Using synthetic signaling to reprogram plants.

Arjun Khakhar

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Eric Klavins, Chair

Georg Seelig, Chair

Herbert Sauro

Program Authorized to Offer Degree:

Department of Bioengineering

University of Washington

Abstract

Using synthetic signaling to reprogram plants

Arjun Khakhar

Chair of the Supervisory Committee:

Eric Klavins, Georg Seelig

Department of Electrical engineering

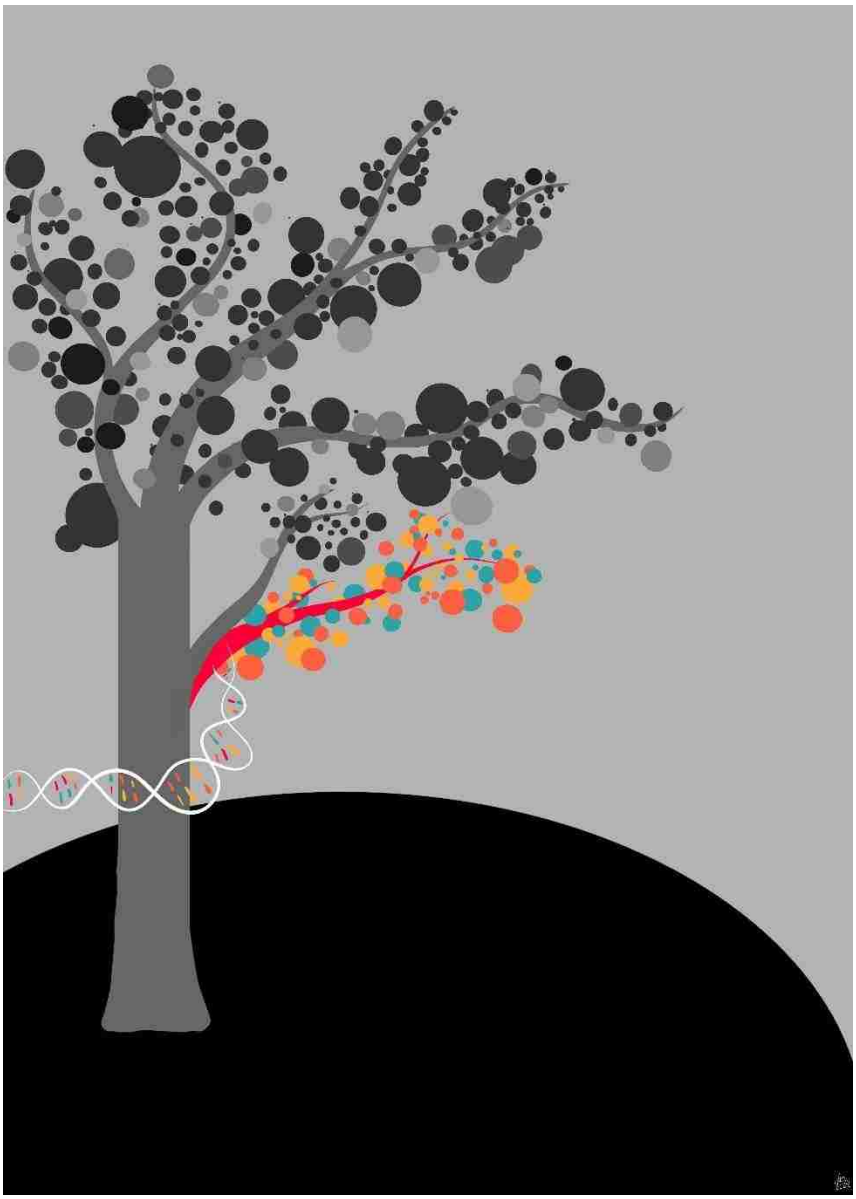
To engineer plants that can address the environmental challenges posed to agriculture we need to be able to rationally design their developmental and stress response phenotypes. To achieve this larger goal, we first need to understand the native mechanisms that control these processes. More specifically, we need to understand the mechanisms behind the flow of information, i.e. some change of state that can propagate through a biological system like a signal through a circuit, in biological systems, as this process is the driver for most multicellular development and stress response phenotypes. Information flow occurs at every level in biological systems, whether from one sub-domain of a protein to another, or from a caterpillar to the plant it is eating. The form this information takes also varies, including phosphorylation events and the production of small molecule hormones. A great deal of work has been done to understand the fundamental molecular mechanisms driving these signal transduction systems and how coordinated signaling events can lead to cell and organism level phenotypes.

This understanding can be leveraged to design novel synthetic systems to either alter or replace these mechanisms to achieve the phenotypes of interest. The field of synthetic biology has made great strides in both developing strategies to reengineer native signaling machinery as well as designing totally synthetic signaling systems. This work has, until recently, been limited to microbes due to their fast generation times which makes prototyping new synthetic signaling systems a more rapid process than is possible in multicellular eukaryotes like plants and animals. However, it has provided a great deal of transferable knowledge and parts, which have made translating these systems into larger organisms a less daunting prospect. There are still major challenges to applying synthetic biology in higher eukaryotes such as relatively more difficult transformation process and heterogeneity across tissues, but we are now poised to leverage these powerful tools for biological engineering.

Here I describe how I used synthetic signaling systems at a range of scales to both learn more about native plant signaling as well as to develop programmable phenotypes in plants. The first few chapters cover work to prototype synthetic signaling systems in vitro and in the more tractable model organism, Saccharomyces cerevisiae. The later chapters describe how we used some of these tools to learn more about the natural signaling systems in the model plant, Arabidopsis thaliana, and to re-program its development and stress response phenotypes.

USING SYNTHETIC SIGNALLING TO REPROGRAM PLANTS

Arjun Khakhar



MAY 1, 2018

THESIS

This work is dedicated to my advisers Georg, Eric and Jennifer
who believed in me enough to let me follow my own path.

Chapter Page

Preface 3

Publication list 6

Introduction 7

Chapter 1 - Characterizing intra-molecular information flow: "A fluorescent readout for the oxidation state of electron transporting proteins in cell free settings"
[Collaborators: Sergii Poshkailov] 13

Chapter 2 - Engineering inter-molecular information flow: "Rewiring MAP kinases in Saccharomyces cerevisiae to regulate novel targets through ubiquitination"
[Collaborators: Ben Groves] 17

Chapter 3 - Engineering flow of information flow from one cell to another: "Cell–cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors"
[Collaborators: Nick Bolten] 49

Chapter 4 - Characterizing the flow of information in a cell: "Using massively parallel assays and machine learning approaches to elucidate ARF-promoter interaction rules"
[Collaborators: Amy Lanctot, Mary Galli] 70

Chapter 5 - Characterizing and re-engineering the flow of information between tissues: "Reprogramming the plant form: Using synthetic hormone signaling to study and reengineer development"
[Collaborators: Andrew Lemmex, Alex Leydon] 78

Chapter 6 - Engineering the flow of information from one organism to another: "Pest triggered immunity: Implementing synthetic hormone-based signal transduction to regulate insect defenses in plants"
[Collaborators: Andrew Lemmex] 106

Concluding thoughts and future work 112

Preface

Food security, the corner stone of a stable and productive society, is being threated across the globe by rapidly changing climates[6], decreasing soil quality[6], and an increasing population[2]. An absurdly large subsection of humanity, one in every eight people, does not have access to sufficient food. The people who face these scarcities are generally people that have the least capacity to deal them, such as children, the elderly and people living near or below the poverty line[3]. Additionally, it has been shown that hunger caused by poverty can reinforce that same poverty, driving societal inequality[4,5]. While a lot of progress has been made in tackling this problem in some parts of the world such as Latin America, other parts of the world, for example South-East Asia, are lagging far behind with approximately one in six people still going hungry[3]. While I had the benefit of growing up having relatively wealthy parents, I was exposed to what hunger looked like on a daily basis, in the faces of the urban poor that populate the streets of Mumbai. The horrendous injustice of that situation, of a child slowly starving on the pavement while I walked past him to my school bus, simply because my parents were wealthy and his were not, has stayed with me to this day. It motivates me to think of ways to try and redress that wrong. To create an environment of plenty where there is such an excess of resources that these disparities need not exist.

Unfortunately, the industrialization of the past hundred years coupled with unsustainable agricultural practices have pushed our environment in opposite direction of what is needed to create this new world of plenty[6]. Droughts, climate change, pesticide resistant pests and soil erosion are all leading to decreases in agricultural yields[6] that we cannot afford. It is projected that to meet the needs of our expanding population, we need to double crop production by 2050. However, projections show that current trends in crop yield increases are going to miss the mark by a large margin [7]. This means the extreme injustice played out on the pavements of Mumbai has the capacity to spread to places that have been largely unaffected by these scarcities, like more affluent western countries.

These are incredibly complex problems which will require multifaceted solutions including improving distribution and storage infrastructures for food supplies as well as reducing corruption in the regulation of food prices. However, one of the key challenges that we must address is how to rapidly increase the productivity of our crop plants and how to make them more resistant to stresses in their environment. One way to do this would be to build a set of tools that would let us rationally alter their biology to control the way plants grow and respond to their environment. This goal can be abstracted to a more general academic question; how can the flow of information be controlled in biological systems? Here I use the term ‘information’ not in the traditional information theory sense, but rather to denote some change of state that can propagate through biological systems like a signal through a circuit. It is on this fundamental process that development and stress response behaviors of all multicellular organisms, including plants, are built.

There has been a great deal of theoretical and biochemical work already done to understand and model the molecular signaling circuitry that underpins the patterning processes in plant development. It has highlighted the importance of a class of chemicals called phytohormones, which form a chemical circuitry that coordinates these processes by regulating certain genes. Additionally, the field of synthetic biology has made considerable advances in the design of tools that facilitate synthetic signaling. However, there has been little done to use this understanding of development in conjunction with the tools offered by synthetic biology to rationally engineer development in a multicellular organism. In this work I demonstrate a set of tools to that can be used to engineer the flow of information at the molecular, cellular, and tissue levels. I then go on to describe a plan to use these tools to rationally perturb developmental pathways with the goal of creating agriculturally productive traits in plants in much more rapid timescales than traditional methods such as would allow.

One of the tools I designed is a novel class of hormone-responsive cas9-based repressors (HACRs) to facilitate reprogramming of the phytohormone circuitry in plants. This tool was first prototyped in the microbe *Saccharomyces cerevisiae* (yeast) that had been engineered to express the signaling pathway for the growth phytohormone, auxin. I demonstrated the HACRs response to auxin and their broad tunability. I then moved this tool into the model plant *Arabidopsis thaliana*. Here I demonstrated its capacity to be easily reprogrammed by simply incorporating a new gRNA into the plant. I also showed how HACRs that are programmed to regulate a reporter gene can be used to visualize phytohormone distributions across plant tissues to further study how the phytohormone circuitry regulates plant growth and behavior.

Finally, I also demonstrated a proof of concept for how this tool could be used to reprogram the plant body plans. There are several mathematical models describing how the pattern of auxin accumulation affects the body plan of plants, specifically the shoot architecture. They identify an important control parameter, the auxin activated expression of PIN1, that is responsible for setting up this pattern. We deployed a HACR to decrease the auxin activated expression of PIN1, which the models predict would lead to a reduced amount of branching and noise in the stereotyped distribution fruits on a branch. Both these phenotypes have agricultural significance, as lowered branching would mean more plants could be fit on a field, and a less noisy distribution of fruits would allow more efficient mechanized harvesting. We were able to demonstrate these predicted phenotypes in lines where the HACR was targeted to regulate PIN1. In addition, we were able to show how this platform could be extended to several other phytohormones, such as jasmonate and gibberellin, and are currently working utilizing this expanded platform to engineer several other agriculturally relevant traits such as flowering time, pest responses and dwarfing.

In addition to HACRs I also developed tools to facilitate the reprogramming of other flows of information such as electron flow in P450s, an important enzyme in secondary metabolism. I also developed an approach to re-wire MAP kinase cascades, which play important roles in plant stress response signaling. Finally, I also have done work on elucidating the design rules for hormone responsive promoters in yeast and using semi-*in vitro* data and deployed them to reprogram shoot architectures in *A. thaliana*.

Thus, this work serves to outline the general approach of engineering the flow of information at the molecular, cellular and tissue scales to engineer the development and behavior of multicellular organisms. It focuses on validating these tools in yeast and then applying them in a model plant system to demonstrate how they could be used to design better crops. It is my hope that this bottom up approach to rationally design new developmental and stress response phenotypes provides a

framework that could be used for rapid development of more robust and productive crops for a plentiful future.

References

1. https://www.wfp.org/hunger/causes
2. http://www.prb.org/pdf14/2014-world-population-data-sheet_eng.pdf
3. http://www.worldhunger.org/2015-world-hunger-and-poverty-facts-and-statistics/
4. Barrett, Christopher B., and Leah EM Bevis. "The self-reinforcing feedback between low soil fertility and chronic poverty." *Nature Geoscience* 8.12 (2015): 907-912.
5. Implementing agriculture for development : World Bank Group agriculture action plan (2013-2015)
6. Ray, Deepak K., et al. "Recent patterns of crop yield growth and stagnation."*Nature communications* 3 (2012): 1293.
7. Ray, Deepak K., et al. "Yield trends are insufficient to double global crop production by 2050." *PloS one* 8.6 (2013): e66428.

Publication list

Paper title	Contribution	Status
A fluorescent readout for the oxidation state of electron transporting proteins in cell free settings	Secondary author – Built and characterized P450 sensor and wrote part of the paper	Published - 2016
Rewiring MAP kinases in Saccharomyces cerevisiae to regulate novel targets through ubiquitination	Co-first author – Built and characterized various MAPK cascades and wrote and edited the paper with co-authors	Published - 2016
Cell–cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors	Co-first author – Built and characterized ADCTF platform and helped characterizing the cell-cell communication and wrote and edited the paper with co-authors	Published - 2015
Synthetic hormone-responsive transcription factors can monitor and re-program plant development	First author – Built and characterized various HACR plant lines and wrote and edited the paper with co-authors	Published - 2018

Introduction

To engineer plants that can address the environmental challenges posed to agriculture we need to be able to rationally design their developmental and stress response phenotypes[1,2]. To achieve this larger goal, we first need to understand the native mechanisms that control these processes. More specifically, we need to understand the mechanisms behind the flow of information, i.e. some change of state that can propagate through a biological system like a signal through a circuit, in biological systems, as this process is the driver for most multicellular development[3] and stress response phenotypes[4,5]. Information flow occurs at every level in biological systems, whether from one sub-domain of a protein to another, or from a caterpillar to the plant it is eating. The form this information takes also varies, including phosphorylation events and the production of small molecule hormones. A great deal of work has been done to understand the fundamental molecular mechanisms driving these signal transduction systems and how coordinated signaling events can lead to cell and organism level phenotypes[3,4,5].

This understanding can be leveraged to design novel synthetic systems to either alter or replace these mechanisms to achieve the phenotypes of interest. The field of synthetic biology has made great strides in both developing strategies to reengineer native signaling machinery as well as designing totally synthetic signaling systems[6,7]. This work has, until recently, been limited to microbes due to their fast generation times which makes prototyping new synthetic signaling systems a more rapid process than is possible in multicellular eukaryotes like plants and animals. However, it has provided a great deal of transferable knowledge and parts, which have made translating these systems into larger organisms a less daunting prospect. There are still major challenges to applying synthetic biology in higher eukaryotes such as relatively more difficult transformation process and heterogeneity across tissues, but we are now poised to leverage these powerful tools for biological engineering.

Here I describe how I used synthetic signaling systems at a range of scales to both learn more about native plant signaling as well as to develop programmable phenotypes in plants. The first few chapters cover work to prototype synthetic signaling systems *in vitro* and in the more tractable model organism, *Saccharomyces cerevisiae*. The later chapters describe how we used some of these tools to learn more about the natural signaling systems in the model plant, *Arabidopsis thaliana,* and to re-program its development and stress response phenotypes.

One of the smallest scales at which information flows in biological systems is between two parts of a single protein. This occurs fairly extensively in nature, with changes in redox state or pH triggering conformational shifts in proteins[8,9]. Being able to rationally design such events would give us exquisite control over how proteins behave. The first chapter discusses how I went about doing this by engineering a novel fusion protein consisting of a mono-oxygenase, cytochrome P450, and a fluorescent protein, GFP. We demonstrate how information about redox state of the heme core of the P450 'flows' into the GFP fused to by deforming its fluorophore and quenching it. Thus the quenching of the GFP signal can be used as a readout for turnover of the P450 *in vitro*. We also propose how this technology could be further refined to potentially build a new class of biosensors for P450 substrates. This is particularly exciting as P450s are a very large family of enzymes with a wide range of substrates that include several important drug compounds[10,11].

In chapter two I zoom out from a single protein and demonstrate engineering the flow of information on a larger scale, namely from one protein to another. There are many examples of this in biology in the form of post-translational modifications such as ubiquination or sumolation, among others[12]. In chapter

two, I describe work demonstrating how mitogen activated protein kinases (MAPKs) can be rewired to regulate any protein of choice through phosphorylation mediated degradation. This work demonstrates the minimum set of modular protein motifs required to create a functional link between a MAPK and a substrate and gives us insight on how such connections may form and break during evolution and disease. We also demonstrate how this rewiring approach can be applied to the MAPK in the yeast mating cascade, Fus3, to create new signaling topologies such as incoherent feed-forward loops. This allows us to rationally re-engineer how yeast cells interpret and respond to a mating pheromone signal in their environment, such as turning what would normally be a signal amplification behavior into a concentration-based bandpass filter. Thus, by engineering the flow of information from outside to inside the cell we are able to engineer how populations of cells can respond to environmental signals.

Zooming out further from the level of proteins to the level of cells, in chapter three I describe work on engineering the flow of information from one cell to another. The capacity to engineer the flow of information between cells would allow us to begin to engineer coordinated multicellular behaviors using normally unicellular organisms—giving insight into the general principles that govern complex multicellular processes like development[13]. Additionally, this technique has potential practical utility in the construction of coordinated synthetic microbial consortia for the more efficient production of metabolic products[14]. Chapter three goes into detail on how we engineered a synthetic cell to cell communication system based on the plant growth hormone, auxin, in yeast. A sender yeast strain was designed to biosynthesize auxin in a tunable fashion. A library of receiver strains which had a novel class of auxin-degradable CRISPR-based transcription factors (ADCTFs) that linked an auxin signal to the transcriptional regulation of a gene of interest were also designed. We demonstrate how we can get different responses to the same signal from the sender cell based on which ADCTF variant was expressed in the receiver strain. This is analogous to a proposed mechanism of auxin response in plants, where cells from different tissues respond to similar auxin signals differently due to differences in the signaling machinery expressed in these cells[15,16].

While engineering the behavior of single cells is an important goal for synthetic biology, my own interests lie in engineering the biological machines which already form a fundamental pillar of human society, namely plants. In plants, developmental events are spatially and temporally regulated, in large part, by the plant growth hormone auxin. The transport of auxin from one tissue to another during development and the signal transduction system that links auxin signals to transcriptional regulation of genes are the two core molecular mechanisms behind this process[17,18]. To be able to rationally reengineer the flow of information between tissues, such as the fluxes of plant hormones, we need to a more complete understanding of how these processes regulate development[18] and stress responses[19]. Auxin regulates both the expression and polarity of a host of auxin import and export pumps which set up the spatiotemporal pattern of auxin distribution in the plant. In each cell a complex network of interacting transcription factors and transcriptional co-repressors, the ARFs and AUX/IAAs respectively, determine how a particular auxin signal will be translated into a transcriptional change[17]. It is thought that based on tissue specific differences in the expression of ARFs and AUX/IAAs[15], similar auxin signals are interpreted and responded to differently.

While a great deal of work has been done to de-convolute the role of specific ARFs in particular developmental events[20,21,22], the rules that govern ARF-promoter interactions and which ARFs regulate which genes and when are still unclear. This knowledge is essential if we hope to reengineer plant development, as it will elucidate control points that synthetic signaling systems could be wired to. In chapter four I detail the results collected so far for ongoing work to use a massively parallel reporter assay to determine these rules and build a predictive model for ARF-promoter interactions. The assay

involves characterizing the expression from a large library of synthetic promoters in a yeast background where the auxin response circuit has been reconstituted[24]. Using the yeast system allows us to functionally isolate single ARFs and pairs of ARFs in a way that a plant system would not. Additionally, by using a library of synthetic promoters we can assay a much larger sequence space than is present in the genome. While synthetic promoters afford us access to a broader sequence space from which to learn ARF-promoter interaction rules from, sufficiently large plant genomes might provide a diverse enough space to train predict machine learning models. The use of convolutional neural networks has shown great promise in discovering the rules behind other complex DNA-protein interactions[23]. I applied this approach to ARF DAP-seq data generated from the maize genome. These models were able to achieve a high degree of predictability, and we were able to demonstrate how ARFs had binding rules that were broadly similar across a clade but differ significantly from clade to clade. We were also able to demonstrate how these models could predict the effects of SNPs in promoters on ARF binding, and how this could be used to predict phenotypic differences, such as herbivore resistance, across different maize land races.

Synthetic promoters have a lot of promise as tools for engineering biology, but there is still a great deal more we need to learn before they can be as programable as a much more well studied synthetic biology workhorse, the synthetic transcription factor. In chapter five describe hormone-responsive cas9-based repressors (HACRs), which we deployed to facilitate studying and reprogramming of the phytohormone circuitry in plants. I demonstrated its capacity to be easily reprogrammed by simply incorporating a new gRNA into the plant. While there is a broad consensus on the spatiotemporal distribution of auxin during development, the current tools available to study these patterns have certain drawbacks that make them non-ideal to study these processes[25]. These include either being indirect or inverted readouts of auxin as well as having limited sensitivity. In chapter five I describe a how I plan to apply the HACR platform in the model plant *Arabidopsis thaliana* to visualize the auxin fluxes during development. I showed how HACRs that are programmed to regulate a reporter gene can be used to visualize phytohormone distributions across plant tissues to further study how the phytohormone circuitry regulates plant growth and behavior. Finally, I also demonstrated a proof of concept for how this tool could be used to reprogram the plant body plans. There are several mathematical models describing how the pattern of auxin accumulation affects the body plan of plants, specifically the shoot architecture[26-30]. They identify an important control parameter, the auxin activated expression of PIN1, that is responsible for setting up this pattern. We deployed a HACR to decrease the auxin activated expression of PIN1, which the models predict would lead to a reduced amount of branching and noise in the stereotyped distribution fruits on a branch. Both these phenotypes have agricultural significance, as lowered branching would mean more plants could be fit on a field, and a less noisy distribution of fruits would allow more efficient mechanized harvesting. We were able to demonstrate these predicted phenotypes in lines where the HACR was targeted to regulate PIN1. In addition, we were able to show how this platform could be extended to several other phytohormones, such as jasmonate and gibberellin, and are currently working utilizing this expanded platform to engineer several other agriculturally relevant traits such as flowering time, pest responses and dwarfing.

Chapter five focused on how the flow of information between tissues, in the form of auxin, could be engineered to alter development. In chapter six I zoom out to the scale of organisms and propose how the flow of information between organisms could be engineered to design new insect herbivory responses in plants[31]. I describe ongoing work using a jasmonate-degradable HACR (JA-HACR) to connect herbivory associated fluxes of the plant hormone jasmonate iso-leucine to the expression of insect specific toxins[32].

Through this body of work, I hope to demonstrate how engineering the flow of information at several scales can be used as a strategy to re-engineer multicellular development and stress response. It is my hope that the work described here will be impactful in the still sparse field of multicellular synthetic biology by demonstrating how synthetic signal transduction systems can be prototyped relatively rapidly in model microbes before being applied to reengineer macro scale phenotypes in multicellular organisms.

References

1. Bengough, G., McKenzie, Hallett & Valentine. Root elongation, water stress, and mechanical impedance: a review of limiting stresses and beneficial root tip traits. *J Exp Bot* **62,** 59–68 (2011).
2. JUROSZEK & von TIEDEMANN. Plant pathogens, insect pests and weeds in a changing global climate: a review of approaches, challenges, research gaps, key studies and concepts. *J Agric Sci* **151,**163–188 (2012).
3. Sparks, Erin, Guy Wachsman, and Philip N. Benfey. "Spatiotemporal signalling in plant development." *Nature Reviews Genetics* 14.9 (2013): 631-644.
4. Wasternack, C., and B. Hause. "Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany." *Annals of botany* 111.6 (2013): 1021-1058.
5. Osakabe, Yuriko, et al. "Sensing the environment: key roles of membrane-localized kinases in plant perception and response to abiotic stress." *Journal of experimental botany* 64.2 (2013): 445-458.
6. Gordley, Russell M., Lukasz J. Bugaj, and Wendell A. Lim. "Modular engineering of cellular signaling proteins and networks." *Current Opinion in Structural Biology* 39 (2016): 106-114.
7. Slusarczyk, Adrian L., Allen Lin, and Ron Weiss. "Foundations for the design and implementation of synthetic genetic circuits." *Nature Reviews Genetics*13.6 (2012): 406-420.
8. Kramer, Jessica R., and Timothy J. Deming. "Glycopolypeptides with a redox-triggered helix-to-coil transition." *Journal of the American Chemical Society*134.9 (2012): 4112-4115.
9. Damberger, Fred F., et al. "Pheromone discrimination by a pH-tuned polymorphism of the Bombyx mori pheromone-binding protein." *Proceedings of the National Academy of Sciences* 110.46 (2013): 18680-18685.
10. Zanger, Ulrich M., and Matthias Schwab. "Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation." *Pharmacology & therapeutics* 138.1 (2013): 103-141.
11. Hamberger, Björn, and Søren Bak. "Plant P450s as versatile drivers for evolution of species-specific chemical diversity." *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1612 (2013): 20120426.

12. Beltrao, Pedro, et al. "Evolution and functional cross-talk of protein post-translational modifications." *Molecular systems biology* 9.1 (2013): 714.

13. Shapiro, James A. "Thinking about bacterial populations as multicellular organisms." *Annual Reviews in Microbiology* 52.1 (1998): 81-104.

14. Brenner, Katie, Lingchong You, and Frances H. Arnold. "Engineering microbial consortia: a new frontier in synthetic biology." *Trends in biotechnology* 26.9 (2008): 483-489.

15. Brady, Siobhan M., et al. "A high-resolution root spatiotemporal map reveals dominant expression patterns." *Science* 318.5851 (2007): 801-806.

16. Wenzel, Carol L., et al. "Dynamics of MONOPTEROS and PIN-FORMED1 expression during leaf vein pattern formation in Arabidopsis thaliana." *The Plant Journal* 49.3 (2007): 387-398.

17. Teale, William D., Ivan A. Paponov, and Klaus Palme. "Auxin in action: signalling, transport and the control of plant growth and development." *Nature Reviews Molecular Cell Biology* 7.11 (2006): 847-859.

18. Lavenus, Julien, et al. "Lateral root development in Arabidopsis: fifty shades of auxin." *Trends in Plant Science* 18.8 (2013): 450-458.

19. Cheong, Yong Hwa, et al. "Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis." *Plant Physiology* 129.2 (2002): 661-677.

20. Overvoorde, Paul, Hidehiro Fukaki, and Tom Beeckman. "Auxin control of root development." *Cold Spring Harbor perspectives in biology* 2.6 (2010): a001537.

21. Guilfoyle, Tom J., and Gretchen Hagen. "Auxin response factors." *Current opinion in plant biology* 10.5 (2007): 453-460.

22. Hardtke, Christian S., and Thomas Berleth. "The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development." *The EMBO journal* 17.5 (1998): 1405-1411.

23. Quang, Daniel, and Xiaohui Xie. "DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences." *Nucleic acids research* (2016): gkw226.

24. Pierre-Jerome, Edith, et al. "Recapitulation of the forward nuclear auxin response pathway in yeast." *Proceedings of the National Academy of Sciences* 111.26 (2014): 9407-9412.

25. Wells, Darren M., et al. "Biosensors for phytohormone quantification: challenges, solutions, and opportunities." *Trends in plant science* 18.5 (2013): 244-249.

26. Traas, Jan. "Phyllotaxis." *Development* 140.2 (2013): 249-253.

27. Smith, Richard S., et al. "A plausible model of phyllotaxis." *Proceedings of the National Academy of Sciences* 103.5 (2006): 1301-1306.

28. Jönsson, Henrik, et al. "An auxin-driven polarized transport model for phyllotaxis." *Proceedings of the National Academy of Sciences* 103.5 (2006): 1633-1638.

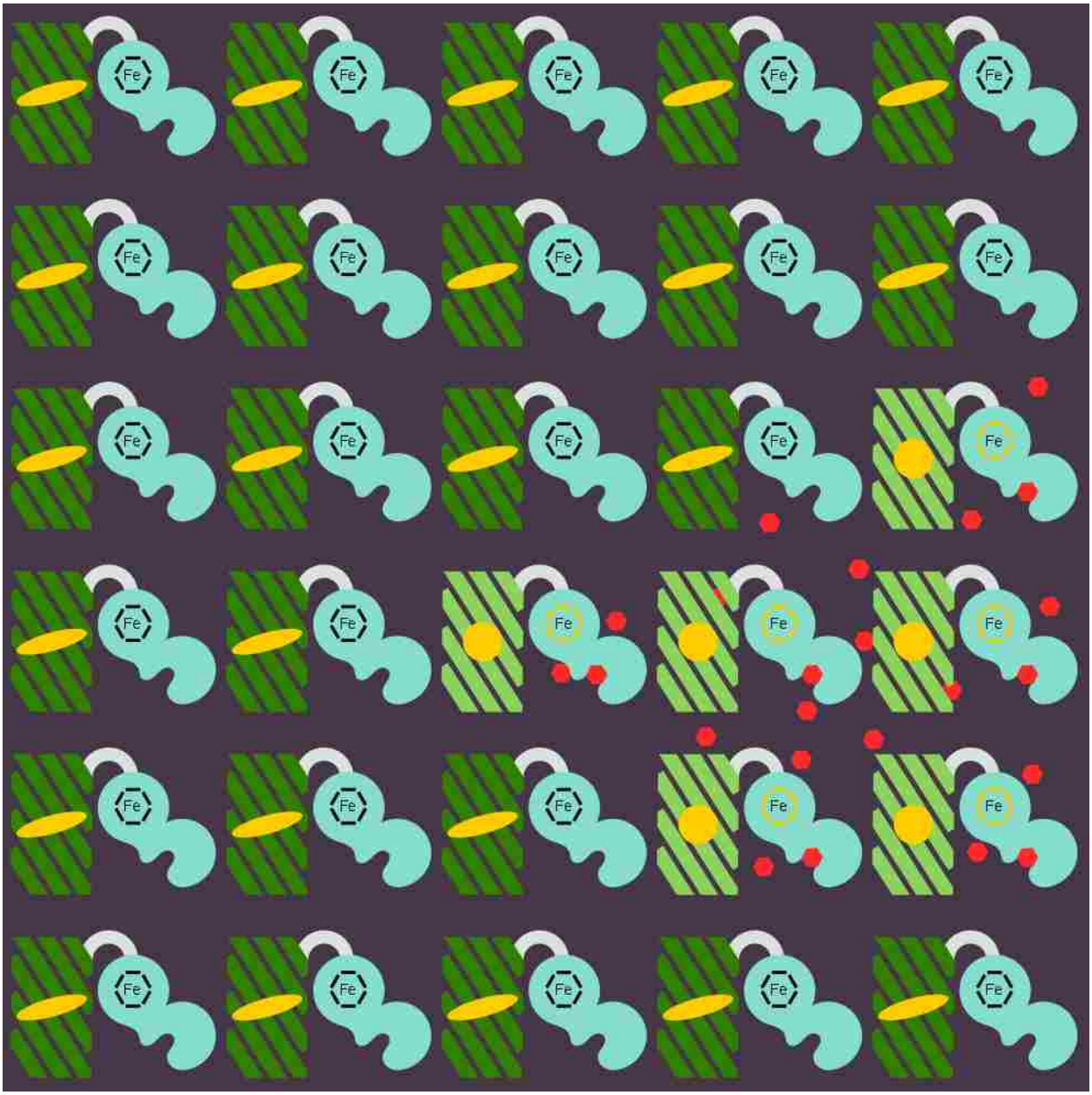
29. Prasad, Kalika, et al. "Arabidopsis PLETHORA transcription factors control phyllotaxis." *Current Biology* 21.13 (2011): 1123-1128.

30. Fozard, J. A., J. R. King, and M. J. Bennett. "Modelling auxin efflux carrier phosphorylation and localization." *Journal of theoretical biology* 319 (2013): 34-49.

31. Howe, Gregg A., and Georg Jander. "Plant immunity to insect herbivores."*Annu. Rev. Plant Biol.* 59 (2008): 41-66.

32. Lövei, Gabor L., David A. Andow, and Salvatore Arpaia. "Transgenic insecticidal crops and natural enemies: a detailed review of laboratory studies." *Environmental entomology* 38.2 (2009): 293-306.

Chapter 1 - Characterizing intra-molecular information flow: "A fluorescent readout for the oxidation state of electron transporting proteins in cell free settings"



Abstract

The smallest unit of information flow, as defined previously, that is useful to think about from a biological perspective is the flow of information from one part of a molecule to another. A classic example of this in biology is conformation change in proteins as a result of an environmental change or some sort of post translational modification. Here I describe work that focused on studying the flow of information in a bacterial enzyme, cytochrome P450, that was fused to green fluorescent protein (GFP). The goal of this work was to design a system whereby the redox state of the P450 enzyme could be read out via effect it had on GFP's fluorescence through the deformation of its chromophore. Cytochrome P450s represent a major class enzyme catalysts in plants, animals and microbes. The ability to accurately read out their activity in a manner that has the capacity to be scaled up and parallelized provides a tool that could potentially be used to screen libraries of these proteins for the discovery of novel enzymatic behavior or biosensors. This work was published as part of a larger publication that looked at using this approach to assay the activity of other electron transport proteins.

Results

GFP-P450 Fusion Proteins Exhibit Redox State Dependent Fluorescent Changes

CYP is a heme iron containing protein that changes its redox state during the catalytic cycle. We set out to investigate whether CYP can influence GFP fluorescence through its altered redox state. We created CYP106A2-GFP fusion proteins and observed fluorescence quenching for both N and C terminal fusions of CYP to GFP (Figure 5A). This quenching was partially relieved when electrons were supplied to the CYP (Figure 5B and Figure S3). The reduced state time was linearly dependent on NADPH concentration. However, the observed absolute fluorescence changes were considerably smaller than those measured for GFP-Adx fusions which had been previously built and tested (around 3% rather than 21%). These smaller changes are likely due to a less favorable configuration of the CYP charge center relative to the GFP fluorophore, possibly a result of the large size of CYP106A2 compared to Adx. Still, this result suggests that direct CYP-GFP fusions can be used to report on CYP activity in real time. This is especially useful for CYPs that cannot accept electrons from Adx or similar nonnative transporters.

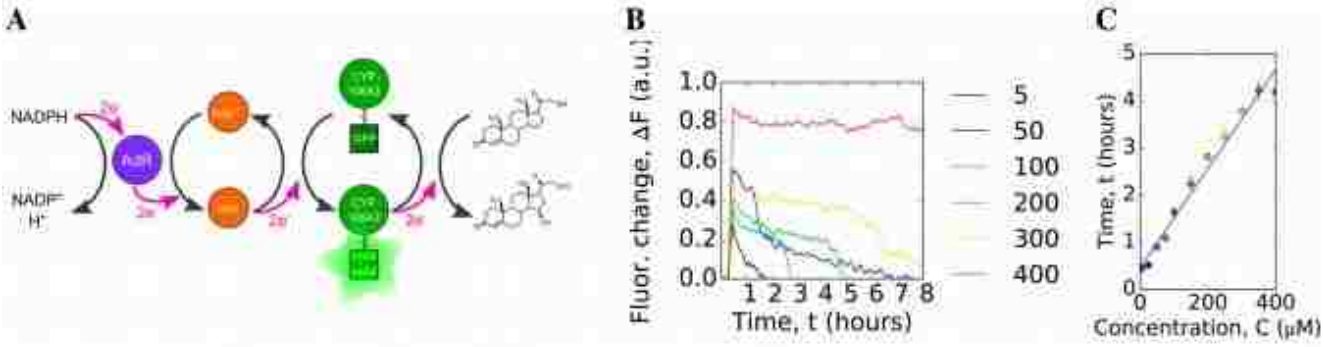


Figure 5. Cytochrome P450 can directly influence GFP fluorescence in a redox state dependent manner. (A) The reaction mixture contains Adx, NADPH, the GFP-CYP106A2 fusion and 21-hydroxyprogesterone as a substrate. Reactions are performed with varying amounts of NADPH. The reaction is activated by addition of AdR. (B) After addition of AdR, CYP106A2 gets reduced, and GFP fluorescence increases. The legend shows the amount of NADPH in μM used for each reaction. Concentrations of other reactants: 200 μM NADPH, 0.4 μM AdR, 4 μM Adx, 5 μM CYP106A2-GFP, 200 μM 21-hydroxyprogesterone. (C) Similarly to the GFP-Adx fusion, the reduced state time is linearly dependent on the NADPH concentration.

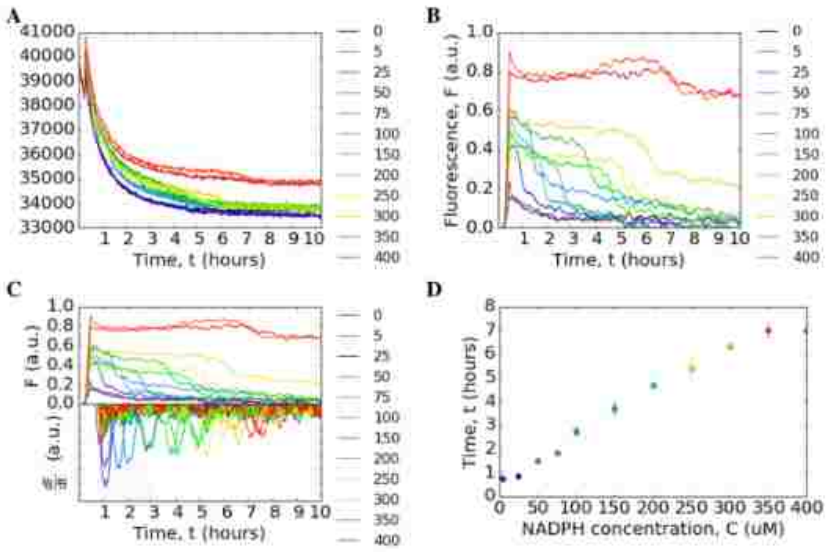


Figure S3: Workflow for processing kinetic experimental data for NADPH titration while measuring fluorescence of CYP106A2-GFP fusion. The figure structure repeats the Fig. S1. This figure supplements the Fig. 2. (**A**) Raw data have baseline shift, which needs to be corrected. (**B**) Correcting baseline shift by subtracting background measurement from the real measurement. The rolling mean operation over 12 data points was performed prior to plotting in order to decrease the noise of the signal. (**C**) Differentiating the corrected kinetic to obtain reduced state time. Although the kinetic lines are quite noisy, differentiation still allowed us to capture the moment at which NADPH runs out. (**D**) Plotting reduced state time versus NADPH concentration.

Discussion

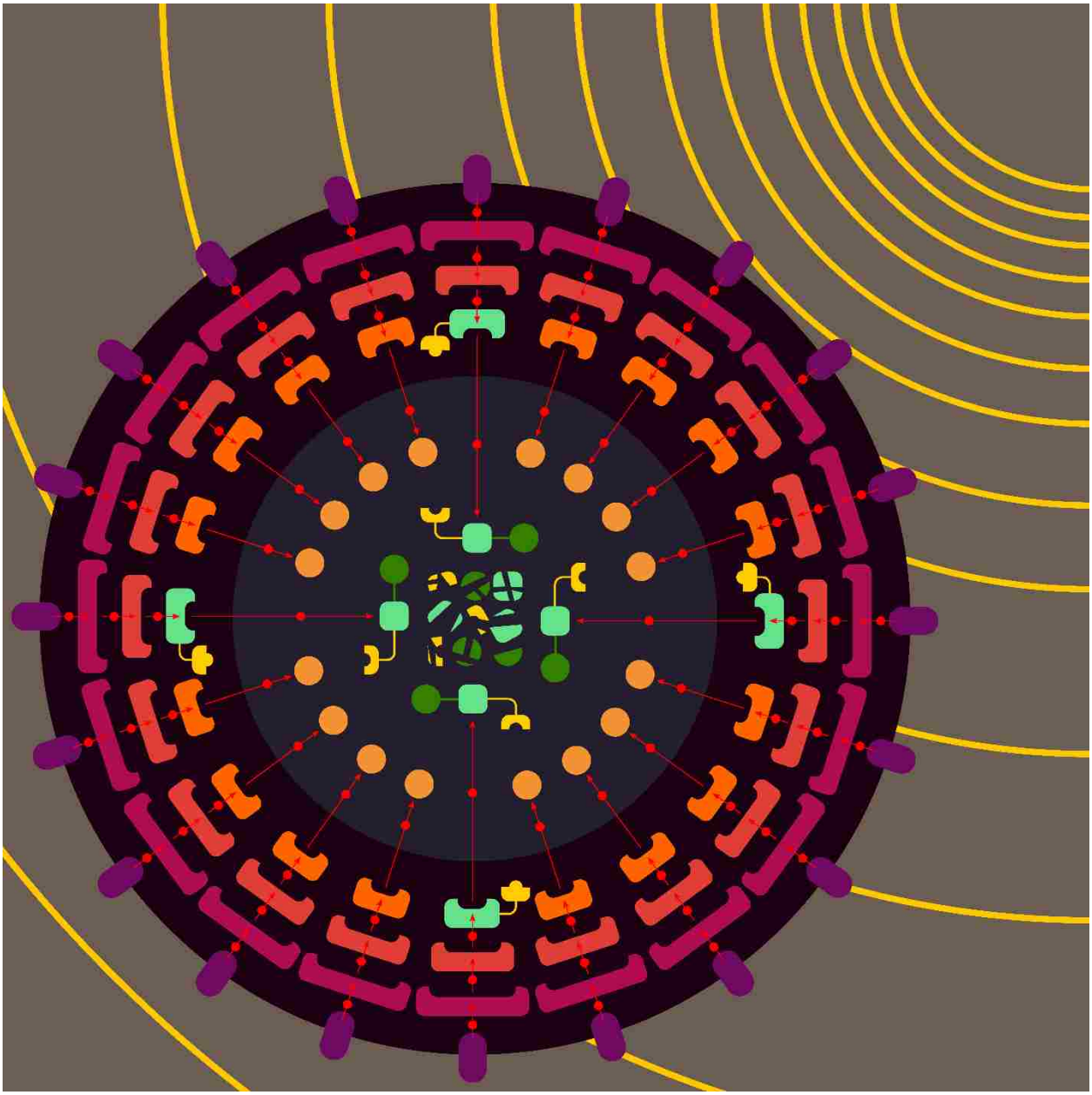
While the CYP-GFP direct fusions did exhibit only a modest fold change *in vitro*, the effect observed gives a real time readout for the redox state of the enzyme. It also demonstrates that this strategy has the potential to yield CYP based biosensors. The minimal fold-change precludes testing this system *in vivo*, so this issue needs to be resolved before any further work can be done to develop these class of fusion-proteins as biosensors. As such I believe there are two avenues that could be explored to optimize the signal from the CYP-GFP fusion. The most straight forward approach would be to use a brighter variant of the fluorescent protein such as Venus, YPet or mCitrine[1]. Another approach would be to optimize the P450-GFP fusion so that the GFP chromophore is located closer to the heme group in the CYP. Instead of using an N or C terminal fusion, circular permutation could be used to generate a library of variants which could then be screened in vivo via FACs. Ultimately, if their signal to noise ratio can be improved, the broad range of cytochrome P450s[2,3] that are very specific to certain substrates make these fusion proteins a promising platform to develop biosensors for small molecules in the future.

References

1. Shaner, Nathan C., Paul A. Steinbach, and Roger Y. Tsien. "A guide to choosing fluorescent proteins." *Nature methods* 2.12 (2005): 905-909.
2. Zanger, Ulrich M., and Matthias Schwab. "Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation." *Pharmacology & therapeutics* 138.1 (2013): 103-141.

3. Hamberger, Björn, and Søren Bak. "Plant P450s as versatile drivers for evolution of species-specific chemical diversity." *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1612 (2013): 20120426.

Chapter 2 - Engineering inter-molecular information flow: "Rewiring MAP kinases in Saccharomyces cerevisiae to regulate novel targets through ubiquitination"



Abstract

Evolution has often copied and repurposed the mitogen-activated protein kinase (MAPK) signaling module. Understanding how connections form during evolution, in disease and across individuals requires knowledge of the basic tenets that govern kinase-substrate interactions. We identify criteria sufficient for establishing regulatory links between a MAPK and a non-native substrate. The yeast MAPK Fus3 and human MAPK ERK2 can be functionally redirected if only two conditions are met: the kinase and substrate contain matching interaction domains and the substrate includes a phospho-motif that can be phosphorylated by the kinase and recruit a downstream effector. We used a panel of interaction domains and phosphorylation-activated degradation motifs to demonstrate modular and scalable retargeting. We applied our approach to reshape the signaling behavior of an existing kinase pathway. Together, our results demonstrate that a MAPK can be largely defined by its interaction domains and compatible phospho-motifs and provide insight into how MAPK-substrate connections form.

Introduction

The MAPK family of proteins is a ubiquitous signaling element in eukaryotes, and is essential to the function of a wide variety of cellular behaviors, from the regulation of differentiation and proliferation to stress responses and more[1]; this diversity of functions has been made possible by the evolutionary expansion of the MAPK repertoire[2,3]. For this expansion of the MAPK signaling module to have been feasible, it needed to be amenable to forming new kinase-substrate regulatory links, while at the same time having the capacity to avoid unwanted crosstalk. However, it still remains unclear what information is sufficient to create an entirely new set of regulatory interactions. One way to understand how potentially large numbers of novel regulatory links can be established is by developing a scalable method to create such links ourselves[4].

What are the core components necessary for the formation of a new – functional – kinase-substrate interaction? Following the association of the kinase and substrate, the amino acids in the immediate vicinity of the phosphorylated residue – together making up the 'phospho-motif' – help to dictate whether the substrate is phosphorylated by the kinase[5,6]. However, it is the site that is phosphorylated – rather than the kinase itself – that mediates the functional outcome of kinase regulation. In particular, the phosphorylated phospho-motif can be recognized by a regulatory protein bearing a phospho-motif binding domain and control protein localization or degradation among many other effects[7,8].

Even before the kinase has a chance to interact with the phospho-motif, the two proteins must be colocalized[9]. Residues apart from the kinase active site are frequently responsible for recognizing a substrate; indeed, several studies have sought to modify or replace these residues in a variety of kinases to redirect them to new – but still related – targets[10–12]. Adaptor proteins, such as synthetic scaffolds, have also been used to steer a kinase towards a particular native substrate[13–16]. And regulation of a modified native substrate by a kinase can also be rescued using a pair of completely heterologous interaction domains[17]. These studies show that by controlling the colocalization of a kinase with a native – or closely related – substrate allows the functional regulation of that target. Taking it a step further, two groups have recently used native MAPK-interacting motifs – 'docking domains' – to allow several types of MAPKs in mammalian cells and yeast to regulate the nuclear localization of fluorescent reporters[18,19]. Although it

is generally accepted that docking domains primarily control colocalization[20], several studies suggest that binding may also serve to allosterically regulate the MAPK[21–25]. As such, the precise role of these interactions remains unclear. Regardless, the question of how completely new and orthogonal regulatory relationships are created remains.

Like the signaling modules that have been expanded in natural systems, engineered genetic circuits also rely on components that are amenable to rewiring. The creation of novel transcription factors has been successful in a large part because the necessary functional characteristics have been identified. Importantly, these characteristics can be embodied in distinct modular DNA and protein domains, such as promoters, transcriptional-regulation domains, and DNA-binding domains – these domains can then be mixed and matched to yield the desired connectivity and regulation[26–30]. Although hurdles to creating large genetic circuits remain[31,32], these parts have allowed scientists to construct and interrogate more complex engineered and naturally occurring genetic systems[33,34]. Unfortunately, our understanding of how to assemble modular post-translational signaling proteins lags behind. At the same time, recent work with engineered modular receptors expressed on T-cells has shown the considerable power of the ability to rationally design even relatively simple post-translational signaling systems[35–37].

Targeting a kinase to a new substrate is an essential step towards creating modular kinase signaling systems. As discussed above, Regot *et al*. and Durandau *et al.* have described an approach wherein a kinase-specific docking domain can be used to direct a particular kinase to a new substrate—a powerful tool for interrogating natural kinase signaling systems[18,19]. However, the number of naturally occurring kinase-substrate docking interactions inherently limits the scalability of the approach. For example, a given kinase 'module' cannot be reused in parallel signaling pathways, because it would not be able to distinguish between downstream targets in one pathway versus another. To overcome this limitation, it would be useful to be able to tease apart the 'targeting' module of the kinase from the 'enzymatic' module—and likewise, the 'targeting' and 'effector' modules of the substrate. If these functions can be defined as separable parts, the enzymatic module of a kinase would be available for reuse in orthogonal pathways, just by pairing it with unique targeting domains.

We have used simple, single-function modular protein domains to explicitly test the requirements for allowing a MAPK to regulate an arbitrary substrate protein. We utilized modular interaction domains to co-localize Fus3 – the terminal MAPK of the mating pathway of the yeast *Saccharomyces cerevisiae* – with a substrate of interest. To link phosphorylation of the substrate to a meaningful regulation event we utilized phosphorylation-activated ubiquitination-based signaling motifs—phosphodegrons. We re-targeted Fus3 to regulate several disparate proteins to determine the flexibility of the substrate design rules. Likewise, to determine whether this approach generalizes to other MAPKs, we retargeted a constitutively active version of the mammalian MAPK, ERK2, to functionally regulate a fluorescent reporter in yeast.

We explored the effect that synthetically implemented post-translational regulatory connections could have on the signaling of an endogenous kinase cascade in yeast. Our results demonstrate that these new connections can be used to alter the natural signaling behaviors, damping signal amplification and even yielding concentration-based band-pass filtering. Taken together, in this paper we define a modular set of scalable components that can be utilized to rewire MAPKs to regulate proteins through ubiquitination. Attempting to rationally design new kinase-substrate regulatory links not only sheds light on the natural

processes, but also serves as the foundation for the construction of synthetic kinase signaling pathways, and with them the control of cell behaviors in biomedical or biotechnological applications.

Results

Targeting a MAPK to phosphorylate and regulate a novel substrate

To test whether a direct interaction – along with a functional phospho-motif – can render an arbitrary protein a substrate for a MAPK, we used the yeast MAPK Fus3 to target and regulate a fluorescent reporter protein. Fus3 is easily triggered using the yeast mating pheromone, α-factor. α-factor signals to the central MAPK kinase cascade via a surface-associated receptor; signaling through the pathway activates Fus3, which in turn mediates signaling to a myriad of downstream effectors, directly regulating protein function and gene expression (Figure 1A)[38].

Given the important role ubiquitin-based degradation plays in signaling[39,40], we decided to use a phosphodegron as the regulated phospho-motif. Upon phosphorylation, the phosphodegron interacts with a specialized F-box protein – Cdc4 – to recruit the E3 ubiquitin ligase machinery (the SCF complex), which then marks the substrate for degradation by covalently attaching a poly-ubiquitin chain (Figure 1A)[41]. A phosphodegron has the added benefit of making a functional phosphorylation event easy to observe: if the substrate protein is a fluorescent reporter, such as YFP, phosphorylation and subsequent ubiquitination is followed by a decrease in YFP fluorescence. Thus, this approach is amenable to high-throughput measurements in a way that changes in localization may not be.

To start, we wanted to use a phosphodegron that was proven to be both functional and compatible with Fus3. The transcription factor Tec1 fulfills these criteria, as it has been shown to be both a substrate for Fus3 and Cdc4[42,43]—thus, we chose a region of Tec1 that encompassed several residues up and downstream of the Cdc4 consensus sequence (37 residues, total)[44,45]. Also, since Cdc4 primarily acts in the nucleus, we added a nuclear localization signal derived from SV40 large T-antigen to the N-terminus of the protein[46,47]. To complete our synthetic substrate, we needed to control its interaction with an engineered kinase. To this end we added the mPDZ domain to the YFP-degron fusion, a modular protein interaction domain that has been used in a variety of different contexts[48–50]. To target Fus3 to the new substrate, we fused the complementary interaction domain, the PDZ ligand, to its C-terminus (Figure 1B). As in all the following experiments, these constructs were integrated as a single copy into the haploid yeast genome. Moreover, since we were only concerned with whether our modified Fus3 construct was able to functionally target our new YFP substrate – and not the behavior of other effectors downstream of the mating pathway – we did not remove the native FUS3 gene—thus, our modified Fus3 construct operated in parallel with the native Fus3.

Following the induction of the mating pathway with 10 μM α-factor, we measured the YFP fluorescence of the cells using flow cytometry—to account for variation caused by cell-to-cell differences in cell size, we normalized the fluorescent signal by cell size (Figure 1 Supplement 1). We observed a ~3.7-fold drop in the yeast strain containing both the Fus3-mPDZ ligand fusion and our new YFP-degron-mPDZ construct (Figure 1C). On the other hand, the drop in fluorescence was not observed when the phospho-acceptor residues in the degron (two threonine residues) were changed to methionine and alanine[43], when the catalytic site of the targeted kinase was inactivated with a K42R mutation[51], or when the interaction domain fused to YFP was changed to an SH3 domain. The latter suggests that the Tec1 degron is not able to directly recruit Fus3 to the YFP construct on its own. Finally, we also found that the drop in YFP fusion

protein level was sensitive to the presence of the proteasomal inhibitor MG132, strongly suggesting that the construct was indeed being tagged and actively degraded (Figure 1 Supplement 2).

We also explored whether our rewiring approach was sensitive to which protein the respective interaction domains were fused to. We built yeast strains in which the interaction domains were flipped—with the Fus3 kinase fused to the mPDZ domain and the YFP-degron substrate fused to a PDZ ligand. Following induction of the mating pathway with α-factor, we measured the YFP fluorescence of the cells using flow cytometry and observed qualitatively similar substrate degradation. However, the fold change observed three hours post induction for the swapped domains was approximately half that of the original orientation (Figure 1 Supplement 3). This is likely due to the fusions affecting either protein expression or sterically interfering with the function of one of the involved enzymes. While these results demonstrate that this retargeting approach is largely modular, they also suggest that other characteristics of the fusions – such as how they effect translation or protein folding – may not be.

We also asked whether endogenous Fus3 could be re-targeted in the fashion described above. We found that by inserting the sequence encoding the PDZ ligand downstream of the native copy of the *FUS3* gene in the yeast genome, the native kinase could just as efficiently cause the degradation of the YFP substrate (Figure 1 Supplement 4). These results – along with those discussed above – imply that an interaction domain and a phospho-motif are necessary and sufficient to target the regulation of a native signal transduction cascade to a substrate of choice.

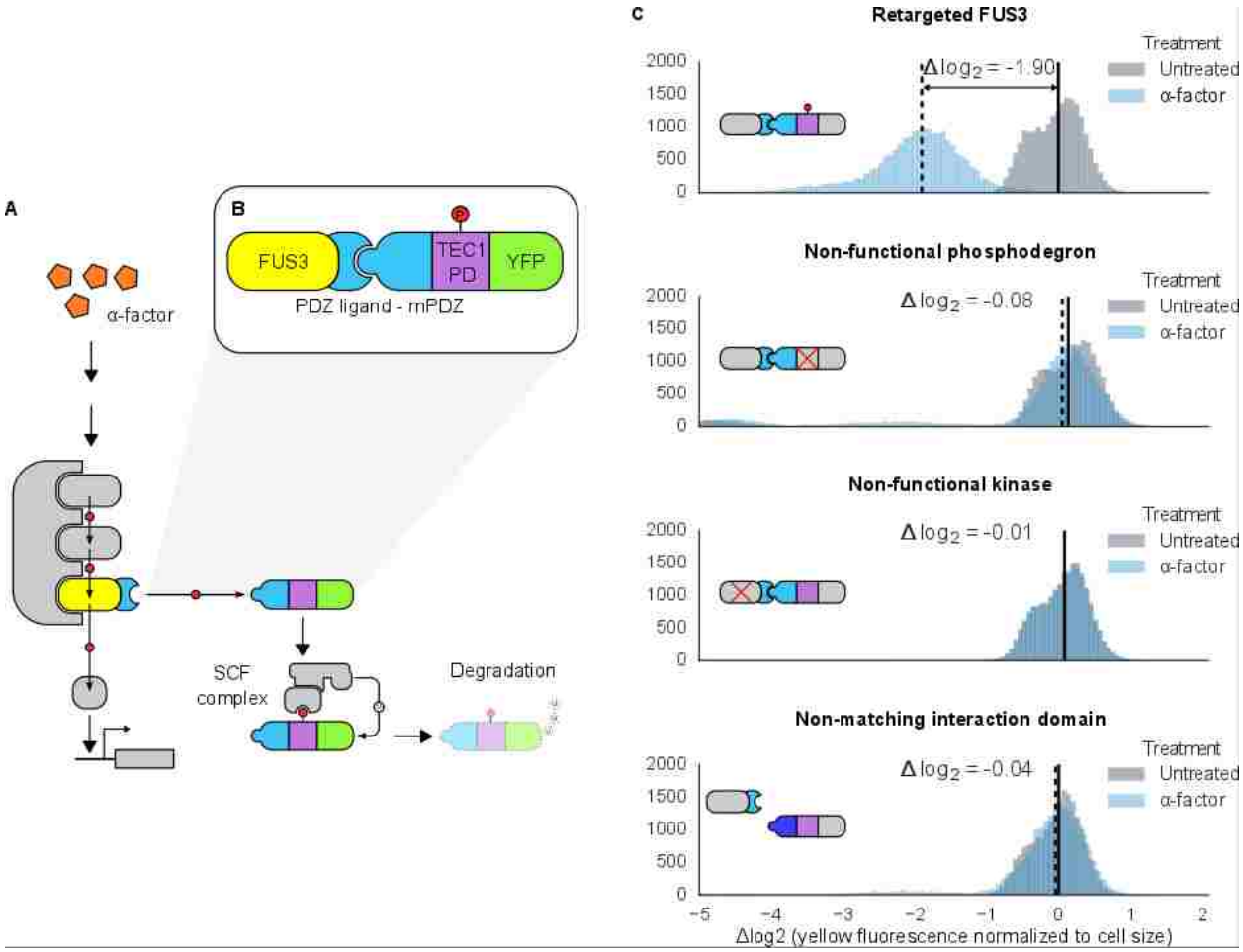


Figure 1. Rewiring the mating cascade MAPK, FUS3, to regulate the degradation of YFP. **A)** The core components of the yeast mating cascade. The yeast mating factor – α-factor – triggers the sequential activation of the kinases Ste11 and Ste7 (rounded

grey rectangles) followed by the MAPK, Fus3 (yellow). Arrows with red circles denote phosphorylation-mediated regulation. All three kinases are organized on the scaffold Ste5 (also grey). Among other effectors, Fus3 activates the transcription factor Ste12 (rounded grey box). **B)** Fus3 targeted regulation of YFP (green). The colocalization was controlled by the addition of the mPDZ domain to YFP and a PDZ ligand to Fus3 (light blue). Degradation was mediated by the addition of a phosphodegron derived from the transcription factor Tec1 (purple). Upon activation of the mating pathway, Fus3 phosphorylates the phosphodegron fused to YFP, resulting in the recruitment of an E3 ubiquitin ligase and the ubiquitination and subsequent degradation of YFP. **C)** Cells bearing the modified Fus3 and either the fully functional system, a reporter construct with a inactivated phosphodegron, a Fus3 with its kinase activity knocked out or an unmatched interaction domain (an SH3 domain instead of mPDZ) were grown to log phase and induced with 10 μM α-factor (blue histograms) or un-induced (gray histograms). The vertical dashed black lines on the histograms represent medians of treated populations and solid black lines represent medians of untreated populations. In all figures, the fluorescence has been normalized to the cell size (see Figure 1 Supplement 1). Full time-course experiments appear in the supplement to Figure 2.

Expanding the repertoire of interaction domains

To determine how general our targeting approach is, we exchanged the mPDZ/PDZ ligand pair for unrelated pairs of modular protein interaction domains, both naturally derived and synthetic. We built variants of our Fus3-substrate pair with the naturally occurring SH3 domain or the synthetic SYNZIP domain (Figure 2A)[48,52]. In both cases we observed significant reporter degradation, ~10.1-fold in the case of SH3 and ~4.7-fold for SYNZIP domains versus a control with a degron in which the two threonine residues in the Cdc4 binding site had been switched to a methionine or alanine (Figures 2B & Figure 2 Supplement 1). These results confirm the flexible nature of the interactions that enable a productive kinase-substrate interaction.

We further tested our approach using a pair of inducible interaction domains derived from a plant hormone-sensing pathway. The association of the protein domains PYL and ABI can be controlled using the small-molecule plant hormone abscisic acid (ABA) (Figure 2A, right side)[53]. When we fused these domains to Fus3 and our YFP-phosphodegron reporter, only when the concentration of ABA was 1 μM did we begin to observe a change in the fluorescent signal (Figure 2C). These results provide additional evidence both that the kinase and substrate are indifferent as to the nature of their interaction, and that the targeting of the kinase to the substrate directly triggers the observed degradation, as the decrease in YFP is correlated with ABA dose. However, it is important to note that the identity of the interaction domain fused to the YFP-phosphodegron target influenced the steady-state fluorescence of the reporter (Figure 2 Supplement 1). Thus, even interaction domains with similar affinities may not have equivalent behaviors when used inside of cells.

We next investigated whether synthetic interaction domains enable multiple MAPKs to target independent substrates in parallel and in an orthogonal manner. We targeted one copy of Fus3 to a mCherry-phosphodegron reporter using a constitutive mPDZ-PDZ ligand interaction and a second copy of Fus3 to a YFP-phosphodegron reporter via the ABA inducible ABI-PYL interaction (Figure 3A). In the presence of α-factor alone only the mCherry signal was reduced, while the YFP value remained unchanged. Only in the presence of both α-factor and ABA did we see a drop in the YFP signal (Figure 3B & C). From this perspective, the two Fus3 variants are analogous to orthologous MAPKs, with each targeting its own substrate.

However, we noted that when there were two parallel MAPK-substrate systems in the same cell the net fold change of the ABA-sensitive YFP-phosphodegron reporter was moderately reduced compared to when it was present on its own—from ~2.15 to ~1.85 fold. We tested whether this decrease in efficiency was due to competition—either between the two copies of Fus3 for the pool of activated upstream MAPK kinase, Ste7, or between the substrates for the ubiquitination/degradation machinery. To examine this question we constructed strains that expressed our standard system – one kinase targeting one substrate

– and added either a competing copy of Fus3 or a competing substrate. In both experiments we observed a diminished response in YFP degradation in the presence of the competitor (Figure 3 Supplement 1 & 2). Thus, it is likely that a confluence of factors – both saturation points as well as the less efficient ABA-induced interaction – contribute to the different levels of degradation observed for the mCherry and YFP reporters in this dual-targeting system.

It is also interesting to note that the behavior of the parallel synthetic kinases mimics the behaviors of a natural pair of yeast MAPKs, Fus3 and Kss1. Fus3 and Kss1 share many of the same targets, but also have distinct substrates, presumably as a result of the specialization of their preferences for related docking domains[54]. The engineered system described above also retains the native targeting of Fus3, but uses distinct heterologous protein interaction domains to recognize unique targets.

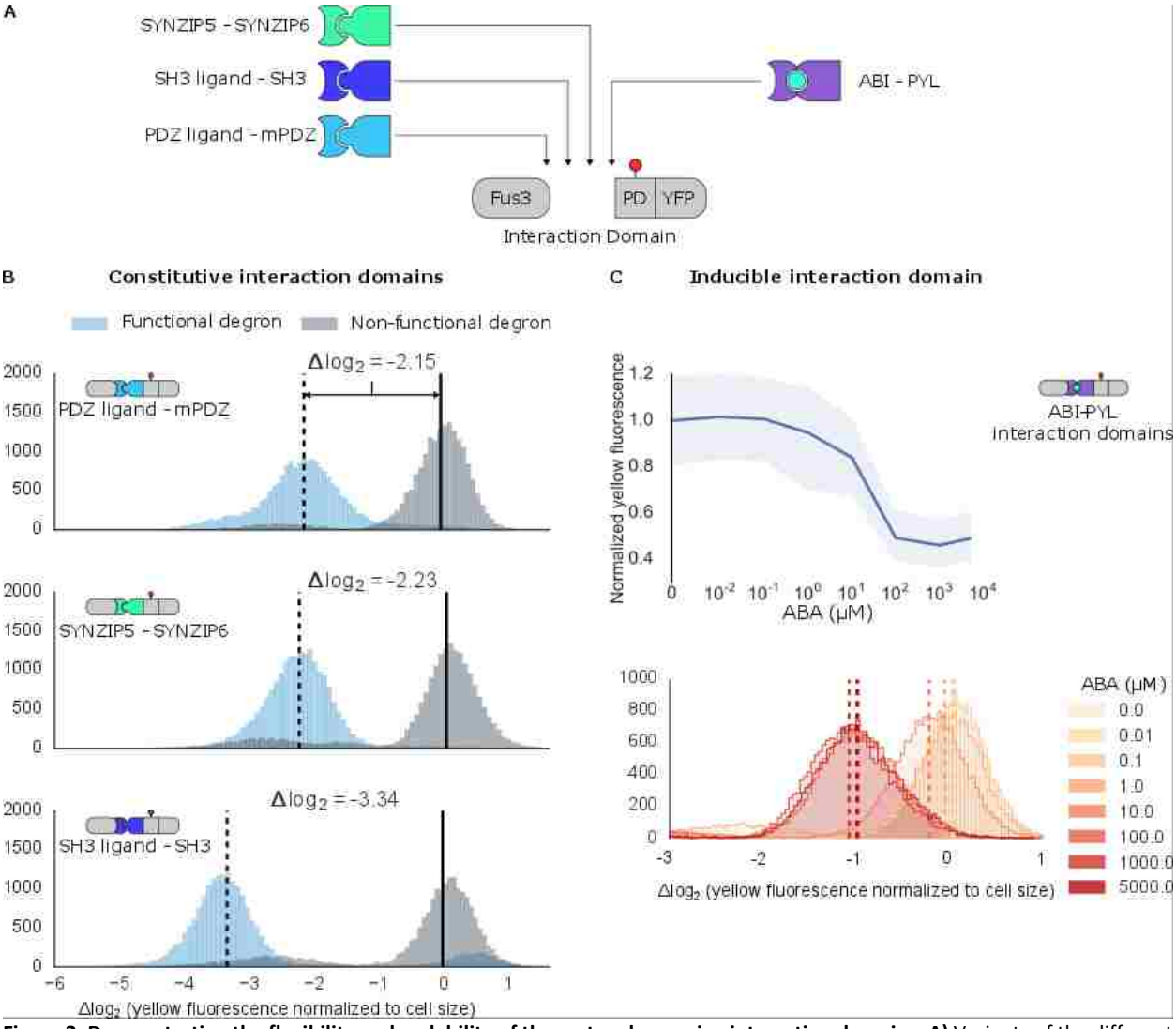


Figure 2. Demonstrating the flexibility and scalability of the system by varying interaction domains. **A)** Variants of the different complementary interaction domain used. The constitutive interaction domains mPDZ, SH3 and SYNZIP are shown on the left; the ABA inducible ABI-PYL interaction domains appear on the right. **B)** Comparison of YFP signal normalized by cell size from constructs bearing the indicated interaction domains along with either a functional (blue histograms) or non-functional (grey histograms) phosphodegron in yeast treated with 10 μM α-factor as in Figure 1C. The vertical dashed black lines on the histograms represent the medians of the populations with functional degrons whereas the solid black lines represent the median of the populations with non-functional degrons. **C)** Median fluorescence – shaded regions cover the first and third quartiles – and

population histograms of the YFP signal normalized to cell size from cells expressing the ABA inducible ABI-PYL interaction domains fused to Fus3 and YFP, respectively for a range of ABA concentrations. The raw time-course data corresponding to these endpoint observations can be found in Figure 2 Supplement 1. All error bars are 95% confidence intervals.

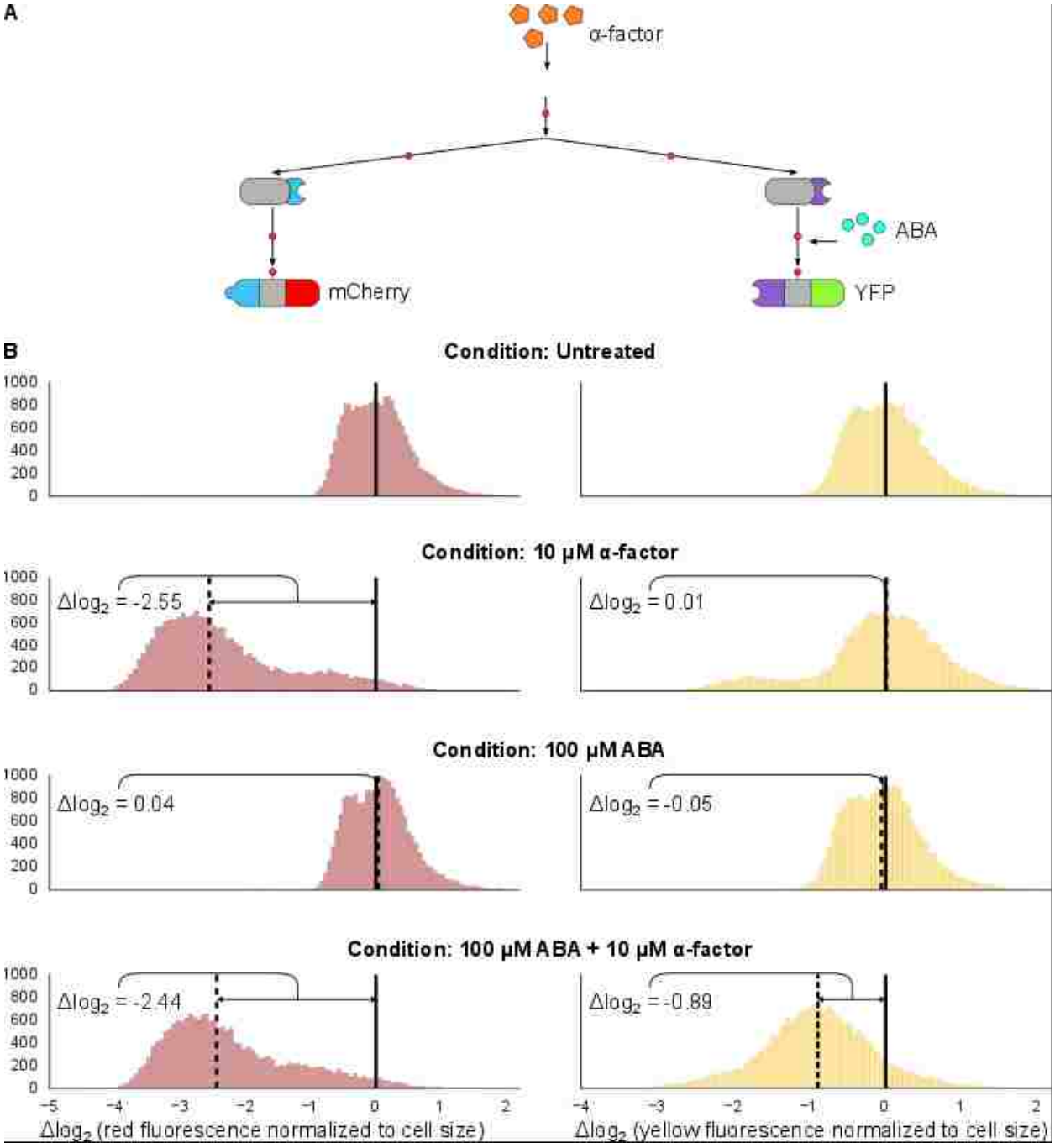


Figure 3. Targeting of orthogonal substrates. A) Cells expressed two distinct forms of modified Fus3 and used either a constitutive interaction domain (left) or the ABA inducible domains (right) to target mCherry or YFP, respectively. **B)** Population histograms of mCherry (left) and YFP (right) fluorescence normalized by cell size for cells under the indicated conditions—i.e. untreated, treated with 10 μM α-factor, treated with 100 μM ABA or both. The solid vertical black lines on the histograms represent the medians of the untreated populations and the dashed black lines represent the medians of the treated populations.

Exploration of alternative phosphodegrons

The ability to modulate the dynamics of MAPK-dependent degradation would be useful for reprogramming cell behaviors. We explored two strategies to modulate the degradation dynamics. First, we varied the number of phosphodegrons fused to the protein (Figure 4A). As we increased the number of phosphodegrons from one to three, we observed a concurrent increase in the rate of degradation of the reporter; adding more than three phosphodegrons to the reporter did not seem to affect the rate of degradation further (Figure 4B). Interestingly, in addition to changing the degradation dynamics, increasing the number of phosphodegrons also decreased the steady state expression of the reporter, possibly by multiplying the weaker interactions of the un-phosphorylated degron(s) with the degradation machinery (Figure 4 Supplement 1).

We found that altering the amino acid sequence of the phosphodegron itself also changed the dynamics of degradation. We constructed two additional variants of the phosphodegron motif that more closely mimicked the amino acid sequence of the published 'consensus motif' for the WD40 domain of Cdc4 (Figure 3A)[44,45]. The sequences of the two variants only differ at one site — two residues N-terminal of the phosphorylated threonine — where the Cdc4 consensus leucine was changed to a proline, an amino acid that is supposed to be preferred by Fus3[5]. Both variants had similar behaviors, with a similarly decreased rate of degradation relative to the phosphodegron derived from Tec1. (Figure 4C). These results suggest that phosphodegron design is flexible, and with more study it may become feasible to rationally tune their degradation dynamics. Moreover, the number of phosphodegrons is not limited to those found in nature. Taken together, these results demonstrate that our approach is applicable to several different phosphodegrons, and lays out a framework for how phosphodegrons may be used to alter degradation dynamics of a protein of interest.

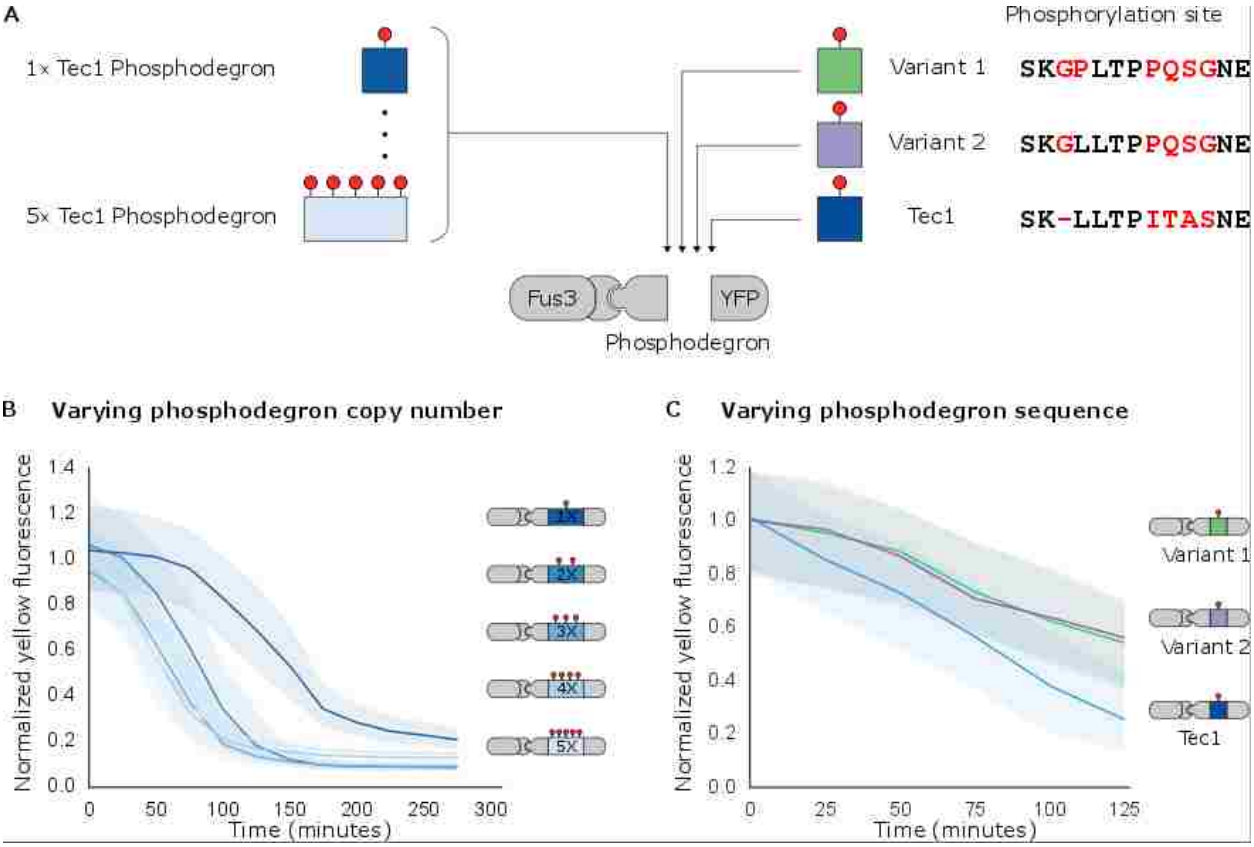


Figure 4. Modulating regulation by altering the number and sequence of phosphodegrons. A) We varied either the phosphodegron number (left) or the sequence (right)—differing residues are red, the phosphorylated residue is highlighted in blue. **B)** Time-course data of strains induced with 10 μM α-factor and expressing Fus3 targeting YFP reporters with one to five phosphodegrons. The fluorescence of each strain was normalized to cell size and then to its initial fluorescence. Data normalized only to cell size can be found in Figure 3 Supplement 1. **C)** Fus3 targeting of YFP substrates with the indicated phosphodegron sequence variants. As in B), the fluorescence of each strain is normalized to cell size and then against its initial fluorescence. Data normalized only to cell size can be found in Figure 3 Supplement 2. The curves indicate the median values, while the shaded regions cover the first and third quartiles.

Retargeting the mammalian MAPK ERK2

We next swapped out the kinase module to test whether other MAPKs are also amenable to rewiring in the same manner. We focused on the human MAPK, ERK2—a widely studied kinase implicated in several pathologies, which has also been previously studied in the context of yeast[55]. Native ERK2 has been shown to regulate protein stability via phosphodegrons; for example, a phosphodegron found in the protein MKP1 is targeted by ERK2 and subsequently tagged by the ubiquitin machinery and degraded[56]. Our engineered substrate consisted of a 64 residue region surrounding the phosphodegron of MKP-1 fused to a YFP reporter. Rather than port the entire ERK2 signaling cascade into yeast, we used a constitutively active version of the MAPK created by fusing the upstream MAPK kinase – MEK1 – to ERK2[57]. To enable the kinase-substrate interaction we fused the mPDZ domain and PDZ ligand to the substrate and MAPK, respectively (Figure 5A). We also included a construct missing the heterologous targeting domains to make sure that targeting was not simply due to direct interactions mediated by sequence elements surrounding the phosphodegron. Since the strains constitutively expressed both the engineered kinase and target, we measured the steady-state YFP fluorescence via flow cytometry. In strains with the active kinase targeting the functional YFP reporter, fluorescence did not rise above background levels (Figure 5B)—suggesting that the substrate is actively targeted, phosphorylated and then degraded. Fluorescence was significantly higher in control strains where the interaction domain, the phosphodegron or both were

missing or inactivated, respectively (Figure 5B). These results indicate that the interaction domains and the phosphodegron are necessary and sufficient for retargeting the regulation of ERK2. Importantly, these results also strongly suggest that this rewiring approach is potentially applicable to a wide range of MAPKs.

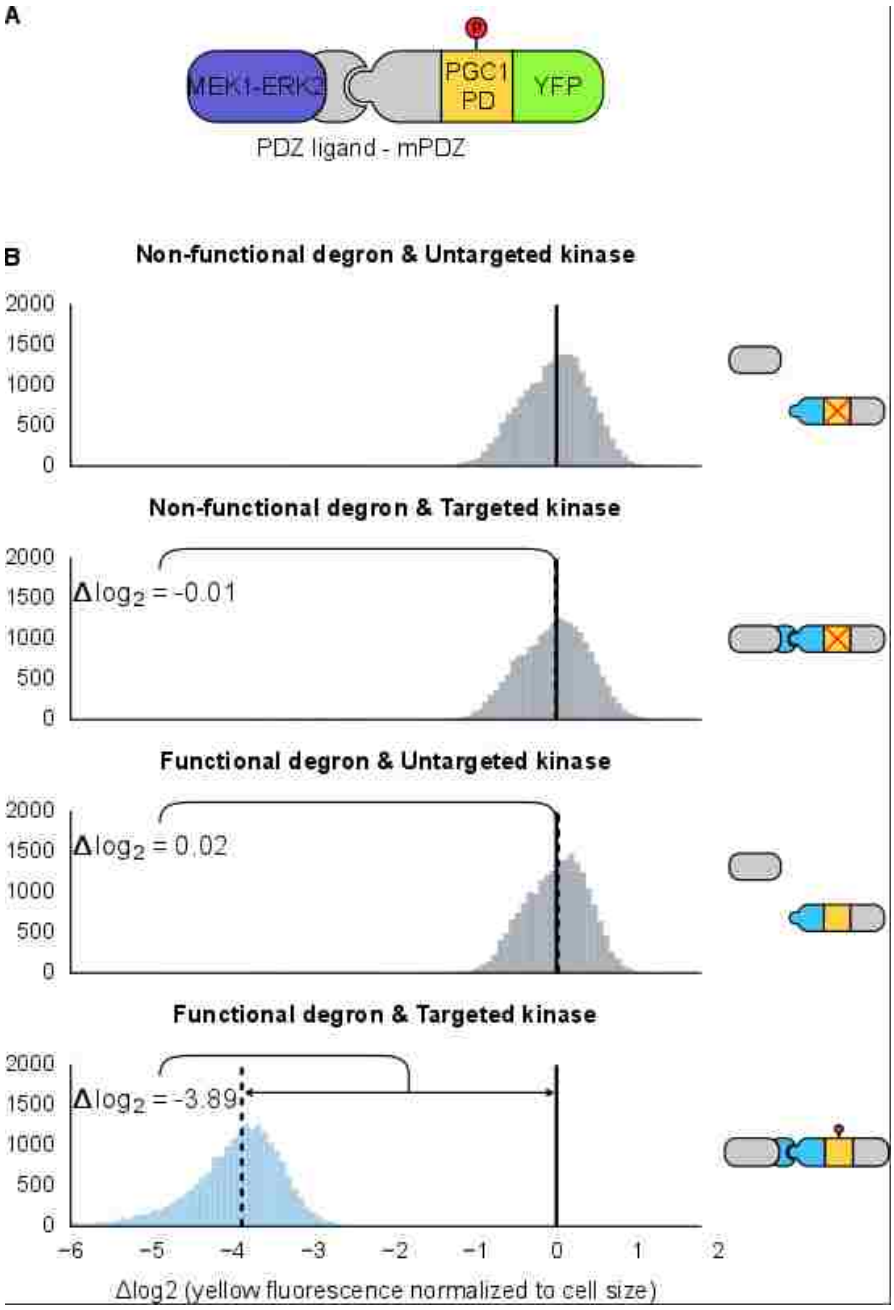


Figure 5. Retargeting the mammalian MAPK, ERK2. A) As with Fus3, the human MAPK, ERK2, was targeted to a YFP reporter (green) via an interaction between the mPDZ domain and the PDZ ligand. A phosphodegron (yellow) fused to the YFP reporter was derived from the mammalian MKP-1. ERK2 was rendered constitutively active by fusing it to a constitutively active form of MEK1 (purple). **B)** Population histograms of YFP fluorescence normalized by cell size of yeast strains in log phase growth with active ERK2 targeted to YFP with a functional phosphodegron (blue histogram). Controls strains with an inactive phosphodegron fused to YFP and/or an untargeted version of the kinase were also tested (gray histograms). The solid vertical black lines on the histograms represent the medians of the first histogram – the untargeted kinase paired with the non-functional degron – and the dashed black lines represent the medians of each subsequent population.

Modifying MAPK cascade signal processing

Thus far, we have described the retargeting of MAPKs to synthetic targets such as fluorescent proteins, which double as the readout for kinase activity. Next, we asked whether MAPKs could be targeted to arbitrary endogenous substrates and – more specifically – whether this approach can be used to modify the response of an existing signaling pathway. To answer these questions, we targeted Fus3 to up- and downstream elements in the yeast mating cascade, including the kinase Ste7, the scaffolding protein Ste5, and the transcription factor Ste12. We built a total of six yeast strains containing the synthetic kinase-substrate pairs. Three of these strains constitutively expressed Fus3 with a PDZ interaction domain, while the other three expressed a version of Fus3 with a non-matching interaction domain. All of the strains included one of the mating cascade proteins – Ste5, Ste7 or Ste12 – fused to a complementary interaction domain and the Tec1 phosphodegron. The interaction domain and phosphodegron were inserted into the native genomic locus of the protein of interest. The Fus1 gene, whose expression is activated by the mating pathway upon induction with α-factor, was fused to YFP to provide an independent readout for pathway activation.

We chose these specific target proteins because their regulation by Fus3 results in interesting regulatory topologies. Specifically, Fus3-medited degradation of Ste7 and Ste5 are examples of negative feedback loops, while the degradation of Ste12 results in an incoherent feed-forward loop (Figure 6A-C). Such regulatory links can be used to fundamentally alter the signal processing properties of native pathways[58,59]. Of note, this is the first time that purely post-translational feedback loops have been used to re-engineer signaling.

To determine the impact of negative feedback, we measured the fluorescence output of the pathway following induction with varying levels of α-factor. Relative to the untargeted controls, the negative feedback through Ste5 or Ste7 reduced the maximal pathway activation in those backgrounds by ~60% and ~45%, respectively. The apparent Hill coefficients (n_H) were also moderately changed with negative feedback compared to the untargeted kinase controls—when Ste5 was the target, n_H increased from 1.5 to 1.9, while when Ste7 was negatively regulated n_H remained 1.6 (Figure 6A & B). These values are qualitatively consistent with but slightly higher than the sensitivities reported previously for a system with negative feedback realized through transcription and recruitment of a phosphatase in the yeast mating cascade[59].

The increase in pathway sensitivity observed for negative feedback applied to Ste5 is surprising[60]. However, the response of a scaffolded signaling cascade is highly sensitive to the concentration of the scaffold protein—with a reduction of the scaffold concentration resulting in an increase in the sensitivity of the cascade[61]. Although a more detailed analysis is required, this observation suggests a potential explanation for the observed increase in the apparent Hill coefficient. However, we also note different fusion proteins are required for each experiment and that these protein modifications alone can result in changes of the pathway sensitivity—e.g. by changing the concentrations of pathway components[55]. For example, the non-feedback controls in the three experiments shown in Figure 6 have apparent Hill coefficients of n_H =1.5, 1.6 and 2.1.

In the incoherent feed-forward loop – created by having Fus3 both activate and inhibit the transcription factor Ste12 – we find that the inhibitory connection dominates at all levels of induction resulting in a complete abolishment of the downstream response (Figure 6C & Figure 6 Supplement 1). However, as we will show next, more interesting behaviors are possible in a slightly more complex incoherent feed-forward loop.

Hybrid regulatory schemes that occur at the level of both transcription and translation are often observed in nature and further enrich the available behaviors in the design of engineered biological circuits[62,63]. By putting the YFP-phosphodegron-mPDZ domain fusion protein under the control of the mating pathway-controllable promoter – pFUS1, we created a simple incoherent feed-forward circuit regulated at the level of both transcription and translation. Such a 'type 3' incoherent feed-forward loop design can produce pulses and other behaviors, depending on the design parameters[64]. A phenomenological model of a hybrid incoherent feed-forward loop is included in Appendix 1. We performed time-course experiments over a range of α-factor concentrations (Figure 7B & Figure 7 Supplement 1). In cells containing the feed-forward loop the fluorescent signal initially increased sharply as the α-factor concentration was increased from 0.1 μM to ~1 μM—however, induction with concentrations of α-factor higher than 1 μM resulted in decreasing levels of YFP fluorescence. The incoherent feed-forward loop thus created a concentration-based band-pass filter for the α-factor input. In a control where the phosphodegron fused to YFP was broken we observe the normal signal amplification behavior of the mating cascade (Figure 7A)—thus, it is the targeted regulation of YFP by Fus3 that caused the band-pass-like behavior.

Taken together, these results demonstrate that re-targeted kinases can be used to modulate the behavior of signaling cascades through variety of circuit designs, including negative feedback and incoherent feed-forward loops. The data also highlight the utility of using this rewiring approach to study the effects of kinase-directed ubiquitination-based regulation, which occur extensively in nature[40,65] and adds to the available tools for the study of this pervasive mode of signaling[66–69].

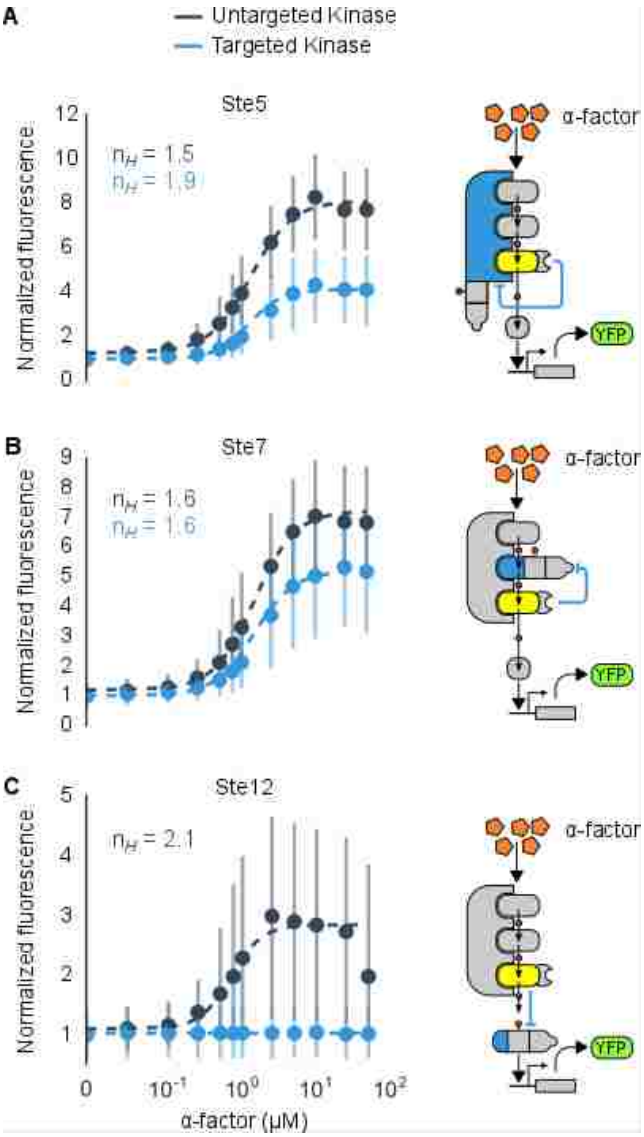


Figure 6. Implementation of negative feedback and feed-forward signaling topologies using a rewired MAPK. A-C) Plots and schematics that depict the relationship between the α-factor input and the YFP reporter for yeast strains with synthetic post-translational negative feedback or feed-forward loops. Fus3 (yellow) was rewired to target (A) the scaffold Ste5, (B) the kinase Ste7 or (C) the transcription factor Ste12 (all depicted in light blue)–in each case, the endogenous copies of these proteins were modified by inserting a phosphodegron and a complementary interaction domain at their C-terminus. Plots of the median fluorescence of the YFP reporter – under the control of the mating-specific pFUS1 promoter – normalized to cell size for increasing concentrations of α-factor. Data from control strains with an untargeted kinase – and thus no feedback/feed-forward control – are shown in dark blue. Points indicate the median values at each α-factor concentration, while the bars cover the first and third quartiles of the data. The data from both the no feedback and feedback conditions were used to determine the parameter values used with the formula: $A + B\frac{[\alpha]^n}{1+C[\alpha]^n}$ – where C was fixed between the two data sets. n and $[\alpha]$ are the hill coefficient and the α-factor concentration, respectively. Fits are plotted as dashed lines. Time courses of the same strains treated with 10 μM α-factor are shown in Figure 6 supplement 1.

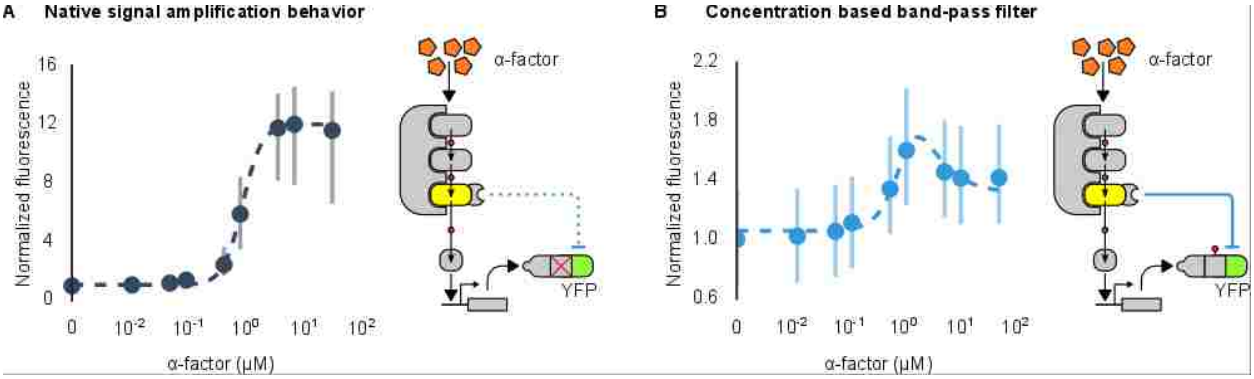


Figure 7. Conversion of a native amplifier to a band-pass filter. A, B) The relationship between the α-factor input and YFP expression – driven by the mating pathway-specific promoter pFUS1 – for strains without and with a synthetic post-translational incoherent feed-forward loop. Induction of the mating pathway activated a modified Fus3 (yellow) that indirectly up-regulates the expression of a YFP reporter (green) fused to a phosphodegron. An interaction between the Fus3 and the YFP-degron reporter was enabled via complementary interaction domains. In one case (A) the phosphodegron was mutated and inactive, while in the other (B) it was fully functional. The points indicate the median YFP signal – normalized by cell size and then to the untreated condition – in yeast strains in log phase growth treated with the indicated concentration of α-factor. The error bars depict the first and third quartiles of the population data. Dashed lines are fits to the equation $A + B\frac{[\alpha]^n}{1+C[\alpha]^n}\frac{1+E[\alpha]}{1+D[\alpha]}$ – model derivation and fitting are described in more detail in Supplementary text. Time-course data is shown in Figure 6 Supplement 1.

Discussion

Here we have demonstrated that MAPK-directed ubiquitin-based signaling can be rewired to regulate a protein of choice. The addition of sets of two modular components is sufficient to rewire a MAPK to regulate any protein of interest—a complementary set of protein interaction domains and a phosphodegron. Natively, MAPKs are co-localized with their substrates via an interaction between the docking peptide of a substrate and a set of residues on the surface of the MAPK; it has been hypothesized that this interaction may be necessary to catalytically unlock the kinase[21–24]. Our results suggest that while these domains may have some allosteric properties, simply co-localizing an active MAPK with a protein bearing a compatible amino acid motif that can be phosphorylated is sufficient for the functional regulation of the protein.

One implication of our results is the that the evolution of new connections within MAPK regulation networks is only constrained by the two criteria discussed above—namely the appearance of 1) an accessible phospho-motif; and, 2) a protein-protein interaction strong enough to co-localize the new kinase-substrate pair. The proteomes of *Saccharomyces cerevisiae* and humans have ~3500 and >50,000 phosphorylation sites, respectively[70–72]. The amino acid composition of the surrounding phospho-motifs are constrained by the residues in and surrounding the kinase active site; as such their length is generally fairly short—on the order of four amino acids on either side of the phosphorylated residue[9]. With such a short length, and given the degeneracy of the recognition requirements[5], the probability that new phospho-motifs will appear by chance is fairly high. Indeed, many human SNPs that are observed or associated with cancer and disease – as well as apparently healthy individual variation – have been observed to create and destroy verified phospho-motifs[70,73,74].

Many protein interactions occur between short, linear stretches of amino acids and protein domains, the classic examples being PDZ and SH3 domains, but which also include the binding of docking domains to the surfaces of MAPKs[75–78]. Like phospho-motifs, these short motifs too can appear spontaneously during evolution[79–81]. Given that both phospho-motifs and short, linear interaction peptides are degenerate,

common and short it is interesting to consider what constrains the formation of a new, functional connection between a kinase and a substrate—i.e. whether it is the formation of phospho-motifs or of protein-protein interactions that is rate limiting? This may be addressed by future studies.

Our creation of modular components for kinase signaling may help recapitulate the success modular transcriptional circuits have enjoyed[27–29,33]. However, while our approach is a powerful tool it does have certain limitations. For instance, our system requires that a phosphodegron be known, and its cognate F-box be expressed for ubiquitination to occur. We demonstrate one way in which this problem may be addressed, i.e. by the design of new phosphodegrons based on the consensus sequences of the MAPK and F-box. Another consideration in any protein-engineering endeavor is the effect that various protein fusions have on expression—indeed we noted in our experiments that fusion of additional domains to MAPKs or their substrates altered the expression levels. These altered expression levels affect the behavior of kinase cascades, and so a sufficiently diverse set of modules need to be defined and characterized to make the desired behavior(s) achievable. Thus, the scalability afforded by the use of modular interaction domains comes at the potential price of altered protein expression. In contrast, using docking domains for co-localization obviates engineering the kinase, but is not a scalable rewiring approach. Finally, more work is necessary to render other kinase families 'engineerable'. Still the flexibility and scalability of kinase-substrate interactions demonstrated through our work lays a comprehensive foundation for future attempts to understand and re-engineer the signaling behavior of cells.

Materials and Methods

Strain construction. All strains were built using a W303a background into which each synthetic construct was integrated at either the URA, HIS, TRP or LEU genomic loci. The plasmids used to generate the strains are listed in Supplementary file 1. The YFP reporter constructs were built by fusing an SV40 nuclear localization tag, an interaction domain and a phosphodegron in tandem to the YFP protein separated by 12 amino acid long glycine-serine linkers. The retargeted kinase constructs were built by fusing a complementary interaction domain to the kinase, also separated by a 12 amino acid glycine-serine linker. The strong constitutive promoter derived from the native TDH3 gene drove all constructs. For all examples of the system that involved the yeast mating cascade, the kinase used was the MAPK FUS3. For the system that demonstrated mammalian MAPK retargeting, the kinase utilized was a constitutively active version of MEK1 fused to ERK2 and an interaction domain. For the feedback and feed-forward strains YFP was fused in tandem with the FUS1 gene, whose expression was activated by the mating pathway, to act as a reporter. These strains also had a FUS3 fused to an interaction domain integrated into the genome. In the case of the negative feedback and the feed-forward strains the genomic copies of Ste5, Ste7 and Ste12 were fused to an interaction domain, a phosphodegron and a mCherry reporter. The incoherent feed-forward strains were identical except that the expression of the YFP-nuclear localization tag-phosphodegron-interaction domain fusion was driven from a FUS1 promoter.

Cytometry. All cytometry measurements in experiments just measuring YFP expression were acquired with an Accuri C6 cytometer with attached CSampler apparatus using 488 nm and 640 nm excitation lasers and a 533 nm (FL-1: YFP/GFP) emission filter (BD Biosciences). In those experiments that included mCherry, we used a MACSQuant VYB (Miltenyi Biotec), with 405, 488 and 561 nm excitation lasers and 561 nm (FSC), 525 nm (YFP) and 615 nm (mCherry) emission filters. Synthetic complete growth medium was used in all experiments. Experiments involving time course data were taken during log phase via the following preparation: 16 hours of overnight growth in synthetic complete medium in a 30°C shaker incubator followed by 1:100 dilution into fresh, room-temperature medium. After 5 hours of growth at 30°C, 100 µL aliquots were read periodically – with 10^4 events collected for every read – until the completion of the experiment. In all cases where FUS3 was being retargeted, the yeast cultures were induced with α-factor 5 hours post-dilution. For experiments involving dose response behavior, cultures were grown overnight, then diluted down in the morning 1:100 in fresh media and grown for 5 hours to log phase. They were then induced with α-factor, as well as other inducers like ABA in some cases, and allowed to grow for between two to six hours depending on the experiment and then read on the cytometer. As the MEK-ERK2 fusion is constitutively active no inducer was necessary[57].

Cytometry data analysis. Data were analyzed using custom python scripts and FCSParser and Seaborn libraries (DOI: 10.5281/zenodo.45133) using the following steps: (1) Anomalies – such as bubbles – were identified by plotting and visually inspecting the FSC-A value versus the time each cell was collected for each well. (2) To prevent the creation of NA values when the data was log transformed any 0 values in the data collected from the Accuri C6 cytometer were converted to 1. Since data collected on the MACSQuant VYB can fall below 0, all the data was normalized by adding the absolute value of the lowest value collected that day to the raw values and then adding 1. (3) To control for the effects of cells size, the fluorescence values for each event were then divided by the FSC-A value for that event. All reported data is the aggregate of at least two technical replicates performed on consecutive days. The fits presented in figures 6 & 7 were performed using custom python scripts.

Supplementary Information

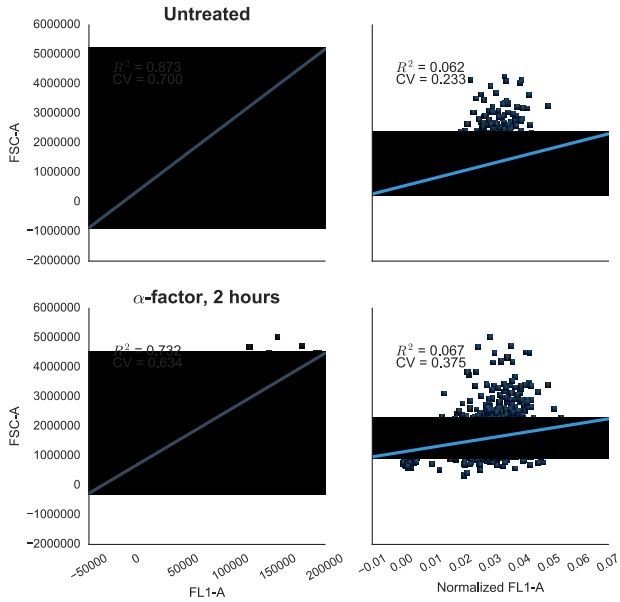


Figure 1 supplement 1. Reducing the variability of single-cell fluorescence by accounting for cell-to-cell variation in cell size. In yeast constitutively expressing a single-copy fluorescent protein inserted in the genome, fluorescence (FL1-A) is strongly correlated with cell size (approximated by FSC-A)—shown by R^2 values. This is true in both cells that are untreated (top left) or treated with 10 µM α-factor (bottom left). The effect is likely due to the way that flow cytometers measure fluorescence, where cells with the same concentration of fluorescent protein – but with different volumes – will have different levels of fluorescence. For example, smaller cells that have just divided will have a lower fluorescence value than larger cells that are just about to divide. Normalizing by cells size – dividing FL1-A by FSC-A – reduces the coefficient of variation (CV) of the fluorescent signal by ~67% or ~41% in untreated or treated cells, respectively (graphs on right).

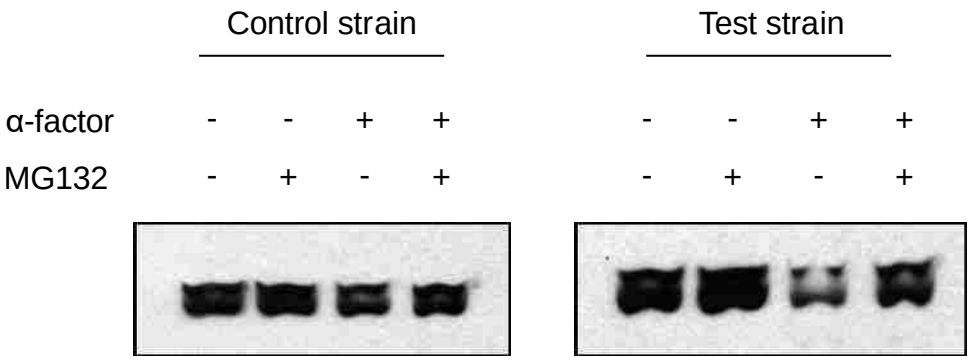


Figure 1 supplement 2. Western Analyses of degradation assays. The representative western blots above show results from degradation assays on a test strain with the entirely functional system described in figure 1 and a control strain that has a non-functional phosphodegron. We observe that

upon treatment with α-factor, the test strain has a significantly fainter band as compared to the untreated lane, whereas the control strain does not. This is consistent with our flow cytometry observations. We also observe that when cells are treated with MG132, a proteasome inhibitor, the alpha-factor triggered degradation in the test strain is prevented, implying that the degradation we observe is indeed due to the proteasome as hypothesized.

Degradation Assays

10 mL cultures of yeast strains expressing untargeted control substrates or targeted test substrates were grown at 30°C in YEPD medium to approximately $1*10^7$ cells/mL. Cells were incubated with DMSO or the proteasome inhibitor MG132 (25 µg/mL) for 30 minutes prior to addition of α-factor or vehicle control for an additional 10 minutes. Cycloheximide was then added to a concentration of 50 µg/mL and cells were harvested by centrifugation at the denoted time points. Pellets were lysed in 200 µL SUMEB buffer (8 M urea, 10 mM MOPS, 10 mM EDTA, 1% SDS, 0.01% bromo- phenol blue, pH 6.8) by vortexing with acid washed beads for 5 minutes at 25°C. Lysate was clarified by centrifugation at 13000 rpm for 5 minutes and subjected to western analysis.

Western analyses

Protein lysates were resolved by SDS-PAGE using 4–20% gradient gels (Lonza). Western analyses were performed with rabbit anti-GFP (1:2500) or mouse anti-ubiquitin antiserum (1:10).

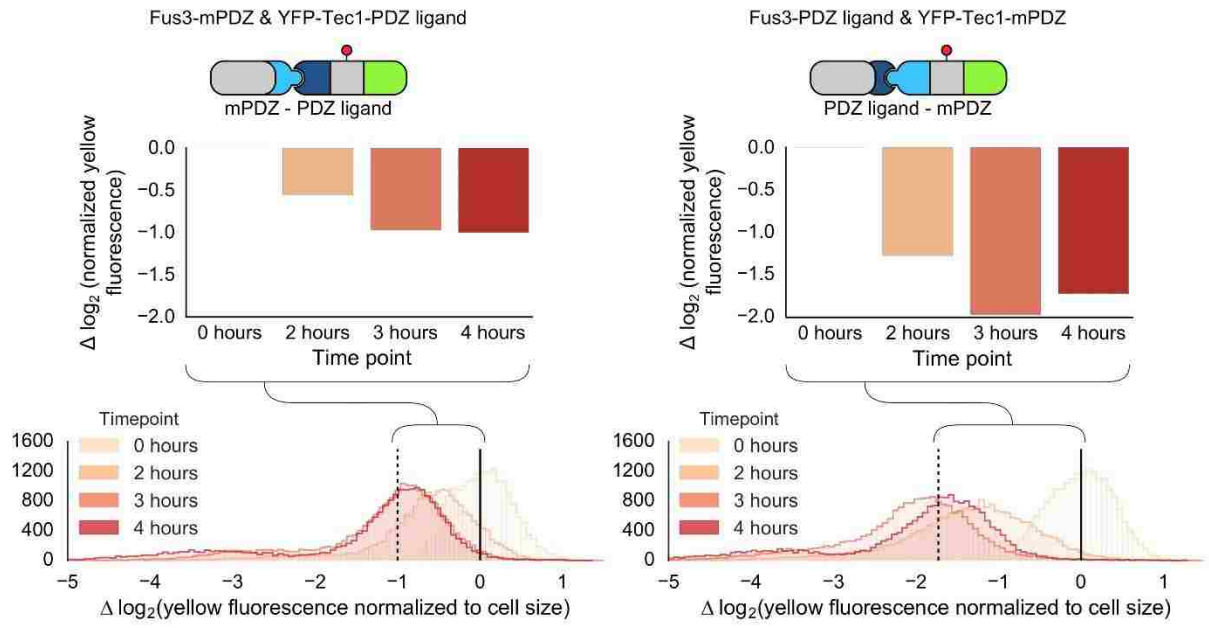


Figure 1 supplement 3. Swapping interaction domains between kinase and substrate. Population histograms and medians of YFP fluorescence signal normalized by cell size for yeast strains that have either a Fus3 kinase fused to a PDZ ligand and YFP fused to a PDZ domain (right) or Fus3 kinase fused to a PDZ domain and YFP fused to a PDZ ligand. These strains were diluted down from saturated overnights and were grown for 5 hours to log phase and then cytometry reads were performed after 2, 3 and 4

hours post induction with α-factor. The dashed black lines represent the median fluorescence at 4 hours and the solid black lines are the media fluorescence at 0 hours.

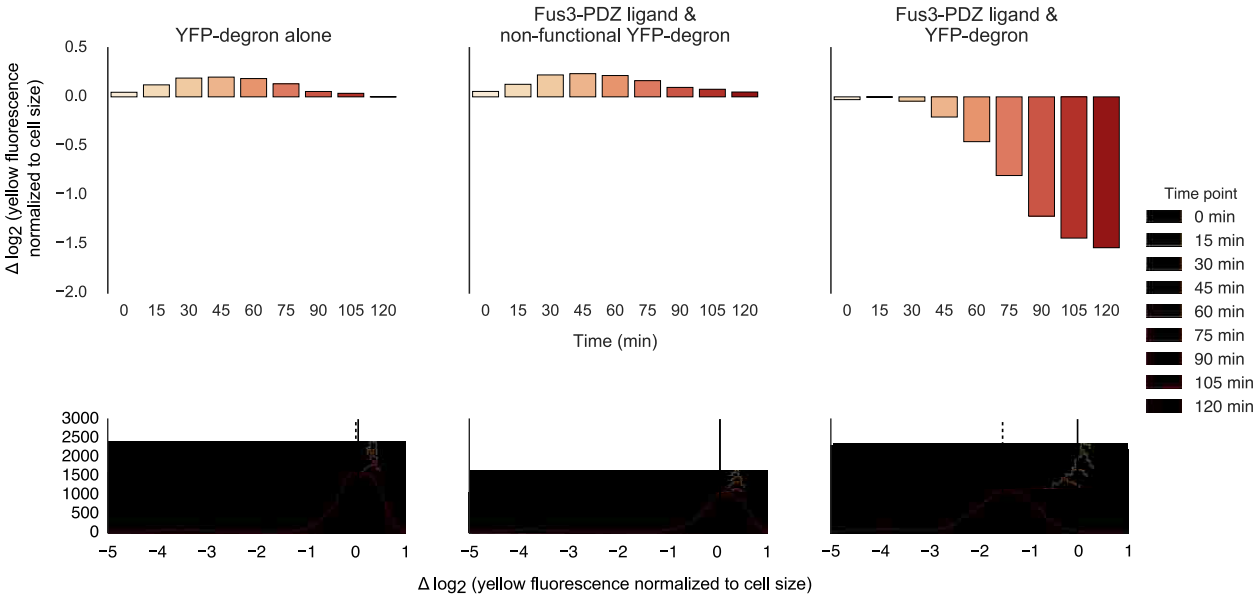


Figure 1 supplement 4. Fusing interaction domain to the native copy of the kinase. Population histograms (bottom) and corresponding medians (top) of YFP fluorescence signal normalized by cell size for yeast strains where the native copy of the kinase is fused to a PDZ ligand. The left most panel describes a strain with just the YFP substrate, the middle one represents a strain where the endogenous copy of Fus3 had an interaction domain fused to it but the YFP substrate has a non-functional degron and the rightmost panel describes a strain that has both an interaction domain on the Fus3 and a functional degron on the YFP substrate. These strains were diluted down from saturated overnights and were grown for 5 hours to log phase and then cytometry reads were performed at 15 minute intervals post induction with α-factor.

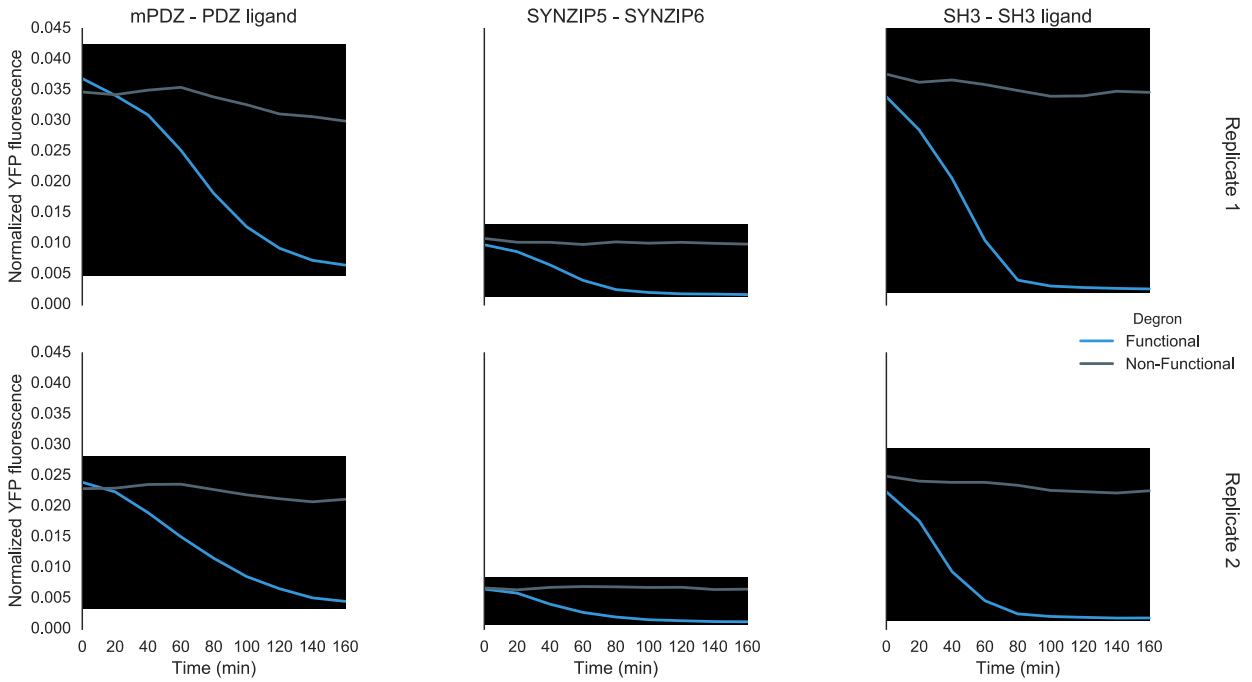


Figure 2 supplement 1. Time course characterization of different interaction domain variants post induction with α-factor. In this plot each subplot is labeled with the name of the interaction domain pair that was used to target the MAPK Fus3 to degrade a YFP reporter in the yeast strain. These yeast strains were grown up to log phase from saturated cultures for 5 hours and then induced with 10 µM α-factor at time 0 to activate the kinase. The dashed lines are uninduced controls. In the case of the ABI-PYL strain two additional induction conditions were assayed, namely with ABA (100 µM) and with ABA and α-factor. For these cultures the medium used to grow the yeast up to log phase had ABA in it. The fluorescence of these cultures was then assayed at regular intervals using flow cytometry. Raw data for two replicates performed on different days under identical conditions is shown.

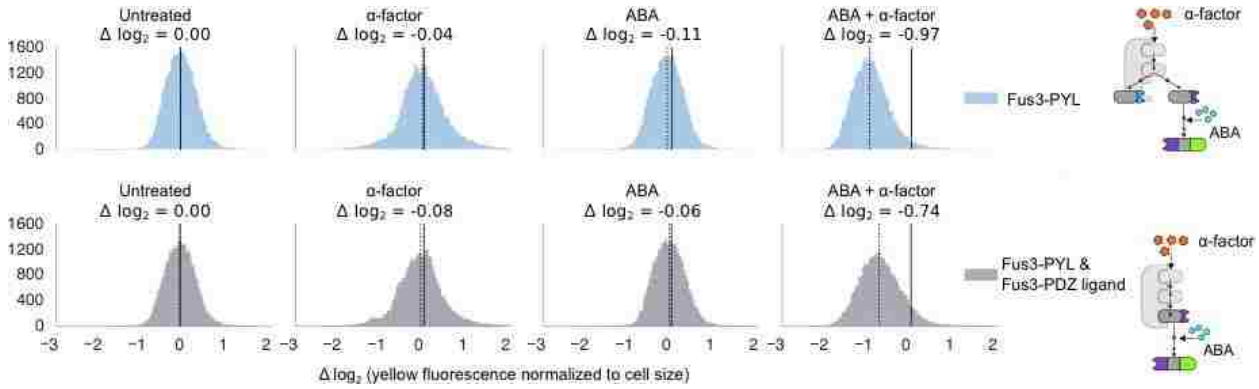


Figure 3 supplement 1. Competition between two Fus3 MAPKS with different interaction domains for MAPKK Ste7. Population histograms of YFP fluorescence signal normalized by cell size for yeast strains that have either a single copy of Fus3 targeted to a YFP substrate via an ABA dependent PYL interaction domain (top) or two copies of Fus3 one with a PDZ ligand and another with a PYL domain (bottom). These strains were diluted down from saturated overnights and were grown for 5 hours to log phase with and without ABA and then cytometry reads were performed after 2 hours post induction with α-factor. Each plot corresponds to a specific induction condition, with the solid black lines indicating the median of the untreated controls and the dashed black lines indicating the median of that particular treatment. The magnitude of log differences between treated and untreated medians for each treatment is displayed the treatment label.

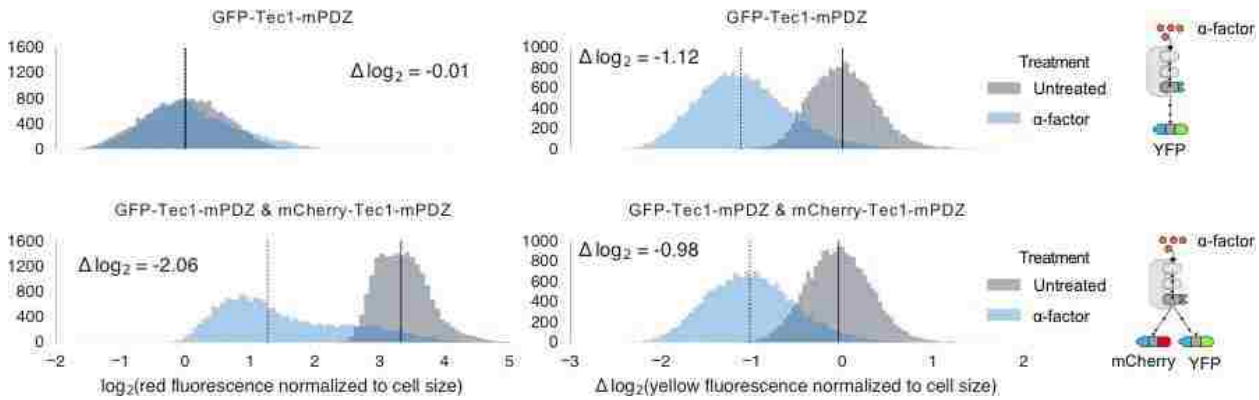


Figure 3 supplement 2. Competition between mCherry and GFP when targeted by the same Fus3. Population histograms of mCherry (left) and YFP (right) fluorescence signal normalized by cell size for yeast strains that have either a single YFP substrate targeted by Fus3 via a PDZ interaction (top) or two substrates, YFP and mCherry, both of which are targeted by Fus3 via a PDZ interaction (bottom). These strains were diluted down from saturated overnights and were grown for 5 hours to log and then cytometry reads were performed after 2 hours post induction with α-factor. The treated populations are depicted as blue histograms and the untreated as gray histograms. The solid black lines indicating the median of the untreated controls and the dashed black lines indicating the median of α-factor treated cells. The magnitude of log differences between treated and untreated medians for each plot.

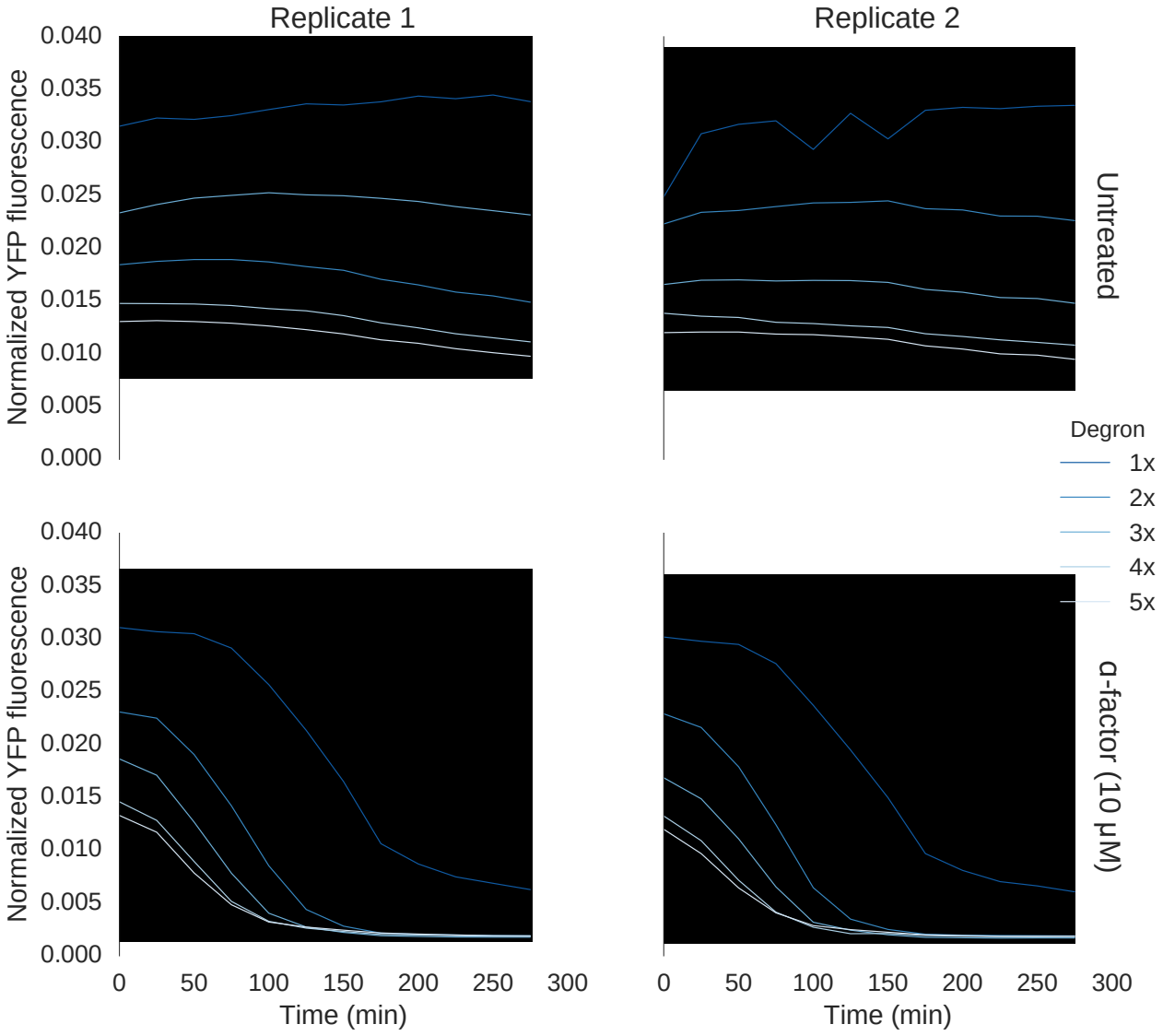


Figure 4 supplement 1. Time course data of reporter variants with different numbers of phosphodegrons normalized by cell size. Each subplot in this plot is a replicate of the experiment performed on different days under identical conditions. Each yeast strain assayed was grown up to log phase from saturated cultures for 5 hours and then induced with 10 μM α-factor at time 0 to activate the kinase. The fluorescence of these cultures was then assayed at regular intervals using flow cytometry. Data for each variant is depicted in a different color with dark blue being one degron and light blue being five tandemly repeating degrons.

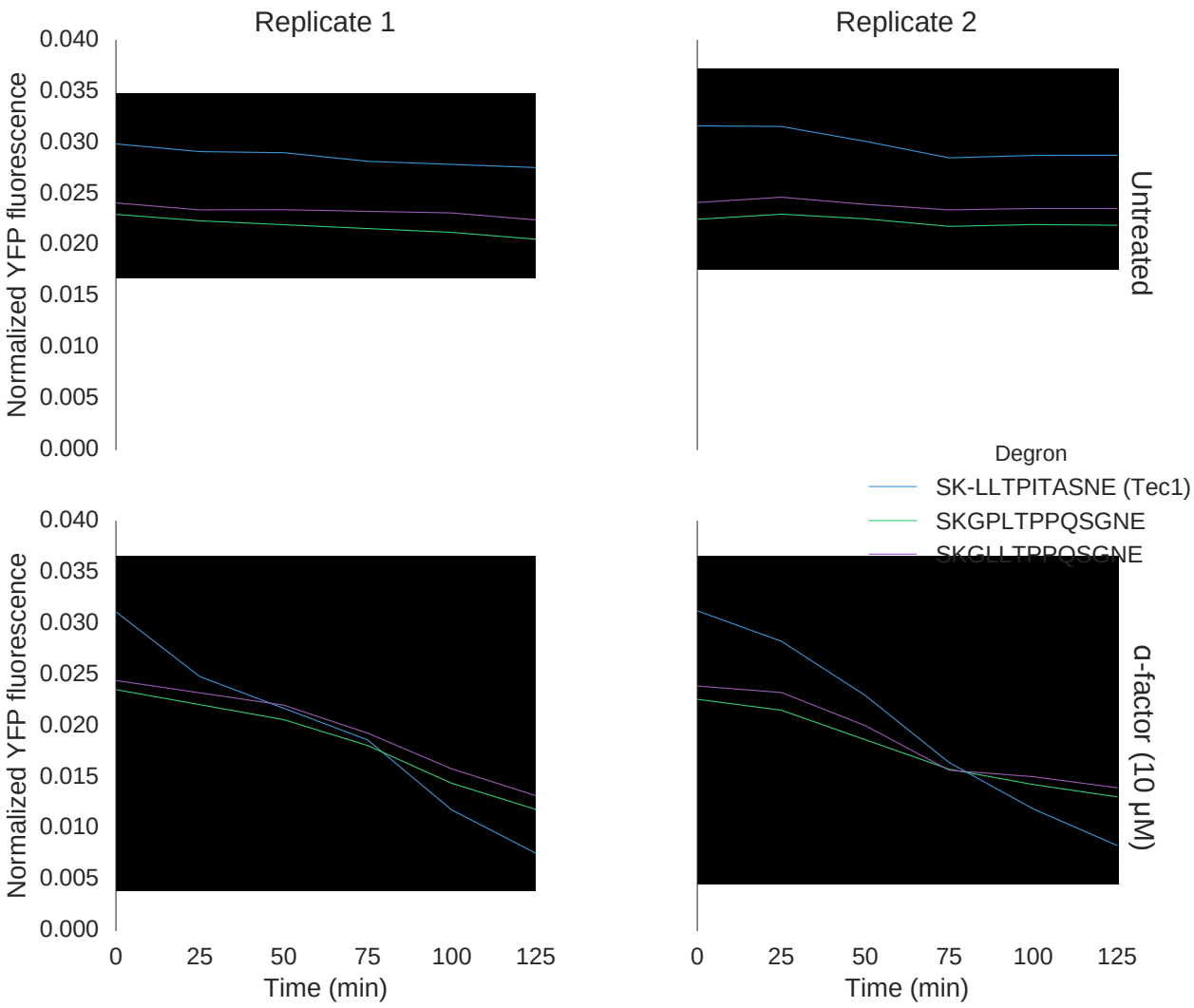


Figure 4 supplement 2. Time course data of reporter variants with different degron sequences normalized by cell size. Each yeast strain assayed was grown up to log phase from saturated cultures for 5 hours and then induced with 10 μM α-factor at time 0 to activate the kinase. The fluorescence of these cultures was then assayed at regular intervals using flow cytometry.

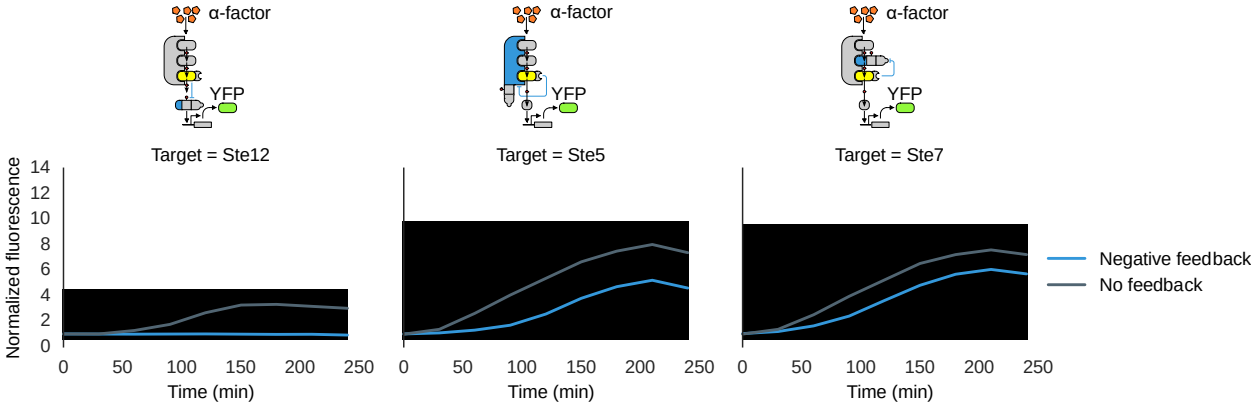


Figure 6 Supplement 1. Time course charecterization of Negative feedback topologies. Timecourse data of YFP fluorescence for populations of yeast cells where the MAPK Fus3 is targeted to phosphorylate and cause the degradation of Ste12, Ste5 or Ste7 (from left to right) after they were induced with 10 µM α-factor at time 0. The blue solid lines depict medians of strains that had a negative feedback topology described in the corresponding cartoon and the gray solid lines are medians for control strains that had a non-frunctional phosphodegrons. The blue and grey ribbions describe the first and third quartiles of the data.

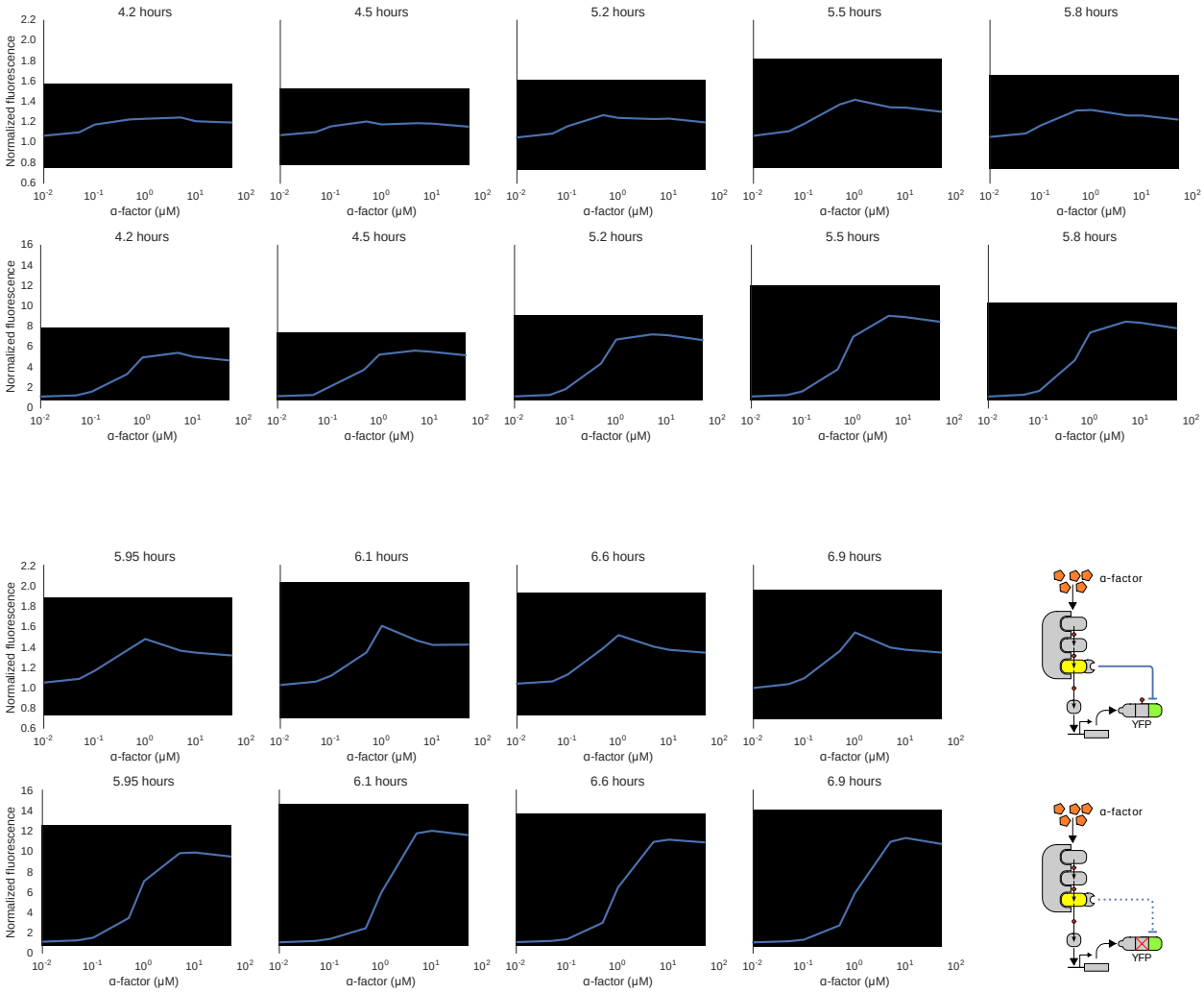


Figure 7 supplement 1. Dose response behavior to α-factor induction over time of a yeast strain with an incoherent feed forward loop implemented on the mating cascade (top row) compared against the behavior of a control system where the loop is broken (bottom row). Each yeast strain assayed was grown up to log phase from saturated cultures for 5 hours and then induced with a range of different α-factor concentrations. The fluorescence of these cultures was then assayed at regular intervals using flow cytometry. Median data (solid blue line) with first and third quartiles (blue ribbon) of the population for two replicates performed on different days under identical conditions is shown.

Acknowledgements

We are grateful to Clare Campbell, Michelle Parks, and Klavins lab technicians for technical support. This work was supported by National Science Foundation (NSF) Awards EFMA 1137266 and CCF 1317653 to GS and an Institute for Protein Design Washington Research Foundation Innovation Fellowship to BG.

References:

1. Cargnello, M. & Roux, P. P. Activation and Function of the MAPKs and Their Substrates , the MAPK-Activated Protein Kinases. *Microbiol. Mol. Biol. Rev.* **75,** 50–83 (2011).

2. Caffrey, D. R., O'Neill, L. A. & Shields, D. C. The evolution of the MAP kinase pathways: coduplication of interacting proteins leads to new signaling cascades. *J. Mol. Evol.* **49,** 567–82 (1999).

3. Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **27,** 514–520 (2002).

4. Elowitz, M. & Lim, W. a. Build life to understand it. *Nature* **468,** 889–890 (2010).

5. Mok, J. *et al.* Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal.* **3,** ra12 (2010).

6. Howard, C. J. *et al.* Ancestral resurrection reveals evolutionary mechanisms of kinase plasticity. *Elife* **3,** 1–22 (2014).

7. Seet, B. T., Dikic, I., Zhou, M.-M. & Pawson, T. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7,** 473–83 (2006).

8. Bhattacharyya, R. P., Reményi, A., Yeh, B. J. & Lim, W. a. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75,** 655–80 (2006).

9. Ubersax, J. a & Ferrell, J. E. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **8,** 530–41 (2007).

10. Skerker, J. M. *et al.* Rewiring the specificity of two-component signal transduction systems. *Cell* **133,** 1043–54 (2008).

11. Won, A. P., Garbarino, J. E. & Lim, W. a. Recruitment interactions can override catalytic interactions in determining the functional identity of a protein kinase. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 9809–14 (2011).

12. Grewal, S., Molina, D. M. & Bardwell, L. Mitogen-activated protein kinase (MAPK)-docking sites in MAPK kinases function as tethers that are crucial for MAPK regulation in vivo. *Cell. Signal.* **18,** 123–134 (2006).

13. Park, S.-H., Zarrinpar, A. & Lim, W. A. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299,** 1061–4 (2003).

14. Whitaker, W. R., Davis, S. a, Arkin, A. P. & Dueber, J. E. Engineering robust control of two-component system phosphotransfer using modular scaffolds. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 18090–5 (2012).

15. Hobert, E. M. & Schepartz, A. Rewiring kinase specificity with a synthetic adaptor protein. *J. Am. Chem. Soc.* **134,** 3976–8 (2012).

16. Harris, K. *et al.* Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins. *Curr. Biol.* **11,** 1815–24 (2001).

17. Yadav, S. S., Yeh, B. J., Craddock, B. P., Lim, W. a & Miller, W. T. Reengineering the signaling properties of a Src family kinase. *Biochemistry* **48,** 10956–62 (2009).

18. Regot, S., Hughey, J. J., Bajar, B. T., Carrasco, S. & Covert, M. W. High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **157,** 1724–34 (2014).

19. Durandau, E., Aymoz, D. & Pelet, S. Dynamic single cell measurements of kinase activity by synthetic kinase activity relocation sensors. *BMC Biol.* **13,** 55 (2015).

20. Sharrocks, A. D., Yang, S. H. & Galanis, A. Docking domains and substrate-specificity determination for MAP kinases. *Trends Biochem. Sci.* **25,** 448–453 (2000).

21. Chang, C. I., Xu, B. E., Akella, R., Cobb, M. H. & Goldsmith, E. J. Crystal structures of MAP kinase p38 complexed to the docking sites on its nuclear substrate MEF2A and activator MKK3b. *Mol. Cell* **9,** 1241–1249 (2002).

22. Heo, Y.-S. *et al.* Structural basis for the selective inhibition of JNK1 by the scaffolding protein JIP1 and SP600125. *EMBO J.* **23,** 2185–2195 (2004).

23. Zhou, T., Sun, L., Humphreys, J. & Goldsmith, E. J. Docking Interactions Induce Exposure of Activation Loop in the MAP Kinase ERK2. *Structure* **14,** 1011–1019 (2006).

24. Tokunaga, Y., Takeuchi, K., Takahashi, H. & Shimada, I. Allosteric enhancement of MAP kinase p38α's activity and substrate selectivity by docking interactions. *Nat. Struct. Mol. Biol.* **21,** 704–711 (2014).

25. Bhattacharyya, R. P. *et al.* The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science* **311,** 822–6 (2006).

26. Khalil, A. S. *et al.* A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150,** 647–58 (2012).

27. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10,** 99–105 (2014).

28. Kiani, S. *et al.* CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods* **11,** 723–6 (2014).

29. Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160,** 339–350 (2015).

30. Khakhar, A., Bolten, N. J., Nemhauser, J. & Klavins, E. Cell-cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors. *ACS Synth. Biol.* 150623113028004 (2015). doi:10.1021/acssynbio.5b00064

31. Brophy, J. a N. & Voigt, C. a. Principles of genetic circuit design. *Nat. Methods* **11,** 508–20 (2014).

32. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology--identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7,** 856–66 (2012).

33. Prindle, A. *et al.* A sensing array of radically coupled genetic 'biopixels'. *Nature* **481,** 39–44 (2012).

34. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661 (2014).

35. Wu, C.-Y., Roybal, K. T., Puchner, E. M., Onuffer, J. & Lim, W. A. Remote control of therapeutic T cells through a small molecule-gated chimeric receptor. *Science (80-.).* **350,** aab4077 (2015).

36. Roybal, K. T. *et al.* Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell* **164,** 770–779 (2016).

37. Morsut, L. *et al.* Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164,** 780–791 (2016).

38. Bardwell, L. A walk-through of the yeast mating pheromone response pathway. *Peptides* **25,** 1465–76 (2004).

39. Hunter, T. The Age of Crosstalk: Phosphorylation, Ubiquitination, and Beyond. *Mol. Cell* **28,** 730–738 (2007).

40. Swaney, D. L. *et al.* Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat. Methods* **10,** 676–682 (2013).

41. Skaar, J. R., Pagan, J. K. & Pagano, M. Mechanisms and function of substrate recruitment by F-box proteins. *Nat. Rev. Mol. Cell Biol.* **14,** 369–381 (2013).

42. Chou, S., Huang, L. & Liu, H. Fus3-regulated Tec1 degradation through SCFCdc4 determines MAPK signaling specificity during mating in yeast. *Cell* **119,** 981–90 (2004).

43. Bao, M. Z., Shock, T. R. & Madhani, H. D. Multisite phosphorylation of the Saccharomyces cerevisiae filamentous growth regulator Tec1 is required for its recognition by the E3 ubiquitin ligase adaptor Cdc4 and its subsequent destruction in vivo. *Eukaryot. Cell* **9,** 31–6 (2010).

44. Nash, P. D. *et al.* Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414,** 514–21 (2001).

45. Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* **112,** 243–56 (2003).

46. Blondel, M. *et al.* Nuclear-specific degradation of Far1 is controlled by the localization of the F-box protein Cdc4. *EMBO J.* **19,** 6085–6097 (2000).

47. Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, a E. A short amino acid sequence able to specify nuclear location. *Cell* **39,** 499–509 (1984).

48. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27,** 753–9 (2009).

49. Moon, T. S., Dueber, J. E., Shiue, E. & Prather, K. L. J. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. *Metab. Eng.* **12,** 298–305 (2010).

50. Ryu, J. & Park, S. Simple synthetic protein scaffolds can create adjustable artificial MAPK circuits in yeast and mammalian cells. *Sci. Signal.* **8,** ra66 (2015).

51. Gartner, A., Nasmyth, K. & Ammerer, G. Signal transduction in Saccharomyces cerevisiae requires tyrosine and threonine phosphorylation of FUS3 and KSS1. *Genes Dev.* **6,** 1280–1292 (1992).

52. Thompson, K. E., Bashor, C. J., Lim, W. a & Keating, A. E. SYNZIP Protein Interaction Toolbox: in Vitro and in Vivo Specifications of Heterospecific Coiled-Coil Interaction Domains. *ACS Synth. Biol.* **1,** 118–129 (2012).

53. Liang, F.-S., Ho, W. Q. & Crabtree, G. R. Engineering the ABA plant stress pathway for regulation of induced proximity. *Sci. Signal.* **4,** rs2 (2011).

54. Reményi, A., Good, M. C., Bhattacharyya, R. P. & Lim, W. a. The role of docking interactions in mediating signaling input, output, and discrimination in the yeast MAPK network. *Mol. Cell* **20,** 951–62 (2005).

55. O'Shaughnessy, E. C., Palani, S., Collins, J. J. & Sarkar, C. A. Tunable signal processing in synthetic MAP kinase cascades. *Cell* **144,** 119–31 (2011).

56. Lin, Y. W. & Yang, J. L. Cooperation of ERK and SCFSkp2 for MKP-1 destruction provides a positive feedback regulation of proliferating signaling. *J. Biol. Chem.* **281,** 915–926 (2006).

57. Robinson, M. J., Stippec, S. a, Goldsmith, E., White, M. a & Cobb, M. H. A constitutively active and nuclear form of the MAP kinase ERK2 is sufficient for neurite outgrowth and cell transformation. *Curr. Biol.* **8,** 1141–1150 (1998).

58. Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8,** 450–61 (2007).

59. Bashor, C. J., Helman, N. C., Yan, S. & Lim, W. a. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* **319,** 1539–43 (2008).

60. Kholodenko, B. N. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur. J. Biochem.* **267,** 1583–8 (2000).

61. Levchenko, A., Bruck, J. & Sternberg, P. W. Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5818–23 (2000).

62. Yeger-Lotem, E. *et al.* Network motifs in integrated cellular networks of transcription-

regulation and protein-protein interaction. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5934–9 (2004).

63. Mishra, D., Rivera, P. M., Lin, A., Del Vecchio, D. & Weiss, R. A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.* (2014). doi:10.1038/nbt.3044

64. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 11980–11985 (2003).

65. Beltrao, P. *et al.* Systematic Functional Prioritization of Protein Posttranslational Modifications. *Cell* **150,** 413–425 (2012).

66. Schneekloth, J. S. *et al.* Chemical genetic control of protein levels: selective in vivo targeted degradation. *J. Am. Chem. Soc.* **126,** 3748–54 (2004).

67. Melchionna, T. & Cattaneo, A. A protein silencing switch by ligand-induced proteasome-targeting intrabodies. *J. Mol. Biol.* **374,** 641–54 (2007).

68. Bonger, K. M., Chen, L., Liu, C. W. & Wandless, T. J. Small-molecule displacement of a cryptic degron causes conditional protein degradation. *Nat. Chem. Biol.* **7,** 531–537 (2011).

69. Neklesa, T. K. *et al.* Small-molecule hydrophobic tagging–induced degradation of HaloTag fusion proteins. *Nat. Chem. Biol.* **7,** 538–543 (2011).

70. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43,** D512–D520 (2015).

71. Hornbeck, P. V *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40,** D261–70 (2012).

72. Beltrao, P. *et al.* Evolution of Phosphoregulation: Comparison of Phosphorylation Patterns across Yeast Species. *PLoS Biol.* **7,** e1000134 (2009).

73. Reimand, J. & Bader, G. D. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol. Syst. Biol.* **9,** 637–637 (2014).

74. Ryu, G. M. *et al.* Genome-wide analysis to predict protein sequence variations that change phosphorylation sites or their corresponding kinases. *Nucleic Acids Res.* **37,** 1297–1307 (2009).

75. Harris, B. Z. & Lim, W. a. Mechanism and role of PDZ domains in signaling complex assembly. *J. Cell Sci.* **114,** 3219–3231 (2001).

76. Mayer, B. J. SH3 domains: complexity in moderation. *J. Cell Sci.* **114,** 1253–1263 (2001).

77. Reményi, A., Good, M. C. & Lim, W. a. Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol.* **16,** 676–85 (2006).

78. Van Roey, K. *et al.* Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem. Rev.* **114,** 6733–6778 (2014).

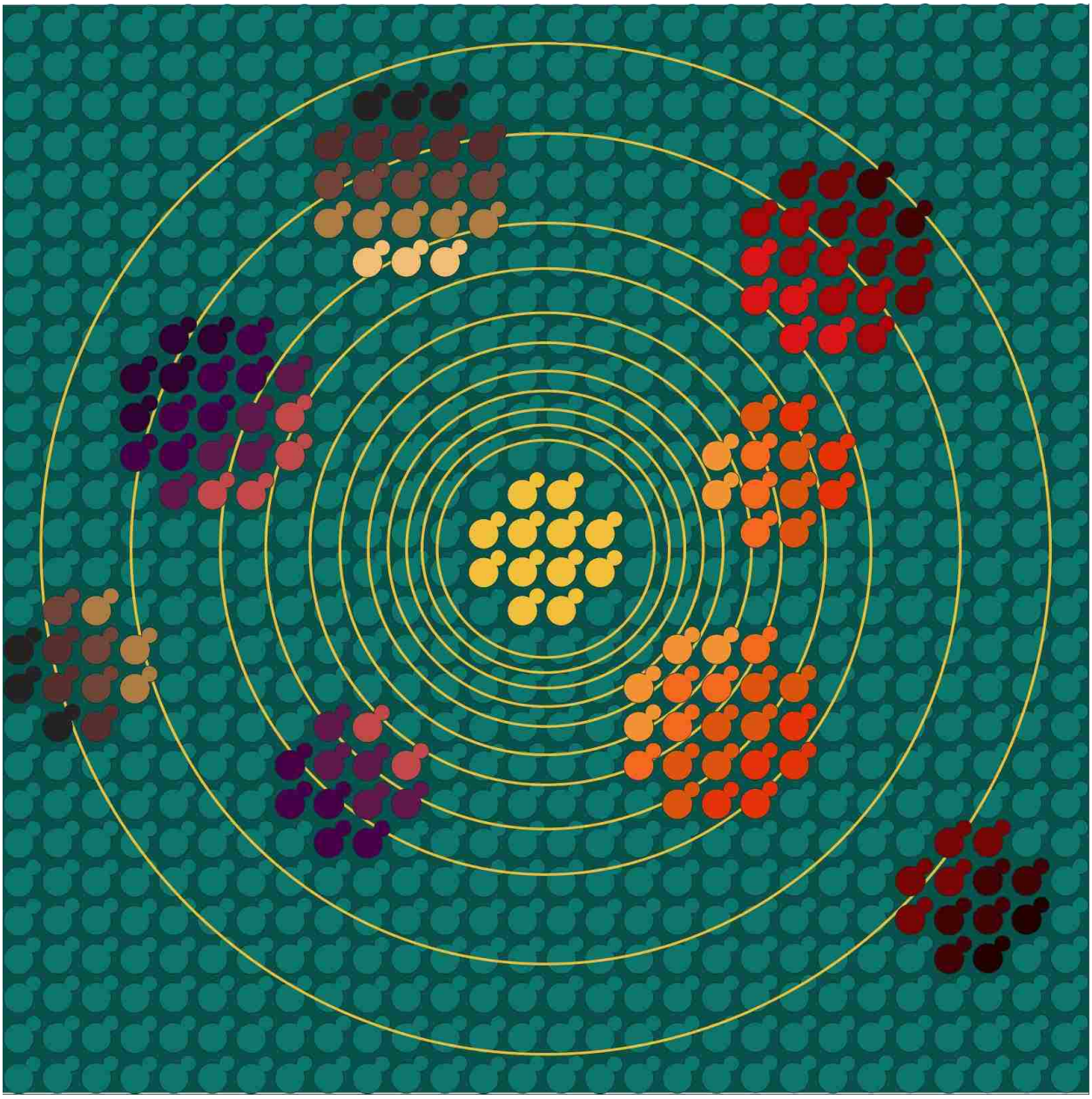
79. Neduva, V. & Russell, R. B. Linear motifs: Evolutionary interaction switches. *FEBS Lett.* **579,** 3342–3345 (2005).

80. Davey, N. E., Cyert, M. S. & Moses, A. M. Short linear motifs – ex nihilo evolution of protein regulation. *Cell Commun. Signal.* **13,** 43 (2015).

81. Beltrao, P. & Serrano, L. Specificity and evolvability in eukaryotic protein interaction networks. *PLoS Comput Biol* **3,** e25–e25 (2007).

82. Shvartsman, S. Y. & Baker, R. E. Mathematical models of morphogen gradients and their effects on gene expression. *Wiley Interdiscip. Rev. Dev. Biol.* **1,** 715–730 (2012).

Chapter 3 - Engineering flow of information flow from one cell to another: "Cell–cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors"



Abstract

An engineering framework for synthetic multicellular systems requires a programmable means of cell-cell communication. Such a communication system would enable complex behaviors, such as pattern formation, division of labor in synthetic microbial communities, and improved modularity in synthetic circuits. However, it remains challenging to build synthetic cellular communication systems in eukaryotes due to a lack of molecular modules that are orthogonal to the host machinery, easy to reconfigure, and scalable. Here, we present a novel cell-to-cell communication system in *Saccharomyces cerevisiae* (yeast) based on CRISPR transcription factors and the plant hormone auxin that exhibits several of these features. Specifically, we engineered a *sender* strain of yeast that converts indole-3-acetamide (IAM) into auxin via the enzyme iaaH from *Agrobacterium tumefaciens*. To sense auxin and regulate transcription in a *receiver* strain, we engineered a reconfigurable library of auxin degradable CRISPR transcription factors (ADCTFs). Auxin-induced degradation is achieved through fusion of an auxin sensitive degron (from IAA co-repressors) to the CRISPR TF and co-expression with an auxin F-box protein. Mirroring the tunability of auxin perception in plants, our family of ADCTFs exhibits a broad range of auxin sensitivities. We characterized the kinetics and steady state behavior of the sender and receiver independently, and in co-cultures where both cell types were exposed to IAM. In the presence of IAM, auxin is produced by the sender cell and triggers de-activation of reporter expression in the receiver cell. The result is an orthogonal, rewireable, tunable, and arguably scalable cell-cell communication system for yeast and other eukaryotic cells.

Introduction

Multicellular systems in nature are capable of incredible feats of distributed computation and self-organization. Examples range from division of labor in filamentous algae[1], to the exquisite sensitivity of the adaptive immune system[2], to morphogenesis and development of tissues and organs. Computer scientists have shown that cells are in principle capable of computing a wide variety of functions[3], generating complex morphologies[4], and of making decisions[5,6]. Experimentally, synthetic multicellular systems have been built to regulate populations[7], synchronize oscillations[8], form patterns[9–11], implement logic functions through distributed computation[3], and cooperate to solve problems[12]. However, a scalable suite of cell-cell communication modules has yet to emerge. In particular, in *Saccharomyces cerevisiae*, strategies that use components of native signal transduction pathways can lead to crosstalk and undesirable phenotypes such as growth arrest[9,13,14]. Such systems are not obviously portable to other eukaryotes, are difficult to reprogram, and require significant changes to the host cell to function correctly[7]. Here, we describe progress toward building an engineering framework for yeast cell-cell communication that is orthogonal to yeast (and many other eukaryotic cells except plants[15]), modular, and tunable.

Orthogonality is crucial for rationally engineering cell-cell communication. Auxin, a plant hormone, does not have measurable effects on laboratory strains of yeast[16,17] when grown in standard conditions. Our receiver cells use elements of the *Arabidopsis thaliana* auxin signaling pathway. Auxin regulates plant development via a system of transcriptional corepressors, the Aux/IAA proteins (referred to as IAAs), which are degraded in the presence of the molecule auxin. Auxin stabilizes the interaction between the degron domain of an IAA and an auxin-signaling F-box protein (AFB). The result is the degradation of the IAA via polyubiquitination[18]. The IAAs exhibit a range of degradation rates and sensitivities to auxin that

are determined, in part, by the sequence of their degron domains and in part by the AFB[16,19]. The degradation dynamics of a large range of auxin degrons with multiple AFBs have been previously studied and thoroughly characterized in yeast[16]. By using this signaling modality as the basis for our communication system, we avoid using any native yeast (or mammalian) signal transduction machinery associated with adverse phenotypes[7]. Additionally, the primary components of the pathway, AFBs and IAAs, have been shown to function in several different mammalian cells[15], suggesting that our system may be broadly portable.

To maximize modularity, we engineered auxin responsiveness into CRISPR transcription factors (CTFs). CTFs consist of a nuclease null Cas9 protein (dCas9) fused to a transcriptional effector domain. The dCas9 can be programmed to target a locus by coexpressing a small guide RNA (gRNA) that has complementarity to the target locus at a site that is adjacent to an 'NGG' sequence, called the PAM sequence. This strategy, as demonstrated by Farzadfard et al.[20] and Qi et al.[21], has the benefit of modularity through easily programmable specificity: dCas9 requires only the expression of a new guide RNA for retargeting. In contrast, zinc finger or TAL DNA binding domains require the design of a new protein for each target[22,23]. These characteristics make CTFs an ideal candidate for signal reception and processing, as they can be targeted to any promoter in the genome that has a suitable PAM site[20], can either activate or repress gene expression, and can be layered to form more complex networks[23,24]. In the present case, CTFs fused to the VP64 strong activator domain were targeted to a promoter upstream of GFP. In addition, these CTFs were fused to Aux/IAA degron domains and co-expressed with AFBs thereby producing auxin-degradable CRISPR transcription factors, or ADCTFs. An ADCTF is thus a modular, coupled sensor-actuator, which should allow cell-to-cell communication to be easily rewired to arbitrary outputs.

Signal production and reception in cell-cell communication is ideally tunable to achieve a broad range of sensitivities and other functions. To implement and tune auxin production in the sender, we integrated the bacterial iaaH gene from *Agrobacterium tumefaciens* into yeast under the control of a constitutive promoter (GPD). Upon the addition of indole-3-acetamide (IAM), sender cells produced a strong enough auxin signal to affect gene regulation via the ADCTFs in co-cultured receiver cells. The concentration of auxin produced can be tuned via the concentration of exogenously added IAM. For increased tunability, we developed a library of ADCTFs, each with a different degron and/or degron location, which displays a range of degradation kinetics and sensitivities to auxin. The sensitivity of the ADCTFs can be further tuned by the selection of the F-box that is coexpressed the ADCTF. Thus, components of the ADCTFs, the auxin degron, and the transcriptional effector domain can all be swapped to obtain, respectively, a range of auxin sensitivities, and repression versus activation.

In summary, the combination of sender and receiver modules described here forms the foundation of an orthogonal, modular, and tunable cell-cell communication framework for yeast. We demonstrate each of these aspects of the system below by describing how the senders and receivers behave in isolation, and that they can be combined in co-culture to form a simple communication channel.

Results

Synthetic, scalable, auxin-modulated transcription factors

To link an auxin sensor to diverse transcriptional responses and targets, we designed auxin degradable CRISPR transcription factors (ADCTFs) with three modular domains (Figure 1A). The core component of the ADCTFs is the CRISPR-based transcription factor described by Farzadfard et al[18], wherein a deactivated Cas9 protein functions as a programmable DNA binding module. The dCas9 was fused to a transcriptional effector domain, in this case the transcriptional activator VP64, and to an IAA degron. In the presence of an AFB, ADCTFs should be ubiquitinated and degraded when exposed to auxin. We tested the ADCTFs by targeting them to activate the expression of EGFP from a minimal CYC1 promoter and observed deactivation of fluorescence upon the addition of auxin. In the absence of auxin, functional ADCTFs significantly activated the production of EGFP as compared to controls lacking a gRNA (Figure 1B). When a functional (coexpressed with gRNA) activator ADCTF was degraded in the presence of auxin, fluorescence dropped to levels at or below the control (no gRNA) levels. Auxin dependent regulation was independent of the promoter being regulated by the ADCTF (Supplementary Figure 1B). The observed effect was also reversible: when auxin was removed from the system, reporter expression returned to its activated state (Supplementary Figure 1A).

One design consideration for building the ADCTFs was the position of the degron within the fusion protein. We hypothesized that degron position could alter accessibility to the AFB or otherwise interfere with protein folding thus modulating auxin sensitivity. We explored several possible positions for the degron relative to the other domains (Figure 1C). In all cases, the degron was flanked by flexible linkers composed of five repeats of the amino acid sequence “GS” to limit fusion-associated misfolding. Changing the position of the degron dramatically altered the sensitivity range, defined as the range of auxin concentrations between which steady-state fluorescence drops from 90% of maximum to 10% (Figure 1D). Position one is sensitive to the lowest levels of auxin, but also saturates earlier than positions two and three. Placing the degron on either side of dCas9 (positions one and two) resulted in higher auxin sensitivity than position three where the degron was placed at the C-terminal end of the fusion. The percentage drop from maximal activation upon auxin induction was directly correlated to auxin sensitivity, with position one dropping to basal levels at steady state, and positions two and three having progressively smaller effects post induction (Supplementary Fig 2). Altering the position of the degron coarsely tuned the upper and lower bounds of the sensitivity range of the ADCTF. However, since the position one variant was the most sensitive to auxin and had the highest fold change, we chose to fuse degrons in all further ADCTF variants at position one.

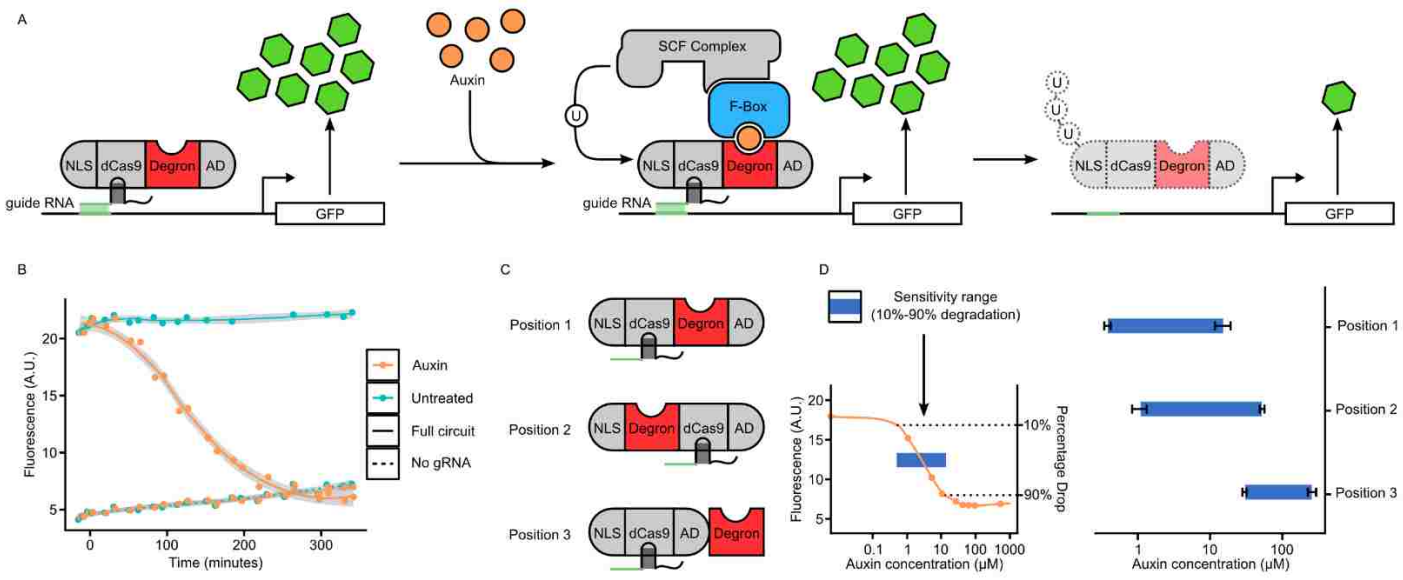


Figure 1. A) The ADCTF design and the molecular mechanism behind its function. An ADCTF is made up of a dCas9 protein fused to an NLS, an activation domain and an auxin sensitive degron. In the presence of auxin, the degron recruits an Auxin Sensing F-box (AFB) protein to form an SCF complex (an E3 ubiquitin ligase). The subsequent ubiquination and degradation of the ADCTF deregulates the gene targeted by the ADCTF. B) Time-lapse cytometry of ADCTF cells with a GFP-producing gRNA target following the addition of auxin or no treatment as well as with and without a guide RNA. The gray ribbon indicates the 95% confidence interval. Following treatment with auxin, the GFP level of the strain expressing gRNA dropped to basal levels (equivalent to a strain with no gRNA). C) Schematic representation of the three fusion proteins tested for the effect of degron position on ADCTF properties. D) Sensitive range characterization of the three degron position variants at steady state. Horizontal bars indicate the range of auxin concentrations between which mean steady-state fluorescence (measured via cytometry) drops from 90% of maximum to 10%. A larger sensitive range correlated with higher maximum fold changes upon induction (Supplementary Figure 2).

Engineered ADCTF variants exhibit a broad range of auxin sensitivities and degradation kinetics

The Aux/IAA family of 29 transcriptional corepressors have been shown to exhibit a large range of degradation rates and sensitivities to auxin in yeast[16]. This range of responses to the same auxin signal is thought to result in part from the sequence of the different IAA degron domains, and in part from the varying activities of different AFBs, each showing different affinities for specific IAAs. We built a library of ADCTFs using degrons from IAA14, IAA15, and IAA17 and coexpressed them with either of two F-boxes (AFB2 or TIR1). These degrons have been previously characterized as encompassing a range of auxin-induced degradation rates. In general, AFB2 promotes faster degradation of IAAs than TIR1. In addition, we included a recently characterized mutant of TIR1, tir1-D170E/M473L (referred to hereafter as tir1-D/M)[25], which has been shown to greatly accelerate auxin-induced TIR1 degradation.

All pairwise combinations of ADCTFs and F-box proteins were tested for their temporal response and dose response to auxin. Temporal responses, all performed with 30 μM auxin induction, exhibited a range of degradation kinetics that depended on both the choice of ADCTF degron and the F-box protein (Figure 2B). The kinetics, characterized by the time to 50% degradation, can be coarsely tuned by the choice of F-Box protein used, with tir1-D/M being the fastest overall, followed by AFB2 and TIR1. Within this coarse tuning, the choice of degron allowed for smaller changes in kinetics. The ADCTF with the degron from IAA15 (ADC15) seemed to have the overall fastest kinetics. The only exception to this trend was the interaction between AFB2 and ADC17, which had the fastest degradation rate. All the ADCTFs had approximately the same percentage change from maximal activation upon auxin induction at steady

state. Thus, tuning kinetics by swapping F-box proteins or degrons had a minimal effect on the steady state response to auxin. Most variants dropped to approximately 75% of maximal activation at steady state with a few between 10% higher or lower than the mean (Supplementary Fig 3D). The ADCTFs exhibited varied sensitivity to auxin that depended on the combination of the degron on the ADCTF and the F-box protein. Swapping F-box proteins allowed for more coarse grain tuning of sensitivity range with TIR1 conferring the broadest sensitivity range overall and tir1-D/M conferring the narrowest (Figure 2C). Swapping degrons allows smaller changes, as was observed within the AFB2 variants wherein there is a progressively narrower sensitivity range from ADC14 to ADC17. The dynamics and steady state behavior of the ADCTFs in response to auxin correspond to the behavior of previously characterized IAA proteins, from which the degrons were taken, in yeast[14]. The only exception being the degron 17 variant, which had much slower degradation kinetics in the ADCTF context in a tir1-D/M background. This result suggests that the auxin responsive behavior may be predictably tuned by swapping degron and F-box protein variants.

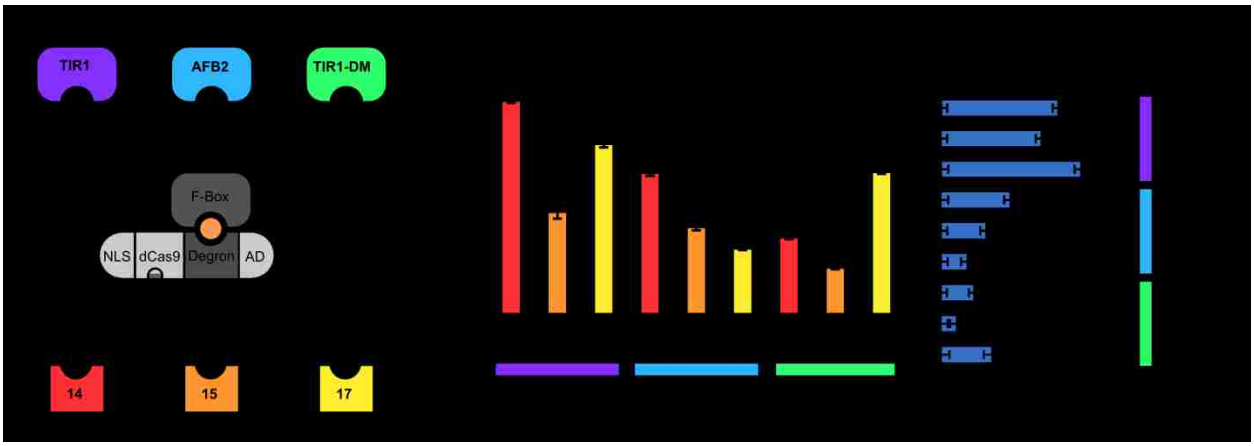


Figure 2. A) The ADCTF (auxin receiver) strain library was generated from all pairwise combinations of three Auxin Sensing F-box protein variants (AFB2, TIR1, tir1-D/M) and three auxin degron variants (from IAA14, IAA15 and IAA17). B) Receiver strain library degradation kinetics measured via time-lapse cytometry. The kinetics of ADCTF responses to auxin were characterized by the time at which fluorescence dropped to fifty percent of maximum: a smaller time implies a faster response. The ADCTF library displays a wide range of degradation kinetics that were modulated by both the choice of F-box protein and the auxin degron. C) Auxin sensitivity ranges for the ADCTF library. Blue bars represent the auxin sensitivity range at steady state as defined in Figure 1. The errors bars correspond to standard deviation of the average population values for individual replicates.

Yeast produce tunable levels of auxin via expression of iaaH from *Agrobacterium tumefaciens*

To generate an auxin producing strain, we integrated half of the IAM pathway from *Agrobacterium tumefaciens* into yeast[26]. The IAM pathway is a two-step enzymatic process that converts tryptophan to IAM and then into auxin. The first step is via tryptophan-2-monooxygenase (iaaM, not examined here). The second step is catalyzed by indole-3-acetamide hydrolase (iaaH). To test whether yeast could produce auxin from IAM using only the second enzyme, we integrated the iaaH gene from the *Agrobacterium tumefaciens* Ti plasmid[27] into an auxin reporting yeast strain (Figure 3A) containing a IAA-YFP fusion protein. After adding IAM, reporter degradation rates were measured via time-lapse cytometry (Figure 3B). Upon the addition of IAM, sender strains produced an auxin response comparable to that of native auxin (Figure 3C). In addition, for a given concentration of IAM, the steady state fluorescence values converge to those of auxin (Figure 3D). There was no significant delay between the addition of IAM and the production of auxin, so the transport and production of auxin from IAM can be assumed to be faster than the reporter's dynamics. We then investigated intercellular auxin

production by coculturing the sender strain with an auxin sensor strain that could be distinguished via its mCherry signal (Figure 4A). Rather than a dose response of IAM, increasing fractions of sender cells were cocultured with sensor strains in constant amount of the auxin precursor (300 µM) to test the dependence of auxin production on sender cell concentration (Figure 4B). Greater concentrations of sender cells produced a greater auxin response in sensor cells. Both the kinetic and steady state behavior suggest there is a lower concentration of auxin in the media than within sender cells.(Figure 4C, Figure 4D).

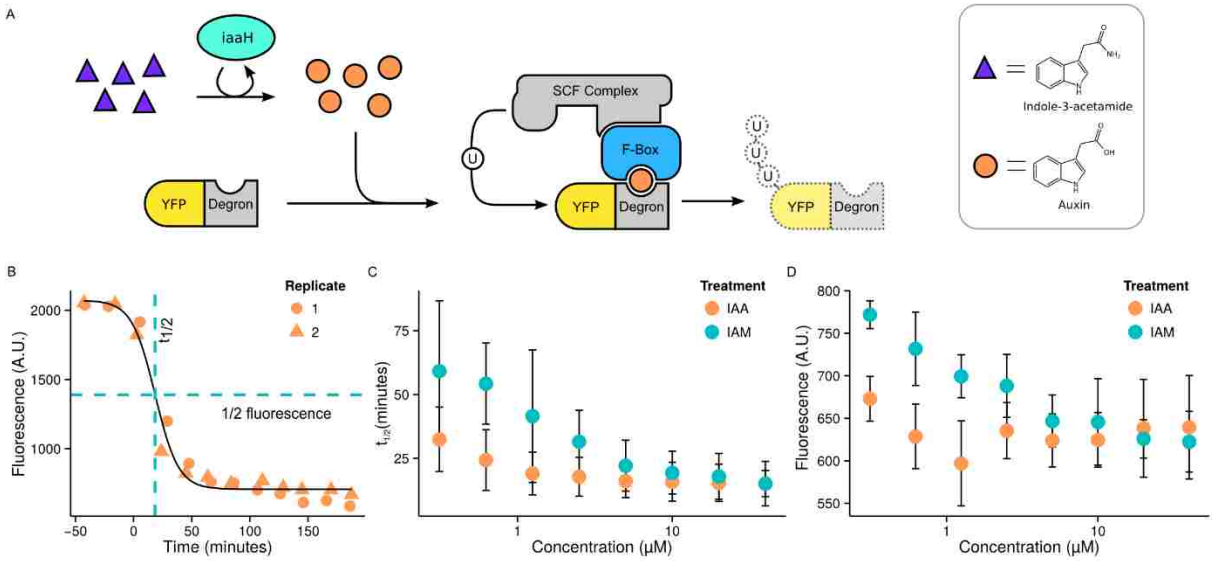


Figure 3. A) Auxin sender strain design. The iaaH enzyme of *Agrobacterium tumefaciens* catalyzes the conversion of indole-3-acetamide (IAM) into auxin, inducing the degradation of proteins fused to an auxin degron. The iaaH enzyme (sender cells) was integrated into an auxin reporting strain (the EYFP-IAA17|AFB2 strain from *Havens et al*[16]) to test for internal auxin production. B) Kinetic auxin response to IAM addition in sender strains. Following the addition of IAM, the fluorescence of sender cells decreased to basal levels. The time to half-maximal ($t_{1/2}$) fluorescence was used to measure the rate of reporter degradation. C) Auxin-induced degradation rate in response to varying doses of either IAM or auxin. Sender cells were treated with either auxin or IAM and read at regular intervals producing time courses as in part B. Nonlinear fitting was used to generate $t_{1/2}$ values. For a given molarity, treatment with IAM produces an auxin-induced degradation similar to, but weaker than, direct treatment with auxin. D) Steady state fluorescence in response to varying doses of either IAM or auxin taken from the same dataset as part C. As the concentration of IAM was increased, a lower steady state fluorescence was produced. The errors bars correspond to standard deviation of the average population values for individual replicates.

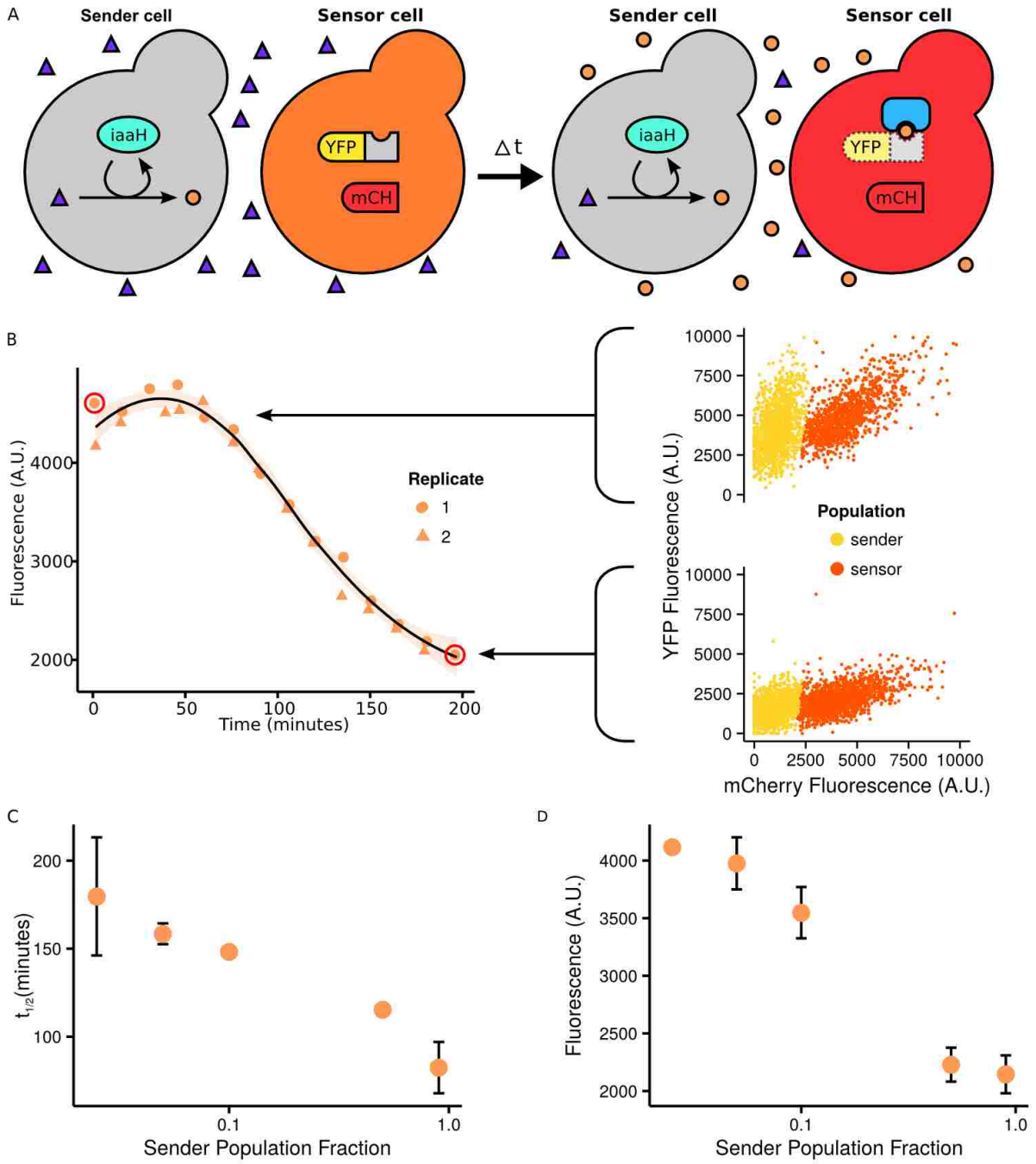


Figure 4. A) Sender-sensor multicellular auxin signaling strains. Sender cells are identical to those in Figure 3 and therefore produce auxin upon the addition of exogenous IAM and sense auxin production via an EYFP-IAA17 reporter. Sensor cells express an EYFP-IAA17 and TIR1 and are distinguished experimentally through the expression of mCherry. In coculture, IAM diffuses into sender cells where it is converted into diffusible auxin that then degrades EYFP-degron proteins in either the sender or sensor cell types. B) Auxin-induced degradation of EYFP-IAA17 in sensor cells cocultured with sender cells in 300 µM IAM. Data for sensor cells can be separated from sender cells via their mCherry signal. The line represents a LOESS fit and the light orange ribbon represents a 95% confidence interval of the fit. C) Sender cell fraction dose response. Each fraction had the same volume, so a larger fraction indicates a larger concentration of sender cells in coculture. As the sender cell population increases, the degradation rates decreases. D) Steady state fluorescence in response to varying doses of either IAM or auxin

taken from the same dataset as part C. As the concentration of sender cells was increased, a lower steady state fluorescence was produced, flatting out at around a 50:50 split. The errors bars correspond to standard deviation of the average population values for individual replicates.

Sender cells produce a tunable auxin response in receiver cells

Sender and receiver cells were cocultured in different ratios to measure the effect of sender cell concentration on auxin signal production. Senders constitutively express iaaH and the receivers expressed an activating ADCTF and a gRNA targeting a minimal CYC1 promoter driving EGFP (Figure 5A). After adding a saturating amount of the IAM and growing the coculture overnight, we observed a reduction in gene activation in the receiver strain comparable to direct addition of auxin (Figure 5B). Three different receiver strains with a range of responses to auxin were tested with the sender strain. All the receiver strains produced an auxin response and behaviors were consistent to those observed via the direct addition of auxin, suggesting that the sender module is compatible with any ADCTF-based receiver module (Supplementary Figure 4). In addition, a 10% fraction of sender cells is sufficient to a significant change in fluorescence in receiver cell at steady state and a 50% fraction produces a nearly saturating signal (Figure 5C, Figure 5D).

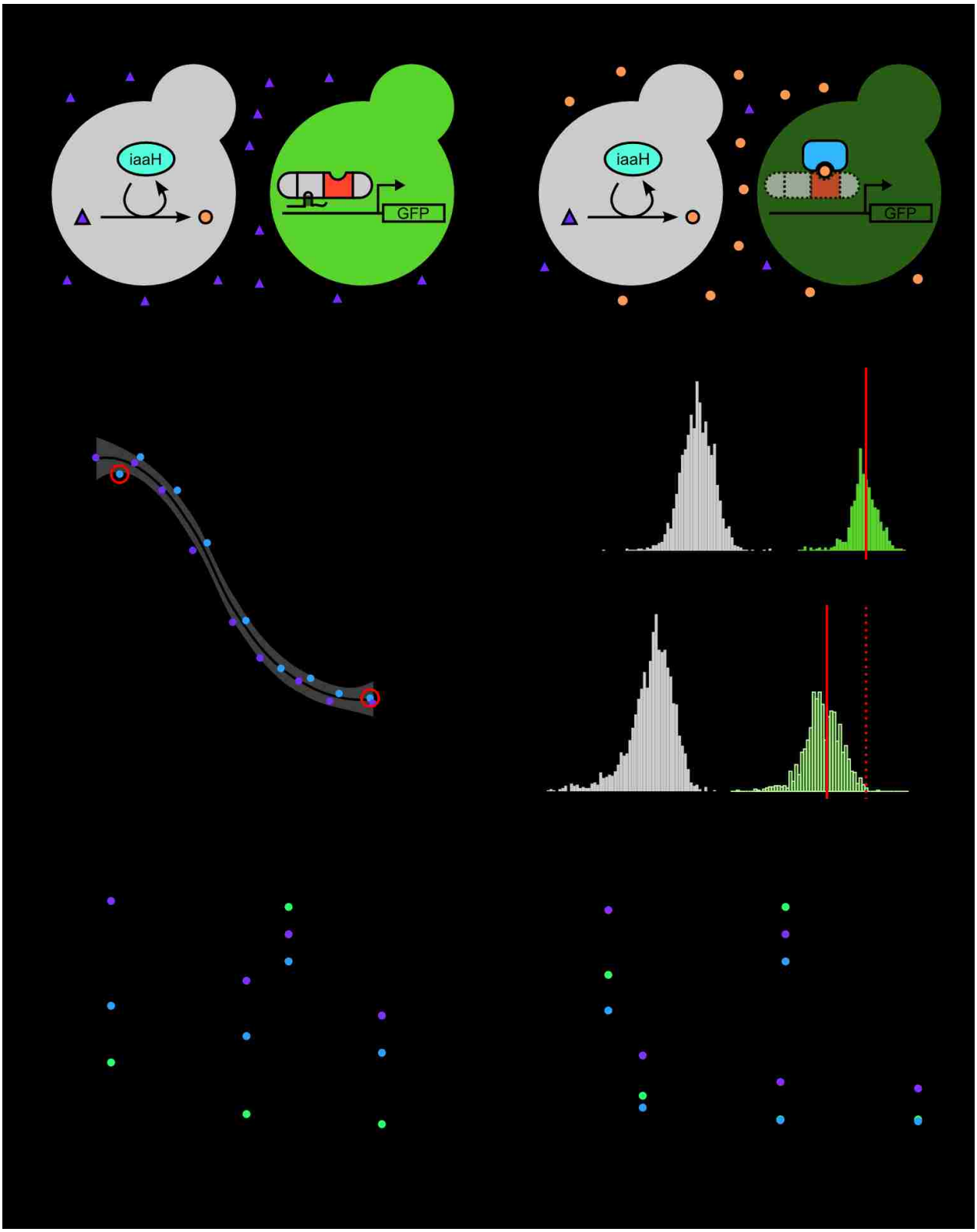


Figure 5. A) Coculture of sender and receiver strains. Sender cells convert IAM into auxin that then diffuses out of sender cells and into receiver cells where it causes the degradation of ADCTFs, producing a drop in fluorescence. B) Time course data for two replicates (shown in blue and purple dots) of a coculture of equal concentrations of sender and receiver cells is plotted on the left. The line represents a LOESS fit and the gray ribbon represents a 95% confidence interval of the fit. On the right,

histograms display distinct populations of sender (gray, left) and receiver (green, right) cells. In the presence of sender cells treated with IAM, receiver cells dropped in fluorescence over time. As in figures 3 and 4, sender cells express also an EYFP-IAA and AFB auxin reporter and therefore also show a decrease in fluorescence. Without IAM, receiver cells did not show a significant decrease in fluorescence. C) Degradation rates (measured as $t_{1/2}$) in receiver strains in response to sender cell concentration. As the fraction of sender cells increased, there is a more dramatic auxin effect in receiver cells that saturates at approximately even fractions of send to receive. The errors bars correspond to standard deviation of the average population values for individual replicates.

Discussion

Our system is based on a signal transduction modality that is unique to plants and so is orthogonal to native yeast signal transduction pathways, as well as to mammalian cells[15]. The simplicity of the system will hopefully allow it to be ported to other contexts, such as mammalian cells. The ADCTF library allows the generation of a range of responses to the same auxin signal, and can in principle be connected to any gene of interest, or to another synthetic gene circuit. Additionally, auxin production levels can also be tuned by titrating in different amounts of IAM. It may also be possible to tune the diffusivity of auxin in yeast[26], or to harness the sequestration and turnover pathways of auxin found in plants. Our approach of detecting small molecules via F-Box mediated degradation of a transcription factor is potentially scalable as there are other plant hormones such as jasmonate that use a very similar signaling pathway[28]. Current work involves building on these characteristics to produce more complex multicellular behaviors. For example, feedback systems can be built through regulation of the iaaH gene via the ADCTFs. More generally, our results form the basis platform for implementing distributed decision making, pattern formation, and other complex cell-to-cell communication based multicellular behaviors.

Methods

Strain construction

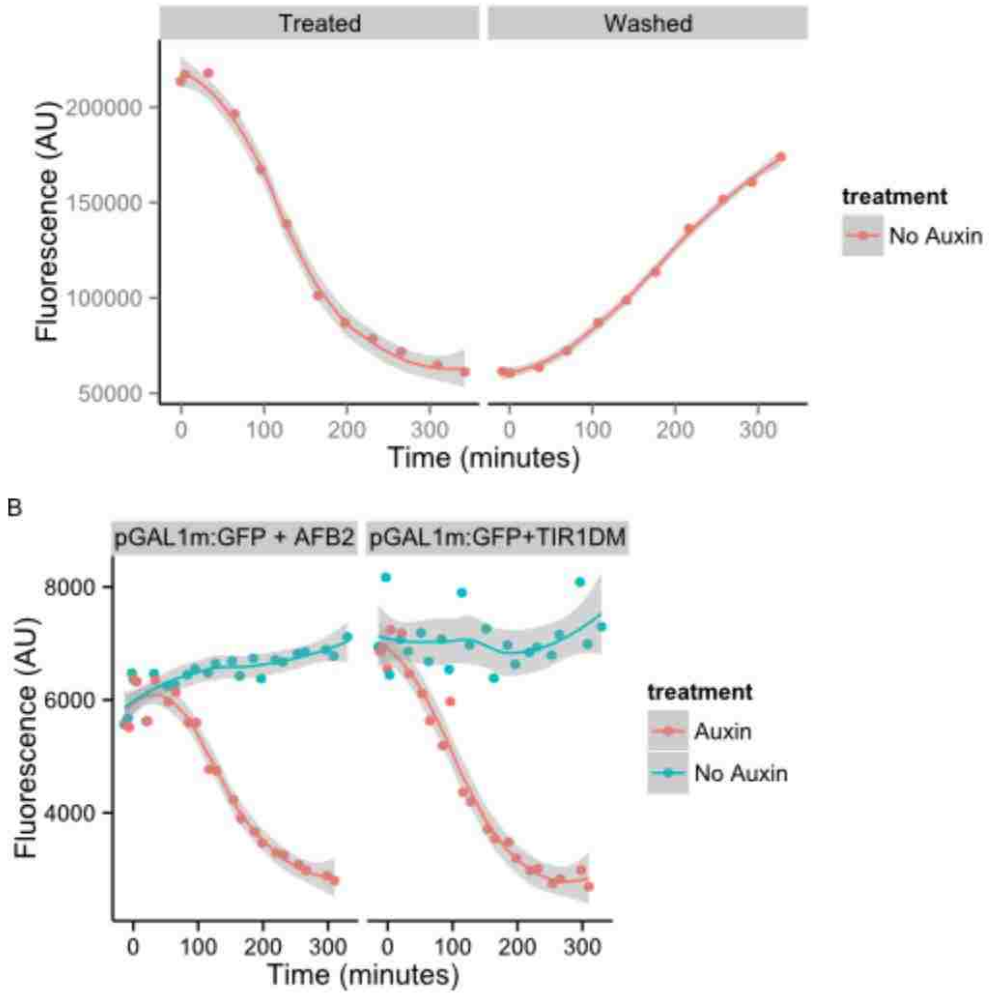
Building off the work of Farzadfard et al[20], the reporter is a yeast-enhanced green fluorescent protein driven by a truncated CYC1 promoter. This reporter was integrated at the URA3 locus in the genome of the W303-1A ADE2 strain of *Saccharomyces cerevisiae* and this reporter strain was used as the parent for all ADCTF strains. All gRNA was driven by an ADH1 promoter driven construct that consists of a gRNA flanked on each side by a hammerhead and an HDV ribozyme, facilitating expression from an RNA polymerase II promoter. All the gRNA constructs were integrated at the HIS3 locus. AFB2, TIR1 and tir1-D/M were integrated, respectively, at the LEU2 locus, and were driven by the GPD promoter. The ADCTFs were constructed by fusing an SV40 nuclear localization tag, a VP64 activation domain, and an auxin degron to a nuclease null version of the Cas9 protein from *Streptococcus pyogenes*. The auxin degron used for all characterization, unless otherwise mentioned, was a truncation of the degron from IAA17 from *Arabidopsis* that was characterized previously to have the fastest speed of degradation in the presence of AFB2 degradation machinery[16]. The other degrons used were the domain two regions from IAA14 and IAA15. The ADCTF is driven by a beta-estradiol inducible version of the GAL1 promoter integrated at the TRP locus in the genome in all strains[29]. The iaaH gene was amplified via PCR from the Ti plasmid of *Agrobacterium tumefaciens* and cloned via the Gateway™ method into a single-integrating HIS3 plasmid behind the strong TDH3 promoter. The integrating plasmid cassette was produced via digestion of the plasmid by PmeI and integrated into an auxin reporter strain via a standard lithium acetate transformation method[30].

Cytometry

All cytometry measurements were acquired with an Accuri C6 cytometer with attached CSampler apparatus using 488 nm and 640 nm excitation lasers and a 533 nm (FL-1: YFP/GFP) emission filter. Experiments involving time course data were taken during log phase via the following preparation: 16 hours of overnight growth in synthetic complete medium in a 30°C shaker incubator followed by 1:100 dilution into fresh, room-temperature medium. After 5 hours of growth under the same incubation conditions, 100 µL aliquots were read periodically until the completion of the experiment. For experiments involving steady state behavior, cultures were grown overnight, then diluted down in the morning 1:100 in fresh media and grown for 5 hours to log phase. They were then induced and allowed to grow for between five and twenty four hours depending on the experiment and then read on the cytometer. Cytometry data were analyzed using custom R scripts and the flowCore[31] package using the following steps: (1) gating for the yeast population, (2) gating for separate sending / receiving strains via the yellow (GFP) and red (mCherry) channels, and the generation of mean fluorescence values.

Supplementary Information

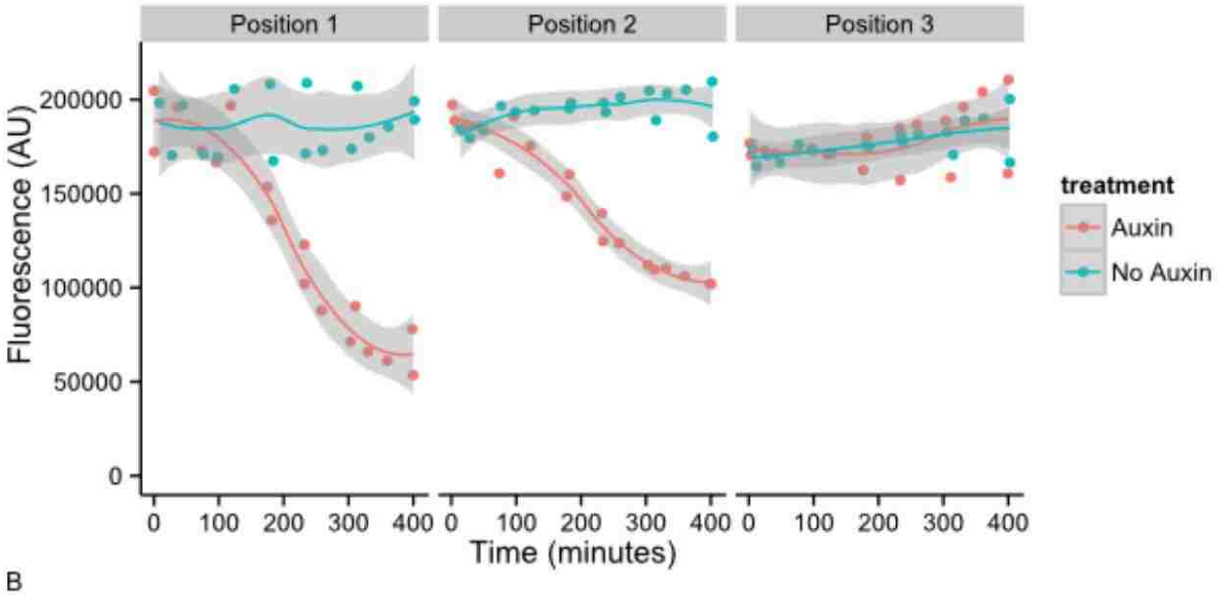
Supplementary Figure 1.



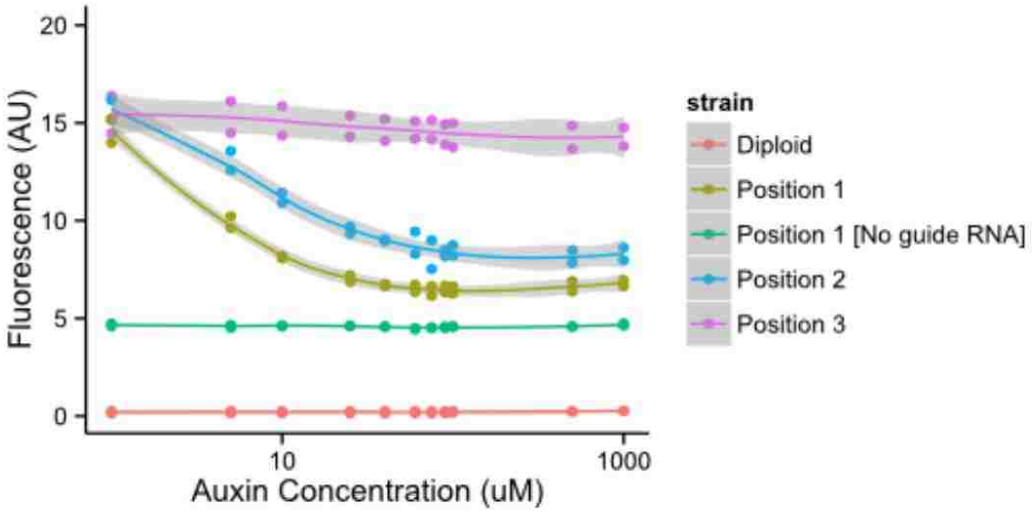
A)Raw data of a time course where receiver cells with ADC17 and AFB2 were treated with Auxin at time zero and then fluorescence was observed until the sample reached steady state, after which they were washed to remove auxin and recovery was observed over a similar period as induction. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.
B) Raw data of two time course replicates of receiver strains with an ADC TF with degron 17 and either AFB2 or TIR1-DM F boxes. Both these strains have a minimal pGAL promoter driving GFP production and a guide RNA that targets this promoter. Cultures treated with 30uM auxin show a comparable release of regulation to when pCYC1 was targeted with the ADC TF. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.

Supplementary Figure 2.

A



B

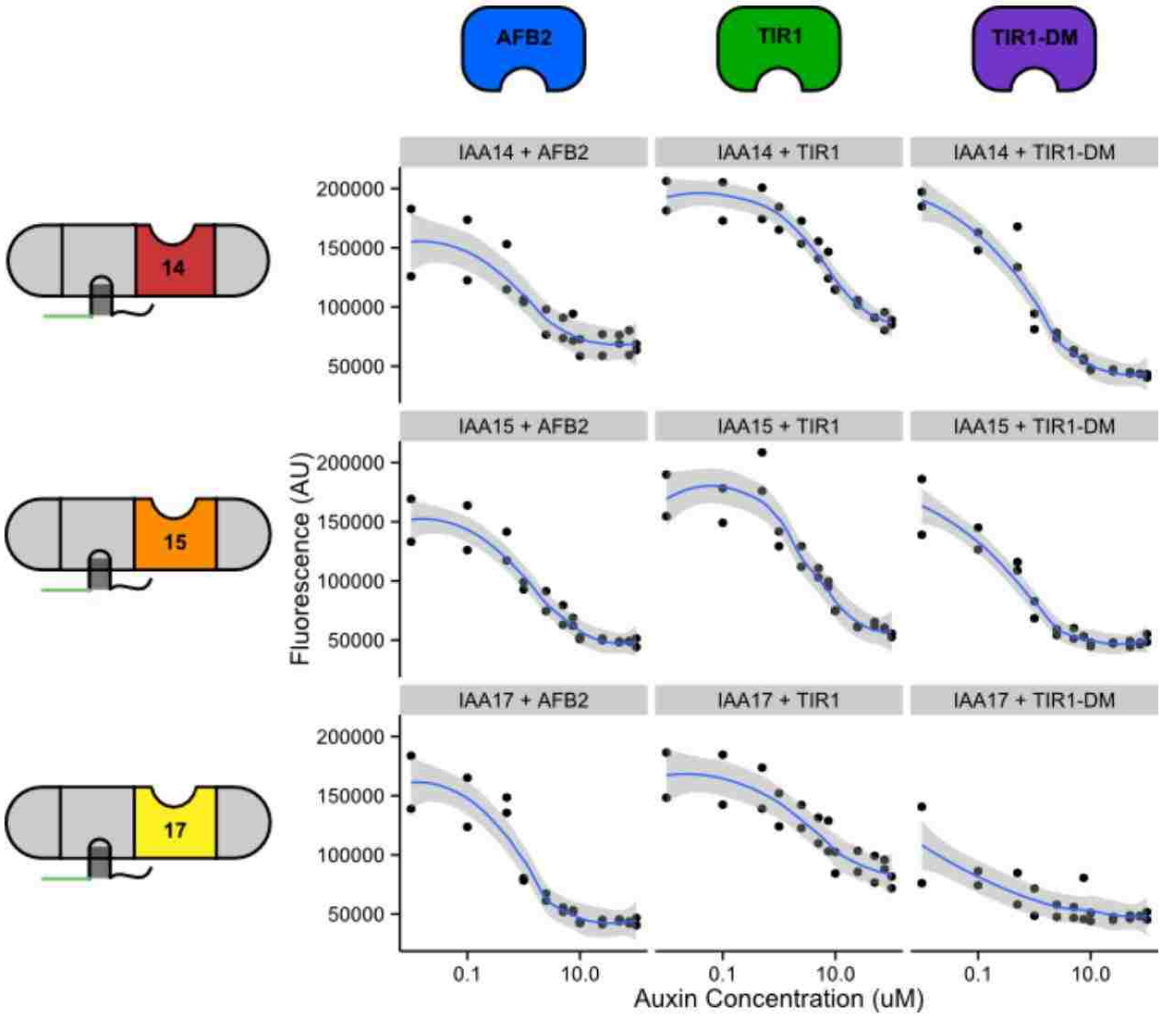


A) Raw data of two time course replicates of the three positional variants of the ADC TFs for cultures that were treated with 10uM auxin, as well as parallel untreated controls. The auxin treated cultures have a consistent drop in fluorescence, with position one having the largest drop at steady state. Position three has no noticeable drop at this auxin concentration. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.

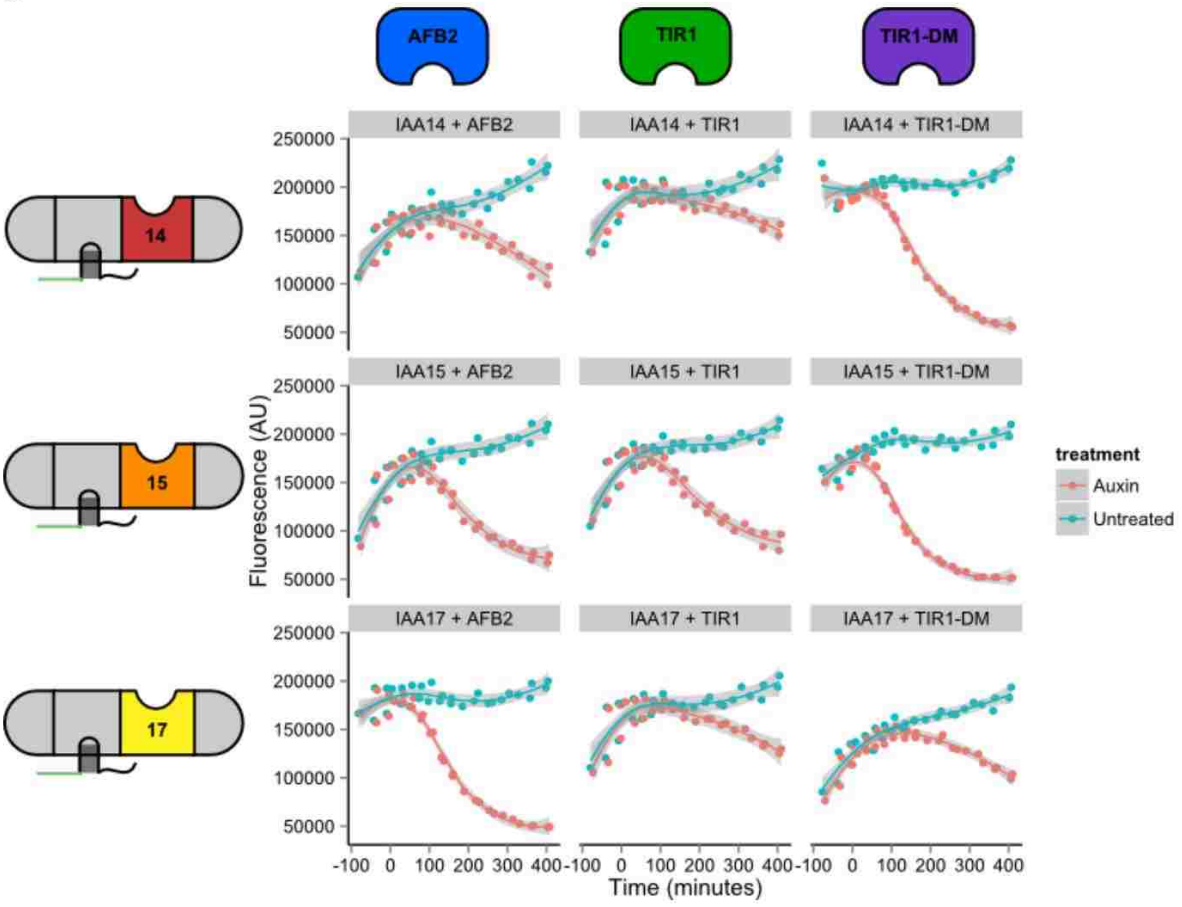
B) Raw data of two dose response replicates of the three positional variants of the ADC TFs five hours after induction with auxin. Position 1 has the highest sensitivity to auxin, and consequently saturates first, followed by position 2 and then position 3. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.

Supplementary Figure 3

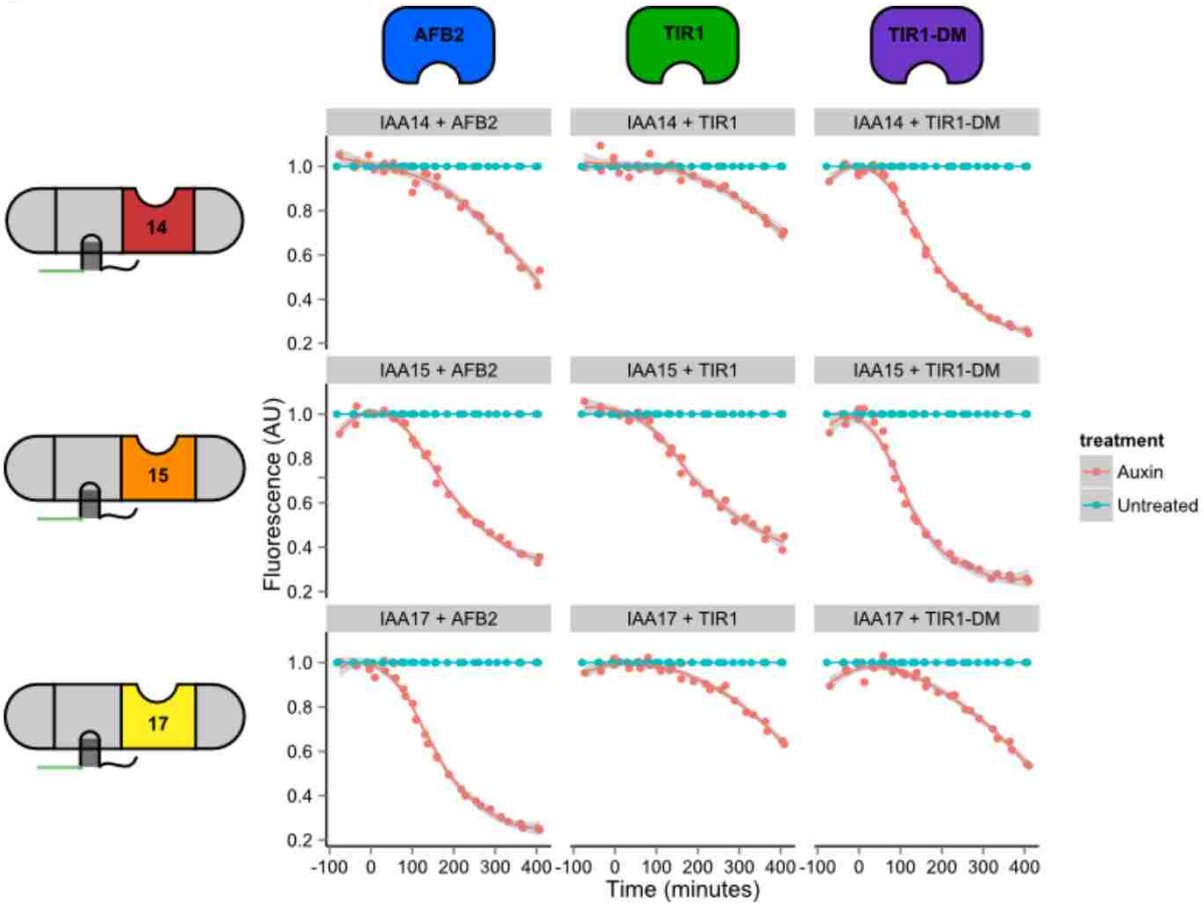
A



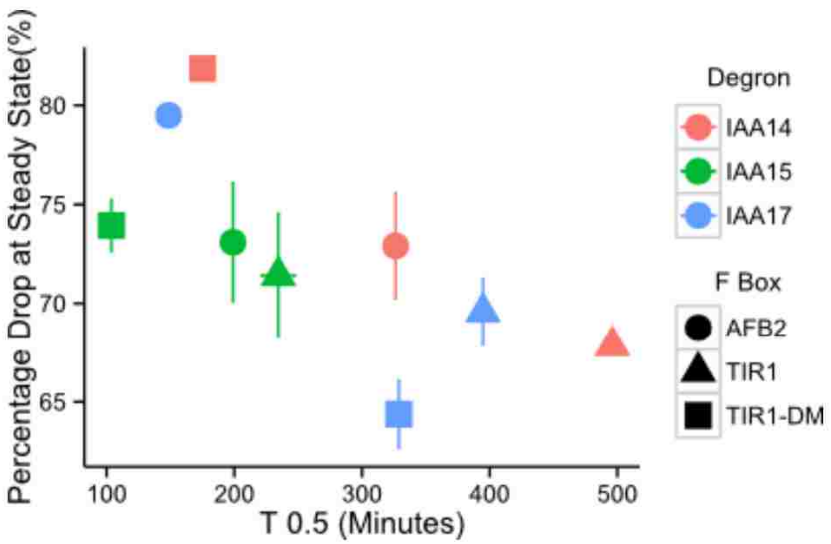
B



C



D

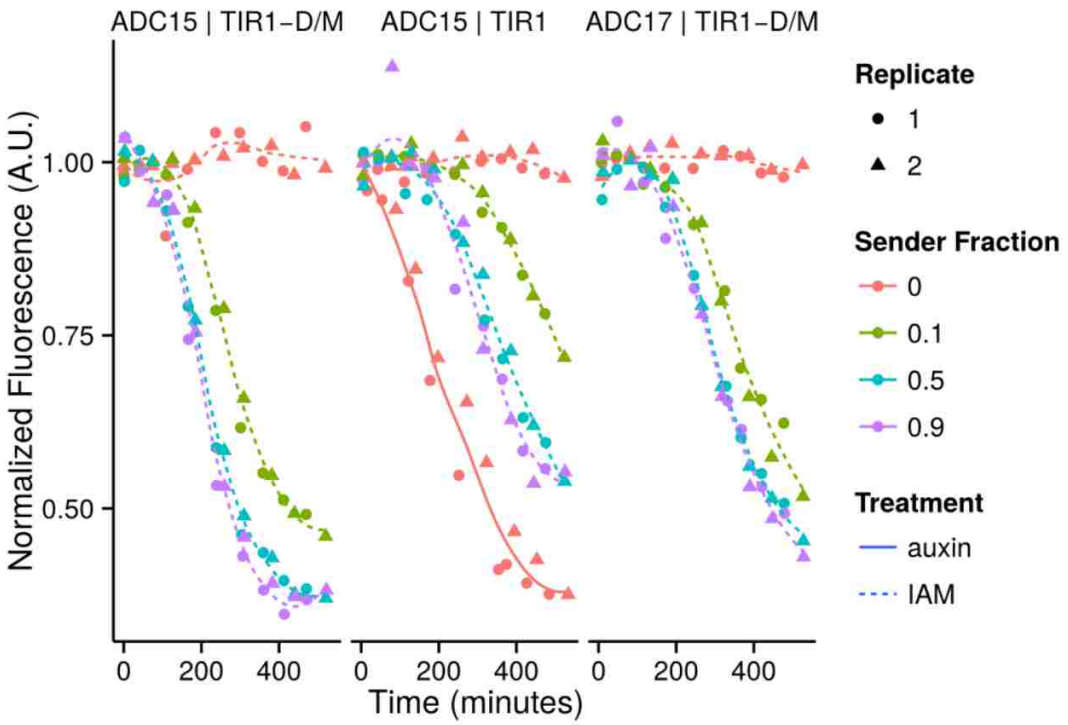


A) Raw data of two dose response replicates of all possible pairwise combination of different degrons on ADC TFs with the three F Boxes AFB2, TIR1 and TIR1-DM twenty four hours after induction with auxin. A range of auxin sensitivities are represented in the library, TIR1 being the most insensitive, and AFB2 and TIR1-DM being much more sensitive, depending on the degron being used. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.

B and C) Raw data and normalized data repectively of two time course replicates of all possible pairwise combination of different degrons on ADC TFs for cultures that were treated with 30uM auxin at time zero. A range of different degradation kinetics are observed with some reaching steady state within 400 minutes such as IAA15 + TIR1-DM, whereas other are still decreasing. The lines are LOESS fits to the data and the gray ribbon represents a 95% confidence interval of the fit.

D) The raw time course data from plot C is summarized here by plotting the T 0.5, time taken for the fluorescence to drop to fifty percent of its maximal drop, versus the percentage drop at steady state, measured twenty four hours after induction. While there appears to be a large range of degradation rates achievable by using different combinations of degrons and ADC-TFs, the steady state percentage changes are all approximately equal, with a variance of 20% between the highest and the lowest, but with most clustering at approximately 75%.

Supplementary Figure 4



Normalized time course data for two replicates of sender receiver coculture experiments corresponding to Figure 5 in the main text. Sender fractions are indicated by different line and point colors. Lines represent LOESS fits to the combined mean fluorescence data from both replicates. Solid lines represent auxin treatment while dashed represent IAM treatment.

Funding

This material is based upon work supported by NSF Grant No. 1411949 (An Auxin Toolbox for Synthetic Multicellular Systems), NSF Grant No. 1137266 (EFRI-MKS: Harnessing Intercellular Signaling to Engineer Pattern Formation), NSF Grant No. 1317653 (Molecular Programming Architectures, Abstractions, Algorithms, and Applications), The Paul Allen Family Foundation (Allen Distinguished Investigator Program).

References

(1) Wilcox, M., Mitchison, G. J., Smith, R. J., and Road, H. (1973) Pattern formation in the blue-green alga, Anabaena I. Basic mechanisms. *J. Cell Sci. 12*, 707–723.

(2) Medzhitov, R. (2007) Recognition of microorganisms and activation of the immune response. *Nature 449*, 819–826.

(3) Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., Hohmann, S., de Nadal, E., Posas, F., and Solé, R. (2011) Distributed biological computation with multicellular engineered networks. *Nature 469*, 207–11.

(4) Abelson, H., Weiss, R., Allen, D., Coore, D., Hanson, C., Homsy, G., Knight, T. F., Nagpal, R., Rauch, E., and Sussman, G. J. (2000) Amorphous computing. *Commun. ACM 43*, 74–82.

(5) Barcena Menendez, D., Senthivel, V. R., and Isalan, M. (2014) Sender-receiver systems and applying information theory for quantitative synthetic biology. *Curr. Opin. Biotechnol. 31C*, 101–107.

(6) Jang, S. S., Oishi, K. T., Egbert, R. G., and Klavins, E. (2012) Specification and simulation of synthetic multicelled behaviors. *ACS Synth. Biol. 1*, 365–374.

(7) You, L., Cox, R. S., Weiss, R., and Arnold, F. H. (2004) Programmed population control by cell-cell communication and regulated killing. *Nature 428*, 868–871.

(8) Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010) A synchronized quorum of genetic clocks. *Nature 463*, 326–30.

(9) Chen, M.-T., and Weiss, R. (2005) Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. *Nat. Biotechnol. 23*, 1551–5.

(10) Sohka, T., Heins, R. a, Phelan, R. M., Greisler, J. M., Townsend, C. a, and Ostermeier, M. (2009) An externally tunable bacterial band-pass filter. *Proc. Natl. Acad. Sci. U. S. A. 106*, 10135–10140.

(11) Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R. (2005) A synthetic multicellular system for programmed pattern formation. *Nature 434*, 1130–4.

(12) Shou, W., Ram, S., and Vilar, J. M. G. (2007) Synthetic cooperation in engineered yeast populations. *Proc. Natl. Acad. Sci. U. S. A. 104*, 1877–82.

(13) Youk, H., and Lim, W. a. (2014) Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science 343*, 1242782.

(14) Zhang, N.-N., Dudgeon, D. D., Paliwal, S., Levchenko, A., Grote, E., and Cunningham, K. W. (2006) Multiple signaling pathways regulate yeast cell death during the response to mating pheromones. *Mol. Biol. Cell 17*, 3409–3422.

(15) Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods 6*, 917–22.

(16) Havens, K. a, Guseman, J. M., Jang, S. S., Pierre-Jerome, E., Bolten, N., Klavins, E., and Nemhauser, J. L. (2012) A synthetic approach reveals extensive tunability of auxin signaling. *Plant Physiol. 160*, 135–42.

(17) Pierre-Jerome, E., Jang, S. S., Havens, K. a, Nemhauser, J. L., and Klavins, E. (2014) Recapitulation of the forward nuclear auxin response pathway in yeast. *Proc. Natl. Acad. Sci. U. S. A. 111*, 9407–12.

(18) Gray, W. M., Kepinski, S., Rouse, D., Leyser, O., and Estelle, M. (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature 414*, 271–276.

(19) Villalobos, C., a, L. I., Lee, S. C., De Oliveira, C., Ivetac, A., Brandt, W., Armitage, L., Sheard, L. B., Tan, X., Parry, G., Mao, H., Zheng, N., Napier, R. M., Kepinski, S., and Estelle, M. (2012) A combinatorial TIR1/AFB–Aux/IAA co-receptor system for differential sensing of auxin *8*.

(20) Farzadfard, F., Perli, S., and Lu, T. (2013) Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol. 2*, 604–13.

(21) Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell 152*, 1173–1183.

(22) Khalil, A. S., Lu, T. K., Bashor, C. J., Ramirez, C. L., Pyenson, N. C., Joung, J. K., and Collins, J. J. (2012) A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions. *Cell 150*, 647–658.

(23) Kiani, S., Beal, J., Ebrahimkhani, M. R., Huh, J., Hall, R. N., Xie, Z., Li, Y., and Weiss, R. (2014) CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods 11*, 723–6.

(24) Nielsen, A. A. K., and Voigt, C. A. (2014) Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol. 10*, 1–12.

(25) Yu, H., Moss, B. L., Jang, S. S., Prigge, M., Klavins, E., Nemhauser, J. L., and Estelle, M. (2013) Mutations in the TIR1 auxin receptor that increase affinity for auxin/indole-3-acetic acid proteins result in auxin hypersensitivity. *Plant Physiol. 162*, 295–303.

(26) Zhao, Y. (2010) Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol. 61*, 49–64.

(27) Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, M. L., and Bellini, C. (2011) Agrobacterium tumefaciens: From crown gall tumors to genetic transformation. *Physiol. Mol. Plant Pathol. 76*, 76–81.

(28) Turner, J. G., Ellis, C., and Devoto, A. (2002) The jasmonate signal pathway. *Plant Cell 14 Suppl*, S153–S164.

(29) McIsaac, R. S., Silverman, S. J., McClean, M. N., Gibney, P. a, Macinskas, J., Hickman, M. J., Petti, A. a, and Botstein, D. (2011) Fast-acting and nearly gratuitous induction of gene expression and protein depletion in Saccharomyces cerevisiae. *Mol. Biol. Cell 22*, 4447–59.

(30) Gietz, R. D., and Woods, R. A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol. 350*, 87–96.

(31) Hahne, F., LeMeur, N., Brinkman, R. R., Ellis, B., Haaland, P., Sarkar, D., Spidlen, J., Strain, E., and Gentleman, R. (2009) flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics 10*, 106.

Chapter 4 - Characterizing the flow of information in a cell: "Using an unbiased deep learning approach to determine ARF-promoter specificity rules"

Abstract

In plants the growth hormone auxin plays a central role in the spatiotemporal coordination of developmental events. Plant cells interpret and respond to auxin signals via a family of transcription factors called ARFs and their transcriptional-corepressors the AUX/IAAs. One of the key challenges to understanding how auxin regulates development is understanding which ARFs are regulating the various genes associated with development. The promoter specificity rules for the ARF family have not been fully described making this a challenging problem. In this work we used a yeast-based competitive growth assay to characterize the ARF activation of a huge range of synthetic plant promoters to elucidate how different promoter architectures and sequences might affect the extent of activation by ARF5 from *Arabidopsis thaliana*. We also tried an alternative approach where we used deep learning algorithms to learn specificity rules for ARF-promoter interactions using DAP-seq data from maize and build predictive models. We demonstrate how these models can be used to predict the effects of SNPs in promoters on gene expression and eventually on the phenotypes that those changes drive.

Introduction

Development in plants is a complex process that consists of constant organogenesis of leaves, fruits, flowers, and lateral roots. These processes are largely coordinated by the hormone auxin[1,2]. In plant cells, auxin signals regulate gene expression via a family of transcription factors, the auxin response factors (ARFs), and their transcriptional co-repressors, the AUX/IAAs. The ARF transcription factors bind to the promoters of certain genes, resulting in either activation or repression of that gene depending on whether it is an activator or repressor ARF. The AUX/IAAs interact with these ARF families and, in the case of activator ARFs, block their regulatory activity in an auxin dependent manner[3,4]. The changes in the transcriptional regulation in response to auxin is thought to be different in cells from different tissues in part due to a different complement of ARFs and AUX/IAAs expressed in these cells[5,6]. Recent work suggests that the AUX/IAAs, being the auxin responsive component of this system, are largely responsible for the differential dynamic response to auxin[7], while the ARFs are thought to control which genes are functionally regulated by auxin, as they directly interact with the DNA[8].

Largely through the study of knockout phenotypes, it has been established that certain ARFs play critical roles in particular developmental and tropic processes, such as the role of ARF2 in leaf senescence[9] or ARF5 in the development of the embryonic root[10]. However, there is no clear consensus on how ARFs can specifically regulate only certain genes. A canonical ARF binding cis-element (AuxRE) was identified long ago[11], however, this is not sufficient to explain the specificity that is observed in plants. Recently the structure of the ARF proteins has been elucidated further and they have been shown to form homo or hetero-dimers[12]. This has led to the hypothesis that these dimers may allow greater degeneracy in one the ARF binding sites which might explain why nearby pairs of AuxREs rarely occur in plant promoters. Additionally, it has been put forward that the length and sequence of the spacers in between the two ARF binding sites might contribute to ARF promoter specificity[13]. While these hypotheses do have a great deal of biochemical and structural evidence to back them up, there is still no set of validated rules that can predict ARF-promoter specificity.

Here we systematically examine the activity of ARFs on a variety of promoter sequences and architectures using a yeast synthetic activation assay. This synthetic activation assay has several

advantages—in yeast we can generate user-defined circuits to look at specific ARF-promoter interactions that are difficult to observe *in planta* due to the co-expression of ARFs in many tissue types. This approach also allows us to test activation on synthetic, standardized promoter variants, as opposed to *in planta* studies that examine only the native promoters in the genome. This approach is scalable, and can be applied to large promoter libraries to test activation on thousands of promoter variants in parallel, allowing us to characterize a much larger and more diverse sequence space than is possible from genomic sequences.

We chose two activator ARFs, ARF5 and ARF19, to examine in depth as they have opposite developmental outcomes, the growth of the primary versus lateral roots. We tested ARF activity on different promoter sequences regulating a fluorescent reporter in yeast by quantifying florescence via flow cytometry. We also used a selective growth assay to further explore the sequence space of ARF-activated promoters. In this study we establish the promoter preferences of ARF5 and ARF19, determine the structure of a minimal ARF-responsive promoter, and find novel ARF-activated promoters.

We also tried a different approach to elucidate the ARF-promoter interaction rules wherein used DAP-seq data collected using *Zea maize* genomic DNA to train convolutional neural net based machine learning models. These have been shown to be able to capture sequence enrichment as well as positional information of transcription factor binding motifs. We implemented these using the keras package in python with a Theano backend[15]. Once trained we explored how similar the binding rules for different maize ARFs by testing if a model trained on data from one ARF could predict binding by a different ARF. We also demonstrated the predictive power of these models by showing that we could predict the changes in ARF binding due to SNPs in promoters from different landraces. We also showed how these predictions could be extended to predictions about phenotypic effects such as herbivore resistance. The elucidation of these ARF-promoter specificity rules would bring us closer to being able to tease apart the complex regulatory mechanisms that control plant development and stress response. Additionally, this work provides an example of how machine learning and large synthetic data sets can be used to learn about cis regulatory motifs and promoter architectures, which has been a problem of general interest across eukaryotic biology.

Results

Testing ARF activity on a randomized library of promoter variants

Previous work from our group and others suggested that promoter context may be a key contributor to whether an ARF can activate on any given AuxRE. To further explore the design space of promoters upon which ARFs could activate transcription, we used a selective growth assay to characterize ARF activation on a large, unbiased set of promoter variants. This experiment let us explore ARF activity on a far more diverse range of promoters than we could with rationally designed variants, and a much larger set of promoters than has ever been assayed previously.

We decided to test ARF5 activation on our unbiased promoter library. From our previous experiments, ARF19 was shown to be a promiscuous activator of transcription, whose activity on various promoters may not be shared with other ARFs with more stringent promoter specificity (as evidenced by its exclusive activation on the single AuxRE TGTCGG). ARF5, on the other hand, seemed to be more sensitive to promoter architecture, in terms of both AuxRE copy number and sequence. We designed a library of promoter variants by inserting 26 semi-degenerate nucleotides into the A1 site of the mutated

IAA19 promoter. For 80% of this library, the first 22 nucleotides were degenerate and the last four nucleotides were the minimal AuxRE GACA. In 10% of the library the first 20 nucleotides were degenerate and the last six nucleotides were the minimal AuxRE TGTCNN. In 10% of the library all 26 nucleotides were degenerate. This ratio was chosen to make our library maximally informative as we assumed a single AuxRE is required for ARF activity. We cloned this promoter library into a 2-micron plasmid upstream of the HIS3 gene, which is essential for histidine biosynthesis in yeast. We estimated library size by colony count, and chose a purified library with approximately 50,000 unique promoter variants. After transformation of the library into yeast constitutively expressing ARF5, we performed a selective growth assay in media lacking histidine, to determine on which promoter variants ARF5 activated transcription, allowing for histidine production. We used next generation sequencing to sequence our pre- and post-selection libraries and compared read counts for individual promoter variants. Before doing the selective growth assay we verified that our previously tested rationally designed promoter variants showed a difference in growth in ARF5-expressing yeast when placed upstream of the HIS3 gene. We also tested a wide range of 3-AT concentrations on these rationally designed variants to optimize our selection for the greatest range of differential growth, but found that no 3-AT addition actually led to an optimal range of growth rates in the different controls.

ARF5 activates on a huge range of promoter variants.

After analysis, our sequencing run resulted in approximately 10,000 promoter variants that were present in both our pre- and post-selection libraries at read counts of five or above. To calculate enrichment of each of these promoter variant, we calculated the log of the read count in the post-selection library divided by the read count in the pre-selection library. To our surprise, we found that enrichment was not correlated with the number of AuxREs within a promoter. We also did not find that enrichment was correlated with AuxRE orientation for those promoter variants that had two AuxREs. Certain combinations of orientation and spacer length did seem to affect enrichment, but overall we did not see the same clear trends of ARF5-mediated activation that we had seen in the rationally designed promoter variants. Likely activation by native yeast transcription factor also contribute to the enrichment of these promoter variants. This makes it difficult to group promoters with similar elements (AuxRE number, orientation, spacer length) as the sequence context of each promoter is so divergent, and different sequences may preferentially be activated by different native yeast factors.

To confirm that the selective growth assay was identifying ARF5-activated promoters, we selected a subset of 46 promoter variants from the assay to test using our fluorescence assay for ARF5-mediated activation. We cloned the 28 most enriched promoters from the screen, as well as the two most enriched promoters with three AuxREs and the two most enriched promoters with two AuxREs in all four orientations (only two of these were in the 28 most enriched promoters) and 10 highly unenriched promoters from the screen into our integrating yeast plasmids regulating a fluorescent reporter. We tested fluorescence in strains expressing and not expressing ARF5, to quantify to what extent activation on each promoter was mediated by ARF5 and to what extent it was mediated by endogenous yeast transcriptional machinery. We found that the selective growth assay was indeed an excellent readout of transcription from a promoter variant in yeast, as the ten highly unenriched promoters did not show fluorescence in either the non-ARF5 or ARF5 expressing strains. Of the highly enriched promoters, varying extents of activation by both the non-ARF5 and the ARF5 expressing strain were seen. Activity in either strain was not correlated with enrichment, suggesting that at least at the upper end of enrichment, scores may give a qualitative but not a quantitative measure of activation. Some promoters'

activation seemed solely mediated by yeast transcriptional machinery, as there was no difference in fluorescence in the ARF5 and non-ARF5 strains, but some showed a substantial increase in activation when ARF5 was present. We quantified ARF5-specific activation by graphing activation in the non-ARF5 strain versus activation in the ARF5 strain and selected strains with high ARF5-specific activation for further study.

To our surprise many of our strains that showed high ARF5-specific activation had a single AuxRE. This result supports our result that ARF19 can activate on a single AuxRE. ARF5 activates on multiple single AuxRE sequences in the context of the 26-base pair degenerate region that it does not activate when the same single AuxRE sequence is placed into the A1 site, suggesting that context is key for ARF activity on single AuxREs. One strain that showed ARF5-specific activation has no recognizable AuxREs. We explored this promoter sequence further in depth and found that the ARF5 activation is derived from nine base pairs of the promoter, TAACCCGGA. Our selective screen identified novel ARF5-activated promoters that would not have been predicted by previous research on ARF-mediated transcription. The role of single-AuxRE and no-AuxRE promoters on auxin responsive transcription in plants has yet to be explored but these results suggest that activation on these promoter variants by ARFs is possible.

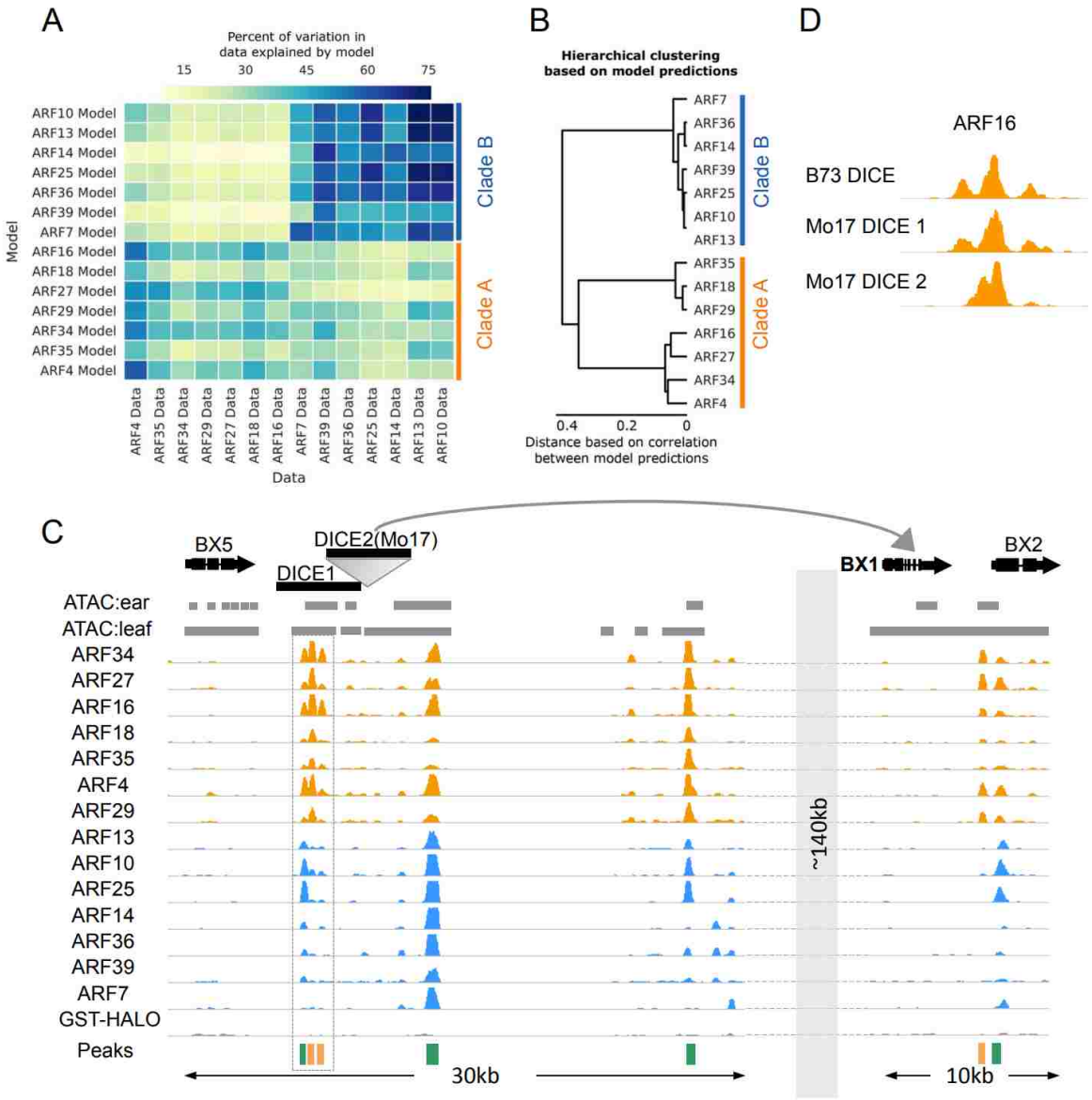
Machine learning predicts Maize ARF binding in an herbivore resistance QTL

To understand whether there were predictable features that discriminated Maize ARF binding patterns from random sequences, we used a supervised machine learning approach previously developed for examining transcription factor DNA binding. We used a subset of each ARF dataset as a training set to develop a binding model and then asked how good each model was at predicting binding strength of test sets from the other ARF datasets. ARF models were able to clearly distinguish between activator and repressor ARFs (Figure A). For the repressor ARFs, robust models that explained up to 75% of the variation seen in other repressor ARFs were generated and hierarchical clustering based on model predictions recapitulated phylogenetic patterns (Figure B). Activator ARF models were also able to predict binding events among fellow activator ARFs with up to ~50% accuracy, but were less predictive than repressor models for sub-clade specificity (Figure B).

We next used these predictive models to assess ARF binding potential at a genomic region associated with an herbivore resistance QTL in the maize inbred line Mo17, that in B73 revealed several strong ARF peaks (Figure S9A)(Betsiashvili et al., 2015). This region (264kb) contains a cluster of at least eight genes belonging to the benzoxazinoid (Bx) biosynthesis pathway, which is responsible for generating the defense compound DIMBOA. DIMBOA confers resistance to both aphids and European corn borer, two highly destructive insects that affect plant fitness and grain yields (Frey et al., 1997). ARF peaks were located in putative regulatory regions of several of the genes in this cluster, implicating ARFs in the control of this important defense compound. Several of the strongest peaks in this cluster were located about 500bp-4kb downstream of the Bx5 gene, in a ~4kb region called DICE (DIstal Cis-Element) which was previously found to influence the expression of the *BX1* gene located about 140kb downstream (Figure C)(Zheng et al., 2015). BX1 is the signature enzyme of the DIMBOA pathway and represents an important branching point from the primary metabolism of free indole and tryptophan to the secondary metabolism of defense compounds (Frey et al., 1997). The maize inbred line Mo17 is known to be one of the highest producers of DIMBOA and consequently displays one of the highest levels of resistance to insect herbivores (Betsiashvili et al., 2015), while B73 is much more susceptible.

The difference in *BX1* expression between these two inbred lines was determined to be caused by the duplicated DICE element in Mo17 (Zheng et al., 2015).

To test whether ARF binding could influence the activity of the Mo17 DICE enhancer, we used the machine learning algorithm to predict the effect of sequence differences between the DICE elements on ARF binding potential. Our ARF binding models suggested that most activator ARFs were likely to bind with a similar affinity to both the highly conserved first copy of DICE as well as the Mo17-specific second DICE element (Figure D,E). This suggests that additional activator ARF binding sites in Mo17 could be a causative feature that leads to higher *BX1* expression in Mo17. To verify the binding of ARFs experimentally we tested Mo17 in DAP-seq experiments with several ARFs and found that they did match the predictions.



ARF display distinct cis-regulatory signatures that are predictive. A. Heatmap showing percent of variation in ARF binding data that can be explained by a machine learning model trained on individual ARF datasets. **B.** Hierarchical clustering of ARFs based on model predictions. **C.** Genome browser view of ARF binding events in the 30kb region surrounding the DICE element (left) and the BX1 genic region located 140kb downstream that is controlled by DICE (right). **D**. Binding score predictions for individual ARFs in the DICE elements **E.** Binding potential of individual ARFs to the B73 DICE, the Mo17 DICE1 and Mo17 DICE2 elements measured as the area under the curve (AUC) of binding scores within the ~4kb DICE elements. Values were normalized by length to account for the slightly different sizes of sequences between the different elements.

Conclusion

The work summarized here demonstrates strategies we could use to elucidate the role of promoter sequence and architecture on function. These insights could then be used to create a set of promoter design rules for the rational creation of synthetic promoters that achieve a target expression profile. The benefit of using a heterologous platform like yeast to prototype rationally designed synthetic promoters is that it allows both parallel characterization of millions of promoters and the capacity to isolate a single signaling pathway without worrying about the native crosstalk from other pathways. It also gives us the capacity to ask questions about activation of a promoter by a transcription factor rather than just binding. The *in vitro* approach, DAP-seq, used later, does have the downside that it characterizes binding rather than activation, however, it does capture the role of genomic sequence variation. Machine learning models trained on this data have been demonstrated to have a predictive capacity, which opens the door to using them for future forward engineering or diagnosis of the impact of SNPs in plant germplasm. While it is harder to glean mechanistic insights from these black box models, their predictive capacity have great promise for future attempts at engineering synthetic promoters.

References

1. Teale, William D., Ivan A. Paponov, and Klaus Palme. "Auxin in action: signalling, transport and the control of plant growth and development." *Nature Reviews Molecular Cell Biology* 7.11 (2006): 847-859.
2. Lavenus, Julien, et al. "Lateral root development in Arabidopsis: fifty shades of auxin." *Trends in Plant Science* 18.8 (2013): 450-458.
3. Guilfoyle, Tom J., and Gretchen Hagen. "Auxin response factors." *Current opinion in plant biology* 10.5 (2007): 453-460.
4. Overvoorde, Paul, Hidehiro Fukaki, and Tom Beeckman. "Auxin control of root development." *Cold Spring Harbor perspectives in biology* 2.6 (2010): a001537.
5. Brady, Siobhan M., et al. "A high-resolution root spatiotemporal map reveals dominant expression patterns." *Science* 318.5851 (2007): 801-806.
6. Hardtke, Christian S., and Thomas Berleth. "The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development." *The EMBO journal* 17.5 (1998): 1405-1411.

7. Havens, Kyle A., et al. "A synthetic approach reveals extensive tunability of auxin signaling." *Plant physiology* 160.1 (2012): 135-142.

8. Boer, D. Roeland, et al. "Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors." *Cell* 156.3 (2014): 577-589.

9. Feng, Guanping, et al. "AINTEGUMENTA negatively regulates age-dependent leaf senescence downstream of AUXIN RESPONSE FACTOR 2 in Arabidopsis thaliana." *Plant Biotechnology* 0 (2016).

10. Crawford, Brian CW, et al. "Genetic control of distal stem cell fate within root and embryonic meristems." *Science* 347.6222 (2015): 655-659.

11. Ulmasov, Tim, et al. "Composite structure of auxin response elements." *The Plant Cell* 7.10 (1995): 1611-1623.

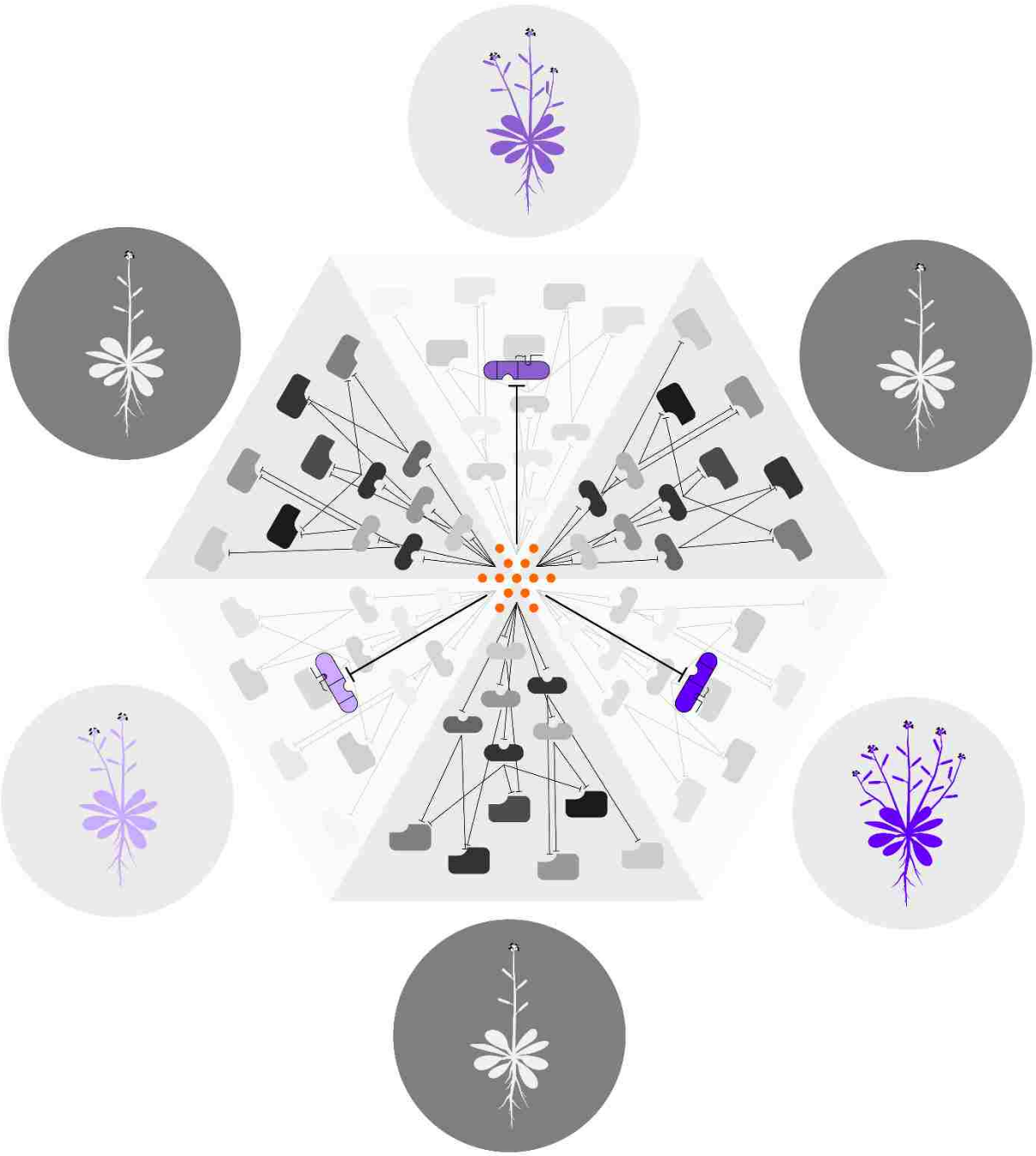
12. Guilfoyle, Tom J. "The PB1 domain in auxin response factor and Aux/IAA proteins: a versatile protein interaction module in the auxin response." *The Plant Cell* 27.1 (2015): 33-43.

13. Smit, Margot E., and Dolf Weijers. "The role of auxin signaling in early embryo pattern formation." *Current opinion in plant biology* 28 (2015): 99-105.

14. Pierre-Jerome, Edith, et al. "Recapitulation of the forward nuclear auxin response pathway in yeast." *Proceedings of the National Academy of Sciences* 111.26 (2014): 9407-9412.

15. Quang, Daniel, and Xiaohui Xie. "DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences." *Nucleic acids research* (2016): gkw226.

Chapter 5 - Characterizing and re-engineering the flow of information between tissues: "Synthetic hormone-responsive transcription factors can monitor and re-program plant development"



Abstract

Developmental programs sculpt plant morphology to meet environmental challenges, and these same programs have been manipulated to increase agricultural productivity[1,2]. Phytohormones coordinate these programs, creating chemical circuitry[3] that has been represented in mathematical models[4,5]; however, model-guided engineering of plant morphology has been limited by a lack of tools[6,7]. Here, we show-case a novel set of synthetic and modular hormone activated Cas9-based repressors (HACRs) in *Arabidopsis thaliana* that respond to three phytohormones: auxin, gibberellins and jasmonates. We demonstrate that HACRs are sensitive to both exogenous hormone treatments and local differences in endogenous hormone levels associated with development. We further show that this capability could be leveraged to reprogram development in an agriculturally relevant manner by changing how the hormonal circuitry regulates genes. By deploying a HACR to re-parameterize the auxin induced expression of the auxin transporter PIN-FORMED1 (PIN1), we decreased shoot branching and phyllotactic noise, as predicted by existing models[4,5].

Introduction

The body plans of plants are inherently plastic, making them amenable to optimization for a wide range of natural or artificial environments. Extrinsic and intrinsic cues are integrated by developmental programs to maximize the fitness of wild plants[3]. Domestication of crops frequently relies on altering such programs to create more productive morphologies for agriculture, such as the dramatic reduction in bushiness of maize[1] or the dwarfing of cereals that drove the green revolution[2].

Developmental programs are coordinated in large part by a set of hormones[3]. Accumulation of a given hormone by *de novo* synthesis or transport influences the expression or activity of developmental master controller genes, analogous to wires in a circuit. Auxin, perhaps the best-studied hormone, controls many developmental programs that drive agriculturally relevant traits[8]. Many mathematical models connecting auxin signaling and transport at the molecular level to specific developmental phenotypes at the whole plant level have been developed[4,5,9]. These models highlight the importance of subtle parameters, like the strength of specific feedback loops in hormone signaling networks, in determining plant morphology.

While the ability of hormones to trigger and tune developmental programs makes altering hormonal signaling an attractive target for re-engineering the plant form, there are significant hurdles to overcome in such approaches. Native hormone signaling pathways are comprised of co-expressed and redundant components, embedded in highly reticulate cross-regulatory relationships with other signaling pathways, and have several layers of feedback[8]. For example, the auxin signaling pathway is comprised of three families of proteins, ARFs, AUX/IAAs, and TIR1/AFBs, all of which have multiple members with redundant regulatory roles and are cross regulated by a plethora of other signals[10,11].

Thus, there is a need for tools that can predictably alter how a specific hormone regulates a gene of interest to facilitate re-wiring plant development[12]. To date, such efforts have been largely limited to reducing or increasing expression of components of the native hormone signaling machinery[7], an approach ill-suited for tuning the strength of connections within a network and easily confounded by redundancy and buffering within a network. In trying to circumvent redundancy, researchers are often forced to construct high order mutants of the multiple genes underlying the function of a single network

hub. This approach reduces the precision of experimental or engineering interventions, as these genes are frequently only partially redundant with one another, and, thus this approach introduces more off-target effects. Chimeric promoters with altered hormonal regulation of a gene of interest have been used with some success[13,14]. However, the paucity of detailed mechanistic maps connecting promoter architecture and chromatin state, and the high heterogeneity in these factors between genes, means that promoter design remains a bespoke approach with an associated high design and development cost for each network of interest. Additionally, these methods often require adding an extra copy of the gene of interest in a novel chromatin context, making it difficult to make definitive mechanistic conclusions. These challenges have made it difficult to study the significance of hormone regulation on specific genes, particularly in regard to the impact of transcriptional feedback loops on differentiation and morphogenesis. For all of these reasons, the potential predictive power of mathematical models has not been fully leveraged in the engineering of morphologies of agronomic interest. To facilitate more sophisticated interventions in plant developmental programs, we designed a set of synthetic and modular hormone-activated Cas9-based repressors (HACRs, pronounced 'hackers').

Results and Discussion

We previously validated the design of similar synthetic auxin-sensitive transcription factors in *Saccharomyces cerevisiae*[15]. Guided by this work, we fused the deactivated Cas9 (dCas9) protein from *Streptococcus pyogenes*[16] to a highly sensitive auxin-induced degron[17] and the first 300 amino acids of the TOPLESS repressor (TPL)[18] (Figure 1A). The dCas9 associates with a guide RNA (gRNA) that targets the HACR to a promoter with sequence complementarity where it can repress transcription. Upon auxin accumulation, the degron sequence targets the HACR for ubiquitination and subsequent proteasomal degradation. Thus, in parallel to the natural auxin response, auxin triggers relief of repression on HACR target genes. Transgenic plants were generated with HACRs and a gRNA targeting a constitutively expressed Venus-Luciferase reporter, and, as expected, auxin treatment increased overall fluorescence (Figure 1 B,C). A time-course using luciferase to quantify de-repression of the reporter supported these results with a significant spike in reporter signal ($p < 0.001$, $n = 10$) peaking approximately 80 minutes post auxin exposure (Figure 1 D,E). A HACR with a stabilized degron[17] showed significantly lower reporter signal upon auxin treatment ($p = 0.01$, $n=10$) (Figure 1F).

The modular nature of HACRs should allow substitution of the degron with any sequence that has a specific degradation cue. We tested this hypothesis by building HACR variants with degrons sensitive to two other plant hormones: jasmonates (JAs)[19] and gibberellins (GAs)[20]. Treatment of transgenic plants with exogenous hormones matched to the expressed variants significantly increased reporter signal as compared to control treatments (Figure 1 H, I, J, Figure 1-figure supplement 1).

To rewire the connections between the hormone circuitry and developmental master controllers, HACRs must be able to respond to local differences in endogenous hormone levels. To visualize subtle differences in HACR sensitivity at the cellular level, we built a ratiometric auxin HACR by combining our previous design with a second reporter (tdTomato) driven by the same UBQ1 promoter driving the Venus reporter, with the only difference being that its gRNA target site was mutated (Figure 2A). An estimation of relative auxin levels was then calculated by normalizing the Venus reporter signal in each cell to that of the tdTomato signal in the same cell, minimizing any effect of differential expression of the UBQ1 promoter in different cell types. Using these lines, we visualized tissues at different developmental stages where auxin distributions had been previously described using auxin reporters

like DII-VENUS or R2D2[21]. Auxin accumulation assayed by the HACR largely matched previous reports, such as the reverse fountain pattern of reporter signal in the root tip[22] (Figure 2B) and higher signal in the vasculature as compared to the epidermis of the elongation zone[22] (Figure 2C). We also observed high reporter signal in emerging lateral root primordia consistent with the auxin accumulation that triggers this developmental event[23] (Figure 2D,E).

To further explore the capacity of HACRs to respond to differences in endogenous hormone levels, we visualized the activity of auxin, GA and JA HACRs targeting a Venus reporter. Auxin accumulates in the apical domain of the early embryo and eventually resolves in later stages to the tips of the developing cotyledons, vasculature, and future root apical meristem[21]– the same patterns that were observed in plants expressing an auxin HACR (Figure 2F-J). In plants expressing a GA HACR, we observed a strong reporter signal in the early endosperm, consistent with the expression of GA biosynthesis enzymes[24] (Figure 2K-M, Figure 2-figure supplement 1). There are few reports of developmental regulation of JA distribution; however, we did detect accumulation of reporter signal in the developing ovule of plants expressing a JA HACR (Figure 2-figure supplement 1). Specifically, reporter signal appeared to be localized to the inner- and outermost layers of the integuments that surround the developing seed. We also observed that the JA HACR reporter was strongly induced in leaves subjected to mechanical damage (Figure 2N-Q), a condition known to induce high levels of JA[19].

Beyond their application as sensors of endogenous hormone distributions, HACRs should also be capable of reprogramming how such signals are translated into plant morphology. To test this, we turned to shoot architecture, an agronomically important trait with a well-established connection to auxin. Fewer side-branches allow for higher density planting[2] and more regular arrangement of lateral organs (phyllotaxy) facilitates efficient mechanized harvest[25]. The molecular mechanisms that control branching and phyllotaxy are well studied and have been mathematically modeled[4,5]. These models predict that a key parameter controlling both these processes is the strength with which auxin promotes its own polar transport[26], which we will refer to as feedback strength. One molecular mechanism that contributes to this feedback is the auxin-induced increase in expression of the auxin transporter PIN-FORMED1 (PIN1)[27]. Thus far, it has been impossible to tune the strength of auxin-mediated transcriptional feedback on PIN1, and thus impossible to fully test its role in regulating shoot architecture or its potential for engineering this trait.

To test whether we could rationally alter shoot architecture by changing feedback strength, we generated transgenic plants with a HACR targeting PIN1 (Figure 3A), as well as a model that produced a qualitative hypothesis of the impact of this intervention (Supplementary note 1). Our model predicts that this perturbation will decrease the activation of expression of PIN1 by auxin and dampen the dose response relationship between auxin and PIN1 expression (Figure 3-figure supplement 1 B,C). Quantitative PCR results on transgenic plants support these predictions, as the modest but significant reduction in PIN1 expression observed in plants expressing a PIN1 gRNA can be erased with exogenous auxin treatment (Figure 3-figure supplement 1 D). Our model and these results highlight the substantial difference between regulation by a hormone-responsive transcription factor and a static repressor. Static repressors would consistently suppress target gene expression at all hormone levels. In contrast, HACRs dampen both the dynamic and steady state dose response relationship between hormone concentration and gene expression akin to modulating the gain in a circuit (Figure 3-figure supplement 1 B,C).

In relation to shoot architecture models, the effect of an auxin-regulated HACR targeting PIN1 should be a reduction in feedback strength. In Prusinkiewicz et al.[5], auxin-regulated feedback is modeled as a post-translational mechanism dependent on the flux of auxin through the cell membrane. The magnitude of this flux is proportional to the recruitment of PIN1 to the membrane. According to their simulations, feedback strength is directly proportional to the number of branches the plant will develop. This effect is hypothesized to result from the reduced ability of lateral buds to establish auxin efflux into the main stem, an essential step in bud outgrowth (Figure 3D). While the transcriptional mode of feedback we are altering with our HACR is not directly encoded in the Prusinkiewicz et al. model, we hypothesized that decreasing transcriptional feedback strength would have qualitatively similar results to decreasing post-translational feedback strength. Thus, we expected a decrease in the number of branches in lines where auxin HACRs were targeted to PIN1. This is exactly what we observed (Figure 3-figure supplement 2,5). In lines with the strongest phenotypes, we observed roughly half the total number of branches per plant (Figure 3E). No difference in the number of branches was observed for lines that had a HACR with a stabilized auxin degron regulating PIN1 expression, suggesting this phenotype was not simply due to repression of PIN1 (Figure 3-figure supplement 3).

Feedback strength is also an important control parameter for the process of phyllotactic patterning. In the inhibition zone model, each primordium (Figure 3F, green circles) creates an inhibition zone around itself by depleting auxin (Figure 3F, shown in orange) from its surroundings, thereby preventing enough auxin to accumulate to form a new primordium. This zone is created by a feedback driven flow of auxin towards the primordium. The cells that are capable of forming new primordia are present in a region called the central zone periphery (Figure 3F, black ring) surrounding the shoot apical meristem (Figure 3F, green circle in the back ring). The overlapping inhibition zones from all the existing nearby primordia leave only certain regions of the central zone periphery capable of forming new primordia (Figure 3F, dashed green circles on yellow arcs). A mathematical model by Refahi et al.[4] divides the central zone periphery into discrete units or cells and calculates a probability for each cell to form a new primordium at every timepoint. This probability is used to simulate the growth of the plant and estimate the expected frequency of phyllotactic patterning errors, such as co-initiation of primordia (Figure 3F, as shown in the grey meristem). This occurs when there is more than one region on the central zone periphery that is competent to form a primordia, leading to two primordia being initiated at the same time. According to the model, the radius of the inhibition zones is inversely proportional to the number of co-initiatiating primordia. In auxin HACR plants with a PIN1 gRNA, we hypothesized that lower feedback strength would lead to a less sharp auxin gradient around each primordium and thus a larger inhibition zone[26] (Figure 3F, as shown in the blue meristem). Consistent with this prediction, plants with a HACR targeting *PIN1* showed a significant reduction in co-initiations (Figure 3G, Figure 3-figure supplement 4).

By making it possible to alter transcriptional feedback strength rather than simply gene expression, the HACR platform enabled exploration of previously inaccessible parameter regimes. This proof-of-concept establishes a new method for modifying a large number of desired traits. Additionally, the modular nature of HACRs allows for independent tuning of hormone sensitivity and repression strength[15], as well as allowing for tissue-specific modulation of target genes. These modifications could substantially extend the range of possible phenotypes and mitigate trade-offs, for example having few branches to fit more plants on a field versus the total number of fruits per plant. The use of HACRs here is among the first examples of utilizing synthetic signaling systems to re-engineer the morphology of a multicellular

organism in a model-driven manner, a long standing goal across the fields of pattern formation and tissue engineering, and this strategy should be extensible to a wide variety of organisms, particularly given the success of implementing the auxin-induced degradation module (AID) in diverse eukaryotes[28]. In agricultural settings, farmers already manipulate development or defense pathways by applying hormones or their synthetic mimics. HACRs could be used to connect these treatments with the expression of genes, such as those involved in defense, to create inducible traits. Additionally, HACRs could be extended to any other hormone that utilizes degradation-based signaling, such as salicyclic acid, strigalactones and karrikins. The wide range of degradation cues, the ease of targeting any gene, and the likely conserved function across angiosperms should mean that HACRs have the capacity to reprogram a plethora of developmental traits in a broad range of crop species.

Methods

Construction of plasmids

Expression cassettes for the gRNAs, HACRs and the reporters were built using Gibson assembly[29]. These were then linearized by restriction enzyme digestion and assembled into a yeast artificial chromosome based plant transformation vector with kanamycin resistance using homologous recombination based assembly in yeast[30]. The PIN1 gRNA expression vector and the additional tdTomato expression vector for the ratiometric lines were built using Golden-Gate assembly[31] into the pGRN backbone[32] with hygromycin resistance.

The gRNA expression cassettes contain a sgRNA driven by the U6 promoter and have a U6 terminator. The HACR expression cassettes are driven by the constitutive UBQ10 (AT4G05320) promoter and have a NOS terminator. All HACR variants contain the same deactivated SpCas9 (dCas9) domain[16] translationally fused at the N-terminus to an SV40 nuclear localization signal. The hormone degron domain and the repressor domain were fused to the C terminus of dCas9, with the respective degron domain in the middle and flexible 6xGS linkers separating the sub-domains. The rapidly degrading NdC truncation of the IAA17 degron[17] was used for all the auxin HACRs described in the paper. The JA HACR contained the degron from the Arabidopsis JAZ9 protein (AT1G70700)[19]. The GA HACRs contained either GAI (At1g14920)[20] or RGA1 (At2g01570)[20] cloned from *Arabidopsis* cDNA. The HACR repression domain was the nucleic acid sequence corresponding to the first 300 amino acids of the TOPLESS repressor (TPL, At1g15750)[18]. We chose this repression domain as TPL is the co-repressor used in native auxin and JA signal transduction pathways. The reporter cassette that was regulated by the HACRs contained a yellow fluorescent protein (Venus) translationally fused to a nuclear localization sequence on its N-terminus and firefly luciferase translationally fused on its C-terminus with flexible linkers. The reporter was driven by a constitutive UBQ1 (AT3G52590) promoter and had a UBQ1 terminator. The additional reporter in the ratiometric lines was identical to these constructs except Venus-Luciferase was replaced with tdTomato and the gRNA target site in the UBQ1 promoter was mutated. The PIN1 gRNA expression vector contained a U6 promoter and terminator.

Construction of plant lines

All HACR reporter lines were built by transforming the yeast artificial chromosome plasmids described above into *Agrobacterium tumefaciens* (GV3101) and using the resulting strains to transform a

Columbia-0 background by floral dip[33]. Transformants were then selected using a light pulse selection[34]. Briefly, this involves exposing the seeds to light for 6 hours after stratification (4°C for 2 days in the dark) followed by a three day dark treatment. Resistant seedlings demonstrate hypocotyl elongation in the case of Hygromycin and leaf greening after 5 days in the case of Kanamycin. After selection seedlings were transplanted to soil and grown in long day conditions at 22°C.

For all the HACR reporter genotypes (Figures 1 and 2) at least three lines were grown to the T2 and tested for their response to the appropriate hormone treatment with n=10 for seedlings. To generate the ratiometric auxin HACR lines the additional tdTomato reporter was transformed into Col0 and then lines that were screened for uniform tdTomato expression were crossed into a line that had the HACR targeted to a Venus reporter.

Three different auxin HACR backgrounds were transformed with a gRNA targeting PIN1. The branching of three independent lines, representing three independent PIN1 gRNA insertion events, in each HACR background was characterized in the T2 at n=5. Several lines were characterized in the T3 at n>20 both with and without selection. The number of co-initiations of three independent lines in one HACR background was characterized in the T2 at n=5. The number of co-initiating siliques of one of these lines was characterized in the T3 at n=25.

Fluorescence Microscopy

For imaging the effects of auxin treatment on root tips we selected plants on 0.5xLS + 0.8% bactoagar containing Kanamycin using the light pulse protocol described above. Four days after the seedlings were removed from the dark we transplanted to fresh 0.5xLS + 0.8% bactoagar without Kanamycin and then imaged on a Leica TCS SP5 II laser scanning confocal microscope on an inverted stand. For auxin induction of root tips, the seedlings were sprayed with a 1:1000 dilution in water of either control (DMSO) or auxin dissolved in DMSO (5µM final concentration) and then mounted on slides in water and imaged after 24 hours.

For the imaging of ratiometric lines seedlings were germinated without selection and then visually screened using a fluorescence microscope for expression of both reporters. These seedlings were then imaged on a confocal microscope at several positions along the primary root to visualize auxin distributions in the root tip, the elongation zone and in developing lateral roots. The images were taken using a Leica TCS SP5 II laser scanning confocal microscope on an inverted stand. The ratiometric images were generated using the calcium imaging calculator in the Leica software, by background subtracting both the tdTomato and Venus signals and then normalizing the Venus signal by the tdTomato signal.

The images of ovules 48 hours after pollination were obtained by emasculating flowers prior to anther dehiscence followed by hand pollination 12 hours after. After 48 hours, the ovules from the pistils of these flowers were dissected using hypodermic needles under a dissection microscope and then mounted on slides in 80mM sorbitol and imaged with confocal microscopy as in Beale *et al*.[35]. To image

the developing embryos, ovules were dissected from siliques at the appropriate developmental stages, individually dissected and mounted onto slides in MS0 media before being analyzed by confocal microscopy. All confocal microscopy images presented in this work are maximum projections of sub-stacks from regions of interest.

Luciferase assays

Luciferase based time course assays were used to characterize the dynamics of HACR response to exogenous or endogenous hormone stimulus. All imaging was done using the NightOWL LB 983 in vivo Imaging System, which uses a CCD camera to visualize bioluminescence. For the data collected for Figure 1 and Figure 3-figure supplement 1, assays were performed on seedlings. Here, T2 plants were selected by Kanamycin selection using the previously described light pulse protocol. These were then transplanted to fresh plates without antibiotic four days after selection and sprayed with luciferin (5µM in water) in the evening. The next morning, after approximately 16 hours, they were sprayed again with luciferin. After 5 hours they were imaged for one hour (10 minute exposure with continuous time points), then sprayed with a control treatment (a 1:1000 dilution of DMSO in water) and then imaged for five hours. These same plates were then re-sprayed with luciferin (5µM in water) and left overnight. The next day these same plates were again imaged with an identical protocol as the previous day, except they were sprayed with a 1:1000 dilution of hormone in water (5µM Indole-3-acetic acid (auxin), 30µM coronatine (JA) or 100µM GA3 post dilution) rather than control. Luminescence of each seedling was recorded over time and reported as values normalized to the time-point prior to treatment. For the mechanical damage assay of the jasmonate HACR in figure 2, plants were treated identically as described above except that instead of being sprayed with hormones, leaves on the plant were mechanically crushed using forceps.

Data Analysis

All the data collected was analyzed and plotted using python (https://github.com/arjunkhakhar/HACR_Data_Analysis). For the luciferase assays, all the time courses were normalized the reading before induction to make them comparable. All p-values reported were calculated in python using the one-way ANOVA function from the SciPy package[36]. (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.f_oneway.html)

Characterizing plant phenotypes

To characterize branching in plant lines with and without an auxin HACR regulating PIN1, we selected T2 transformants for lines that had a gRNA targeting PIN1 and the parental HACR background that had no gRNA. The plants that passed the selection were transplanted onto soil and then characterized as adults at the point that there were on average 4 stems on the no gRNA control lines. In all cases the parental controls that lack a gRNA and the lines derived from them, by transforming with a gRNA targeting PIN1, were all grown in parallel and phenotyped on the same day to ensure the data collected was comparable. Additionally, while we do not believe that the selection would have a significant effect on

the phenotyping data as we collected it more than a month after the plants had been transplanted off selection plates onto soil, both the lines with a PIN1 targeting gRNA and the parental controls they were compared to were selected in parallel to control for any confounding effect. Phenotyping involved counting the number of branches on the plant. We quantified the number of branches on five T2 plants for three different lines with a HACR targeted to regulate PIN1 in two different HACR backgrounds, in parallel with the parental HACR background. The line with the strongest phenotype was propagated to the T3 generation with its parental HACR background and the same experiment was repeated with an n=25. To quantify the number of co-initiating siliques we measured the internode length between the first 20 siliques on a single axillary stem and every instance of two siliques emerging from the same point on the stem (an internode length less than 1 mm which we found to be the threshold for visual discrimination) was considered a co-initiation. The line that showed the strongest phenotype was propagated to the T3 generation with its parental HACR background and the same experiment was repeated with an n=25.

To prove the phenotypes we were observing were independent of selection conditions we also characterized branching of T2 and T3 plant lines that were not selected on antibiotic selections. These plant lines were transplanted off 0.5x LS plates ten days after germination. They were then grown till adulthood and then phenotyped and genotyped for the presence of the HACR and PIN1 gRNA.

All plants that were phenotyped were grown in long day conditions on Sunshine #4 mix soil in rose pots and watered every other day on a watering table.

qPCR assays

All qPCR assays were performed on seedlings seven days after they been selected using the light pulse procedure (fifteen days post germination). For each biological replicate 5 seedlings that passed selection were transplanted off the selection plate and into 4ml of 0.5xLS with either mock of 50nM 2-4D. They were then incubated in well lit, humidity-controlled conditions for 3 hours and then the seedlings were blotted and flash frozen in liquid nitrogen. The RNA was extracted from these seedlings using the Illustra RNAspin Mini Kit from GE. cDNA was then prepared from 1ug of RNA using the iScript cDNA synthesis kit from Biorad and then used to run a qPCR with the iQ SYBR Green Supermix also from Biorad on a Biorad qPCR machine. Each sample was analyzed for expression of PIN1 and PP2A which was used to normalize PIN1 levels. A standard curve was generated using the pooled samples for each primer set to determine amplification efficiency. The primers used are listed below:

PIN1_q_R: AACATAGCCATGCCTAGACC

PIN1_q_F: CGTGGAGAGGGAAGAGTTTA

PP2A_q_R: AACCGCTTGGTCGACTATCG

PP2A_q_F: AACGTGGCCAAAATGATGC

Plant genotype list

Plant genotype	Used in the following figure
ABS44 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-NdC_IAA17-TPLRD2-tNos)	Figure 1B-F , Figure 2F-J , Figure 3B,E,G,H , Supplement Figure 1,2,4,6,7
PHD5 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-Jas9-TPLRD2-tNos)	Figure 1H, Figure 2N-Q, Supplement Figure 1,2
PHD3 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-GAI1-TPLRD2-tNos)	Figure 1J, Figure 2K-M, Supplement Figure 1,2
PHD6 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-RGA1-TPLRD2-tNos)	Figure 1I, Supplement Figure 1,2
ABS44 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-NdC_IAA17-TPLRD2-tNos) + pGRN_H-pU6:pPIN1_gRNA_Target1-tU6	Figure 3C,E,G,H, Supplement Figure 4,6,7
ABS50 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-IAA28_DegronDead-TPLRD2-tNos)	Figure 1D,F, Supplement Figure 5
ABS50 (p2301Y-tOCS-pUBQ1:NLS-Venus-LucPlus-tUBQ1-pU6:pUBQ1_gRNA_Target1-tU6-pUBQ10:dCas9-IAA28_DegronDead-TPLRD2-tNos) + pGRN_H-pU6:pPIN1_gRNA_Target1-tU6	Supplement Figure 5

Plasmid Maps

ABS44 - https://benchling.com/s/yXKJkba5

ABS50 - https://benchling.com/s/897tnlX2

PHD5 - https://benchling.com/s/HnODIKMV

PHD3 - https://benchling.com/s/HOEPc5FA

PHD6 - https://benchling.com/s/Ge8pztYw

pGRN_H-pU6:pPIN1_gRNA_Target1-tU6 - https://benchling.com/s/3RBYAIkF

pGRN_H-pUBQ1_AlteredGrnaTargetSite:NLS-tdTomato-tUBQ1 - https://benchling.com/s/Pd0Ms4Qs

References

1. Doebley, Stec & Hubbard. The evolution of apical dominance in maize. *Nature* **386,** 485–488 (1997).

2. Khush. Green revolution: the way forward. *Nature Reviews Genetics* **2,** 815–822 (2001).

3. Vanstraelen & Benková. Hormonal Interactions in the Regulation of Plant Development. *Annual Review of Cell and Developmental Biology* **28,** 463–487 (2012).

4. Refahi *et al.* A stochastic multicellular model identifies biological watermarks from disorders in self-organized patterns of phyllotaxis. *eLife* **5,** (2016).

5. Prusinkiewicz *et al.* Control of bud activation by an auxin transport switch. *Proceedings of the National Academy of Sciences* **106,** 17431–17436 (2009).

6. Parry *et al.* Mutation discovery for crop improvement. *Journal of Experimental Botany* **60,** (2009).

7. Voytas & Gao. Precision Genome Engineering and Agriculture: Opportunities and Regulatory Challenges. *PLoS Biology* **12,** e1001877 (2014).

8. Weijers & Wagner. Transcriptional Responses to the Auxin Hormone. *Annual Review of Plant Biology* **67,** 1–36 (2015).

9. Smith *et al.* A plausible model of phyllotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 1301–1306 (2006).

10. Koltai. Cellular events of strigolactone signalling and their crosstalk with auxin in roots. Journal of Experimental Botany 66, 4855–4861 (2015).

11. Naseem, Kaltdorf & Dandekar. The nexus between growth and defence signalling: auxin and cytokinin modulate plant immune response pathways. Journal of Experimental Botany 66, 4885–4896 (2015).

12. Brophy, LaRue & Dinneny. Understanding and Engineering Plant Form. *Seminars in Cell & Developmental Biology* (2017). doi:10.1016/j.semcdb.2017.08.051

13. Ulmasov, Murfett, Hagen & Guilfoyle. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant Cell Online 9, 1963–1971 (1997).

14. Rushton, Reinstädler, Lipka & Lippok. Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen-and wound-induced signaling. (2002).

15. Khakhar, Bolten, Nemhauser & Klavins. Cell–Cell Communication in Yeast Using Auxin Biosynthesis and Auxin Responsive CRISPR Transcription Factors. ACS Synthetic Biology 5, 279–286 (2016).

16. Gilbert et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell 154, (2013).

17. Moss et al. Rate Motifs Tune Auxin/Indole-3-Acetic Acid Degradation Dynamics. Plant Physiology 169, 803–813 (2015).

18. Pierre-Jerome, Jang, Havens, Nemhauser & Klavins. Recapitulation of the forward nuclear auxin response pathway in yeast. Proceedings of the National Academy of Sciences 111, 9407–9412 (2014).

19. Katsir, Chung, Koo & Howe. Jasmonate signaling: a conserved mechanism of hormone sensing. Current Opinion in Plant Biology 11, 428–435 (2008).

20. Murase, Hirano, Sun & Hakoshima. Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. Nature 456, 459–463 (2008).

21. Liao et al. Reporters for sensitive and quantitative measurement of auxin response. Nature methods 12, (2015).

22. Band et al. Systems Analysis of Auxin Transport in the Arabidopsis Root Apex. The Plant Cell Online 26, 862–875 (2014).

23. Dubrovsky et al. Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. Proceedings of the National Academy of Sciences 105, 8790–8794 (2008).

24. Hu et al. Potential Sites of Bioactive Gibberellin Production during Reproductive Growth in Arabidopsis. The Plant Cell 20, 320–336 (2008).

25. Burks et al. Engineering and horticultural aspects of robotic fruit harvesting: Opportunities and constraints. HortTechnology 15, (2005).

26. Bennett, Hines & Leyser. Canalization: what the flux? Trends in Genetics 30, 41–48 (2014).

27. Vieten, Anne, et al. "Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression." *Development* 132.20 (2005).

28. Nishimura, Fukagawa, Takisawa, Kakimoto & Kanemaki. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature Methods 6, 917–922 (2009).

29. Gibson et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods 6, 343–345 (2009).

30. Shih et al. A robust gene-stacking method utilizing yeast assembly for plant synthetic biology. Nature Communications 7, 13215 (2016).

31. Engler, Kandzia & Marillonnet. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS ONE 3, e3647 (2008).

32. Hellens, Edwards, Leyland, Bean & Mullineaux. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Molecular Biology 42, 819–832 (2000).

33. Clough & Bent. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal 16, 735–743 (1998).

34. Harrison et al. A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. Plant Methods 2, 1–7 (2006).

35. Beale, Leydon & Johnson. Gamete Fusion Is Required to Block Multiple Pollen Tubes from Entering an Arabidopsis Ovule. Current Biology 22, 1090–1094 (2012).

36. Oliphant. Python for scientific computing. Computing in Science & Engineering (2007). doi:10.1109/mcse.2007.58?crawler=true

37. Khakhar A. 2018. HACR_Data_Analysis. Github. https://github.com/arjunkhakhar/HACR_Data_Analysis

Figures

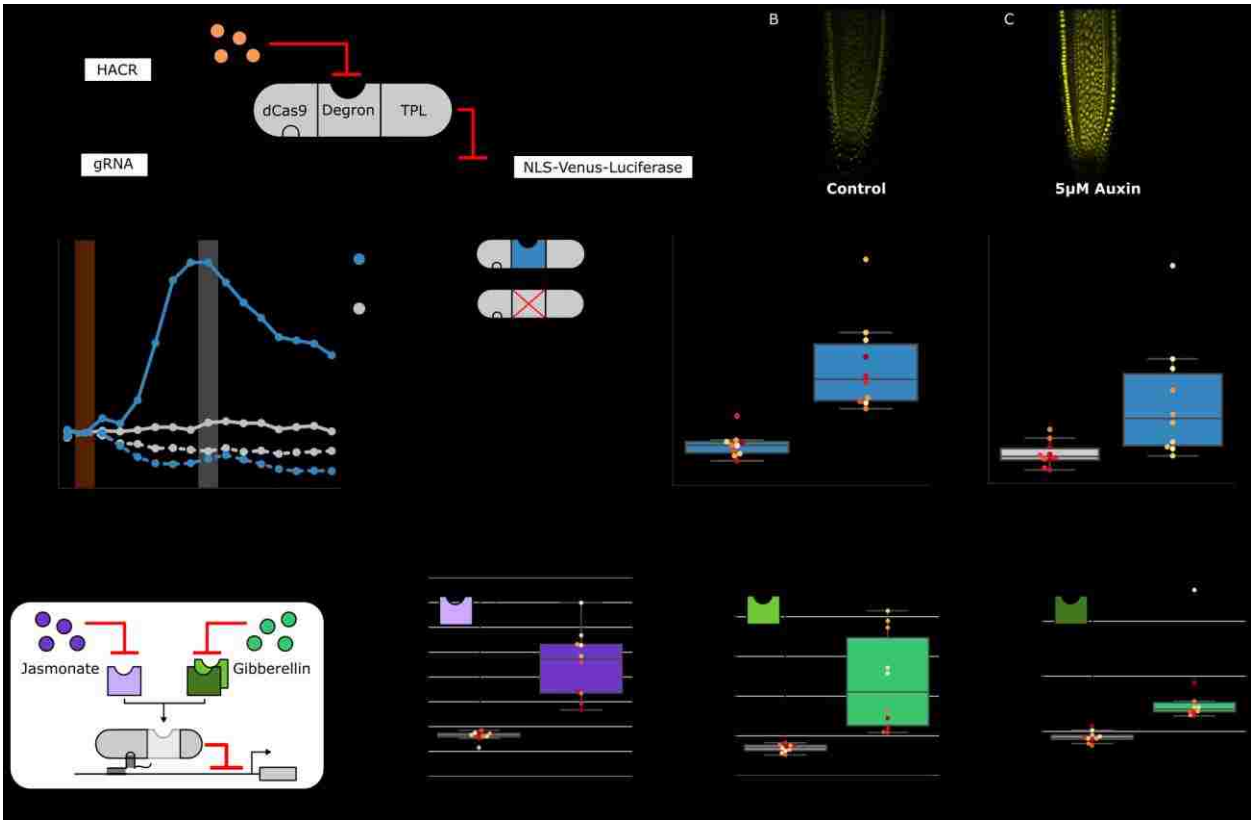


Figure 1: HACRs modulate gene expression upon exogenous hormone treatment. A) A general schematic of the constructs transformed into *Arabidopsis thaliana* to test HACR hormone response. B,C) Confocal microscopy images of root tips from plant lines with an auxin HACR regulating a Venus reporter 24 hours after treatment with (B) control or (C) 5μM auxin. D) An example of a luciferase based time course assay testing whole seedlings of an auxin HACR line treated with auxin (solid blue line) and a control (dashed blue line). The timepoint of auxin induction is highlighted with an orange bar. The time point of maximum auxin response is highlighted by the grey bar. E) The difference between auxin and control induction at the time of maximum auxin response for the tested seedlings (n = 10) is summarized in the box plot. Every seedling is represented as a different colored dot. F) A HACR variant line with a stabilized auxin degron was also assayed (D, solid and dashed grey lines) and the response to auxin of these seedlings compared to seedlings of the line with a functional auxin degron at the time of maximum auxin response are summarized in box plot in F. G) A schematic of how the hormone specificity of HACRs were altered by swapping the hormone degron. H,I,J) These box plots summarize the response of transgenic seedlings carrying these constructs (n=10) to treatment with either control or the appropriate hormone. The degron used in the HACR is specified in the top left corner of the plot. Every seedling is represented as a different colored dot. All p-values reported were calculated using a one-way ANOVA.

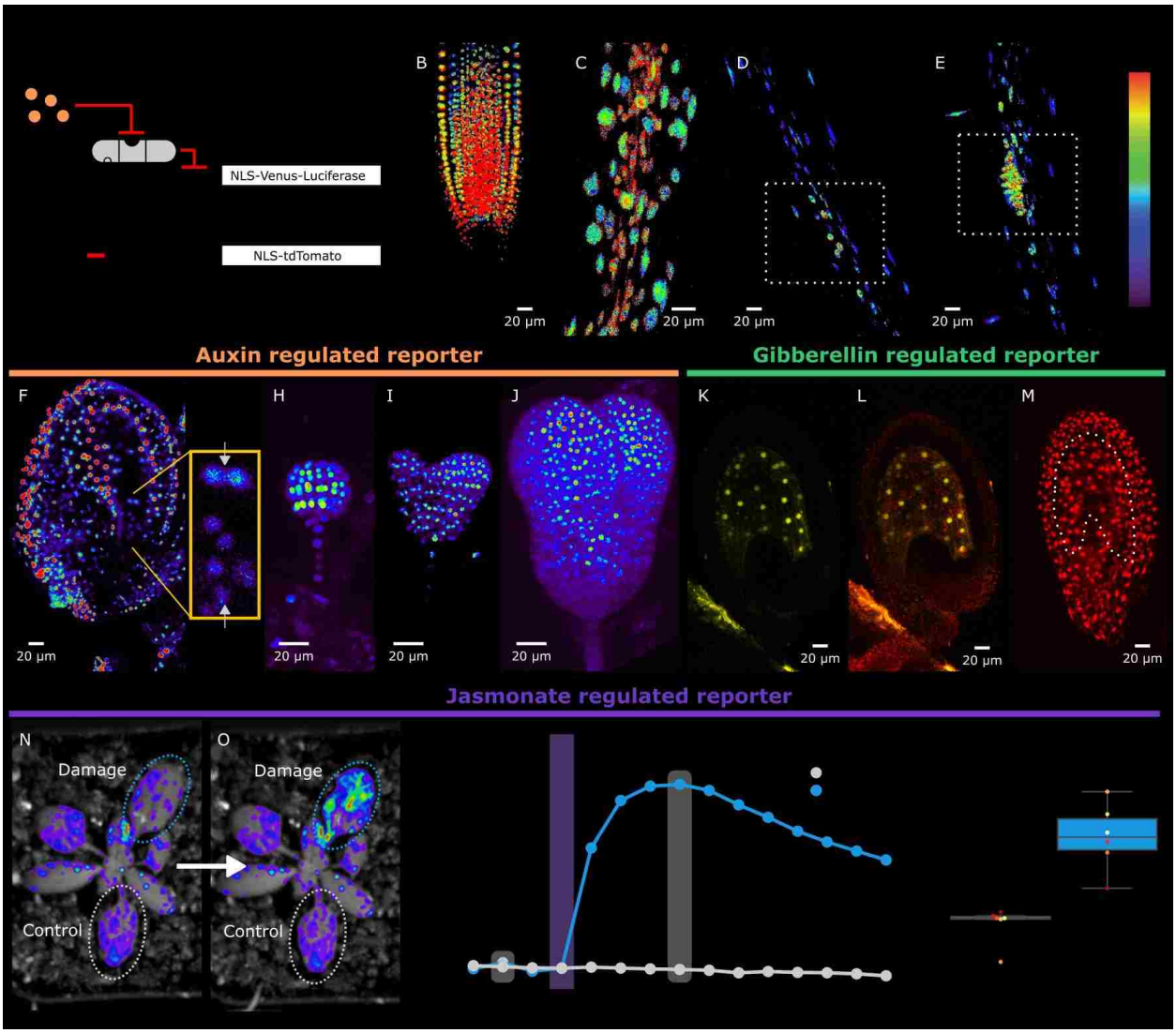


Figure 2: HACRs respond to endogenous hormone signals and can be used to study development. A) Schematic of the genetic circuit used to build ratiometric lines of auxin responsive HACRs. In addition to an auxin HACR regulating a nuclear localized Venus-luciferase reporter the lines also have a nuclear localized tdTomato reporter being driven by a version of the UBQ1 promoter with the gRNA target site mutated. B-E) Confocal microscopy images of roots of seedlings from lines described in A. Reporter signal in images is the background subtracted Venus signal normalized by the background subtracted tdTomato signal. Warmer colors correspond to higher normalized reporter signal. B) The stereotypical reverse fountain pattern of auxin distribution is observed in the root tip. C) Higher reporter signal is observed in the vasculature compared to the epidermis of the elongation zone of the root, consistent with auxin being trafficked along the vasculature. The dashed white boxes highlight high reporter signal in (D) the founder cells of lateral roots and in (E) a developing lateral root primordium. F-J) Confocal microscopy images visualizing reporter signal of a non-ratiometric auxin HACR regulated reporter (F) in the ovule 48 hours post pollination, (G) in the two-cells embryo, (H) in the globular embryo, (I) in the heart stage embryo and (J) in the early torpedo stage embryo. Warmer colors correspond to higher

reporter signal. K-M) Confocal microscopy images visualizing reporter signal of a GA HACR regulated reporter (K) in the ovule 48 hours post pollination, (L) reporter signal merged with red auto-fluorescence to highlight the endosperm region and (M) an unregulated tdTomato reporter, with the endosperm highlighted with a dashed white line, for comparison. N-Q) Visualization of JA HACR regulated reporter expression in leaves in response to mechanical damage using a luciferase-based assay. Images of leaves overlaid with the luciferase signal before (N) and after damage (O) are shown to the left of a representative plot of the normalized reporter signal over time (P). Q) Box plot summarizing the maximum fold change at 70 minutes for control and damaged leaves. Points of the same color represent leaves from the same plant.

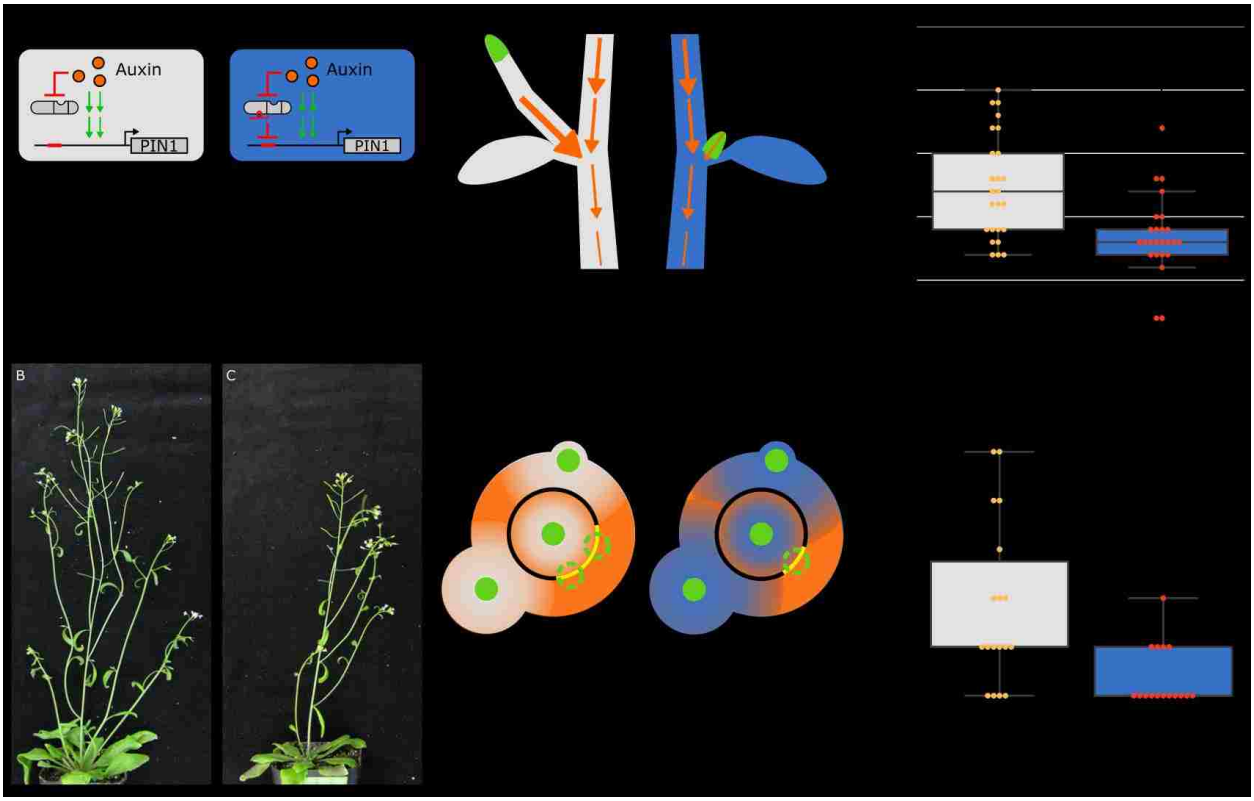


Figure 3: The developmental circuit regulating branching can be rewired using auxin HACRs. A) Schematics of auxin driven PIN1 expression, which is one of the drivers of transcriptional feedback. In the box on the right we show how we decreased PIN1 transcriptional feedback strength by targeting an auxin HACR to regulate PIN1. B,C) Representative pictures of T3 plants of the same age without (B) and with (C) a gRNA targeting an auxin HACR to regulate PIN1. D) Schematic of the mechanism behind the predicted decrease in branching from decreasing transcriptional feedback strength. In plants without a HACR targeted to PIN1 (grey), the stronger transcriptional feedback allows the lateral bud (green) to drain auxin (orange arrows) into the central vasculature. In plants with a HACR targeted to PIN1 leading to reduced transcriptional feedback (blue), the bud is not able to drain its auxin, preventing branch formation. E) Box plots summarizing the number of branches of adult T3 plant lines (n = 25) with a HACR targeted to regulate PIN1 expression (blue boxes), compared to control lines that did not have a gRNA targeting PIN1 (grey boxes). Every dot represents an individual plant. F) Schematic depicting the role of transcriptional feedback in the pattern of formation of new primordia (green circles) around the shoot apical meristem. We hypothesize that in the shoot apex of lines without a HACR targeting PIN1 (grey) the stronger transcriptional feedback leads to smaller zones of auxin depletion around primordia compared to lines that have a HACR targeting PIN1 (blue). This leads to a broader zone where auxin can accumulate (orange) and create new primordia (dashed green circles) which increases chances of phyllotactic defects. G) Box plots summarizing the number of co-initiations in T3 plant lines (n = 25) with a HACR targeted to regulate PIN1 expression (blue boxes), compared to parental control lines that did not have a gRNA targeting PIN1 (grey boxes). Every dot represents an individual plant. All p-values reported were calculated using a one-way ANOVA.

Author information

Authors declare no competing interests.

Acknowledgements

We thank Dr. Takato Imazumi for sharing resources and advice, particularly on luciferase imaging; Dr. Dominic Loque for materials to build the yeast artificial chromosomes ahead of publication; Ms. Mrunmayee Shete and the Aquarium team for technical assistance; Dr. Mallorie Taylor-Teeples for resources and advice related to phyllotactic patterning; and Dr. Hardik Gala for advice on microscopy. We would like to thank Mr. Randolph Lopez and Dr. Clay Wright for commenting on the manuscript, and all members of the Klavins, Seelig and Nemhauser labs for their advice and input. This work was supported by shared grants to EK and JLN from the National Science Foundation (MCB-1411949) and the National Institutes of Health (R01-GM107084), as well as support from the Howard Hughes Medical Institute Faculty Scholars Program to JLN.

Supplementary Materials

Supplementary note 1

We built a model that captured the transcriptional activation of PIN1 by auxin and its repression by an auxin responsive HACR at the mRNA and protein levels. In this model PIN1 expression is activated proportional to the auxin concentration and repressed proportional to the auxinHACR concentration. Auxin causes the degradation of the auxinHACR protein. In addition to the passive diffusion of auxin in and out of the cell, auxin is actively transported out at a rate proportional to the concentration of PIN1. While the quantitative behavior of the model is dependent on the parameter set chosen, such as the repression strength, as we intend to use the model to make purely qualitative predictions all parameter values were chosen to generate biologically plausible behavior of the wildtype and have arbitrary units. The fact that the relative expression levels of PIN1 seem to agree with the qualitative predictions of model (Figure 3-figure supplement 1) implies that the parameter set is plausible. This model allows us to make qualitative predictions of how we would expect a HACR to perturb PIN1 expression. It also serves to highlight the significant differences that hormone responsive and static repression have on both the dynamic and steady state expression of PIN1 in response to auxin. The equations used to build the model, as well as the parameter values are listed below.

$\varphi_{PIN1_{mRNA}} = 1$

$\theta_{PIN1_{mRNA}} = 1$

$\varphi_{AuxinHACR_{mRNA}} = 1$

$K_{Repression\ strength} = 10$

$\delta_{PIN1} = 1$

$\mu_{PIN1} = 0.1$

$\delta_{PIN1} = 2$

$\mu_{PIN1} = 0.1$

$K_{degradation\ rate} = 5$

$K_{Auxin\ diffusion\ in} = 1$

$K_{Auxin\ diffusion\ out} = 0.01$

$K_{PIN1\ transport\ efficiency} = 1$

$K_{Auxin\ activation} = 1$

$$\frac{d[PIN1_{mRNA}]}{dt} = \varphi_{PIN1_{mRNA}} \times \left(\frac{K_{Auxin\ activation} \times [\text{Auxin}]}{K_{Auxin\ activation} \times [\text{Auxin}] + \theta_{PIN1_{mRNA}} + K_{Repression\ strength} * [\text{AuxinHACR}]} - [\text{PIN1}_{mRNA}] \right)$$

$$\frac{d[AuxinHACR_{mRNA}]}{dt} = \varphi_{AuxinHACR_{mRNA}} \times (1 - [\text{AuxinHACR}_{mRNA}])$$

$$\frac{d[PIN1]}{dt} = \delta_{PIN1} \times [\text{PIN1}_{\text{mRNA}}] - \mu_{PIN1} \times [\text{PIN1}]$$

$$\frac{d[AuxinHACR]}{dt} = \delta_{AuxinHACR} \times [\text{AuxinHACR}_{\text{mRNA}}] - \mu_{AuxinHACR} \times [\text{AuxinHACR}] - K_{degradation\ rate} \times [\text{Auxin}] \times [\text{AuxinHACR}]$$

$$\frac{d[Auxin]}{dt} = K_{Auxin\ diffusion\ in} - K_{Auxin\ diffusion\ out} * [\text{Auxin}] - [\text{PIN1}] \times K_{PIN1\ transport\ efficiency} \times [\text{Auxin}]$$

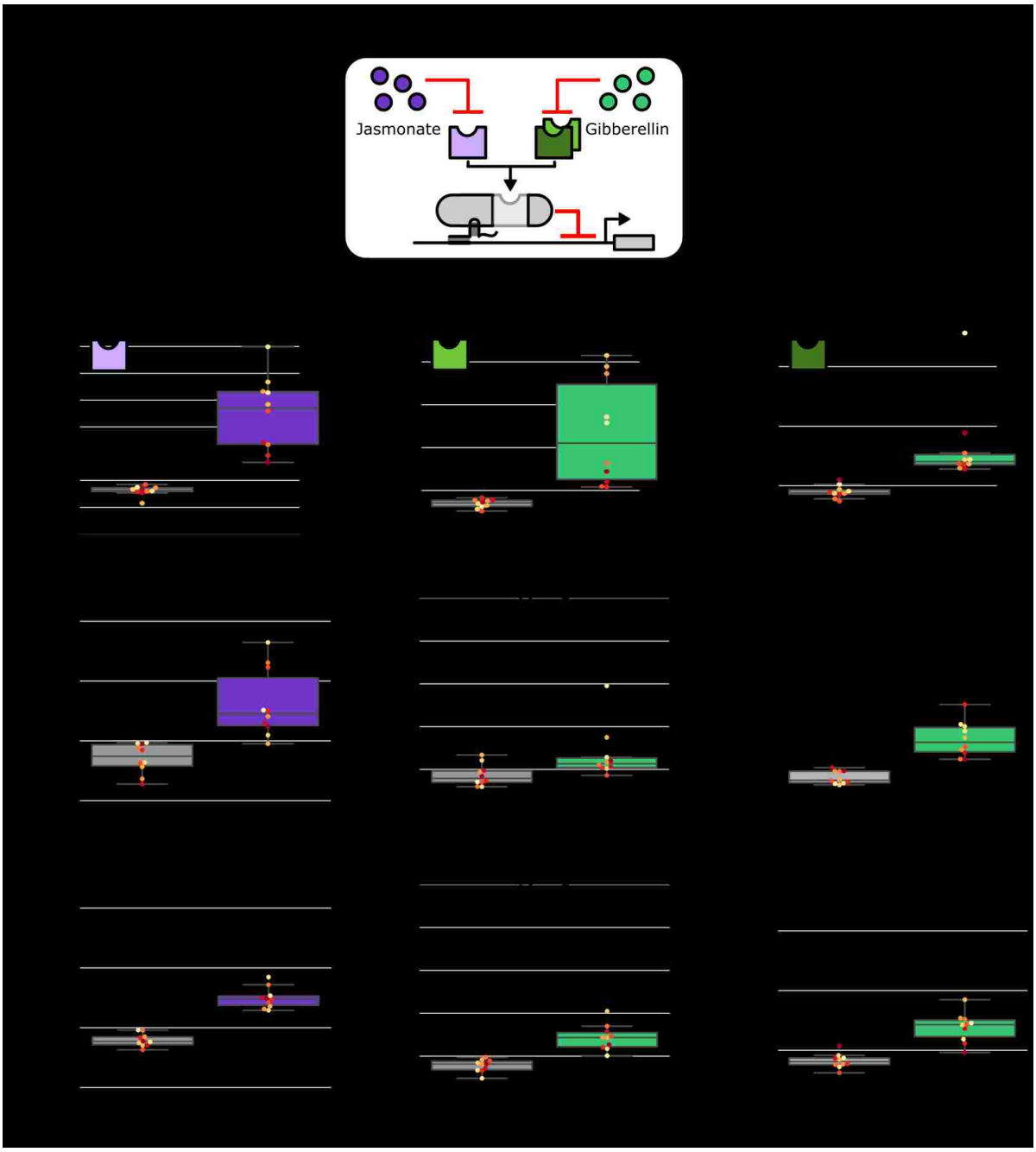


Figure 1-figure supplement 1: The hormone specificity of HACR response can be predictably altered by including different hormone responsive sequences. A) A schematic of how the hormone specificity of HACRs were altered by replacing the hormone degron with the jasmonate responsive degron from JAZ9[19] or the gibberellin responsive RGA1 or GAI proteins[20]. B,C,D) These box plots summarize the response of T2 transgenic seedlings of the indicated genotype (n=10) to treatment with either control or the appropriate hormone. The degron used in the HACR is specified in the top left corner of every column of plots. Every seedling is represented as a different colored dot. Three different lines representing three different insertion events of the HACR variant are shown in each column.

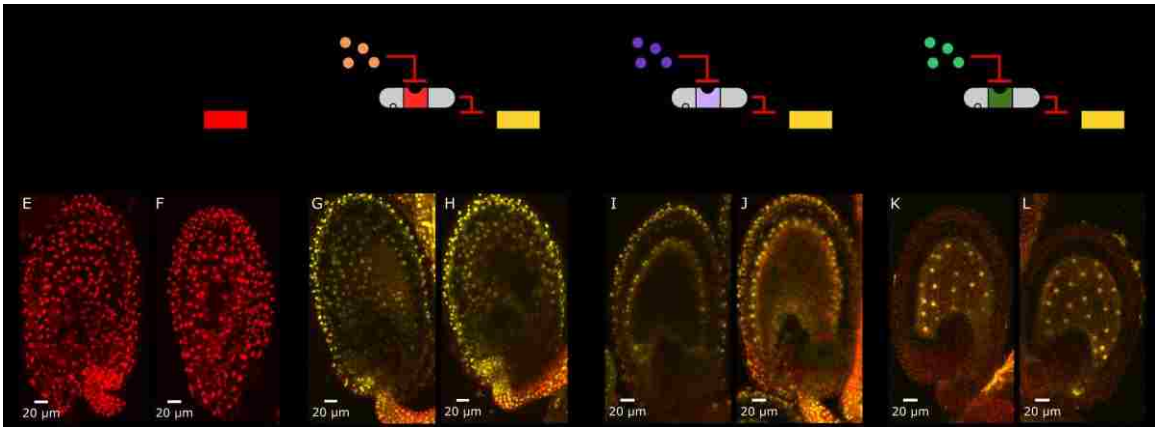


Figure 2-figure supplement 1: Distributions of auxin, gibberellin and jasmonate during early embryo development can be mapped out using HACRs. A-D) These schematics depict the HACR circuits in the plant lines whose ovules have been imaged using confocal microscopy directly below them. E-L) Confocal microscopy images of ovules from these reporter lines 48 hours post pollination visualizing either the tdTomato (in the case of the unregulated reporter) reporter signal or the Venus (in the case of the HACR lines) reporter signal. We observe ubiquitous reporter signal in the unregulated reporter (E,F) confirming that the UBQ1 promoter does not have differential expression across the ovule that could confound the observations from the HACR reporters. The auxin HACR reporter signal is present in the developing embryo, as reported previously. The JA HACR reporter signal is potentially localized to specific cell files (inner- and outer-layers) in the integument of the ovule (I,J). The GA HACR reporter signal is specifically localized to the endosperm, as we would expect based on the expression of GA biosynthesis genes (K,L).

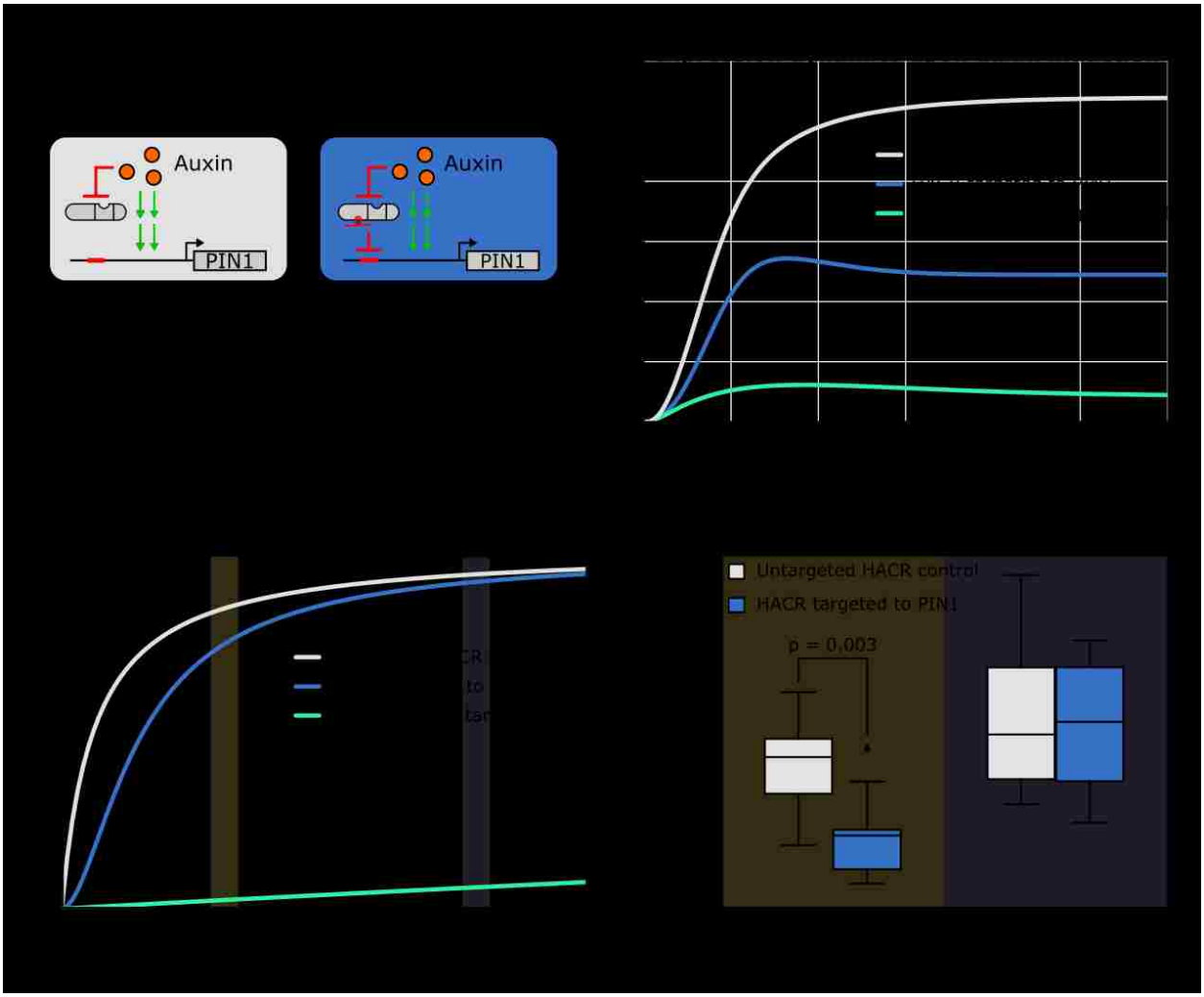


Figure 3-figure supplement 1: The effect of targeting a HACR to regulate PIN1 can be predicted using ordinary differential equations and qualitatively validated using qPCR. A) Schematics of auxin driven PIN1 expression with and without a HACR targeted to regulate PIN1 expression. B) Simulations of the expression of PIN1 in response to auxin induction at time 0. The grey line depicts the auxin activation of PIN1 in a cell with an untargeted HACR, while the blue line depicts the damped activation when a HACR is regulating PIN1. The green line depicts PIN1 activation with a repressor that has no auxin degron targeting PIN1. It demonstrates how not having an auxin degron present leads to qualitatively different activation dynamics as compared to a HACR. C) A plot depicting the dose response relationship between auxin and PIN1. As in panel B the grey line depicts the wild type, the blue line depicts the case where a HACR is regulating PIN1 and the green line depicts the case where a repressor without a degron is regulating PIN1. The model predicts a damped dose response relationship when a HACR is regulating PIN1, with larger damping at lower auxin concentrations and little to no damping at higher auxin concentrations. D) Plot depicting the relative expression levels of PIN1 measured by qPCR, normalized to the constitutively expressed gene PP2A, for seedlings of lines that either had a gRNA targeting a HACR to

PIN1 (Blue box plots) or did not (Grey box plots), with either mock or 50nM 2-4D (synthetic auxin) treatment for three hours. The yellow and purple bars in C correspond to approximately where in the dose response relationship we hypothesize these treatment conditions to be. We observe that the seedlings with a HACR targeted to PIN1 have significantly lower PIN1 expression compared to the untargeted control ($p=0.003$, $n=9$). We observe no significant difference between the two genotypes under 50nM 2-4D induction ($n=4$).

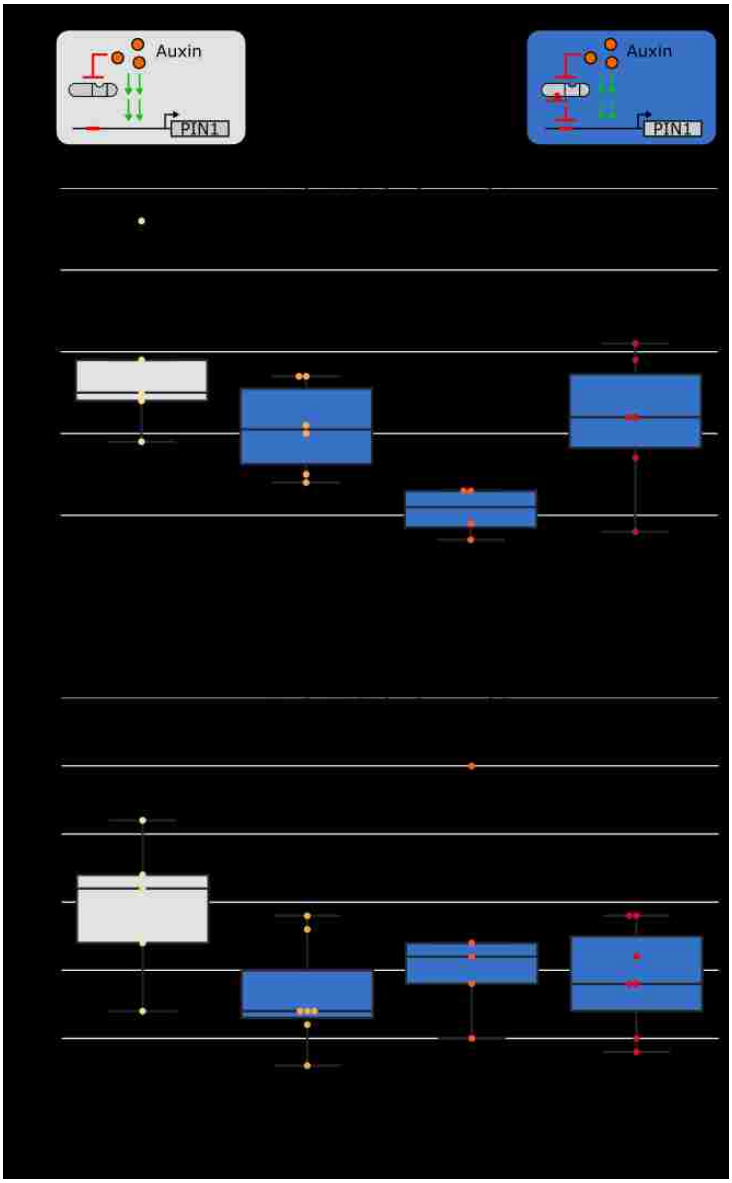


Figure 3-figure supplement 2: The developmental circuit regulating branching can be predictably rewired using auxin HACRs. A) A schematic that illustrates how transcriptional feedback strength was decreased by targeting an auxin responsive HACR to regulate PIN1. B,C) Box plots characterizing the number of branches of adult T2 plant lines with the HACR targeted to regulate PIN1 expression (blue boxes), compared to a parental control line that did not have a gRNA targeting PIN1 (grey boxes). HACR backgrounds (B & C) indicate lines generated from independent insertion events of the same HACR constructs and every dot represents an individual T2 transgenic plant. Each auxin HACR targeted to PIN1 line is a different PIN1 targeting gRNA insertion event into the same HACR background (Control). All data was collected from plants that were grown in parallel.

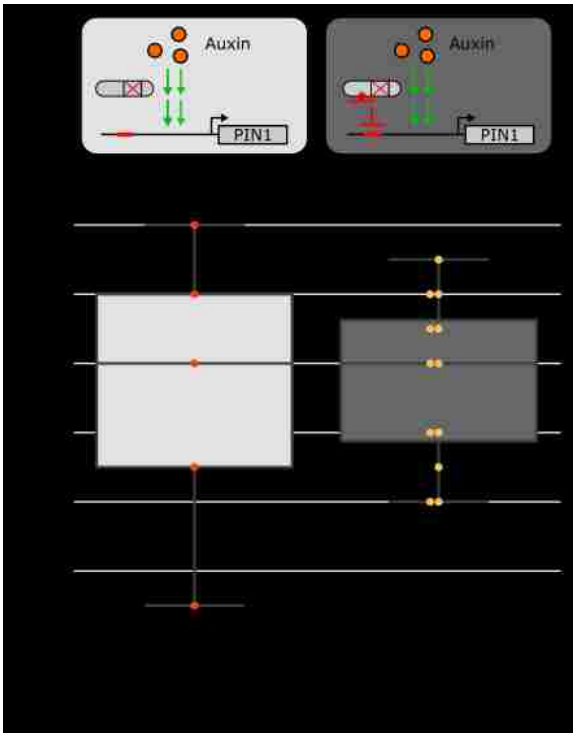


Figure 3-figure supplement 3: A HACR variant with a stabilized auxin degron does not produce a reduced branching phenotype when targeted to PIN1. The box plots summarize the number of branches on plant lines with the genetic circuit shown in the corresponding schematic above them. The light grey box plot depicts data for a line with a HACR that has a stabilized auxin degron[17] but no gRNA targeting PIN1, where each dot is an individual T2 transgenic plant. The dark grey box plot depicts data for first-generation transgenic plants where each dot represents a different insertion of the PIN1 gRNA into the previously described HACR background. All data was collected from adult plants that were grown in parallel. We observe no change in the median number of branches when the HCAR with the stabilized degron is targeted to PIN1. This is consistent with the hypothesis that it is a change in the auxin-activated expression of PIN1 (canalization) that is causing the branching phenotype and not simply the repression of PIN1.

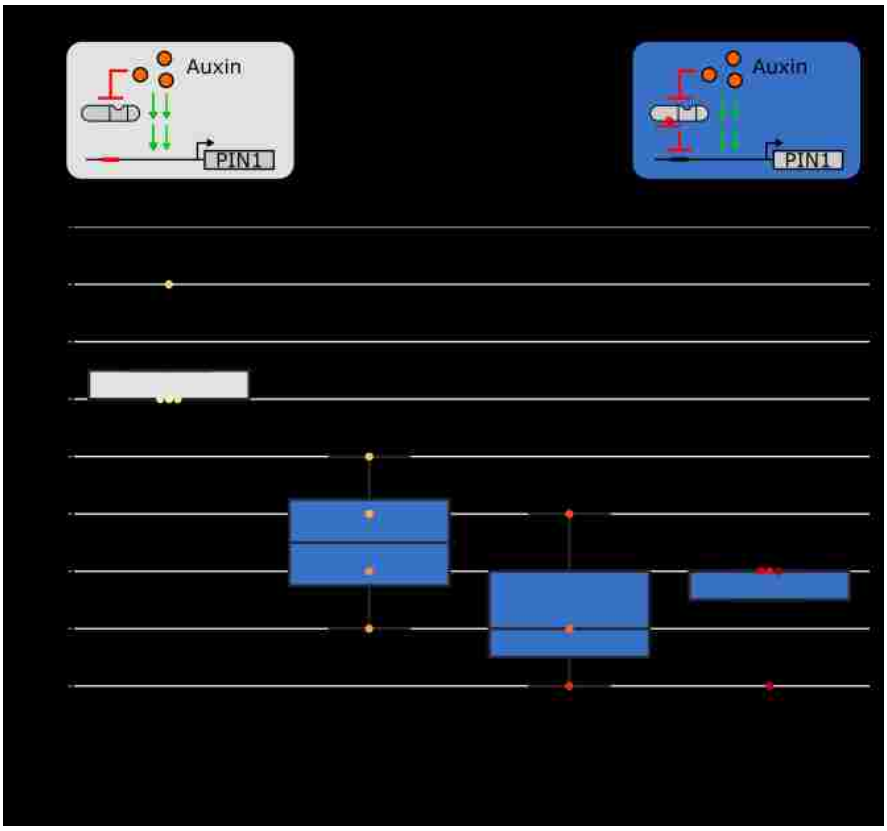


Figure 3-figure supplement 4: The developmental circuit regulating phyllotaxy can be predictably rewired using HACRs. A) A schematic that illustrates how we decreased transcriptional feedback strength by targeting an auxin responsive HACR to regulate PIN1. B) Box plots summarizing the number of co-initiations of lateral organs. Data was collected from adult T2 transgenic plants carrying a HACR targeted to regulate PIN1 expression (blue boxes), compared to parental control lines that lacked the PIN1 gRNA (grey boxes). All data was collected from plants that were grown in parallel. The number of co-initiations is an established metric of phyllotactic noise and we expect decreasing transcriptional feedback strength would decrease the noise. Every dot represents an individual plant and each line is a different PIN1-targeting gRNA insertion event into the parental control background.

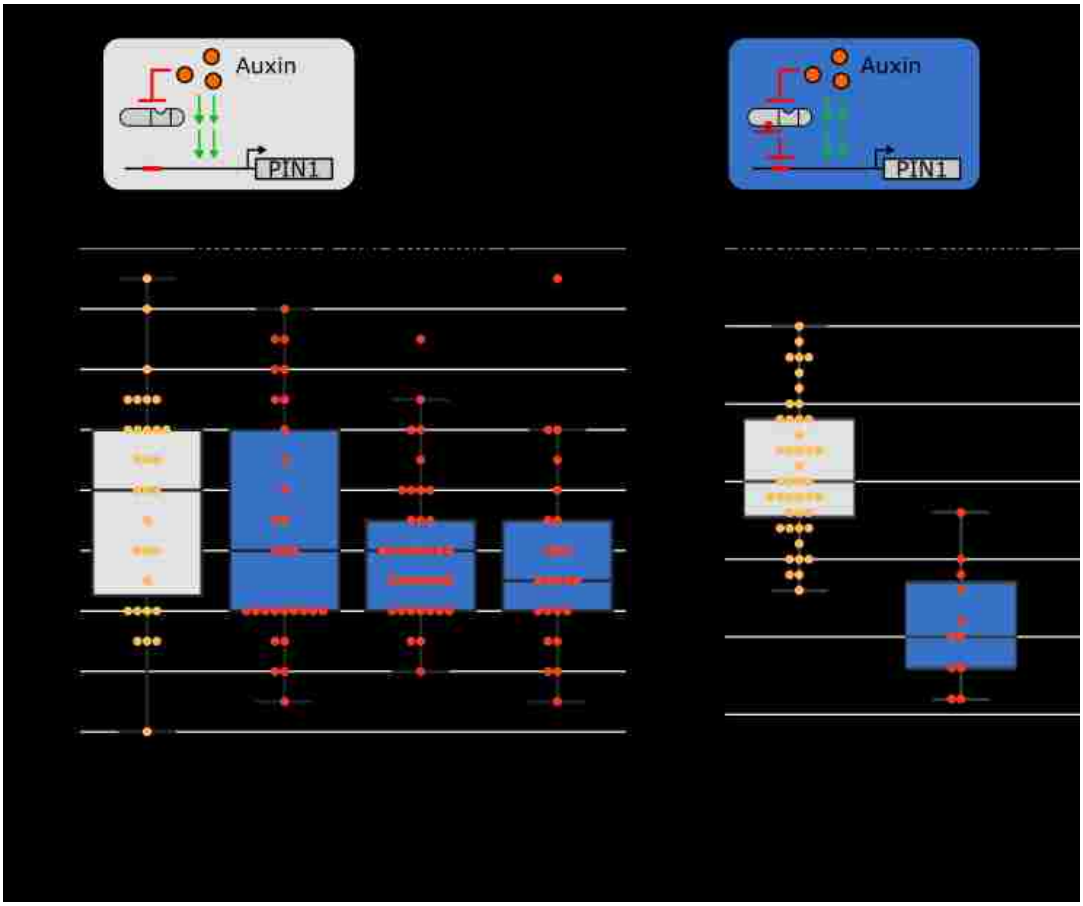
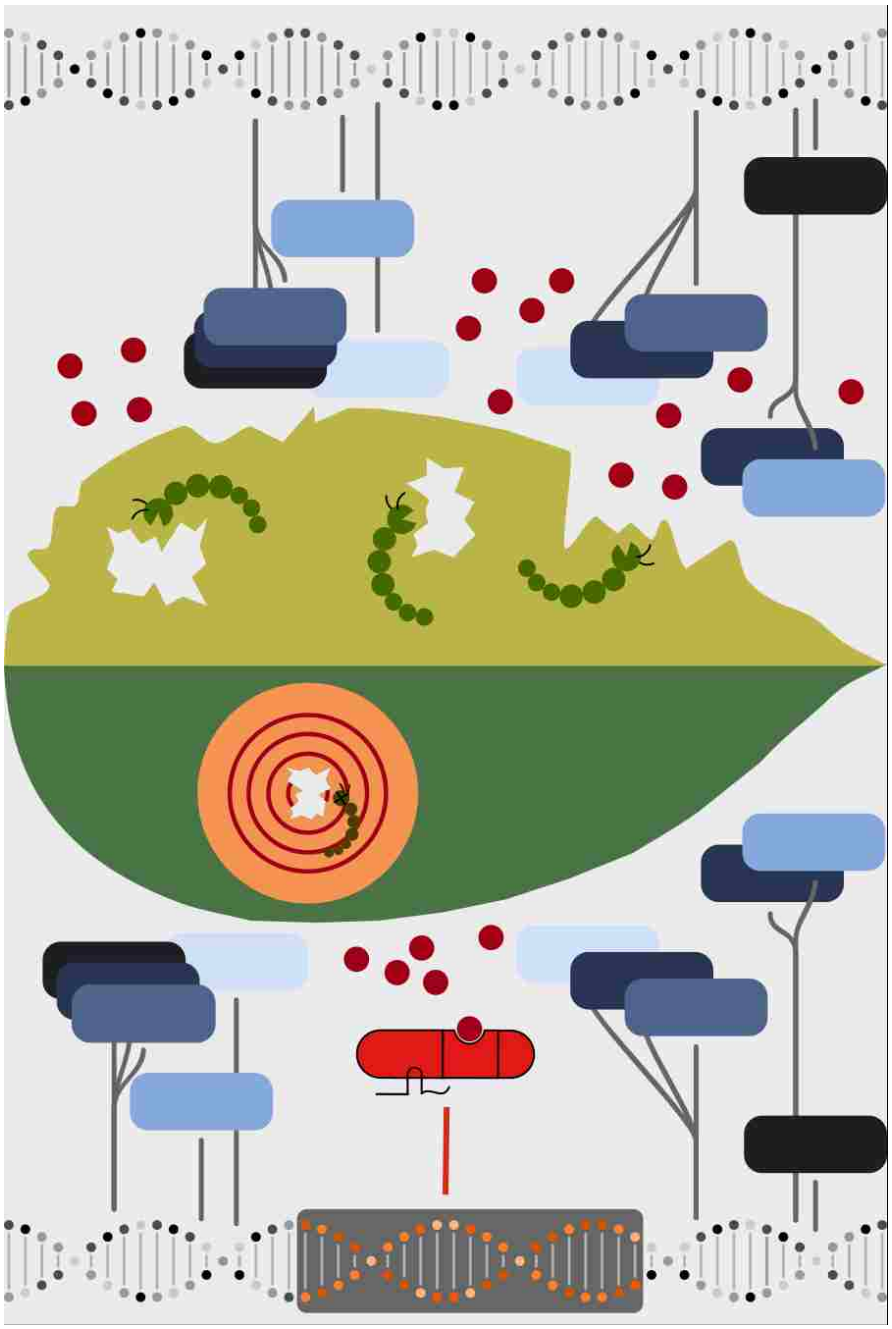


Figure 3-figure supplement 5: The shoot architecture phenotypes generated by targeting a HACR to regulate PIN1 are not due to antibiotic selection and can be observed in multiple different lines and Auxin HACR backgrounds. A) Schematic of how we reduced the transcriptional feedback strength by using a HACR to regulate PIN1. B) Box plot summarizing data collected from T3 plants of lines that have either just a HACR (grey box plot) or T3 plants of that same HACR background transformed with a gRNA targeting PIN1. These plants were not selected using antibiotic, but rather genotyped for the presence of the HACR and the gRNA by PCR after they were phenotyped. All data reported here is from plants that passed genotyping. We observe the same trends as before with a lower number of branches in lines with the PIN1 gRNA as compared to the control. There was a significant decrease for lines 4 and 5 ($p<0.05$). C) Box plots summarizing the number of branches observed on T2 plants from lines in a third auxin HACR background. These plants were also selected with genotyping post phenotyping rather than with antibiotics. We observe a significant decrease in the number of branches in the line that has the gRNA (blue box plot), as compared to the control (grey box plot) ($p<0.05$).

Chapter 6 - Engineering the flow of information from one organism to another: "Pest triggered immunity: Implementing synthetic hormone-based signal transduction to regulate insect defenses in plants"



Abstract

Insect pests present one of the most pressing threats to crop plants, destroying approximately 15% of the world's crops, with some crops such as cotton experiencing losses of up to 80%. This translates to massive economic losses both globally and specifically in the United States. Insect related crop losses are also a serious threat to food security in a world where a 40% increase in crop yields is needed by 2050 to match the current rate of population growth. One strategy that has had major success in preventing these losses is the constitutive expression of insecticidal toxins such as *Bacillus thuringiensis* endotoxin (BT). However, this strategy been shown to have detrimental effects on crop yield due to continuous diversion of resources to the production of toxins, even when they are not required. Additionally, the ubiquitous expression of these toxins has led to a more rapid development of BT resistant pests, which continuously necessitates the development of new toxins. We propose a strategy to address these draw backs by linking the expression of insect resistance mechanisms to the endogenous hormonal cues associated with insect attack using synthetic phytohormone-based signal transduction machinery. By fusing hormone regulated degradation motifs to CRIPSR-based transcription factors we propose to create of a platform for pest triggered immunity that can be ported from model systems to crops. The inducible nature of our proposed system has the potential to increase yields as well slow down the development of insecticide resistance in pests, resulting in safer and more productive crops.

Introduction

Insect pests cause an enormous amount of damage to crops worldwide, accounting for, on average, a 15% loss of yield across all crops[1] with some such as cotton experiencing severe damage of up to 80% loss of yields[2]. These losses translate to a huge economic impact on countries such as Brazil, which lost approximately $18 billion to crop damage by insects in 2014[3], or the US, where insect damage to just maize yields resulted in a loss $800 million dollars in 2013[4]. In the constant battle against agricultural pests[2], one of our most effective weapons to date has been plants engineered to synthesize insecticidal proteins. This is especially true in developing countries[3]. However, while this technology has been transformative to agriculture, it has major drawbacks associated with the constitutive expression of the insecticidal proteins. These include environmental contamination with these insecticidal proteins through pollen produced by the engineered plants[5], and the consequent development of resistance in pests[6] exposed to sub-lethal doses. Chillcutt et al demonstrated that pollen from BT maize contaminates nearby refuges and results in hybrid maize that has a significantly decreased BT concentration. Low concentration BT maize has been identified as the source of the BT resistant western corn rootworm outbreak in Iowa by Gassmann et al. If this sort of uncontrolled herbivory by insecticide resistant pests becomes more widespread it will be devastating to world food supplies.

Additionally, the constant expression of these proteins in every tissue comes at the cost of a decrease in optimal yield[7,8,9,10] from the plant due to a constant diversion of resources. There is a large body of work that has demonstrated the detrimental effect of constitutive expression of plant defense genes. Notably, Tian et al. and Purrington et al. demonstrated a 9% and 26% reduction in seed yield respectively in field trials of *A. thaliana* constitutively expressing different resistance genes. In a similar side by side trial conducted by Halfhill et al. of *Brassica Rapa* constitutively expressing BT, a decrease in vegetative dry weight of 33% was observed as compared to a wild type control. With an ever increasing world population, it is projected that a 40% increase in crop yields will be necessary to meet food demands by 2050[1], making maximizing the crop yields very important for food security.

The work proposed here seeks to address these problems by linking the plants perception of pests, in the form of insect herbivory associated hormonal profiles, to the expression of defense genes creating a pest triggered immunity system. We plan to build this synthetic insect triggered defense mechanism in plants in a modular manner by fusing phytohormone regulated degradation motifs to CRISPR-based transcription factors (CTFs) that regulate the expression of defense mechanisms. CTFs have recently become a widely used technology in a variety of organisms, including *Arabidopsis thaliana*[11,12,13,14]. The dCas9 protein in complex with a guide RNA (gRNA) acts as an easily reprogrammable DNA binding domain. To make these transcription factors phytohormone regulated we will fuse phytohormone regulated degradation domains (degrons) to the dCas9 protein The degrons we propose to use are the jasmonate isoleucine (JA) responsive degrons from the JAZ protein family[15]. The JA degrons function by forming a complex with an adapter F-Box protein (COI1) that recruits the E3-ubiquitin ligase machinery in the presence of JA, leading to ubiquination and subsequent degradation of the degron tagged protein[15]. Previous work has demonstrated that CTFs fused to auxin (a plant growth hormone) responsive degrons co-expressed with the relevant auxin sensitive F-Box protein allow auxin signal transduction to be recapitulated in *Saccharomyces cerevisiae* and *Arabidopsis thaliana*.

We used these phytohormone degradable TFs to link the flux of phytohormones produced within minutes of an insect attack on a plant to the expression of defense mechanisms. The relative concentration of these different hormones tends to be plant and pest specific, with different kinds of pests, such as phloem feeders like aphids versus leaf chewers like caterpillars, eliciting different hormonal profiles[16]. However, certain trends are conserved across most plants. It has been widely shown that mechanical damage to tissues by pests causes a systemic JA response in the plant within minutes, starting from the point of attack[17,18]. Previous attempts to link the plants perception of insect herbivory to the expression of defense responses have focused on designing synthetic promoters with cis regulatory elements that are thought to activate transcription upon biotic stress. This approach is not ideal as we still do not understand the architecture of plant promoters well. Additionally, stress response pathways in plants tend to be heavily cross regulated resulting in non-specific activation of defense responses. As a result, it remains difficult to use this approach to design stress response systems that function robustly in the field.

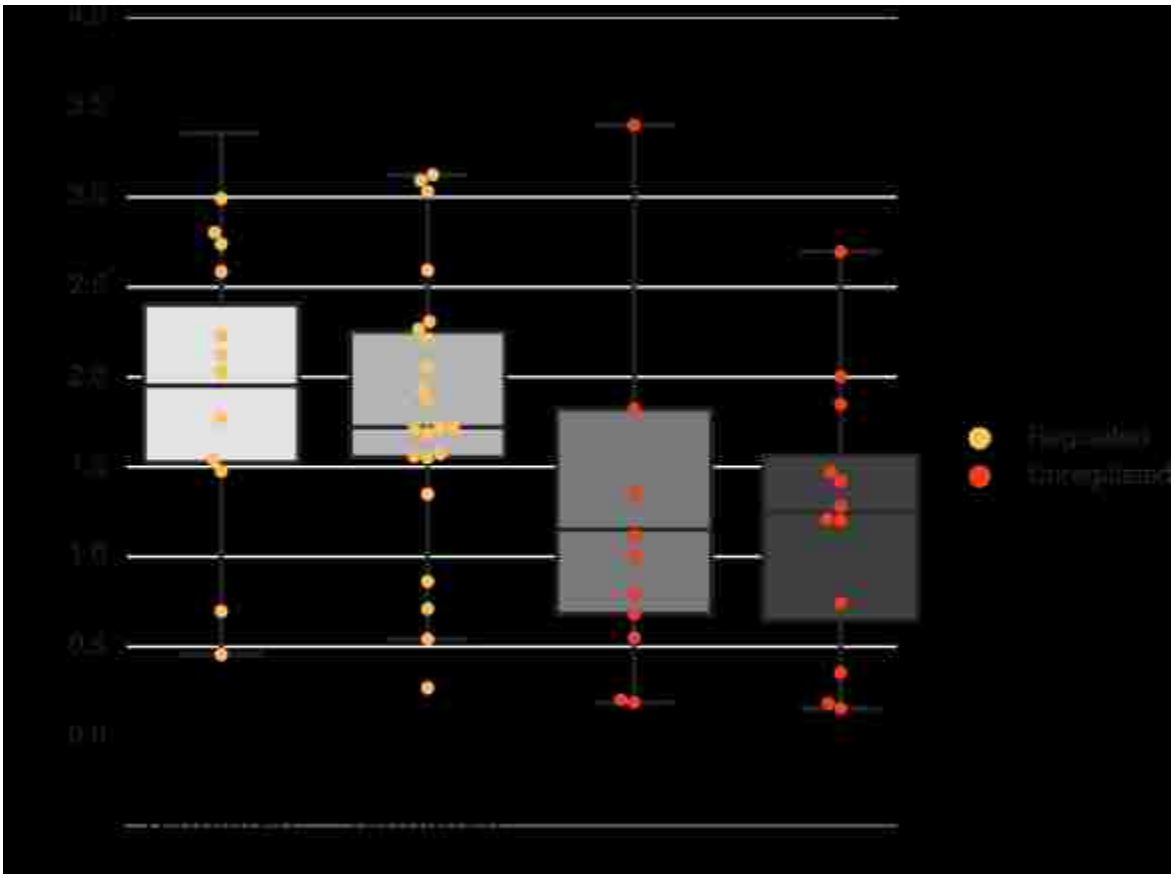
We plan to address this problem by using our synthetic signal transduction system to link the pest induced jasmonate response to the production of a BT, Cry1Ab, which has been shown to be very effective against lepidopteran pests[19]. The synthetic nature of our system will potentially insulate it against cross regulation by other plant pathways, due to the absence of multiple regulation motifs that are typically present on native transcription factors. Easy retargeting of the CTFs to regulate arbitrary promoters and ability of phytohormone degrons to perceive insect associated JA fluxes that are largely conserved among different plants will make our proposed system easy to port from model systems to relevant crop plants. While our transcription factor would need to be constitutively expressed it would need to be at a much lower concentration than BT to be effective so we still expect a yield benefit.

By spatially and temporally restricting the expression of the defense genes to the plant being attacked by pests we would minimize diversion of resources and potentially maximize yield. This system could dramatically reduce the environmental contamination with insecticidal proteins and thereby potentially slow the development of resistance[30]. Thus, this technology would represent a major step forward in agricultural pest management as it has the potential to improve food security by slowing the development of resistance in pests as well as improving yields.

Results

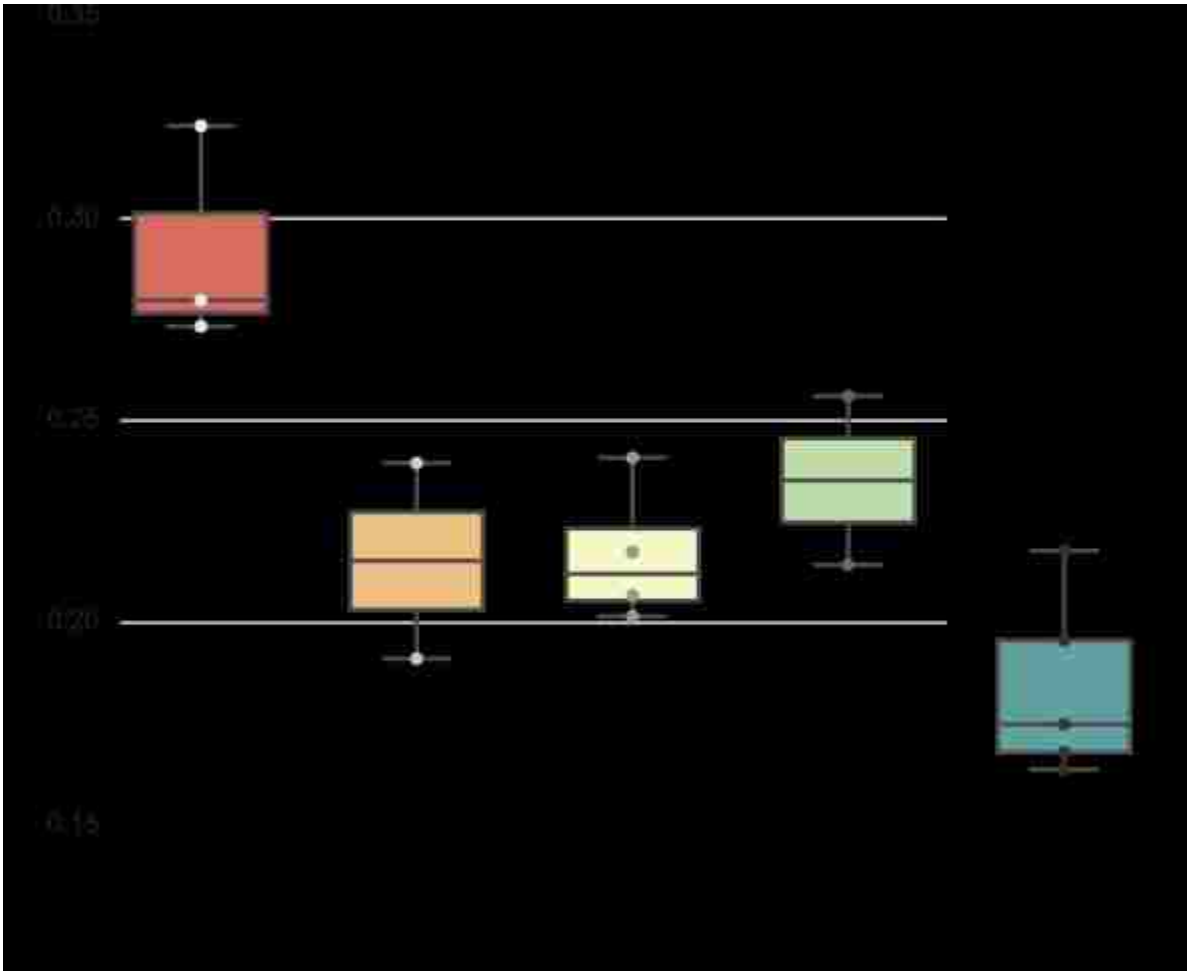
Using a JA-HACR to regulate Cry1AC reduces the metabolic load of Cry1AC expression and leads to increased biomass accumulation

Previous field experiments in maize and other plants have characterized a metabolic load cost associated with the expression of BT in plants. We hypothesize that by using a JA-HACR to conditionally repress BT expression we would be able to reduce this metabolic load effect on the plants. To test this we designed transgenic lines of *Arabidopsis thaliana* that have a HACR which can respond to the phytohormone jasmonate iso-lucine (JA-HACR) targeted to regulate a *Bacillus thuringiensis* endotoxin gene, Cry1AC, being constitutively expressed from the UBQ1 promoter. We also generated transgenic lines that had the Cry1AC gene being driven by a version of the UBQ1 promoter with its gRNA target site mutated. This would prevent the JA-HACR from regulating Cry1AC and thus give us the capacity to compare the biomass accumulation for both JA-HACR regulated and unregulated expression of Cry1AC. T2 plants from these lines were grown in parallel till a week after bolting and then weighed to characterize biomass accumulation. We observed that the lines which had a regulated copy of Cry1AC did result in increased biomass accumulation as compared to unregulated lines.



Plants with JA-HACR regulated Cry1AC are as effective at pest defense as plants with constitutively expressed Cry1AC.

While the plant lines with a regulated copy of Cry1AC have an improved capacity to accumulate biomass thanks, it is possible that this could come at the cost of a reduced capacity to defend the plants against pests. This might be due to the delay associated with transcription and translation associated with the JA-HACR. We compared the pest defense capacity of the regulated and unregulated lines described previously using the herbivore *Manduca sexta*. As *M. sexta* is not a *A. thaliana* specialist we also grew lines that had the JA-HACR but lacked the Cry1AC gene to act as a control for the native glycosylate based defenses of *A. thaliana*. We grew these lines to the rosette stage in parallel and then added the same number of *M. sexta* larvae at the third instar stage to each pot. The larvae were allowed to feed on the leaves for five days after which they were harvested and weighed. The reduction in weight compared to the control line without the Cry1AC gene is a metric for the pest defense capacity of a plant line. We observed that there was an approximately similar reduction in insect weight post feeding compared to the control for the regulated and unregulated lines. This demonstrates that the regulated lines that have a metabolic load advantage are as effective at pest defense as the unregulated lines.



References

1. Maxmen, A. Crop pests: Under attack. *Nature* (2013).

2. Oerke, EC. Crop losses to pests. The Journal of Agricultural Science (2006).
3. Oliveira, CM, Auad, AM, Mendes, SM & Frizzas, MR. Crop losses and the economic impact of insect pests on Brazilian agriculture. *Crop Protection* **56,** 50–54 (2014).
4. Implementing agriculture for development : World Bank Group agriculture action plan (2013-2015)
5. Chilcutt, C. & Tabashnik, B. Contamination of refuges by Bacillus thuringiensis toxin genes from transgenic maize. Proceedings of the National Academy of Sciences of the United States of America 101, 7526–7529 (2004).
6. Gassmann, A. J. et al. Field-evolved resistance by western corn rootworm to multiple Bacillus thuringiensis toxins in transgenic maize. Proc. Natl. Acad. Sci. U.S.A. 111, 5141–6 (2014).
7. Purrington, C. & Bergelson, J. Fitness consequences of genetically engineered herbicide and antibiotic resistance in Arabidopsis thaliana. Genetics 145, 807–14 (1997).
8. Brown, J. Yield penalties of disease resistance in crops. Current Opinion in Plant Biology 5, (2002).
9. Tian, D., Traw, M., Chen, J., Kreitman, M. & Bergelson, J. Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana. Nature 423, 74–77 (2003).
10. Halfhill, M. D. *et al.* Growth, productivity, and competitiveness of introgressed weedy Brassica rapa hybrids selected for the presence of Bt cry1Ac and gfp transgenes.*Molecular Ecology* **14,** 3177–3189 (2005).
11. Gilbert, L. et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell 154, (2013).
12. Farzadfard, F., Perli, S. & Lu, T. Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. ACS Synthetic Biology 2, 604–613 (2013).
13. Qi, L. et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152, (2013).
14. Piatek, A. et al. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. Plant Biotechnology Journal (2014).
15. Browse, J. Jasmonate passes muster: a receptor and targets for the defense hormone. Annual review of plant biology 60, 183–205 (2009).
16. Joo, S., Liu, Y., Lueth, A. & Zhang, S. MAPK phosphorylation induced stabilization of ACS6 protein is mediated by the non catalytic C terminal domain, which also contains the cis determinant for rapid degradation by the 26S proteasome pathway. The Plant Journal 54, 129–140 (2008).
17. Herde, M., Koo, A. J. & Howe, G. A. Elicitation of jasmonate-mediated defense responses by mechanical wounding and insect herbivory. 51–61 (2013).
18. Erb, M., Meldau, S. & Howe, G. Role of phytohormones in insect-specific plant reactions. Trends in Plant Science 17, (2012).
19. Cao, J, Tang, JD, Strizhov, N, Shelton, AM & Earle, ED. Transgenic broccoli with high levels of Bacillus thuringiensis Cry1C protein control diamondback moth larvae resistant to Cry1A or Cry1C. Molecular Breeding (1999).

Concluding thoughts and future work

In conclusion, this body of work demonstrates one key idea: the behavior and development of multicellular organism can be reprogramed by engineering the way cells in that organism sense, interpret and respond to signals from their environment. My graduate work both provides a frame work for how new signal transduction systems could be engineered in the future, such as new hormone responsive HACRs or synthetic promoters. In addition, the novel signaling systems presented here provide a broad set of tools to both study and re-engineer agriculturally and industrially relevant phenotypes in plants.

We have only scratched the surface of the kinds of systems that could be built with them in chapters five and six and I hope to build on this work in the future. During my post-doctoral fellowship at the University of Minnesota I hope to translate the HACR platform into crop plants to create crops that are pre-engineered with the infrastructure for epigenetic reprogramming. In parallel hope to design plant virus-based vectors that could systemically deliver RNA scaffolds to these crops that could direct the epigenetic reprogramming by deploying the HACRs. This would allow crop plants to be dynamically reengineered in the field with new traits to meet the ever-changing needs of a rapidly changing agricultural environment. It also has the potential to decentralize the power currently held by big agricultural companies, by empowering farmers to create new germplasm as required, thereby leveraging their in-depth knowledge of their local environment, which is generally not leveraged. In this way, I believe this work takes us one step further in the journey to leverage biology to create a more plentiful world.

~Fin~