12. doi:10.1242/dev.01043.
4. Schier, Alexander F. "Nodal Morphogens." *Cold Spring Harbor Perspectives in Biology* 1, no. 5 (November 2009). doi:10.1101/cshperspect.a003459.
5. Weng, Wei, and Derek L. Stemple. "Nodal Signaling and Vertebrate Germ Layer Formation." *Birth Defects Research. Part C, Embryo Today: Reviews* 69, no. 4 (November 2003): 325–32. doi:10.1002/bdrc.10027.
6. Kutejova, Eva, James Briscoe, and Anna Kicheva. "Temporal Dynamics of Patterning by Morphogen Gradients." *Current Opinion in Genetics & Development* 19, no. 4 (August 2009): 315–22. doi:10.1016/j.gde.2009.05.004.
7. Gibson, Daniel G., Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison, and Hamilton O. Smith. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." *Nature Methods* 6, no. 5 (May 2009): 343–45. doi:10.1038/nmeth.1318.
8. Shahian, Tina, and Geeta J. Narlikar. "Analysis of Changes in Nucleosome Conformation Using Fluorescence Resonance Energy Transfer." *Methods in Molecular Biology (Clifton, N.J.)* 833 (2012): 337–49. doi:10.1007/978-1-61779-477-3_20.
9. Schmid, Benjamin, Gopi Shah, Nico Scherf, Michael Weber, Konstantin Thierbach, Citlali Pérez Campos, Ingo Roeder, Pia Aanstad, and Jan Huisken. "High-Speed Panoramic Light-

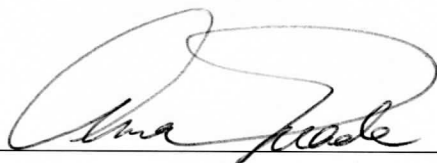
Sheet Microscopy Reveals Global Endodermal Cell Dynamics.” Nature Communications 4
(July 25, 2013). doi:10.1038/ncomms3207.

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.

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 _____



Date _____

06/08/15